Multiple Sclerosis Biomarkers – moving towards clinical applications

by Margareta Anna Clarke, BSc

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

September 2019

DECLARATION

I, Margareta Anna Clarke, confirm that this work is my own. Where work and help has been contributed by my colleagues and from other sources, I indicate this in the specific chapter and acknowledgements sections.

ACKNOWLEDGEMENTS

I would like to acknowledge many individuals without whom, it would not have been possible to carry out the work described in this thesis.

First and foremost, I would like to thank my supervisors, Dr. Nikos Evangelou and Dr. Alain Pitiot for their guidance and support. Thank you for making this experience as enjoyable as it could be, for your knowledge and good sense of humour throughout! Special thanks also go out to my internal examiner, Dr. Jan Derrfuss, for his critical review of the initial reports of the findings presented in this thesis and his practical advice.

I would also like to thank all my colleagues from the Department of Clinical Neurology at the University of Nottingham. Specifically, I would like to thank Dr. Amal Samaraweera, who first trained me to spot the central vein sign and patiently answered all my questions, Dr. Yasser Falah, Dr. Chris Allen, Joe Paul for interesting discussions and useful feedback. I will miss our tea breaks!

I would also like to express my thanks to Dr. Chris Tench, for his solid statistical advice, Dr. Paul S Morgan and Dr. Rob Dineen for their help with various aspects of this project and invaluable input with regards to the practical implications of the findings reported in this thesis.

Most of the work described in this thesis would not have been possible without collaborations. I would like to thank Dr. Jens Wuerfel and Dr. Tim Sinnecker for

being great collaborators on the MAGNIMS project and showing me the beauty of Switzerland. I would also like to thank Dr. Alex Rovira, Dr. Deborah Pareto and the rest of the Barcelona team for their invaluable support, training and an opportunity to experience your sunny city.

Special thanks go to Graham White and the Neuroscience Support Group, who provided funding for my PhD. Your financial support was invaluable to me.

Lastly, I would like to thank my family for always supporting me in this endeavor. In particular, I would like to thank my husband, Reece, who was with me through all the lows and celebrated all the ups of this PhD.

ABSTRACT

Background: Previously proposed multiple sclerosis (MS) biomarkers, including the central vein sign (CVS), which can help differentiate between MS and non-MS, and brain atrophy, which has been shown to predict disability, hold the promise of better diagnosis and monitoring of MS patients. However, findings from research MRI studies need further validation on conventional hospital MRI, before they can be incorporated into clinical practice.

Methods: Using 3T MRI, I tested the diagnostic usefulness of the CVS in three studies: (1) a large, cross-sectional, multi-centre study of patients with MS and mimicking disorders, (2) a prospective single-centre study of patients with atypical first presentation of a clinically isolated syndrome (CIS) and (3) a prospective single-centre study of patients with a typical CIS and MS mimicking disorders. I also explored the utility of routine, non-standardised MRI scans in atrophy estimation and proposed a novel “Frankenstein” approach for atrophy estimation in scans with missing slices.

Results: The CVS achieved high sensitivity (up to 100%) and specificity (up to 90%) in diagnosing MS against its imaging mimics and in predicting a final diagnosis of MS in patients presenting with both typical and atypical CIS. Moreover, the detection of the CVS was largely independent of the MR sequence or scanner used. Brain atrophy estimations from suboptimal clinical MRI scans improved with the “Frankenstein” approach, increasing successful processing rate from 0% to 59%.

Conclusions: The CVS appears to be a highly sensitive and specific marker of MS lesions at 3T and its diagnostic value, as a stand-alone test and in combination with the current diagnostic criteria, should be assessed in large, prospective, multi-centre studies. Assessment of brain atrophy remains challenging in clinical practice. The "Frankenstein" approach appears promising in reducing atrophy estimation failures but future studies using a larger cohort of patients are required to further validate its performance.

TABLE OF CONTENTS

Declaration **1**

Acknowledgements **2**

Abstract **4**

Table of Contents **6**

Table of Figures **12**

Table of Tables **14**

Abbreviations **15**

Chapter 1: Introduction **18**

1.1 An overview of multiple sclerosis 18

1.2 Motivation for the work carried out 21

1.2.1 WM lesions *22*

1.2.2 Brain atrophy *22*

1.3 Thesis aims 23

1.4 Peer-reviewed publications related to this thesis 25

1.5 Poster Presentations related to this thesis 25

1.6 Other contributions during this PhD 26

1.7 Awards 28

1.8 Thesis outline 29

Chapter 2: Literature review **30**

2.1 Multiple Sclerosis 30

2.1.1 Clinical course *30*

2.1.2 Diagnosis *32*

2.2 The Pathology of MS 35

2.2.1 White matter lesions *35*
2.2.1.1 Classification and formation 35
2.2.1.2 Perivenous nature of WM lesions 40
2.2.1.3 Correlation with clinical parameters 45
2.2.2 Brain atrophy *46*
2.2.2.1 Brain atrophy in healthy ageing and MS 46
2.2.2.2 Pathogenesis of brain atrophy 49
2.2.2.3 Brain atrophy: regional or global? 54
2.2.2.4 Automated assessment of brain atrophy 58
2.2.2.5 Factors which can affect atrophy assessment 64
2.3 IMAGING MS PATHOLOGY 71
2.3.1 Ultra-high field MRI *73*
2.3.2 Clinical MRI *75*
2.3.3 Patient registries *77*

Chapter 3: The Central Vein Sign in MS and mimicking Disorders; A multi-centre MAGNIMS study 81

3.1 INTRODUCTION 82
3.2 AIMS 84
3.3 METHODS 84
3.3.1 Ethical approval *84*
3.3.2 Patient sample *85*
3.3.3 MRI acquisition *85*
3.3.4 Image pre-processing *87*
3.3.5 Lesion analysis *87*
3.3.6 Statistics *90*
3.4 RESULTS 93
3.4.1 Patient sample *93*
3.4.2 WM lesion and CV analysis *94*
3.4.3 Comparison between WM lesion and CV numbers in different disorders *98*

3.4.4 Simplifying the central vein diagnostic rule 98

3.4.5 Agreement between the raters 103

3.4.6 Relationship between MS patient clinical characteristics and the proportion of WM lesions with CVs 103

3.4.7 Effect of scanning parameters on diagnostic accuracy 103

3.5 DISCUSSION 104

3.5.1 The central vein sign as a diagnostic marker of MS 104

3.5.2 Simplifying the rule 105

3.5.3 The central vein sign in MS and other disorders 107

3.5.4 Inter-rater reliability 108

3.5.5 Clinical and MRI effects on the detectability of CVs 109

3.5.6 Comparison of the results 110

3.5.7 Study strengths 111

3.5.8 Study limitations 111

3.5.9 Conclusion 112

Chapter 4: STAR-MS: Single Test to ARrive at MS diagnosis; a prospective pilot study assessing the accuracy of the central vein sign in predicting multiple sclerosis in cases of diagnostic uncertainty 113

4.1 INTRODUCTION 113

4.2 AIMS 115

4.3 METHODS 115

4.3.1 Patient sample 116

4.3.2 MRI acquisition 116

4.3.3 Lesion analysis 117

4.3.4 Establishing the diagnoses 117

4.3.5 Statistics 118

4.4 RESULTS 118

4.4.1 Demographics 118

4.4.2 Lesion analysis 120

4.4.3 Agreement between the diagnoses 120

4.4.4 Rate of patient recruitment 126

4.5 DISCUSSION 128

4.5.1 Usefulness of the CV sign in predicting MS in patients with atypical presentation 128

4.5.2 Comparison of the CV sign to OCBs testing 130

4.5.3 Inter-rater agreement 131

4.5.4 Rate of patient recruitment 132

4.5.5 Study strengths 132

4.5.6 Study limitations 132

4.5.7 Conclusion 133

Chapter 5: Diagnosing Multiple Sclerosis Using a Single, 3T Susceptibility-Weighted MRI Scan 135

5.1 INTRODUCTION 135

5.2 AIMS 137

5.3 METHODS 137

5.3.1 Ethics 137

5.3.2 Patient sample 137

5.3.3 IgG oligoclonal bands 139

5.3.4 MRI acquisition 139

5.3.5 Image registration 139

5.3.6 Lesion analysis 139

5.3.6.1 Part 1: Incidence of CVs and PIRs in CIS and non-MS patients 141

5.3.7 Diagnostic performance of central veins and iron rims 141

5.3.7.1 Part 2: Diagnostic value of CVs and PIRs in predicting conversion to MS according to the 2017 McDonald criteria 141

5.3.7.2 Part 3: DIS+ analysis 141

5.3.8 Statistics 142

5.4 RESULTS 143
5.4.1 Patient demographics 143
5.4.2 Lesion analysis 143
5.4.2.1 Part 1: Incidence of CVs and PIRs in CIS and non-MS patients 147
5.4.3 Diagnostic performance of central veins and iron rims 151
5.4.3.1 Part 2: Diagnostic value of CVs and PIRs in predicting conversion to MS according to the 2017 McDonald criteria 151
5.4.3.2 Part 3: DIS+ analysis 151
5.5 DISCUSSION 153
5.5.1 Diagnostic utility of the central vein sign and iron rims in CIS 153
5.5.2 Usefulness of SWI in MS diagnosis 153
5.5.3 Alternatives to DIT 154
5.5.4 Study strengths 155
5.5.5 Study limitations 156
5.5.6 Conclusions 156

Chapter 6: Brain Atrophy Estimation from Routine, Non-Standardised, Clinical Scans with Incomplete Brain Coverage – a Method Validation 158

6.1 INTRODUCTION 158
6.2 AIMS 160
6.3 METHODS 160
6.3.1 Ethical consent 160
6.3.2 Patient sample 161
6.3.3 Audit of clinical MRI protocols 161
6.3.4 Research MRI protocols 162
6.3.5 Atrophy estimation from clinical scans with missing slices and research scans 162
6.3.6 The Frankenstein method 165
6.3.7 Frankenstein validation 168
6.3.8 Statistics 168

6.4 RESULTS 169
6.4.1 Variability of clinical protocols 169
6.4.2 Patient demographics 169
6.4.3 Atrophy estimation in incompletely covered clinical scans 172
6.4.4 Frankenstein validation 172
6.4.5 Atrophy estimation from "Frankenstein" scans 175
6.4.6 Atrophy estimation from research scans 175
6.5 DISCUSSION 180
6.5.1 Assessment of the variability of clinical protocols in the East Midlands 180
6.5.2 The Frankenstein approach 181
6.5.3 Previous studies of atrophy assessment in the clinic 182
6.5.4 Atrophy estimation failures using high resolution research scans 183
6.5.5 Study limitations 184
6.5.6 Future analyses 185
6.5.7 Conclusions 186
Chapter 7: Thesis conclusions 187
7.1 THE USEFULNESS OF THE CENTRAL VEIN SIGN IN THE DIAGNOSIS OF MULTIPLE SCLEROSIS 187
7.2 ATROPHY ESTIMATION FROM ROUTINE, NON-STANDARDISED, CLINICAL MRI 190
7.3 FUTURE RESEARCH 192
Appendix A 194
Appendix B 197
Appendix C 200
Appendix D 202
Appendix E 203
References 210

TABLE OF FIGURES

Figure 2-1 Dawson fingers as seen on sagittal FLAIR MRI. 37

Figure 2-2 MS lesions (outlined in colour) visible on a FLAIR scan (left) and their corresponding appearance on a T2*-weighted 3T MRI scan (right). 42

Figure 2-3 Clinical scans of an MS patient taken seven years apart demonstrate brain atrophy. 48

Figure 3-1 The parcelation of the brain into eight equally-sized blocks. 88

Figure 3-2 Images used to identify WM lesions and CVs. 89

Figure 3-3 Location-specific proportion of WM lesions with CVs in non-MS and CIS/MS individuals expressed as percentage of al the lesions included in the study. 96

Figure 3-4 Patientwise proportion of WM lesions with CVs for each disease group. 97

Figure 3-5 The receiver operator characteristics (ROC) curve (black) showing the sensitivity plotted against 100-specificity for the CVS to differentiate between MS and non-MS 100

Figure 4-1 Patient flow diagram comparing the diagnoses made using the central vein sign (index test) and McDonald criteria (reference standard)..... 119

Figure 4-2 Typical appearance of lesions in MS (left) and non-MS (right). 124

Figure 4-3 Distribution of the index test results (percentage of WM lesions with CVs), assessed by the two raters, as compared to the reference standard (clinical diagnosis). 125

Figure 4-4 The ROC curve (blue) showing the diagnostic accuracy of the central vein sign in our study. 127

Figure 5-1 Patient flow diagram showing how patients were selected for the study and reasons for exclusion. 144
Figure 5-2 Lesion appearance on axial FLAIR (top) and the corresponding SWI (bottom) images in a patient who was subsequently diagnosed with Sjögren disease (A and C) and MS (B and D). 146
Figure 5-3 The location of lesions with central veins reported separately for the CIS and non-MS groups 148
Figure 5-4 Summary of the incidence of lesions with central veins in the CIS and non-MS groups. 149
Figure 5-5 Summary of the lesions with iron rings in the CIS and non-MS groups. 150
Figure 6-1 An example 2D hospital scan taken in the coronal plane. 163
Figure 6-2 3D MPRAGE MRI scans with low (left), medium (middle) and high (right) atrophy levels. 166
Figure 6-3 A “Frankenstein” brain MR 167
Figure 6-4 Patient flow diagram 171
Figure 6-5 Extracted brain tissue mask (orange) from an incompletely acquired clinical MRI scan. 173
Figure 6-6 Improvement in brain extraction using SIENA, following the application of the Frankenstein approach (top row), compared to to pre-Frankenstein results (bottom row). 176
Figure 6-7 Issues with brain extraction from standardised, 7T T1-weighted scans. 179

TABLE OF TABLES

Table 3-1 The MR scanning parameters for the vein (T2*/SWI/SWAN) and lesion sequences used by the eight MAGNIMS centres. 86
Table 3-2 Lesion-wise summary classified by disease group 95
Table 3-3 Sensitivity, specificity, positive and negative LR for different thresholds of WM lesions with CVs 99
Table 3-4 Sensitivity and specificity values for a range of simplified central vein sign diagnostic criteria using a subset of randomly picked WM lesions 101
Table 3-5 Sensitivity and specificity values for a range of simplified CV criteria using a fixed number of lesions with CVs 102
Table 3-6 Sensitivity and specificity values for a range of simplified CV criteria using a fixed number of subcortical WM lesions 102
Table 4-1 Cross-tabulation of results from the CV assessment and clinical diagnoses (sorted by %WMLs with CVs) with the results of OCBs for comparison. 121
Table 5-1 Demographic and clinical data of the patients included in the study 145
Table 5-2 The diagnostic properties and hazard ratios for radiological and clinical conversion to MS over time 152
Table 6-1 Protocol information for the T1-weighted research scans 164
Table 6-2 Acquisition parameters for selected common sequences in a sample of 78 patients from the Nottingham MS Clinic. 170
Table 6-3 Results of the Frankenstein validation 174
Table 6-4 Results from atrophy estimation using SIENA following the application of the Frankenstein approach 177

ABBREVIATIONS

ACTRIMS	American Committee for Treatment and Research in Multiple Sclerosis
ANOVA	Analysis Of Variance
AUC	Area Under the Curve
BBB	Brain Blood Barrier
BPF	Brain Parenchymal Fraction
BVL	Brain Volume Loss
CDMS	Clinically Defined Multiple Sclerosis
CI	Confidence Interval
CIS	Clinically Isolated Syndrome
CMSC	Consortium of MS Centers
CNS	Central Nervous System
CSF	Cerebrospinal Fluid
CT	Computer Tomography
CV	Central Vein
DIR	Double Inversion Recovery
DIS	Dissemination In Space
DIT	Dissemination In Time
DMT	Disease Modifying Therapy
EAE	Experimental Autoimmune Encephalomyelitis
ECTRIMS	European Committee for Treatment and Research in Multiple Sclerosis
EDSS	Expanded Disability Status Scale
EPI	Echo Planar Imaging
FDA	Food and Drug Administration
FLAIR	Fluid-Attenuated Inversion Recovery
FLIRT	FMRIB's Linear Image Registration Tool
FSL	FMRIB Software Library
GM	Grey Matter
GRE	Gradient-Echo
IR	inversion Recovery

LASSO	Least Absolute Shrinkage and Selection Operator
MAGNIMS	Magnetic Resonance Imaging In Multiple Sclerosis
MPRAGE	Magnetisation-Prepared Rapid Gradient-Echo
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
MS	Multiple Sclerosis
MSSS	Multiple Sclerosis Severity Scale
MTR	Magnetisation Transfer Ratio
NEDA	No Evidence Of Disease Activity
NMOSD	Neuromyelitis Optica Spectrum Disorder
NPV	Negative Predictive Value
OCB	Oligoclonal Band
PBVC	Percentage Brain Volume Change
PET	Positron Emission Tomography
PP	Primary Progressive
PPV	Positive Predictive Value
PIR	Paramagnetic Iron Rim
PSIR	Phase-Sensitive Inversion Recovery
RF	Radio-frequency
RIS	Radiologically Isolated Syndrome
ROC	Receiver Operating Characteristic
ROI	Region Of Interest
RR	Relapsing-Remitting
SAD	Systemic Autoimmune Disease
SAR	Specific Absorption Rate
SD	Standard Deviation
SDMT	Single Digit Modality Test
SIENA	Structural Image Evaluation, using Normalisation, of Atrophy
SLE	Systemic Lupus Erythematosus
SNR	Signal-to-Noise Ratio
SP	Secondary Progressive
SPM	Statistical Parametric Mapping
STAR-MS	Single Test To Arrive At Multiple Sclerosis Diagnosis
SWAN	Susceptibility-weighted Angiography

SWI	Susceptibility-weighted Imaging
T	Tesla
TE	Echo Time
TI	Inversion Time
TIRM	Turbo Inversion Recovery Magnitude
TR	Repetition Time
UHF	Ultra-High Field
WM	White Matter

CHAPTER 1: INTRODUCTION

1.1 An overview of multiple sclerosis

Multiple sclerosis (MS) is considered a multifactorial disorder caused by a combination of genetic and environmental factors. It leads to chronic demyelination of the central nervous system (CNS) which primarily affects young adults (Rosati, 2001). The condition is typically diagnosed between the ages of 20 and 50 and there are an estimated 127,000 people with MS living in the UK, leaving approximately one in 500 with an MS diagnosis (Mackenzie et al., 2014). Worldwide, an estimated 2,500,000 people are affected by MS (Compston and Coles, 2002) with women being two to three times more likely to develop the disorder, and current epidemiological evidence suggests that both the gender gap and frequency of MS keep increasing (Pugliatti et al., 2002, Trojano et al., 2012, Kingwell et al., 2013) .

MS can lead to a variety of symptoms, depending on the CNS site affected. As a result, there is vast heterogeneity in the clinical manifestation of the disease between patients. Some of the most common symptoms are sensations of tingling, unsteadiness and problems with balance, motor weakness, vision loss, bladder dysfunction, or a mixture of them (Lublin, 2005). Most MS patients experience alternating episodes of disease-free periods followed by attacks; this is termed relapsing-remitting MS (Goldenberg, 2012). A small proportion of patients will experience a steady decrease in neurological function with no (or few) attacks; this is termed progressive MS (Dutta and Trapp, 2014).

As MS is typically diagnosed in young adults, it has a significant socio-economic impact with a mean annual cost per person ranging from 11,400 GBP in patients with little or no disability, to 36,500 GBP for the most disabled patients (Thompson et al., 2017). According to data from the MS Register, the largest UK database of people with MS, only 42% of all MS individuals remain in employment up to four years following their diagnosis, with just over a quarter still being employed five to nine years post-diagnosis (Ford et al., 2012). In comparison, over 75% of the UK's population aged over 16 are considered economically active (Office for National Statistics, 2018). The high unemployment rate among MS patients has been found to be predicted by symptoms worsening, including problems with mobility, fatigue and cognition (Julian et al., 2008). This demonstrates the need for a deeper understanding of MS to improve the quality of life of people affected by it.

MS takes its name from the latin *sclerosis multiplex* meaning 'multiple scars'. It is a reference to the demyelinating lesions in the white matter (WM) of the CNS, a hallmark of the disease first described in the earliest depiction of MS in 19th century (Charcot, 1868 cited in Zalc, 2018). The appearance of new WM lesions on magnetic resonance imaging (MRI) can indicate new, on-going inflammation (Gadolinium-enhancing and T_2 lesions) and permanent tissue destruction (T_1 black holes). Due to the high sensitivity of conventional MRI in their detection, WM lesion numbers and/or volume have long been used as an outcome measure in clinical trials. More recently, they have also become one of the components of 'no evidence of disease activity' (NEDA), a treatment goal in MS characterised by no new or enlarging T_2 or Gad-enhancing lesions, as well as no relapses or disability (measured by expanded disability status scale [EDSS]) (Havrdova et al., 2009).

MS WM lesions form preferentially around blood vessels, findings known since the earliest post mortem studies of MS (Cruveilhier, 1829-1842, Charcot, 1868, Charcot and Marie, 1886, Dow and Berglund, 1942), and, more recently, visualised using MRI (Tan et al., 2000, Sati et al., 2016). This phenomenon, termed *the central vein (CV) sign*, is currently being explored as an MS-specific diagnostic biomarker with studies suggesting nearly half of all WM lesions at 1.5 and 3T have a CV whilst hardly any WM lesions in disorders, which can mimic MS, have them on MRI (Sati et al., 2016). This is an important area of research as misdiagnosis is a significant problem in MS with an estimated one in five patients being incorrectly given a diagnosis of MS (Kaisey et al., 2019).

Apart from WM lesions, brains of people with MS are also characterised by accelerated cerebral atrophy (loss of brain tissue). Whilst brain atrophy has been known to scientists since the earliest descriptions of MS (Charcot, 1868, Charcot and Marie, 1886), it was only recently that we have been able to properly quantify it *in vivo* using imaging. Atrophy has been found to be a useful biomarker of future disability (Popescu et al., 2013b), cognitive decline (Zivadinov et al., 2001, Calabrese et al., 2009) and disease progression (transition from the relapsing form of MS to a progressive form of the disease termed secondary progressive MS) (Tomic et al., 2018) suggesting it could be used to predict the future course of the disease in patients.

1.2 Motivation for the work carried out

MS is a vastly heterogeneous disorder. Clinically, some patients will develop debilitating disability and require care, whilst others will have few attacks and continue being able to carry out the majority of their everyday activities independently (Sartori et al., 2017). Similarly, from a radiological perspective, some patients will develop many WM lesions, whilst others might have more pronounced pathology of the grey matter (GM) (Bo et al., 2003, Zarei et al., 2003). Overall, there is consensus that MS is characterised predominantly by multifocal WM lesions and accelerated cerebral atrophy that can help explain this variability in disease activity and progression between patients (Popescu et al., 2013b, Cortese et al., 2019).

Slowing, stopping or reversing the accumulation of disability has been identified as the top research priority by MS clinicians, patients and carers (MS-Society, 2012). In order to achieve this goal, clinicians need to be able to quickly and reliably diagnose MS as mounting evidence suggests that early treatment is beneficial to reducing long-term disability (Cerqueira et al., 2018). Secondly, identifying patients who have a higher risk of developing debilitating disability would be another crucial step towards reducing the burden of disability in MS. Already WM lesions are used in the diagnosis of MS and, together with brain atrophy, as outcome measures in clinical trials of new disease modifying therapies (DMTs) (Tur et al., 2018). More recently, brain atrophy has also been suggested as the most viable candidates to be incorporated into the future revisions of NEDA (Giovannoni et al., 2018), which already incorporates information about WM lesions in the assessment of patients.

In this thesis, I explore the application of routine, clinical brain MRI and its ability to reliably visualise the central vein sign and quantify brain atrophy in MS patients to provide meaningful clinically-relevant results that could be used in the diagnosis and monitoring of patients. Currently, post mortem studies, using histology, are considered gold standard in MS research whilst ultra-high field (UHF) MRI (7T and higher) provides unparalleled visualisation of MS-related, pathological changes *in vivo*. Nevertheless, in order to be of clinical value, any findings reported post mortem or using UHF MRI must be confirmed using 1.5 and/or 3T imaging routinely used in radiology departments.

1.2.1 WM lesions

WM lesions are the most well described feature of MS and are already used in the diagnosis and monitoring of patients in routine, clinical practice. In this work, I focused on the characteristic perivenular distribution of MS lesions, and further explored the role of the CV sign in MS diagnosis. Specifically, I explored the detection of the CV sign in MS and mimicking disorders using different gradient echo (GRE) protocols, used by eight different European MS centres. Moreover, I also carried out the first prospective studies of the CV sign using routinely available susceptibility imaging at 3T to determine its accuracy in diagnosing patients with both typical and atypical first presentation of MS.

1.2.2 Brain atrophy

Cerebral atrophy is considered a marker of global neurodegeneration and previous studies have found it to be a significant predictor of patient clinical outcomes. Nevertheless, it is crucial to determine if atrophy assessment is feasible in MS

clinics, where patients are often scanned using different scanners and protocols. In this thesis, I explore the utility of routine, non-standardised MRI scans in atrophy estimation and validated a novel approach for atrophy estimation using scans with missing slices.

1.3 Thesis aims

My research aims to facilitate the translation of research findings into clinical practice and test if brain atrophy and WM lesions, two imaging markers, could help better diagnose, monitor and assess MS patients. Specifically, I aim to:

1. Assess the detection of the CV sign on variable, clinical-quality MR scans, and determine its usefulness in diagnosing MS (Chapter 3)
2. Determine any effects of clinical and imaging characteristics on the detectability of the CV sign (Chapter 3)
3. Test the usefulness of the CV sign in diagnosing MS in patients with an atypical presentation of a demyelinating event and determine an ideal threshold for an MS diagnosis in this cohort (Chapter 4)
4. Test the usefulness of the CV sign in diagnosing MS in patients with a typical presentation of a demyelinating event and determine an ideal number of lesions with CVs for an MS diagnosis in this cohort (Chapter 5)
5. Test the diagnostic value of the CV sign in combination with evidence of dissemination in space (DIS), compared to the currently used DIS and dissemination in time (DIT), and DIS and positive oligoclonal bands (OCBs) (Chapter 6)

6. Examine if, and under what circumstances, variable, clinical-quality MRI scans can be used to reliably detect atrophy (Chapter 7)

Based on these aims, I hypothesised the following:

1. The CV sign will be able to successfully differentiate between MS and non-MS cases although optimised protocols will be more successful at this
2. Older age will lead to a reduction in the proportion of WM lesions with CVs, due to the appearance of age-related, non-specific WM changes, thus leading to a reduced diagnostic sensitivity of the CV sign in older populations. On the other hand, higher MRI spatial resolution will enable better differentiation between MS and non-MS cases using the CV sign
3. The CV sign, detected on a single, baseline MRI scan, will have high sensitivity and specificity in differentiating between patients with atypical first presentation, who subsequently develop MS and those who do not with
4. The CV sign, detected on a single, baseline MRI scan, will have high sensitivity and specificity compared to the 2017 McDonald criteria in patients with a typical first presentation of MS
5. In combination with evidence of DIS, the CV sign will achieve similar sensitivity and specificity compared to DIS and DIT, and DIS and positive OCBs, currently used to diagnose MS.
6. A proportion of routine, clinical scans will be able to successfully estimate atrophy, providing values in line with those obtained from standardised scans and results previously reported in literature.

1.4 Peer-reviewed publications related to this thesis

Sinnecker T, **Clarke MA**, Meier D, Enzinger C, Calabrese M, De Stefano N, Pitiot A, Giorgio A, Schoonheim MM, Friedemann P, Pawlak MA, Reinhold S, Kappos L, Montalban X, Rovira A, Evangelou N, Wuerfel. Evaluation of the Central Vein Sign as a Diagnostic Imaging Biomarker in Multiple Sclerosis. JAMA Neurol. Published online August 19, 2019. doi:10.1001/jamaneurol.2019.2478

Clarke MA, Samaraweera APR, Falah Y, Pitiot A, Allen C, Dineen RA, Tench CR, Morgan PS, Evangelou N. Single Test to ARrive at MS diagnosis (STAR MS); a prospective study assessing the accuracy of a 3T MRI scan in predicting multiple sclerosis in cases of diagnostic uncertainty. Accepted for publication in Multiple Sclerosis Journal.

1.5 Poster Presentations related to this thesis

M. Clarke, L. Pessini-Ferreira, D. Pareto, G. Arrambide, M. Alberich, F. Crescenzo, A. Garcia-Vidal, M. Tintore, C. Auger, N. Evangelou & A. Rovira. *The diagnostic value of susceptibility-weighted imaging in multiple sclerosis: the central vein sign and iron rims are specific markers of multiple sclerosis lesions*. Presented at the 35th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS) in Stockholm, September 2019.

M. Clarke, A.P.R. Samaraweera, Y. Falah, A. Pitiot, R.A. Dineen, P.S. Morgan, C.R. Tench, N. Evangelou. '*STAR MS: Single Test to ARrive at MS diagnosis; a prospective, investigator blinded pilot study assessing the accuracy of the ´Central*

Vein Sign´ using a single 3T MRI scan in diagnosing MS in cases of diagnostic uncertainty (NCT02485223)'. Presented at the 34th Congress of ECTRIMS in Berlin, October 2018.

A. Pitiot, **M. Clarke**, P.S. Morgan, N. Evangelou. '*Brain atrophy estimation from incomplete clinically-acquired MRI scans - a validation of the MS "Frankenstein" approach'*. Presented at the 34th Congress of ECTRIMS in Berlin, October 2018.

M. Clarke, T. Sinnecker, A. Pitiot, D. Meier, A. Rovira, N. Evangelou & J. Wuerfel. *'The central vein sign in multiple sclerosis: a real-life multi-centre MAGNIMS study*'. Presented at the 25th Annual Meeting of the Charcot Foundation in Baveno, November 2017.

A. Pitiot, **M. Clarke**, J.Paul, P.S. Morgan, N. Evangelou. *'Estimating brain atrophy from incompletely acquired clinical MR scans in MS*'. Presented at the 7th Joint ECTRIMS-ACTRIMS Meeting in Paris, October 2017.

1.6 Other contributions during this PhD

E. Barratt, P. Tewarie, **M.A. Clarke**, E. Hall, P. Gowland, P. Morris, S.T. Francis, N. Evangelou, M.J. Brookes, 2017. *Abnormal Task Driven Neural Oscillations in Multiple Sclerosis: A Visuo-Motor MEG Study*. Human Brain Mapping (2017);38(5):2441- 2453

A. Samaraweera, **M.A. Clarke**, A. Whitehead, Y. Falah, I.D. Driver, R.A. Dineen, P.S. Morgan and N. Evangelou, 2017. *The Central Vein Sign in Multiple Sclerosis Lesions Is Present Irrespective of the T2* Sequence at 3 T*. J Neuroimaging. 2017 Jan;27(1):114-121

T. Sinnecker, **M. Clarke**, D. Meier, C. Enzinger, M. Calabrese, N. de Stefano, A. Pitiot, A. Giorgio, M. Schoonheim, F. Paul, M. Pawlak, R. Schmidt, L. Kappos, A. Rovira, N. Evangelou, J. Wuerfel. '*Evaluation of the "three lesions with a central vein sign" criteria for the differentiation between multiple sclerosis and its imaging mimics - a 3T MRI multi-center MAGNIMS study*'. Presented at 71st Annual Academy of American Neurology Meeting in Philadelphia, USA in May 2019.

E. Peake, **M. Clarke**, C.R. Tench, N. Evangelou, P.S. Morgan. '*Diagnosing MS using Central Veins at Clinically Relevant Image Resolutions*'. Presented at the 27th Annual International Society for Magnetic Resonance in Medicine (ISMRM) Meeting in Montreal, Canada in May 2019.

M. Clarke, A.P.R. Samaraweera, O. Mougin, M.South, R.A.Dineen, P.S.Morgan, N.Evangelou. '*Utility of the central vein sign to diagnose multiple sclerosis; A comparison of 3T T2* and FLAIR-SWI*'. Presented at 32nd ECTRIMS Meeting in London, UK in October 2016.

A.P. Samaraweera, O. Mougin, **M. Clarke**, R.A. Dineen, P.S. Morgan, N. Evangelou. '*T2* at 3T detects a higher proportion of white matter lesions with central veins compared to FLAIR-SWI in patients with MS*.' Presented at 32nd ECTRIMS Meeting in London, UK 2016

E. Barratt, **M. Clarke**, N. Evangelou, P. Gowland, P. Morris, S. Francis and M. Brookes. '*Reduced Visual Gamma Oscillations in Multiple Sclerosis Patients'*. Presented at MEG UK 2016 in York, UK and at 20th International BIOMAG Conference in Seoul, South Korea.

A. Samaraweera, **M. Clarke**, Y. Falah, O. Mougin, A.Pitiot, C.R. Tench, R.A. Dineen, P.S. Morgan, N. Evangelou. '*STAR-MS Study: Single Test to ARrive at MS diagnosis. A prospective, investigator blinded pilot study, assessing the accuracy of a single 3 Tesla MRI scan in predicting multiple sclerosis in cases of diagnostic uncertainty: Study protocol*.' Presented at the NAIMS Meeting in LA, USA in November 2015 and 31st ECTRIMS Meeting in Barcelona, Spain 2015

M. Clarke, A. Samaraweera, A. Connolly, O. Mougin, R.A. Dineen, P.S. Morgan, N. Evangelou. '*Diagnosing multiple sclerosis using the central vein sign. What is the agreement between a blinded rater and the final diagnosis using T2* weighted sequences at 3T MRI*?' Presented at the 31st ECTRIMS Meeting in Barcelona, Spain 2015.

A. Samaraweera, **M. Clarke**, O. Mougin, C.R. Tench, R.A. Dineen, P.S. Morgan, N. Evangelou. '*A comparison of FLAIR* and T2*- weighted imaging in detecting white matter lesions and central veins in patients with MS and ischaemic lesions at 3T*'. A poster presented at the 23rd ISMRM Annual Meeting 2015 in Toronto, Canada.

1.7 Awards

Top Score Poster at 35th Congress of ECTRIMS, Stockholm, September 2019.

Highly Commended Postgraduate Teaching Award, University of Nottingham, May 2017.

Poster Presentation Winner Division of Clinical Neuroscience Research Day, University of Nottingham, April 2017.

1st Poster Prize at the Division of Clinical Neuroscience Research and Poster Day, University of Nottingham, January 2016.

1.8 Thesis outline

The second chapter of this thesis provides a detailed literature review of MS and its pathophysiology including the two hallmarks of the disease: WM lesions and cerebral atrophy. It then goes on to discuss the tools used to visualize and monitor the pathology of MS including clinical MRI.

Results chapters three, four and five explore the role of the CV sign in MS diagnosis. Chapter three describes the results of a European, multi-centre study involving over 600 subjects with MS and mimicking disorders to determine whether the CV sign can be reliably used to classify patients as MS and non-MS using variable, non-standardised, imaging protocols. Chapters four and five are prospective studies of the CV sign in patients with an atypical first presentation of a demyelinating event (chapter four) and typical presentation (chapter five). Chapter six is the last results chapter, which explores cerebral atrophy calculation from clinical scans and proposes a novel approach to deal with errors in atrophy estimation when using routine, non-standardised, clinical MRI with incomplete coverage of the brain.

The last section of this thesis summarises the main findings from my PhD, discusses their implications and provides recommendations for future study.

CHAPTER 2: LITERATURE REVIEW

2.1 Multiple Sclerosis

2.1.1 Clinical course

Clinically isolated syndrome (CIS) is frequently the first presentation of the disease before a diagnosis of MS, or another demyelinating condition, can be made. As of 2013, CIS is considered to be on the spectrum of MS (Lublin et al., 2014). It is typically characterised by quick-onset clinical symptoms, which last for more than 24 hours, and the presence of inflammatory lesion(s). Recovery can happen spontaneously or can be aided through drug therapy consisting of steroids. One of the most common presentations of CIS is unilateral inflammation of the optic nerve, optic neuritis, which can cause symptoms ranging from slight blurring of vision to a complete loss of sight. Other common manifestations of CIS include transverse myelitis, Lhermitte's sign (a sensation of electric shocks going down one´s back and spreading out to the limbs) and cerebellar and brainstem syndromes (Brownlee and Miller, 2014).

It has been reported the three year conversion rate from CIS to clinically definite MS (CDMS), which is characterised by evidence of at least two separate clinical attacks, is around 40%, which increases to over 80% over a seven year follow-up (Moraal et al., 2009, Patrucco et al., 2013). Individuals with abnormal MRI at the time of CIS are more likely to convert to CDMS compared to those without brain lesions as demonstrated by Fisniku and colleagues, who found that 82% of CIS patients with T_2 lesions at baseline MRI had converted to CDMS 20 years later, compared to only 21% of those with a normal baseline MRI (Fisniku et al., 2008).

In line with the latest revisions carried out by the International Advisory Committee on Clinical Trials of MS in 2013 (Lublin et al., 2014), once a person is diagnosed with MS, their disease can be classified as either relapsing-remitting (RR) or primary progressive (PP). RRMS consists of disease-free periods, followed by relapses and subsequent remissions that affect about 80-85% of the entire MS population (Confavreux et al., 1980). PPMS, on the other hand, is characterised by a steady accumulation of disability from disease onset, typically with no clinical relapses and, up until 2018, no DMTs were available for people with PPMS in the UK. Ocrelizumab, a monoclonal antibody treatment, was approved by the US Food and Drug Administration (FDA) as the first ever treatment for PPMS (and RRMS) in the US in 2017 and was later granted approval by the European Medicines Agency (EMA) in early 2018 (Mulero et al., 2018).

Pathologically the main difference between the two subtypes is thought to be the degree of the inflammatory component of the disease, which is more prominent in the RR course. As the disease progresses, so does the neurodegenerative component and it is estimated that after 15 years of disease duration, over 50% of RR patients will enter the secondary progressive (SP) stage (Scalfari et al., 2014). SPMS development is characterised by a gradual shift in disease activity characterised by fewer or no relapses and steady accumulation of neurological impairment. Overall, around nine in ten RRMS patients will experience accumulation of disability over a 25 year period and enter the SP course (Trojano et al., 2003, Scalfari et al., 2010).

In 2013, Lublin et al. further classified the existing MS phenotypes based on disease activity and progression (Lublin et al., 2014). According to the consensus criteria, active disease is characterised by relapses and MRI activity including gadolinium-enhancing lesions, and/or new or enlarging T_2 lesions, as detected through recommended annual assessments; in comparison, inactive disease is free of relapses and MRI activity. Similarly, recommended annual assessments of disease progression in progressive patients, independent of relapses, were also introduced. As such CIS, relapsing and progressive patients can be classed as either active or inactive whilst progressive patients can additionally be classed in terms of their disease progression (disease worsening).

2.1.2 Diagnosis

The current diagnostic criteria for MS are the so-called 2017 McDonald criteria (Thompson et al., 2018) and, like their predecessors, incorporate the concepts of DIS and DIT, first proposed by Schumacher (Schumacher et al., 1965). DIS can be established through the presence of at least one T_2 lesion in two or more CNS areas including the periventricular, juxtacortical, infratentorial, spinal cord and, following the latest revisions, cortical. DIT can be confirmed using a single MRI scan if simultaneous presence of Gadolinium-enhancing and non-enhancing lesions (early, acute and older, chronic lesions respectively) can be demonstrated; alternatively, the presence of one or more new T_2- or Gadolinium-enhancing lesion(s) on a follow-up MRI scan, with reference to a baseline scan, can also demonstrate DIT. Finally, the most recent modifications to the McDonald criteria allow a positive lumbar puncture with cerebrospinal fluid (CSF)-specific bands to replace the need to demonstrate DIT. OCBs testing involves the analysis of the CSF taken from the patient's lumbar

region; the immune system produces IgG antibodies which are released into the CSF in response to the inflammation in the CNS resulting in multiple bands (two or more) when tested using electrophoresis. Although OCBs testing can be used to support a diagnosis of MS, only around 80% of CDMS patients have positive OCBs (Lechner-Scott et al., 2012) and other diseases can also lead to a positive test including Lyme disease (Hansen et al., 1990), neuromyelitis optica spectrum disorder (NMOSD) (Bergamaschi et al., 2004) and systemic lupus erythematosus (SLE) (Ernerudh et al., 1985).

Finally, unlike the 2010 criteria, the 2017 McDonald criteria make no distinctions between symptomatic and asymptomatic lesions, thus simplifying the diagnostic process. Previously, symptomatic lesions were not counted towards the demonstration of DIT or DIS, whilst they can now be used to support a diagnosis in patients presenting with WM pathologies in the brainstem or spine (but not in the optic nerve).

Although MRI is frequently used to support the diagnosis of MS, this can be done on clinical grounds alone if evidence of at least two typical clinical episodes, also known as relapses, exacerbations and flare-ups, can be documented. A relapse can be defined as a sudden-onset of symptoms, or worsening of existing symptoms, which lasts at least 24 hours and occurs at least 30 days after the start of a previous episode. There must also be no other explanation for the appearance of symptoms, including presence of an infection or increased body temperature, which could account for the development of symptoms. A relapse can either be detected

retrospectively, looking at the patients' medical history, or through neurological examination at the time of the episode.

When the above criteria are fulfilled, patients are given the RRMS diagnosis. In order to be diagnosed with PPMS, patients need to exhibit at least one year of sustained disease progression from disease onset (retrospectively or prospectively) without relapses, and two of the following: (1) established DIS in the brain (2) and/or spinal cord, and/or (3) positive OCBs in the CSF. On the other hand, SPMS is usually diagnosed in RRMS patients around 10 to 20 years post-diagnosis, with most people with MS (PwMS) entering the SP course closer to the 20 year mark (Weinshenker et al., 1989, Myhr et al., 2001, Eriksson et al., 2003). A diagnosis of SPMS is based on retrospective assessment of disease progression over six to 12 months, characterised by a significant reduction or complete cessation of relapses and build-up of neurological disability.

It is important to note that alternative diagnoses of other demyelinating disorders need to be excluded before a diagnosis of MS can be made, as the current criteria incorporate the concept of 'no better explanation' for patient symptoms. It is also worth noting that the McDonald criteria were developed and validated in populations of patients with typical presentations of demyelinating disease suggestive of MS (Thompson et al., 2018). Consequently, they should only be used in patients with a high probability of MS at presentation. Atypical presentations of demyelination, both clinically and on imaging, can therefore lead to delays in diagnosis, as further clinical follow-up and tests need to be carried out in order to gather sufficient evidence to accept or reject an MS diagnosis. As a result, there is a lot of interest in biomarkers,

which could successfully distinguish between MS and mimicking disorders. For example, NMOSD can result in clinical symptoms and MRI appearance with similarities to MS, and the discovery of an association between serum aquaporin-4 immunoglobulin G antibodies (AQP4-IgG) and NMOSD has led to significant improvements in the differential diagnosis of NMOSD (Wingerchuk et al., 2015). For now the search for MS-specific biomarkers, which could help successfully differentiate it from mimicking disorders, continues.

2.2 The Pathology of MS

Historically, MS used to be considered a disease characterised predominantly by focal WM pathology (Charcot, 1868). Evidence from pathological studies and advancing MR technology have allowed scientists to identify more subtle pathologies related to MS, such as diffuse WM damage, termed “normal appearing” WM (NAWM), as well as cortical lesions. Moreover, MS used to be considered a two-phase disease with an in initial inflammatory component, followed by irreversible neurodegeneration in the latter stages of the disease. However, it is now widely accepted that CNS tissue loss is present from the start of the disease. This overview summarises the main pathological hallmarks of MS, including WM lesions and cerebral atrophy, and their contribution to disability accumulation in patients.

2.2.1 White matter lesions

2.2.1.1 Classification and formation

In MS the myelin sheaths, which insulate and protect axons, become targets of immune attacks, leading to their destruction. Moreover, oligodendrocytes, supporting glial cells which produce myelin in the CNS, are also lost. This leads to a reduction

in, and in some cases a complete loss of, myelin which can cause damage to the axons themselves as they are left exposed. As a result, MS is considered a demyelinating disease characterised by astrocytic scarring with relative sparing of axons.

MS WM lesions can range in size from less than 1 mm to spanning across large portions of the brain; the latter is termed confluent lesions. MS WM lesions are most commonly found in the optic nerve, corpus callosum, cerebellum, around the ventricles, in juxtacortical WM and in the spinal cord (Evangelou et al., 2000b, Trip and Miller, 2005, Popescu and Lucchinetti, 2012b) . They tend to be round or oval in shape and periventricular lesions can often form the so-called ‘Dawson fingers’ (see Figure 2-1). These finger-like areas of inflammation form perpendicularly to the lateral ventricles and are MS-specific (Dawson, 1916, Traboulsee and Li, 2006); MS plaques, including Dawson fingers, often surround and follow the path of blood vessels.

Depending on the stage of the disease, different types of WM lesions have been identified, all of which are characterised by inflammation with higher levels of inflammation detected in the earlier stages of the disease (Frischer et al., 2009). Acute plaques, also known as active plaques/lesions, are most common in the early stages of relapsing MS. Whilst active lesions are characterised by the relative sparing of axons, the fate of oligodendrocytes is less clear with some studies suggesting they are preferentially targeted in early lesions (Prineas et al., 1984) and others reporting their relative sparing and evidence of remyelination (Bruck et al., 1995). There is some evidence to suggest inter-patient pathological heterogeneity

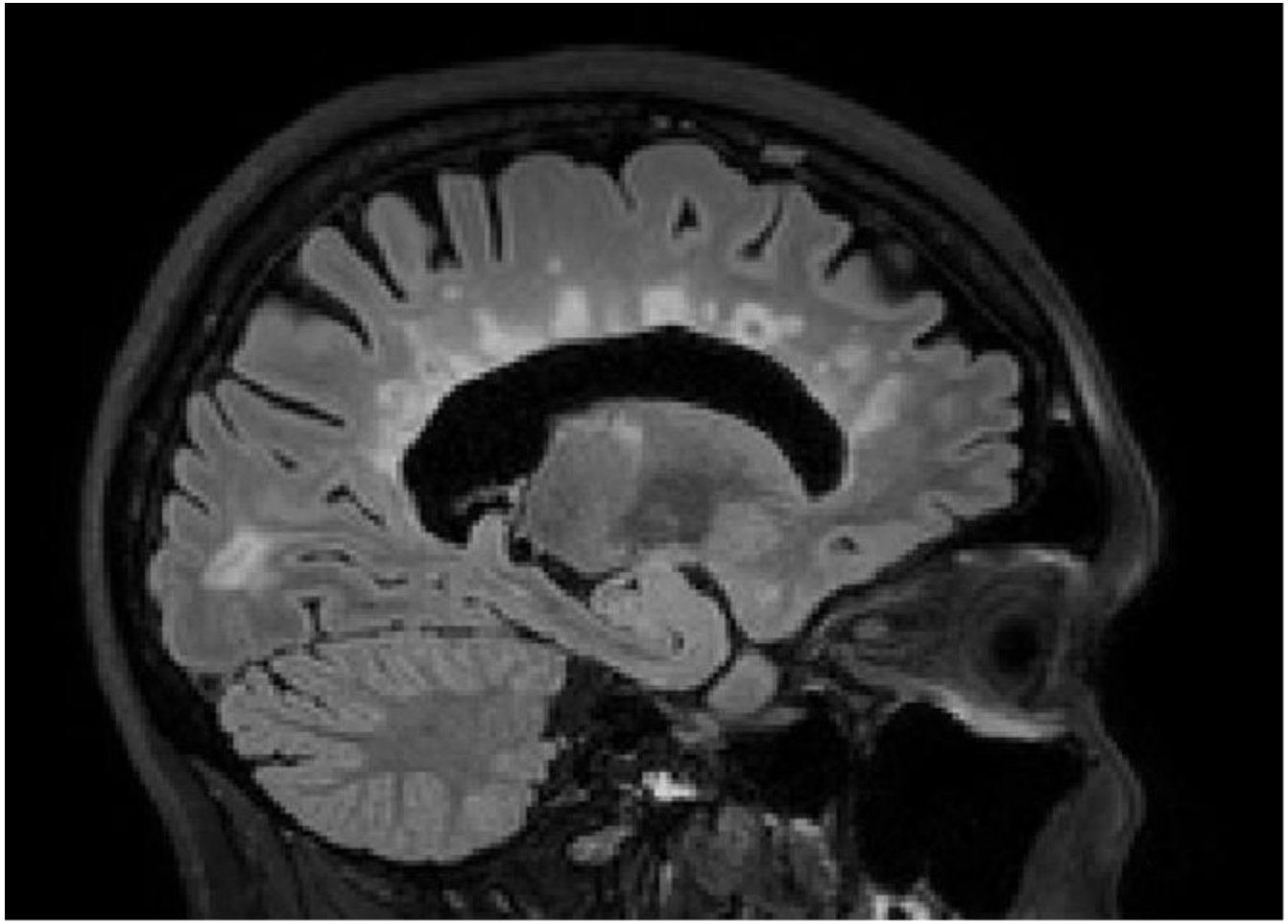


Figure 2-1 Dawson fingers as seen on sagittal FLAIR MRI.

These finger-like projections, found perpendicularly to the lateral ventricles, are MS-specific.

with regards to the fate of oligodendrocytes in early MS lesions (Lucchinetti et al., 1999). Furthermore, acute lesions are infiltrated by macrophages, which are present across the lesional area, as demyelination takes places in all parts of the lesion simultaneously (Bruck et al., 1995, Bruck et al., 1996, Popescu and Lucchinetti, 2012a). Hypertrophic astrocytes with large nuclei and lymphocytes, most notably CD-8-positive T lymphocytes with some CD-4-positive T cells and B cells, are also present (Wu and Raine, 1992, Frischer et al., 2009, Ludwin et al., 2016). The breakdown of the blood brain barrier (BBB), which allows the inflammatory infiltrates to enter the CNS, also allows these relatively new, active lesions to be detected on MRI.

Chronic lesions are more common in the progressive stage of the disease and can be further subdivided into chronic active and chronic inactive lesions. Chronic active plaques are characterised by a loss of myelin, with relative axonal preservation (Popescu and Lucchinetti, 2012a). There is a marked shift in the neuroinflammatory profile of the lesion, which is reduced compared to active lesions and present mainly on the expanding outer edges of the lesion (Lassmann, 1998).

Once the lesions become inactive, there is a marked reduction in the expression of inflammatory agents, macrophages and a marked reduction of oedema; instead, astrocytic scarring appears (Lassmann, 1998, Popescu and Lucchinetti, 2012a). Chronic inactive lesions are clearly demarcated and characterised by axonal loss within the lesion centre of up to 80% (Frischer et al., 2009). Many MS lesions are characterised by T_1-hypointensity on MR images, which reflects tissue destruction within lesions; it is estimated that around 40% of T_1-hypointense plaques, also

known as ‘black holes’, will become chronic as a result of multiple demyelinating events or persistent chronic active inflammation (Sahraian et al., 2010).
Whilst many axons undergo demyelination, there is also evidence of active remyelination in some MS lesions. This is supported by the presence of new myelin sheaths and oligodendrocyte precursor cells in formerly demyelinated areas (Popescu and Lucchinetti, 2012a). Lesions with complete remyelination are often called *shadow plaques* and have reduced myelin density compared to NAWM. It is estimated that around 40% of lesions will undergo remyelination (Barkhof et al., 2003).

Over the years, many studies have investigated the chronology of events which lead to WM lesion formation in MS. There is general consensus that myelin sheaths become targets of immunological attacks, which is the basis of MS pathogenesis (Frohman et al., 2006, Wu and Alvarez, 2011). It is also clear that inflammation, demyelination and neurodegeneration are involved, however, the sequence of events in this complex process is still largely debated. Several studies investigating early lesion formation using frequent MR (more than once a month) reported Gadolinium enhancement in all new lesions (Barkhof et al., 1992, Lai et al., 1996, Guttmann et al., 2016) concluding that the breakdown of the BBB is an essential step in this process. In contrast, other studies reported T_2, magnetic resonance spectroscopy (MRS), magnetisation transfer ratio (MTR), diffusion and perfusion changes before contrast enhancement (Filippi et al., 1998, Narayana et al., 1998, Werring et al., 2000, Wuerfel et al., 2004) suggesting that the breakdown of the BBB might be a secondary event preceded by a variety of changes in both cerebral tissue and blood. More recently, the existence of (p)reactive lesions, also termed NAWM,

was proposed (De Groot et al., 2001). They are characterised by microglial and lymphocytic presence with no apparent loss of myelin suggesting that inflammatory processes may precede demyelination. Interestingly, many of the preactive lesions have been found to resolve without becoming demyelinating lesions, and so understanding their pathophysiology could shed more light on the aetiology of WM lesion formation (van der Valk and Amor, 2009).

WM lesions continue to have a pivotal role in patient diagnosis, monitoring, and as an outcome in clinical trials as they are easily identified even at low field strength. More recently, there has been considerable interest in exploring the perivenous nature of WM lesions as a diagnostic tool in MS.

2.2.1.2 Perivenous nature of WM lesions

In the 1820s, French anatomist Jean Cruveilhier provided the first account of the perivenous distribution of MS plaques, comparing them to ‘areas of embolism’ (Cruveilhier, 1829-1842 cited in Olek, 2005). Approximately 60 years later, Charcot, who first described MS, noted that WM plaques formed alongside blood vessels, although their contribution to MS pathogenesis was dismissed and considered ‘secondary’ (Charcot and Marie, 1886 cited in Zalc, 2018) . Then in 1942, Dow and Berglund studied five post-mortem brains of people with MS and found that 66% of the lesions had CVs and six out of the 20 lesions without a CV had multiple veins (Dow and Berglund, 1942) thus confirming and quantifying the preferential formation of MS plaques around blood vessels.

It was only in 2000 that the perivenous nature of MS lesions *in vivo* was demonstrated for the first time using contrast enhancement and venography at 1.5T (Tan et al., 2000). Veins can also be visualised using GRE imaging without the need for a contrast agent as deoxygenated haemoglobin has paramagnetic properties, unlike oxygenated blood, which is diamagnetic (Ogawa et al., 1990). As a result, susceptibility-weighted sequences have been used to visualise CVs, which appear dark against the bright lesions on those types of images.

Many studies have investigated the so-called CV sign in MS, suggesting that it could be used as a diagnostic biomarker which can successfully distinguish between MS and other conditions resulting in WM lesions. Studies using ultra-high field MRI have reported around 80% of lesions as having a CV, with periventricular lesions having the highest number- a total of 94% (Tallantyre et al., 2009, Tallantyre et al., 2011, Kuchling et al., 2014). Perivenous lesions have been reported in all MS phenotypes (Kilsdonk et al., 2014a) and all brain regions including the thalamus, cerebellum, pons and the cortex (Sati et al., 2012, Kilsdonk et al., 2014a) , although the proportion of lesions with CVs decreases with distance away from the ventricles (Tan et al., Tallantyre et al., 2008, Kuchling et al., 2014, Lane et al., 2015) and patient age (Kilsdonk et al., 2014a).

Studies using clinically available lower field strength scanners, including 1.5T and 3T, have also successfully visualised the CV sign (see Figure 2-2 for a typical appearance of the CV sign on a 3T scan). They consistently reported a 40-45% threshold, which has been shown to be able to successfully differentiate between MS, which is characterised by a higher proportion of CVs, and other conditions,

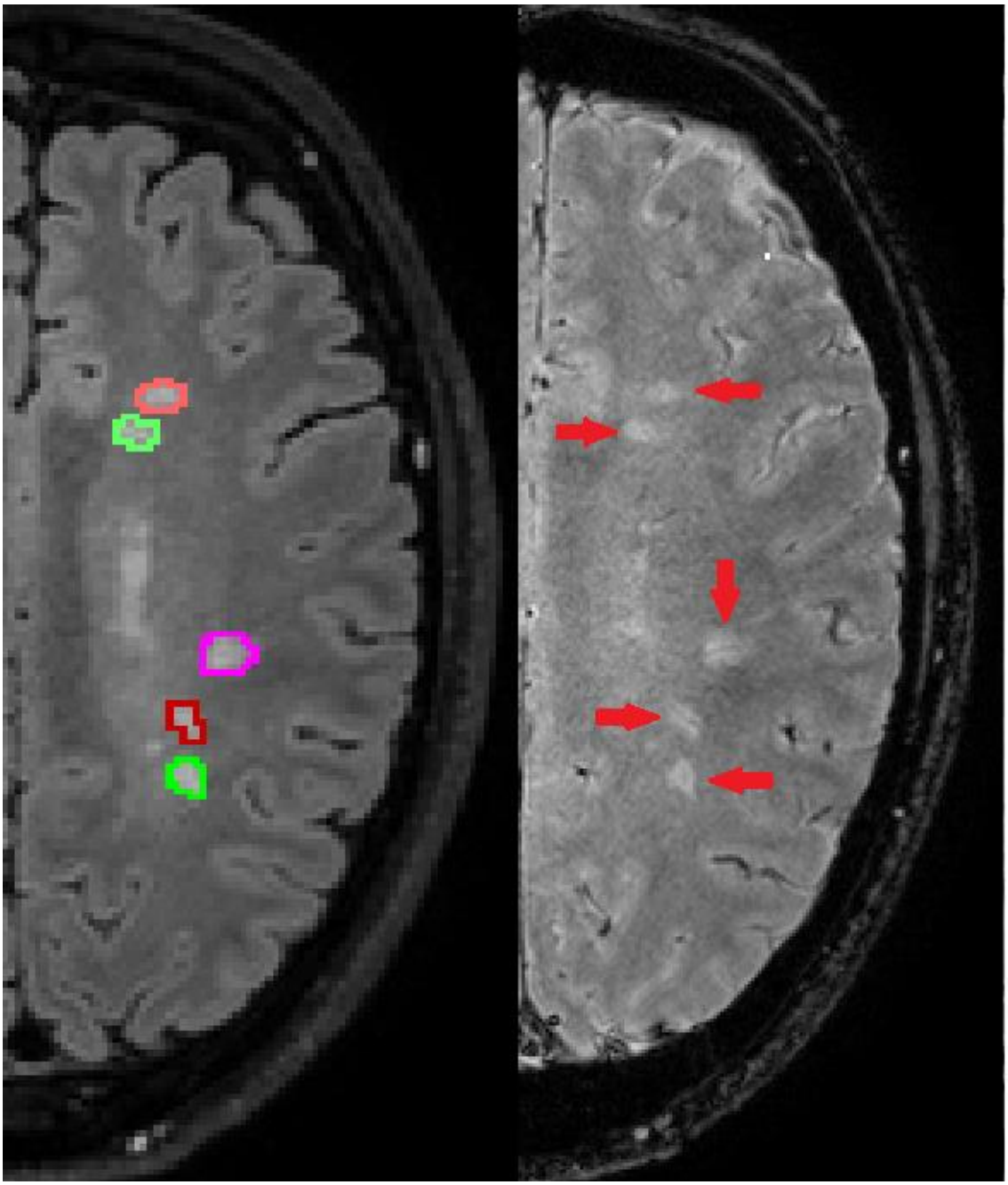


Figure 2-2 MS lesions (outlined in colour) visible on a FLAIR scan (left) and their corresponding appearance on a T2*-weighted 3T MRI scan (right).

The bright WM lesions contrast with the dark veins that pass through them. Typically more than 40% of all WM lesions have a visible CV at 3T.

which can mimic its appearance on MRI (Tallantyre et al., 2009, Campion et al., 2017). So far studies have investigated the CV sign in MS and NMOSD (Sinnecker et al., 2012a, Kister et al., 2013, Cortese et al., 2018), small vessel disease (Tallantyre et al., 2011, Kilsdonk et al., 2014b, Lane et al., 2015, Mistry et al., 2015, Samaraweera et al., 2016, Sparacia et al., 2018) Susac syndrome (Wuerfel et al., 2012), migraine (Solomon et al., 2015), dementia (Lane et al., 2015) and systemic autoimmune diseases (SADs) such as Behçet syndrome and SLE (Massacesi, 2016, Maggi et al., 2018). All of the studies have confirmed the 40-45% cut-off rule, with small vessel disease and Berçet syndrome patients having the highest median proportion of CVs, 45% and 40% respectively.

The increasing interest in the application of the CV sign, as a diagnostic biomarker in MS, has prompted the release of official guidelines by both the Magnetic Resonance Imaging in MS (MAGNIMS) (Filippi et al., 2016) and the Consortium of MS Centres (CMSC) groups (Traboulsee et al., 2016) requesting further validation before it could be incorporated as a clinical tool in MS. Furthermore, the North American Imaging in MS (NAIMS) Cooperative released official guidelines and consensus opinion concerning the identification of the CV sign radiologically (Sati et al., 2016) to promote uniform practice. The three criteria that need to be satisfied in order to identify a lesion as having a CV are: (1) the appearance of the vein as a dark line or dot, (2) the vein should be detected in at least two orthogonal planes and (3) the vein must be located centrally irrespective of the lesion's shape.

In order to be of clinical value, the CV sign needs to be readily identifiable in a clinical setting where time constraints mean it may not be possible to count all the

lesions, especially in patients with a high lesion load. As a result, various groups have proposed simplified rules and/or algorithms to instead identify a proportion of lesions with CVs, in order to make a diagnosis of MS or exclude it. For example, Mistry and colleagues proposed a ‘rule of six’ whereby if an MR scan contains six or more lesions with CVs, and at least six of them have a CV, the diagnosis is MS; if there are fewer than six lesions with CVs and at least half of them have a CV, the diagnosis is MS; otherwise an MS diagnosis cannot be made (Mistry et al., 2015). They reported that the test takes around two minutes and it was able to correctly identify 20 out of 20 patients in the study with MS and small vessel disease using a 3T MRI. Similarly, Solomon proposed an algorithm which uses 3T FLAIR* images and allows three lesions based in the subcortical, deep WM or juxtacortical regions to make a diagnosis of MS based on the presence of a CV (Solomon et al., 2018). They obtained good specificity, sensitivity and fair inter-rater reliability, demonstrating that simplified algorithms may be useful in the clinical application of the CV sign.

As evidenced, the CV sign has the potential to become an MS-specific biomarker used for diagnostic purposes, however further evidence is needed to establish whether it can successfully differentiate MS from its mimics in a clinical setting. So far, there have been no large, prospective, multi-centre studies to assess the accuracy of the CV sign using MRI data with variable protocols, to ensure it can be reliably detected in a typical clinical setting.

2.2.1.3 Correlation with clinical parameters

One of the most fundamental studies of WM lesions in MS was carried out by Fisniku and colleagues who followed-up CIS patients for 20 years. They found that individuals with abnormal MRI at baseline were more likely to develop CDMS and that WM lesion volume also correlated with disability (R_2 values ranging from 0.48 to 0.67, $p<0.001$ depending on time point measured). Moreover, those who converted to SPMS had a significantly higher annual lesion growth compared to those who remained RRMS (almost 3 cm_3/year compared to 0.80 cm_3/year) (Fisniku et al., 2008). In addition to this, early (first five years) increases in T_2 lesion volume correlated strongly with increased disability ($R=0.69$, $p<0.001$) in those who developed MS. Another study involving almost 400 patients looked at the prognostic value of new T_2 lesion development in predicting poor response to first line treatments (interferon beta) and found that each new lesion on MRI increased the augmented risk of being a poor responder 10-fold (Prosperini et al., 2009), providing further evidence that high lesion load is associated with poorer clinical outcomes. Finally, a meta-analysis of a collection of prospective studies looking at the predictive value of WM lesions in determining relapse rates in the first year after diagnosis, found a moderate correlation between the number of Gadolinium-enhancing lesions in the first six months and subsequent clinical events (Kappos et al., 1999). These findings were further expanded in another meta-analysis looking at treatment efficacy in randomised clinical trials and it was demonstrated that short-term (6-9 month) follow-up of WM lesions predicted the annual relapse rate long-term (year one to two; slope=0.52; R_2=0.71) (Sormani and Bruzzi, 2013). All of the above provide strong evidence for the role of WM lesions in predicting future disease progression in MS patients.

As a result, information about WM lesion load can be helpful when assessing new patients in the clinic. For example, the choice of DMTs in newly diagnosed patients is often based on the appearance of their baseline MRI with first line treatments being offered to those with few lesions, and more effective medications offered to those with more abnormal MRIs. Nevertheless, the correlation between lesions detected on MRI and clinical outcomes remains moderate at best, which has led to the formation of the term clinico-radiological paradox. It is likely that this discrepancy between individual MRI biomarkers, such as WM lesion load, and clinical outcomes exists because of the vast heterogeneity of the pathophysiological processes taking place simultaneously in the MS brain. For example, T_2 lesions provide information about the ongoing, focal neuroinflammatory processes going on in the brain but do not reflect the irreversible axonal damage caused by the disease. As a result, using multiple MRI biomarkers, such as WM lesions and global atrophy, is likely to resolve the clinico-radiological paradox as a more complete picture of the disease is presented.

2.2.2 Brain atrophy

2.2.2.1 Brain atrophy in healthy ageing and MS

Brain atrophy is the gradual decline of brain tissue over time and its rate is increased in disorders such as Alzheimer's disease or MS. It leads to reduced brain weight and volume, thinning of the gyri and expansion of the ventricles and sulci (Kemper, 1994). A meta-analysis, using over 2,000 healthy participants, found that the human brain grows around 1% annually up until early teenage years before experiencing no apparent annual change in volume in young adults (Hedman et al., 2012). After 35

years of age the annual rate of atrophy is around 0.2% per annum (Raz et al., 2004, Hedman et al., 2012). Atrophy rates further accelerate with advancing age; the brain shrinks about 5% every decade after the age of 40 and this increases further after the age of 70 (Svennerholm et al., 1997, Hedman et al., 2012).

Brain shrinking is accelerated in MS (see Figure 2-3) and it is estimated that people with MS lose between 0.46 to 1.35% of their brain volume every year (De Stefano et al., 2014, Vollmer et al., 2015). A recent study found that a cut-off point of -0.37% could reliably differentiate between healthy controls and people with MS, resulting in 67% sensitivity and 80% specificity (De Stefano et al., 2016).

Cerebral atrophy affects both progressive and relapsing patients (Kalkers et al., 2002, De Stefano et al., 2015) and can be observed as soon as CIS and radiologically isolated syndrome (RIS), whereby patients demonstrate demyelinating lesions with no clinical symptoms (Anderson et al., 2007b, De Stefano et al., 2011). Indeed, CIS individuals with higher annual atrophy are at a greater risk of converting to CDMS (Di Filippo et al., 2010) and evidence suggests that early treatment with DMTs, which reduce the rate of atrophy, can reduce conversion rate to CDMS (Jacobs et al., 2000).

Various studies have also examined the differences in atrophy rates between the different subtypes of MS; there is evidence to suggest that the rate of brain atrophy is similar for both relapsing and progressive patients, including both PP and SP patients. For example, a study involving 83 patients with RRMS, SPMS and PPMS demonstrated that the rate of brain atrophy progressed at a similar rate irrespective

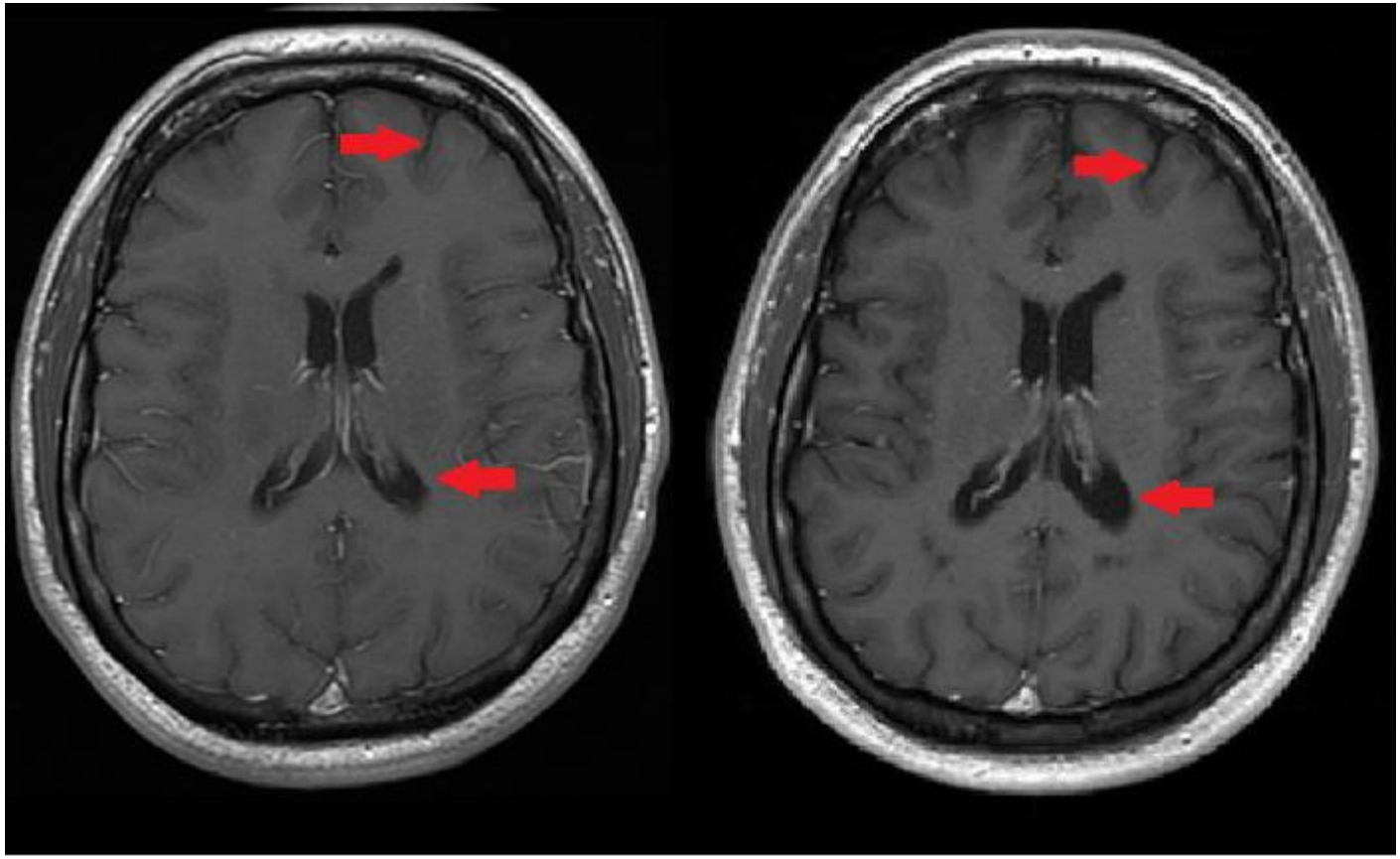


Figure 2-3 Clinical scans of an MS patient taken seven years apart demonstrate brain atrophy.

There is a marked expansion of the sulci and thinning of the gyri (top arrow) and enlargement of the lateral ventricles (bottom arrow)

of disease type when adjusted for baseline brain volume (Kalkers et al., 2002) and a large, longitudinal study of over 900 untreated patients echoed these findings (De Stefano et al., 2010). On the other hand, Tedeschi and colleagues reported higher levels of WM and GM atrophy in SPMS patients than RRMS (Tedeschi et al., 2005). Finally, Vollmer's meta-analysis of 38 longitudinal studies, published between 2003 and 2013, investigated brain volume change in patients and reported annual mean PBVC ranging from -1.32% to -0.22% in RRMS, -2.1% to -0.59% in SPMS and -0.94% to -0.63% for PPMS, thus suggesting that despite the large overlap in annual atrophy rates between the different subtypes of MS, there are subtle differences with SPMS patients showing the highest levels of atrophy (Vollmer et al., 2015). As atrophy measurements involve small changes, especially when considering annual measurements of brain volume, these findings must be interpreted with caution and it is unlikely that atrophy measurements could be used to reliably differentiate between the different MS subtypes.

Furthermore, atrophy does not affect all brain regions in MS at the same rate, and several structures have been found to be particularly susceptible to it. These include the frontal and temporal lobes, with the hippocampi being particularly vulnerable, and the ventricles which expand as the brain shrinks (Raz et al., 1997, Scahill et al., 2003, Carmichael et al., 2007, Frisoni et al., 2008).

2.2.2.2 Pathogenesis of brain atrophy

Historically, brain atrophy in MS was believed to be an end-stage result of the disease, however more recent evidence has shown that brain tissue loss is present from the very start of the disease (De Stefano et al., 2010, Perez-Miralles et al.,

2013). Accelerated brain atrophy in MS is thought to reflect both the direct, irreversible damage and death of the glial cells, axons and neurons of the CNS and the indirect, pathological changes in the lesioned and normal appearing tissues.

Initially, neuroinflammation is believed to lead to the destruction of myelin sheaths which can be seen as areas of hyperintense signal on T_2-weighted scans. These are thought to reflect the ongoing inflammation, which can disappear and, especially in the early stages of the disease, exhibit spontaneous or DMT-induced remyelination. Over time, some of the damage can lead to axonal and neuronal death, which in turn can be seen as hypointense 'black holes' on T_1-weighted scans, areas thought to represent the irreversible tissue damage. Axonal loss within WM lesions was described as early as Charcot's initial account of MS (Charcot, 1868). More recent studies have found that there is a direct correlation between neuroinflammation and axonal damage (Ferguson et al., 1997, Trapp et al., 1998, Evangelou et al., 2000a), with around 60% reduction in axonal density within demyelinated lesions (Mews et al., 1998). These findings have been reported both in active as well as chronic lesions (Ferguson et al., 1997, Trapp et al., 1998).

Axonal transection and loss is not limited to MS lesions, but has been found to be significantly increased in the NAWM too. Fu and colleagues used MRS to detect N-acetylaspartate (NAA), a marker of axonal injury, and calculated the NAA to creatine/phosphocreatine (NAA:Cr) ratio both within lesions and in the NAWM. They found reduced NAA:Cr ratio in WM lesions compared to the NAWM, however it was the decrease in NAWM that correlated strongly with changes in disability over a period of 30 months (Fu et al., 1998). Evangelou and colleagues furthered this and

quantified the number of axons within NAWM, using histological sections of the corpus callosum taken from MS patients and comparing them to sections taken from healthy controls (Evangelou et al., 2000b). They reported a reduction of both axonal density and volume in MS, with patients having less than half the number of axons as healthy controls demonstrating that axonal loss is also prominent outside of MS lesions. In an effort to further understand axonal loss, several studies have attempted to evaluate the characteristics of the axons more prone to injury. Data from research groups investigating these effects in the anterior optic pathways, cervical spinal cord and the brain found that smaller axons seem more susceptible to injury (Ganter et al., 1999, Bitsch et al., 2000, Evangelou et al., 2000a, Lovas et al., 2000, Evangelou et al., 2001).

It has long been known that both neuroinflammation and axonal injury can lead to neuronal apoptosis (Blinzinger and Kreutzberg, 1968, Al-Abdulla et al., 1998), and it therefore comes as no surprise that evidence of apoptotic neurones has been found both within MS lesions and NAGM. Peterson and colleagues studied cortical neurones in MS and found that there was evidence of apoptosis in 10 out of the 22 GM lesions studied (Peterson et al., 2001). Just under half of chronic inactive lesions and over half of chronic active lesions were characterised by apoptosis but no active lesions had any evidence of apoptosis, although only two lesions of that type were studied. Contrary to preferential targeting of smaller axons in case of axonal transection in the WM, apoptotic neurones were generally large. Moreover, GM lesions were less inflamed compared to WM lesions in this study. The findings therefore suggest that neuronal loss might not be primarily driven by neuroinflammation and demyelination, as GM lesions have significantly lower levels

of pro-inflammatory cytokines and contain less myelin than WM, but other pathological processes might also be involved.

Another consequence of axonal injury is Wallerian degeneration, a process whereby the part of the axon distal to the damaged area degenerates and dies. Evangelou and colleagues found that there was a direct inverse relationship between regional lesion volume and axonal density ($R=–0.67$, $p<0.001$) and numbers ($R=-0.66$, $p<0.001$) in parts of the corpus callosum where the axons were projecting to (Evangelou et al., 2000b). Other studies using a variety of visualisation techniques including histology, structural MRI and DTI have provided further supporting evidence that Wallerian degeneration contributes to axonal loss in MS, even in very early stages of MS (Simon et al., 2000, Bjartmar et al., 2001, Ciccarelli et al., 2003, Dziedzic et al., 2010). If that is the case, this would suggest that anti-inflammatory therapies have the potential to reduce neurodegeneration not only by reducing inflammation within WM lesions, but also by having an indirect effect on the NAWM.

The above findings demonstrate that axonal loss is prevalent both within pathological as well as normal-appearing tissue in MS and could, at least in part, explain disability accumulation over time. Nevertheless, the relationship between neuroinflammation, demyelination and neurodegeneration remains unclear and requires further study. Despite the strong evidence of increased axonal loss within lesions, axonal injury in the NAWM is also apparent. There is evidence that NAWM of MS patients has increased levels of pro-inflammatory molecules (Kutzelnigg et al., 2005), giving further support to the theory suggesting a causal link between neuroinflammation and axonal loss. However, other evidence points to a conclusion that demyelination

and neurodegeneration are, at least in part, independent. Bitsch et al. used biopsy samples to study the relationship between neuroinflammation, demyelination and axonal loss and found that there was a correlation between axonal injury, indicated by increased levels of amyloid precursos protein, and macrophages, which promote a pro-inflammatory state (Bitsch et al., 2000). However, no correlation was found between reduced axonal densities and tumor necrosis factor α (alpha) and inducible nitric oxide synthase, two markers of demyelination. Further evidence to suggest neurodegeneration is driven by mechanisms other than demyelination comes from patients with extensive axonal loss and low lesion load (De Stefano et al., 2002).

Apart from neurones, glial cells are also targeted in MS and considering the fact that they outnumber neurones, the loss of glial cells might considerably contribute to the overall brain volume loss in MS. Mews and colleagues found that whilst there was a 60% reduction in axonal density within demyelinated lesions, virtually all oligodendrocytes were lost (Mews et al., 1998). Loss of astrocytes has also previously been reported in MS (Brosnan and Raine, 2013); as glial cells have a role in supporting and nourishing neurones, their loss could be a significant factor contributing to neurodegeneration in MS.

Neurodegeneration is a complex mechanism which, for now, remains poorly understood. What is clear is that it is present from the very start of the disease and reflects irreversible damage to brain tissue, which has been found to correlate well with disability levels in patients. Demyelination, neuroinflammation, Wallerian degeneration and reduction in the number of glial cells all seem to contribute to neurodegeneration, however further studies are required to understand the

sequence of events in this complex process. This could in turn help us understand the aetiology of MS and design more effective DMTs, which could help slow down or effectively stop the irreversible damage in MS.

2.2.2.3 Brain atrophy: regional or global?

The brain is made up of three main compartments: WM, GM and the CSF. Changes in any of those tissues, whether physiological or disease-driven, can contribute to changes in the overall brain volume. Various studies have assessed the individual contribution of GM, WM and the CSF as well as global brain volume, to patient disability and progression in MS, aiming to identify structures which reflect MS-related damage most accurately and whether they could be easily measured in a clinical setting.

WM volume varies across lifespan, with its volume increasing throughout childhood and early adolescence before then starting to decrease between the ages of 30 and 40 (Bartzokis et al., 2003, Bartzokis et al., 2004). The exact mechanisms that lead to WM atrophy remain to be further studied; so far a variety of theories have been proposed, including increased risk of WM lesion incidence with age (Sachdev et al., 2007), vascular risk factors which lead to a decrease in the amount of blood supplied to deep WM (Pantoni and Garcia, 1995), or age-related changes in the myelin structure (Madden et al., 2012). In MS, it is likely that WM volume decreases are driven by axonal transection and apoptosis. However, even the NAWM in MS has been found to experience up to 50% axonal loss and reduced axonal density, suggesting that there is substantial tissue loss not only in areas of WM lesions, including black holes characterised by irreversible tissue loss, but also in non-lesioned tissue (Evangelou et al., 2000a). Nevertheless, many studies, both cross-

sectional and longitudinal, have failed to find significant reductions in WM volume which suggests that tissue swelling and oedema associated with neuroinflammation could effectively mask any evidence of WM loss on MRI (Ge et al., 2001, Dalton et al., 2004, Tedeschi et al., 2005).

Unlike WM, GM appears to start declining early on. Both deep and cortical GM have been consistently found to show a linear negative relationship with age (Ge et al., 2002, Walhovd et al., 2005, Giorgio et al., 2010). In MS, GM atrophy has been found to be present from the start of the disease and progressed at a rate of around 0.7% annually (Uher et al., 2014, Zivadinov et al., 2016b). More recently, deep GM atrophy, especially thalamic atrophy, has been the subject of many studies investigating markers of neurodegeneration, and as an end-point in clinical trials (Rocca et al., 2010, Azevedo et al., 2018, Schoonheim and Ciccarelli, 2018). However, more longitudinal studies are needed to establish its role as a sensitive marker of progressing pathology.

Another potential marker of neurodegeneration is the ventricular system which consists of four interconnected CSF-filled cavities, also called ventricles, involved in the production of the CSF; lateral ventricles are the largest and are connected to the third ventricle via the interventricular foramina. The third ventricle is bounded by the thalami and it communicates with the fourth ventricle, a diamond-shape cavity in the pons, via the cerebral aqueduct. The enlargement of the ventricles accompanies brain atrophy and can be easily visualised on MRI scans, even on 2D scans with low contrast and resolution, especially in patients with long-standing disease (Vrenken et al., 2013). Ventricular enlargement can be observed at the earliest stages of MS and

people who convert to MS had larger ventricular volumes than those who did not convert to CDMS (Brex et al., 2000). Another study reported that the annual rate of ventricular enlargement in MS patients was five times that of healthy controls (1.6mL vs. 0.3mL) (Fox et al., 2000).

Various approaches have been proposed for the optimal measurement of ventricular enlargement, including calculating ventricular volumes and/or ventricular width, and the best structure to choose (lateral, third and/or fourth ventricles) to establish the most sensitive methods for atrophy assessment. Turner and colleagues compared 2D (width) and 3D (volume) measurements of the ventricles and found a strong positive correlation (R_2=0.81) between third ventricle measurements, but not for fourth and lateral (Turner et al., 2001).Third ventricle volume was also deemed the most sensitive measure of symptom duration, disability and disease type whilst third ventricle width could also successfully differentiate between MS and healthy controls. Other studies have provided further supporting evidence for the role of the third ventricle as a sensitive marker of disability and disease progression (Berg et al., 2000, Müller et al., 2013), which is likely a result of its easy-to-measure shape, being insensitive to magnetic field inhomogeneities and good tissue contrast (Miller et al., 2002). Lateral ventricle volumes, on the other hand, have been found to correlate strongly with cortical atrophy and can help to differentiate between patients who were stable and those experiencing disease progression, as well as whole brain atrophy (Jones et al., 2013, Zivadinov et al., 2016b).

Ventricular enlargement is a proposed indirect marker of neurodegeneration, which is simple to calculate and provides good correlations with clinical measures. It has

been shown to be relatively insensitive to variable acquisition protocols (Dwyer et al., 2015a), however calculation of ventricular enlargement requires expertise when performed manually. So far only a few software packages have been developed to help calculate ventricular volume automatically, using scans which are not 3D and with low contrast (Vrenken et al., 2014).

Finally, whole brain volume measurements are thought to reflect all of the pathological processes taking place in MS brains and provide a global measure of neurodegeneration. Brain atrophy has been suggested as a valid marker of long-term disease progression in MS (Maghzi et al., 2014) and one of the most promising outcomes of neuroprotection in clinical trials (Barkhof et al., 2009). Annual atrophy rate was found to be a good predictor of physical disability two (Sormani et al., 2014), five (Zivadinov et al., 2013) and eight years later, with patients with the highest decrease in brain volume being four times more likely to reach an EDSS score of 6 (characterised by the need for walking assistance) or greater, eight years on (Fisher et al., 2002). More recently, a 10 year follow up study further confirmed that patients with confirmed disability progression experienced greater annual atrophy rate, with percentage brain volume change (PBVC) at year two being predictive of disability at year 10 (Zivadinov et al., 2016b).

Early atrophy could potentially also act as a marker of cognitive dysfunction, currently often overlooked during clinical assessment. Maghzi and colleagues found that every 1% decrease in brain volume was associated with a 1.14 point decrease in the symbol digit modalities test (SDMT) and so brain atrophy has the potential to

capture changes in information processing speed, the most vulnerable cognitive domain in MS (Denney et al., 2008, Maghzi et al., 2014).

As there are currently no existing MRI surrogate markers of permanent damage and long-term disability, efforts have been made to help address this through the inclusion of brain volume loss as a forth component of NEDA. However, further validation is needed before NEDA-4 can be fully-established as a marker of disease progression (Kappos et al., 2015).

Nevertheless, caution must be exercised as global atrophy is a non-specific marker of damage which can be influenced by various disease-related processes, such as oedema, as well as fluctuations to due physiological changes such as hydration level of the patient (Zivadinov et al., 2016a). As a result, it is recommended that the follow-up period is at least six months and any atrophy measurements should be interpreted with caution (Sastre-Garriga et al., 2017). Despite issues with interpretation of atrophy values in individual patients, it is clear that global atrophy is a reliable biomarker of treatment effects, future disability and long-term disease progression when used on a group level.

2.2.2.4 Automated assessment of brain atrophy

The reproducibility of MR brain volume measurements is an important aspect of MS research as many multi-centre phase III clinical trials rely on accurate brain volume measurements to assess new drugs. Brain volume loss can be measured using two approaches. Segmentation techniques are based on tissue classification, most commonly into CSF, GM, WM and sometimes WM lesions, and are typically used in

cross-sectional studies. They frequently involve normalisation of brain volume to account for individual head size and facilitate comparisons across subjects. More recently, region specific analyses, such as volumetric measurements of the thalamus, have also been used to aid our understanding of future disability in MS (Rocca et al., 2010).

Registration-based methods on the other hand, involve longitudinal follow-up as images from the same patient from two time points are used to calculate the changes in brain volume. One of the main advantages of using registration-based methods to assess atrophy is the reduced measurement error, as each individual's brain is compared to itself and so any morphological differences between the two images are reduced. This is not the case in segmentation-based approaches as each brain is registered onto an atlas (an average of hundreds of MR brain images) to help correctly classify the different tissues.

Thanks to the good contrast and high resolution of MR images, various neuroimaging software packages have been designed to automatically calculate brain volume and/or atrophy. Some of the most widely used packages in MS research are the FMRIB Software Library (FSL) (Smith et al., 2004), Statistical parametric Mapping (SPM) (Penny et al., 2011) and FreeSurfer (Dale et al., 1999). These provide a way to reproducibly analyse large quantities of MR images and make comparisons across different patient groups. In medical image analysis, automated neuroimaging software packages are often preferred to manual methods, which are considered time-consuming and often result in poor reproducibility.

Below, some of the most commonly used neuroimaging software packages in MS are outlined.

FSL is a library of image analysis tools for both structural and functional data (Smith et al., 2004). FAST (FMRIB Automated Segmentation Tool) is one of FSL's tools for tissue segmentation and intensity correction (Zhang et al., 2001). It represents a step forward in image segmentation as its model relies not only on information from tissue intensity histograms to correctly classify each tissue type (WM, GM, CSF), but also takes into account the location of neighbouring voxels, thus improving the robustness of the software package. The MRI image is first stripped of non-brain tissue, including the skull, using BET (Brain Extraction Tool) and classed into three tissue types before the volume of each is calculated to allow across-individual comparisons. Another FSL package, FLIRT (FMRIB's Linear Image Registration Tool), allows linear (affine) image registration for the longitudinal study of brain atrophy, using 12 degrees of freedom (translation, rotation, scaling, and skew) (Jenkinson et al., 2002). FLIRT relies on the skull as the constraining factor under the assumption that skull size does not change across lifespan and works in half-space between the two input images to reduce interpolation blurring.

SIENAX and SIENA are two well-validated FSL packages respectively designed for cross-sectional and longitudinal assessment of neurodegeneration using a number of FSL programs (Smith et al., 2002). SIENAX uses BET, FLIRT and FAST, and it registers the brain image to standard space to obtain normalised brain volume values. SIENA also uses BET, FLIRT and FAST before estimating atrophy using edge motion; the boundaries between brain and non-brain tissues from two time

points are found and the shift between them is calculated to estimate the apparent tissue ‘growth’ and/or ‘atrophy’. The values are then converted into PBVC to calculate the rate of brain loss over time. This approach is similar to the boundary shift integral (BSI) method, another registration based method, which helps capture volumetric changes taking into account the motion of the boundaries of the brain (Freeborough and Fox, 1997). However, the BSI has not been adopted as widely in the MS research community and SIENA remains the preferred method of atrophy estimation.

SPM is another software package specifically designed for the analysis of medical images (Ashburner and Friston, 2005, Penny et al., 2011). SPM’s unified segmentation algorithm performs tissue classification, bias correction and registration. It utilises *apriori* information from tissue probability maps, including tissue maps for non-brain structures such as bone, based on intensity values to correctly classify WM, GM and CSF and avoid misclassification of non-brain tissue as brain tissue. Each tissue type is segmented using partial volume segmentation and their respective volumes are calculated. The total intracranial volume can then be calculated by adding the three values together. Longitudinal atrophy can be measured by calculating the PBVC between any two images from the same individual.

FreeSurfer is a tool developed for the analysis of human MR images (Dale et al., 1999). Similarly to SPM, FreeSurfer is a cross-sectional tool which can be used for longitudinal assessment of atrophy and involves motion correction, affine registration of the MR image onto standard space (12 degrees of freedom), non-uniformity

correction and removal of non-brain tissue (Fischl et al., 2002, Fischl et al., 2004). It also utilises location information to facilitate the correct classification of tissues before calculating the intracranial volume.

The software packages described above are freely available, well-documented and long-established tools designed for neuroimaging research in MS and other disorders. More recently, NeuroQuant® was introduced as the first commercial, FDA-approved and CE-marked automated tool for brain volume measurement in the clinic (Brewer, 2009). NeuroQuant® allows cross-sectional volumetric assessment of the brain and its individual structures using 3D T_1-weighted images and compares their values to those from a normative database adjusted for age, gender and intracranial volume (Luo et al., 2015). It also allows longitudinal assessment of atrophy in individual patients. NeuroQuant® appears to perform well compared to semi-manual segmentations (Brewer, 2009) and other automated software packages (Ochs et al., 2015, Wang et al., 2016, Ross et al., 2018), and represents a real step forward towards routine atrophy assessment in the clinics. Nevertheless, for now NeuroQuant® can only provide volumetric measurements using 3D images and it is not known how the results are affected by the use of different scanners. As a result, its clinical use remains out of reach for many MS centres.

As the various proposed approaches vary in the way they calculate the overall brain loss across time, they produce similar but not identical values. This is problematic, especially when assessing cerebral atrophy across short periods of time, where small pathological changes may be under- or over-calculated, depending on the software used. Various studies have attempted to quantify the error associated with

each software package by comparing their results to the gold standard (manual segmentations), or scanning the same individuals multiple times on the same day, in order to determine which of the tools for atrophy assessment is least prone to error.

Heinen and colleagues scanned a group of healthy individuals on the same day using two scanners, 1.5T and 3T, and used FSL (SIENAX, version 2.6), SPM (version 12) and FreeSurfer (version 5.3.0) to determine which software package was the most robust across different field strengths and when compared to manual segmentations (Heinen et al., 2016). They reported that when comparing volumes across different field strengths, calculating total brain volume was least prone to error with both Freesurfer and FSL, but not SPM, resulting in a mean absolute percentage difference (expressed as a percentage of the mean volume at 3T) of less than or equal to 1%. Moreover, the assessment of the accuracy of the automated software packages showed that the mean absolute difference between manual and automated measurements (expressed as a percentage of the mean manual measurement) was smallest for the SPM and FSL tools when computed using total brain volume values (3%). FSL was just as accurate when the results were calculated using intracranial volume. Considering the typical annual atrophy rates reported in MS patients, these findings demonstrate that atrophy estimates should ideally be performed over a period of a few years in order to minimise the measurement errors associated with the software packages.

Another study, looking at longitudinal brain volume loss measurements over one year (three time points) using two 1.5T scanners, also compared the most commonly used packages for atrophy estimation including SPM, FreeSurfer, SIENA and

SIENAX (Durand-Dubief et al., 2012). The authors found that registration based methods, including SIENA, significantly outperformed segmentation based tools when images from the two different scanners were used to calculate atrophy, resulting in less variability. Finally, a study which aimed to determine the sample size required for placebo-controlled trials, using brain atrophy as one of the end-points, found that SIENA was the most robust software, requiring the smallest sample size to detect significant treatment effects (Anderson et al., 2007a).

Out of the most widely used and validated packages it appears that SIENA is the most accurate and robust software for atrophy assessment due to its registration-based approach. Previous validation studies estimated that the median brain scan-rescan error for SIENA is 0.15% (Smith et al., 2002), suggesting that it is a reliable tool that can be useful for the detection of pathological levels of atrophy (De Stefano et al., 2016). Nevertheless, whilst SIENA has been validated and successfully applied in research studies, its performance on clinical-quality images requires further study. Routinely acquired MRI scans are often of suboptimal quality, compared to research scans, due to time constraints of hospitals. This can result in incomplete head coverage, lower spatial resolution and presence of artefacts on clinical MR images, which may negatively impact its performance.

2.2.2.5 Factors which can affect atrophy assessment

There is a growing interest in performing atrophy measurements as they contribute to our understanding of normal ageing and disease. The brain is a dynamic environment and as a result, various physiological effects can lead to temporary changes in brain volume and affect its accurate measurement. These include internal factors, such as the hydration status of the individual, and external factors

such as the software used to calculate atrophy. Below are described some of the most important factors which can affect brain volume measurements.

Disease-related changes

Previous research suggests that a variety of MS-related, pathological processes can lead to subtle changes in brain volume. These include oedema and inflammation, which can, in fact, lead to a temporary increase in brain volume. To my knowledge, the only study to test the short-term effects of inflammation by following patients on a monthly basis, identified that each T_2/FLAIR and/or Gadolinium-enhancing lesion was associated with a transient 0.1% increase in brain volume (Cheriyan et al., 2012). This suggests that brain atrophy measurements in patients with a high lesion load might potentially be masked and underestimated.

DMTs

The use of steroids and DMTs has been previously associated with *pseudoatrophy*, a perceived shrinkage of the brain as oedema and inflammation are reduced (Zivadinov et al., 2008). Vidal-Jordana and colleagues studied a group of patients on natalizumab who had their PBVC measured at year 1 and 2 (Vidal-Jordana et al., 2013); they found that on average, the rate of atrophy in the first year was twice as high as PBVC calculated in year two (-1.10% vs. -0.51%), an effect thought to be mediated by active WM lesions. Similar findings were reported in another study of natalizumab involving more frequent, six monthly scans, which showed that the PBVC in natalizumab-treated patients decreased as the study went on, with -1.7% change in the first six months, through -0.74% between months six through to 12 and, finally, -0.50% at months 12-18 (Magraner et al., 2012). Lastly, Miller and

colleagues studied almost 1,000 natalizumab clinical trial patients and demonstrated that during the first year, patients treated using the DMT experienced a greater reduction in brain parenchymal fraction (BPF; ratio of brain parenchyma to total intracranial volume) than placebo treated patients (-0.56% vs. 0.40%; p=0.0002) (Miller et al., 2007). This however changed in year two as natalizumab patients experienced an average reduction of their BPF equal to 0.24%, compared to 0.43% in the placebo group. These findings suggest that as the inflammation subsides following treatment with a DMT, there is a marked perceived brain volume loss, which far surpasses the typical values reported in MS literature. With the continuation of the treatment, the drug's protective effects can be detected in year two, as the atrophy rate falls below the levels of the placebo group.

Similar findings have been reported following the use of interferons with a mean change from baseline to 3 months equal to -0.95% (Dwyer et al., 2015b), much higher than the typical atrophy values reported in the literature. Another study showed that following treatment initiation with interferons, the mean PBVC from baseline to year 1 was -1.35%, whilst the change between baseline and year 2 and baseline and year 3 were equal to -1.48% and -1.68%, respectively (Frank et al., 2004). This indicates that the reduction in brain volume was most striking across the first year with more gradual decline over the following years (Frank et al., 2004). These findings were confirmed by another study of interferons, which showed that the relative BPF change was most pronounced in the first year (-0.69%) compared to later years (year 2= -0.38%; year 3=0.38%), with the highest atrophy rates experienced in the first four months after treatment initiation. This also coincided with

a reduction in the volume of T_2-weighted and gad-enhancing lesions (Hardmeier et al., 2005).

A steroid-mediated effect on brain atrophy measures has also been described as soon as one week post-treatment (Fox et al., 2005). However, this effect persisted in SP patients only over the eight week study, not RRMS, resulting in the most pronounced effects in SP patients at week 2, causing an almost one percent loss in BPF before decreasing to -0.5% at week eight.

DMT-mediated effects are very pronounced in the first year of treatment, with highest values reported in the first few months. They are believed to be mediated through changes in the WM and are associated with a reduction in the volume of WM lesions and oedema. As a result, it has been suggested that atrophy should be assessed at least six to 12 months post-DMT initiation (Zivadinov et al., 2016a), with another study suggesting that inflammation-mediated effects may confound atrophy assessment up to two years following the initiation of DMTs (Sastre-Garriga et al., 2015).

Diurnal fluctuations

Data from over 9,000 MR scans from MS and Alzheimer's patients demonstrated that the time of day has a significant effect on brain volume, with larger volumes reported in the morning (a change in BPF equivalent to 0.18% in MS and 0.44% in AD) (Nakamura et al., 2015). Another study of a small group of young healthy controls found that on average brain volume decreased at a rate of 0.1% every hour resulting in a total decrease of around 1.15% of from morning to evening, an effect

found to be driven by changes in the CSF, GM and WM (Trefler et al., 2016). As a result, patients in research studies, especially small studies, should ideally be scanned at the same time of day or be randomised to account for diurnal variations in brain volume.

Hydration

The level of hydration can lead to significant effects on brain volume, with severe dehydration leading to temporary volumetric changes in the WM and ventricles resembling those found in Alzheimer's disease (Streitbürger et al., 2012). Previous studies suggest that dehydration can lead to a decrease in brain volume of around 0.55% (Duning et al., 2005), equivalent to the expected volume changes in MS over the course of a year, although many other studies failed to find significant changes following periods of dehydration in healthy controls (Kempton et al., 2007, Streitbürger et al., 2012, Nakamura et al., 2014). On the other hand, increased hydration seems to have a more pronounced effect on brain volume, leading to an increase in brain volume equivalent to 0.36-0.72% (Duning et al., 2005, Nakamura et al., 2014). Based on the results of these studies, hydration and dehydration can lead to significant temporary volume changes and as such, it is recommended that patients do not consume excessive amounts of fluids or perform strenuous tasks at least 24 hours prior to scanning.

Scanner

Takao and colleagues studied the effect of inter-scanner variability on brain volume changes over a period of two years (Takao et al., 2011). 208 healthy subjects were scanned using 3T GE scanners of the same model and were randomly allocated to

be either scanned using the same scanner for their baseline and follow-up scan or the two different scanners. The results showed that even using the exact same scanner model could cancel out any longitudinal brain volume changes and this is most likely due to tissue contrast differences. However, using SIENA to correct for any differences in image geometry significantly reduced those effects resulting in a mean -0.46% change over 2 years, which is consistent with previous studies. Huppertz and colleagues investigated the inter-scanner variability on a single subject using a variety of different MR scanners. They found that the inter-scanner variability for whole brain measurements was consistently higher than that for intra-scanner variability, and the mean coefficient of variation for whole brain volume was equal to 3.78%, compared to 0.5% for intra-scanner comparisons (Huppertz et al., 2010). Nevertheless, a study of 62 Azheimer's disease patients and 74 healthy controls using six different scanners found that the scanners had less of an effect than the groups did, suggesting that with large enough samples it is possible to overcome the effects of inter-scanner variability (Stonnington et al., 2008). A similar study, using 3T scanners from the three most well-known scanner manufacturers, Siemens, Philips and GE, scanned six MS patients six times over the course of three weeks (Biberacher et al., 2016). They found that there was a significant effect of scanner on PBVC and that these effects were random, rather than systematic, across the different scanners. The authors concluded that reliable atrophy assessment in individual patients could only be possible over periods exceeding five years.

Head motion

Head movement during scanning can lead to the appearance of ghosting artefacts and blurring on MR images, which can negatively affect atrophy measurements.

Reuter and colleagues studied the effects of various head motions, including nodding (forward-backward movements), shaking (side-to-side movements) and free movements (in all directions) on PBVC, assessed by SIENA (Reuter et al., 2015). They scanned subjects twice on the same day with one scan being movement-free and the other subject to one of the movements during scanning. A control condition involved motion-free scans. They found that whilst scans taken when the subjects were still resulted in a mean PBVC equal to -0.1%, nodding and shaking resulted in a mean PBVC of around -1.0% and free movements resulted in mean PBVC of -2.4%. This shows that head movements can lead to an overestimation of brain atrophy equivalent to values expected to take place across two to three years.
Two other studies which simulated excess motion on MR scans found similar findings, with increasing levels of motion leading to higher overestimation of atrophy of up to -1.0% (Camara-Rey et al., 2006, Preboske et al., 2006).

Brain volume change in MS is the net effect of the pathological changes within the brain and the additional, biological and external factors that can lead to its over or under-estimation. Even before a patient is placed in the scanner, various factors, such as their hydration level or movement inside the scanner, can lead to a temporary increase or decrease in their brain volume, which will be detectable on the MRI scan. Finally, the choice of scanner and the software used to calculate atrophy could add to the variability, taking the final PBVC value further away from the ground truth. As a result, the variability introduced by these many factors challenges the idea to assess atrophy in individual patients in the clinic. For now, it appears that atrophy assessments should be performed using longitudinal approaches, at a group level, and ideally performed over a period of a few years in order to minimise error

and overcome the influence of temporary changes in the biological tissue. Despite the limitations, strong correlations between atrophy, clinical outcomes and treatment effects exist, making it a promising marker of future disability and disease progression in MS patients.

2.3 Imaging MS Pathology

Post-mortem studies used to be the only way of studying MS pathology. In fact, our fundamental understanding of MS disease processes is largely based on findings from the earliest studies using post-mortem tissue. Histology is still considered the 'gold standard' in MS research, yet, histopathological studies suffer from bias as the tissue used comes from autopsies thus representing only a snapshot of the overall pathology experienced by patients throughout their disease duration. Such tissue is typically characterised by end-stage disease processes as the samples often come from older patients with more aggressive disease (Ludwin, 2000). As a result, MRI offers an invaluable opportunity to study MS *in vivo.*

MRI is, arguably, one of the most important tools in aiding our understanding of MS which has transformed both clinical practice and research. Indeed, MRI has identified MS related changes in virtually all brain compartments including deep and cortical GM, WM, ventricles, and cerebellum as well as individual brain structures, most notably corpus callosum, hippocampus and thalamus (Ge, 2006, Popescu et al., 2013a).

Developed in 1970s, this technique utilises magnetic fields and radio waves to generate 2D and 3D images of internal organs, including the brain. Unlike computer tomography (CT) or positron emission tomography (PET), it is radiation-free.

MRI makes use of the properties of hydrogen atoms, which are found abundantly throughout the human body, especially in water and fat. Due to the unpaired proton in its nucleus (non-zero charge), hydrogen atoms spin randomly when no external magnetic field is applied causing their magnetisation vector to be equal to zero. Hydrogen atoms also have another property, known as spin angular momentum, which allows their protons and neutrons to precess (orbit around the nucleus) in a way resembling the shape of a cone. The rate at which the nuclei precess, is defined by the Larmor equation, and states that the stronger the external magnetic field, the higher the precession frequency. When placed in a strong magnetic field (called B_0), such as that created by MRI scanners, the magnetic field of the protons aligns along the scanner. Almost all of the hydrogen nuclei will position themselves parallel (low energy state), whilst few anti-parallel (high energy state), to the external magnetic field, resulting in a net magnetisation vector termed longitudinal magnetization. When a radio frequency (RF) wave is applied, this causes the protons to enter a high energy state, thus reducing longitudinal magnetisation, and start precessing in phase (phase coherence) resulting in transversal magnetisation. The Larmor equation specifies the frequency of the radiofrequency (RF) pulse to be applied, which needs to be the same as the frequency of the protons, to enable an effective energy transfer between the RF pulse and the protons; this is termed resonance. When the RF pulse is turned off, the longitudinal magnetisation increases (described by a constant T_1) and transverse magnetisation decreases (described by a constant T_2). T_1 is influenced by rate of energy transfer back to the surroundings as protons return

to their preferred low energy state whilst T_2 is affected by the inhomogeneities of the local and external magnetic fields, which cause the protons to get out of phase (precess at different frequencies). The released energy induces a current in the receiver coils in the scanner and the signal can then be amplified and processed to reconstruct an image of the scanned organ. As different tissues have different compositions that can facilitate or impede energy exchange and phase coherence, T_1 and T_2 values vary depending on the tissue imaged. Using multiple RF pulses (time to repeat; TR) and modifying the time interval (TI) between them can result in different image contrasts that can be used to extenuate differences between various tissues in the brain, e.g. normal-appearing and lesioned WM (Schild, 1992)

2.3.1 Ultra-high field MRI

UHF MRI includes scanners with field strength of 7T and more recently 9.4T, which are primarily used for research purposes. As of 2014, there were approximately 55 centres with access to UHF scanners worldwide (van Osch and Webb, 2014) and this number is expected to have increased since.

The field strength of a scanner refers to the static magnetic field created by its magnet; this can be measured using units of Tesla (T). Increasing the scanner's magnetic field strength results in increased quality of the MR images. This is because with increasing field strengths, the signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) increase, resulting in MR images with higher resolution, smaller voxel sizes and improved contrast. 7T MRI also leads to decreased scanning times (Trattnig et al., 2016).

UHF MRI offers a unique opportunity to study MS pathology in greater detail than possible through clinically available scanners. Previous research has found that 23% more WM lesions were detected on 7T scans compared to 1.5T scans (Kollia et al., 2009). Moreover, 7T studies of the CV sign have consistently found MS patients to have veins in over 80% of all WM lesions, which decreases to over 40% when detected on 1.5 or 3T scanners (Tallantyre et al., 2009, Sati et al., 2016). This is most likely due to the lower resolution of images and higher noise at lower field strength which impedes the detection of small vessels. 7T scanners have also proven unparalleled at cortical lesion detection compared to conventional scanners (Kilsdonk et al., 2016). Up until recently, it was believed that cortical lesions remained largely undetected on MRI, even when using UHF scanners, however, a 7T study combining histopathology and MRI demonstrated that with long enough MRI scanning times it is possible to visualise the vast majority of cortical lesions present on the corresponding histopathology sections (Lassmann, 2017).

Nevertheless, UHF imaging suffers from drawbacks too, besides the obvious high cost of around 10 million USD (approximately 7.7 million GBP) (Balchandani and Naidich, 2015) per scanner. High specific absorption rates (SARs) and field inhomogeneities limit the applications of 7T MRI. SAR provides a measure of the energy absorption rate by the human body in response to an RF electromagnetic pulse. Put simply, some of the RF energy released during scanning can be absorbed by the human body and cause tissue overheating. SAR increases with increasing field strength and various MR sequences which use 180°RF pulses, such T_2-weighted images, can quickly reach the SAR safety limits calculated by scanner manufacturers (Biller et al., 2014). Field inhomogeneities arise both within the main

magnetic field (B_0) and the RF field (B_1) leading to regional signal variation and/or the appearance of susceptibility on the image produced.

Currently the clinical use of 7T scanners remains extremely limited. Due to the high cost of UHF MRI machines and the high sensitivity of 1.5 and 3T scanners in detecting brain abnormalities, it does not appear that 7T scanners will be commonly used in patient clinics. It is therefore important that any research findings at UHF can be demonstrated and applied at lower field strengths in a typical hospital setting.

2.3.2 Clinical MRI

Since 2001 MRI has been used in MS diagnosis through demonstration of DIT and DIS and exclusion of other mimicking disorders (McDonald et al., 2001). More recently, these so called McDonald criteria have been simplified to allow a single scan to support a diagnosis of MS (Polman et al., 2011). MRI continues to be a fundamental tool in the monitoring of patients and as a secondary outcome measure in many phase III clinical trials of DMTs, which use MRI-based markers of disease activity, such the number of Gadolinium-enhancing and T_2 lesions, to assess treatment efficacy of new drugs (van Munster and Uitdehaag, 2017).

Almost all hospitals operate 1.5T or 3T scanners which are sensitive to the detection of WM lesions and signs of atrophy. The consensus guidelines released by the MAGNIMS group recommend that each patient suspected of having MS is scanned using the following mandatory MRI sequences: (1) axial proton density and/or T_2-FLAIR/T_2-weighted, (2) sagittal 2D or 3D FLAIR and (3) 2D or 3D Gad-enhanced (0.1 mmol/kg body weight) T_1-weighted, with 2D scans not exceeding slice thickness

of 3mm (Rovira et al., 2015). Other, optional sequences such as high resolution, isotropic T_1-weighted and/or diffusion-weighted images are also recommended, although the whole scanning session should not exceed 30 minutes. In a typical MS clinic at the Nottingham University Hospitals NHS Trust, the standard MR protocol used in patients suspected of MS consists of 1.5T or 3T coronal T_1-weighted spin echo (to assess the gross anatomy of the brain and aid lesion identification and characterisation), axial T_2-weighted turbo spin echo (to identify lesions), diffusion-weighted imaging (to help differentiate between MS and ischaemic lesions) and 2D sagittal FLAIR (to identify lesions, especially juxtacortical and periventricular).

The MAGNIMS recommendations aim to better define and optimise the use of MRI in MS and reflect the unmet need for protocol standardisation across different MS centres. This is important in light of ongoing discussions about automated, individual patient assessment of WM lesion load and cerebral atrophy in the clinic (Jain et al., 2015, Smeets et al., 2016). Nevertheless, despite the clear benefits to MS patients, conventional MRI fails to detect the more subtle MS pathologies, such as cortical lesions, leading to an incomplete picture of the disease. Moreover, despite its ability to visualise newly-developed lesions with the help of Gadolinium contrast agents, or the end-stage destruction of tissue in hypointense areas on T_1 images, MRI fails to distinguish between the different stages of lesion destruction in detail. For example, oedema, gliosis, mild and severe demyelination as well as remyelination all appear as areas of hyperintensity on T_2-weighted images (Katz et al., 1993, van Waesberghe et al., 1999, De Groot et al., 2001, Meier et al., 2007). Furthermore, the lack of protocol standardisation across and sometimes within clinics lead to MR variability making comparisons of scans across different clinics and individuals difficult.

Despite the lack of specificity of clinical MRI, biomarkers of disease activity such as WM lesions, can easily be identified on conventional MR images. Together with patient clinical information, they can be used to help guide clinical decisions to ensure the best available care is given to patients.

2.3.3 Patient registries

Patient registries are increasingly more common in MS, both clinically and in research, due to their ability to capture rich, large-scale, real-life data; this is important as there is a growing interest in performing studies using 'big data'. MS registries often allow long-term follow-up of a heterogeneous population of patients providing detailed demographic statistics which can then be correlated with clinical data, MR images or biological samples to help us better understand issues such as MS risk factors or biomarkers of disease activity and disability progression. A recent systematic review found that there were 15 major MS patient registries involving a total of over 113,000 patients (Aggarwal et al., 2016) whilst another study looking at European registries identified a total of 20 national registries (Flachenecker et al., 2014).

MSBase, based in Australia, is one of the largest worldwide registries of over 60,000 MS patients from 39 countries (MSBase, 2018a). Equipped with an electronic data entry system, this initiative uses prospectively collected clinical data to answer epidemiological questions in MS. MSBase makes use of the iMed software, used by many MS clinics, to identify patients who have consented for their data to be shared anonymously. The uploaded data includes information on patient age, sex, EDSS,

DMTs, multiple sclerosis severity score (MSSS), disease course and duration, age at onset and diagnosis, and relapse count. More recently, information about patient cognitive test scores and MRI images can also be uploaded onto the register and used for research. As a result the MSBase published 10 peer-reviewed journal articles in 2017 delving into topics including prognostic factors in CIS patients, long-term disability prediction or treatment effectiveness demonstrating that there is much that can be learnt from observational, clinical data (MSBase, 2018b).

In Europe, some of the most well-characterised MS registries are the Danish and the Swedish MS registries. The Danish MS Registry was created in 1956 and contains data on all patients diagnosed with MS since 1949, which as of January 2012, was equal to over 20,000 patients (Bronnum-Hansen et al., 2011). The high level of completeness (91%) and validity (94%) of the registry data (Bronnum-Hansen et al., 2011) is likely a result of an agreement between the Registry and Danish neurology departments and rehabilitation hospitals which send mandatory information reporting on patient personal and disease data as well as results of any paraclinical tests (MRI, CSF and evoked potentials); as lumbar punctures are a requirement for an MS diagnosis in Denmark, virtually no patients can be omitted from the register (Flachenecker et al., 2014, Koch-Henriksen et al., 2015). In 1996, with the introduction of DMTs in Denmark, the Danish MS Treatment Register was established. Its main purpose is to monitor the safety and quality of DMTs. The register holds information from over 8,000 patients whose data is updated at the start of the treatment, three months post-treatment initiation and every six months thereafter (Magyari et al., 2016). Together with the Danish MS Registry, these

databases are invaluable for epidemiological research of virtually the entire Danish MS population.

Similarly, the Swedish MS Register, established in 2001, is a government-funded initiative used in all neurology departments in Sweden which contains data from an estimated over 80% of the entire country's MS population (Hillert and Stawiarz, 2015). The register forms an integrated part of the clinical documentation system and its main purpose is to provide the best possible care for MS patients, including informed decisions with regards to DMTs.

In the UK, the UK MS Society piloted the MS Register in 2009 and there are now over 17,000 MS patients from 29 participating MS clinics (UK MS Register, 2018), including, as of August 2018, 764 patients from the Nottingham University Hospitals NHS Trust. The number of MS Register consented patients represents over 13% of the estimated UK MS population (Mackenzie et al., 2014, UK MS Register, 2018). The register incorporates both medical information recorded by healthcare professionals as well as self-reported information recorded through a web-based questionnaire platform. Data on patient demographics (age, gender, postcode and ethnicity), diagnosis (type, date, first clinical symptoms), education, employment and disability (EDSS) are recorded and can be shared anonymously thus providing a platform for big-scale research studies.

With increasing opportunities for data sharing and developments in machine learning, there is no doubt that large-scale, multi-centre studies will become more common in MS in the near future. Patient registries are important tools , which facilitate collection of anonymised patient data, which would be incredibly costly to

do in a research setting, and can be explored to answer important questions in MS that could lead to changes in clinical practice. Moreover, established registries are likely to play a crucial role in the long-term monitoring of DMT safety and efficacy as practical and economic reasons currently impede long-term follow-up of patients in clinical trials.

CHAPTER 3: THE CENTRAL VEIN SIGN IN MS AND MIMICKING DISORDERS; A MULTI-CENTRE MAGNIMS STUDY

This project was a European collaboration carried out by the MAGNIMS group with principal investigators Dr. Jens Wuerfel form the Medical Image Analysis Center (MIAC) in Basel, Switzerland; Dr. Nikos Evangelou, from the University of Nottingham, UK and Prof Alex Rovira from the Vall d´Hebron Research Institute in Barcelona, Spain.

Other contributors include: Dr T. Sinnecker (MIAC, Basel, Switzerland), Dr A. Pitiot (Laboratory of Image & Data Analysis, Ilixa Ltd., Nottingham, UK), Dr D. Meier (MIAC, Basel, Switzerland), Dr M. Calabrese (Department of Neuroscience, Biomedicine and Movement Sciences, Neurology Unit, University of Verona, Verona, Italy), Dr N. De Stefano & Dr A. Giorgio (Department of Medicine, Surgery and Neuroscience, University of Siena, Siena, Italy), Dr C. Enzinger & Dr R. Schmidt (Department of Neurology, Medical University of Graz, Graz, Austria), Dr L. Kappos (Department of Neurology, University Hospital, Basel, Switzerland), Dr X. Montalban (Division of Neurology, St. Michael's Hospital, University of Toronto, Canada and Department of Neurology-Neuroimmunology, Centre d'Esclerosi Multiple de Catalunya (Cemcat), Hospital Universitario Vall d'Hebron, Barcelona, Spain) Dr F. Paul (NeuroCure Clinical Research Center, Universitätsmedizin Berlin, Berlin, Germany), Dr M. A. Pawlak (Department of Neurology and Cerebrovascular Disorders, Poznan University of Medical Sciences, Poznan, Poland), Dr M. M. Schoonheim (Department of Anatomy and Neurosciences, VU University Medical Center, Amsterdam, the Netherlands).

Lesion analysis was largely carried out by two main raters, Dr Tim Sinnecker (Basel) and myself, who each analysed around a third of all the MRI scans in this study. Ann-Christin Pehrs, Manuel Huerbin and Kerstin Bendfeldt also contributed to the lesion analysis and altogether analysed the remaining third of the images. Statistical analysis of the data was carried out independently by the two teams, Basel (Dr Sinnecker and Dr Meier) and Nottingham (myself). The results presented in this thesis contain only the analyses conducted by myself. A comparison of the findings reported by Dr Sinnecker and myself is explored in the discussion section of this chapter.

3.1 Introduction

The CV sign is a proposed diagnostic biomarker of MS lesions, which has been visualised using 7T, 3T and 1.5T MRI and a variety of different sequences including venography, SWI, T2* and the newly proposed FLAIR* which combines FLAIR and T2* in an extra post-processing step (Tan et al., 2000, Tallantyre et al., 2009, Sati et al., 2012, Kau et al., 2013, Mistry et al., 2015, Samaraweera et al., 2016). This reflects the variability of MR protocols used in different hospitals and MS centres, which although similar, can result in variable image contrast, slice thickness and level of noise. This could, potentially, lead to difficulties translating the CV sign findings conducted in one centre with a specific set of protocols, to other MS clinics. In fact, inter-scanner variation can lead to a significant variability in lesion load measurements with inter-scanner coefficient of variation values ranging between 7 and 20%, depending on the protocol used (Filippi et al., 1999). Moreover, considering the small size of the cerebral blood vessels it is vital to test what effect different sequences and/or protocols have on the detectability of the CV sign.

Previously, Samaraweera and colleagues tested the detectability of the CV sign in a group of 10 MS and 10 small vessel disease patients using two different 3T scanners, a GE MR750 and a Philips Achieva, and four different protocols including T2*-weighted scans on both the Philips and the GE, a SWAN scan performed on the GE and a T2*-weighted MRI with high echo-planar imaging (EPI) factor of 15 echos on the Philips (Samaraweera et al., 2016). They found significant differences in the mean number of WM lesions detected in MS patients both between scanners and when using different sequences on the same scanner with the highest differences reported between the T2*-weighted GE protocol and T2* with high EPI Philips sequence, 15 and 66 WM lesions respectively. The Philips T2* with high EPI protocol overall detected the highest numbers of WM lesions, CVs and proportion of WM lesions with CVs, out of all the protocols. Despite the differences, the overall proportion of WM lesions with CVs could still be used to successfully class patients as MS and non-MS, irrespective of the scanner and protocol used. Although promising, this study was limited to a small sample of patients from a single centre. In fact, most of the studies of CVs so far have been from single centres, included small patient numbers and typically compared between MS and a group of patients with another condition that can mimic MS. In order to rigorously test its diagnostic value, the detection of the CV sign should be assessed in a large cohort of patients with a wide range of disorders, which can mimic MS, using clinically available scanners, including different scanner manufacturers, models and imaging protocols. To be of clinical value, the CV sign should be easily detectable across scanners and protocols; alternatively, the most optimal sequence for the detection of CVs should be identified.

Here, I present the first multi-centre study involving a large sample of patients to test the usefulness of the CV sign in differentiating MS from multiple other conditions that cause WM brain lesions and can be potentially confused with MS. In this study, we used 3T MR images from European MS centres with variable protocols consisting of co-registered FLAIR, for the detection of WM lesions, and SWI-, SWAN- or T2*-weighted images, for the detection of veins.

3.2 Aims

1. To compare the expression of the CV sign in patients with CIS, MS and MS-mimicking disorders and identify the optimal threshold of WM lesions with CVs to reliably differentiate between MS and non-MS disorders
2. To test whether a simplified approach involving small numbers of WM lesions with CVs can be reliably used to correctly classify patients as MS and non-MS
3. To test the inter-rater reliability of the CV assessment using a subset of lesions
4. To assess the effect of clinical and demographic characteristics on the expression of CVs in MS patients
5. To examine the effect of scanning parameters on the agreement between the clinical and CV diagnoses

3.3 Methods

3.3.1 Ethical approval

Local ethics approval was obtained at all the participating sites and all patients gave informed consent for their data to be shared anonymously.

3.3.2 Patient sample

Patients were recruited from eight MAGNIMS centres including Vall d'Hebron University Hospital (Barcelona), Universitätsmedizin (Berlin), Medical University of Graz (Graz), Poznan University of Medical Sciences (Poznan), University of Nottingham (Nottingham), University of Siena (Siena), University of Verona (Verona), VU University Medical Center Amsterdam (Amsterdam).

Patients between the ages of 18 and 65 with a variety of conditions that result in brain WM lesions were retrospectively included in the study. These included typical CIS with documented evidence of one demyelinating episode (excluding optic neuritis), RRMS; cerebral vasculitis; cluster headache; dementia; episodic migraine; metachromatic leukodystrophy; AQP4+ NMOSD assessed using one of the recently publicised assays (Huh et al., 2014, Geraldes et al., 2018); small vessel disease; SLE; Sjögren vasculitis and diabetes mellitus. Clinical disability in MS and CIS patients was assessed using EDSS.

3.3.3 MRI acquisition

Each patient had a 3T MRI brain scan consisting of an inversion recovery (IR) sequence, 3D FLAIR or TIRM (turbo inversion recovery magnitude), for the detection of lesions, and a GRE sequence consisting of susceptibility-weighted imaging (SWI), T_{2^*} or susceptibility-weighted angiography (SWAN), for the detection of veins. The scans were acquired using different scanners and protocols, reflecting the multi-centre setting of the study. Table 3-1 shows the individual specifications of the IR and GRE protocols included in the study.

Table 3-1 The MR scanning parameters for the vein (T2*/SWI/SWAN) and lesion sequences used by the eight MAGNIMS centres.

Centre	Number of analysable scan pairs contributed	Vein (GRE) sequence				Lesion (IR) sequence				
		Type	TE (ms)	TR (ms)	Spatial resolution (mm)	Type	TE (ms)	TR (ms)	TI (ms)	Spatial resolution (mm)
Amsterdam	40	SWI	23	31	0.49x0.49x3.0	3D FLAIR	125	8000	2350	0.98x0.98x1.2
Barcelona I	24	SWI	24.6	33	0.65x0.65x3.0	TIRM	93	9000	2500	0.49x0.49x2.99
Barcelona II	58	SWI	24.6	32	0.78x0.78x3.0	TIRM	93	9000	2500	0.49x0.49x2.99
Berlin	74	SWI	24.6	33	0.78x0.78x3.0	3D FLAIR	300	6000	2100	0.98x0.98x1.0
Graz	246	SWAN	59	68	0.90x0.90x4.0	TIRM	69-70	10000	2500	0.86x0.86x3.0
Nottingham	30	T2*	25	150	0.55x0.55x1.05 (with 1.05 gap)	3D FLAIR	289	4800	1650	1.0x1.0x1.0
Poznan	113	SWI	20	28	0.86x0.86x1.5	3D FLAIR	397	5000	1800	0.49x0.49x1.0
Siena	29	SWI	13.4	31	0.3x0.3x1.0	3D FLAIR	253	8000	2400	1.0x1.0x1.0
Verona	34	SWI	29	51	0.55x0.55x0.55	3D FLAIR	292	5500	1650	1.0x1.0x1.0

GRE – gradient echo; IR – inverse recovery; TE – echo time; TR – repetition time; TI – inversion time

3.3.4 Image pre-processing

Each IR image was registered to the GRE image using 3D Slicer version 4.6 (Fedorov et al., 2012) on MacOS X (10.11.4) on 5K Retina Display screens by Dr Meier (Basel). Image registration was based on the Insight Segmentation and Registration Toolkit (ITK) library (Johnson et al., 2007) and used rigid, affine and B-Spline (7x7x7 grid) registration.

Each co-registered pair of images was then divided into eight equally-sized, 3D blocks using the fslroi function of FSLUTILS, FSL (Smith et al., 2004) by Dr Pitiot (Nottingham) (see Figure 3-1). The parcelation of the brain was performed in order to limit the field of view and thus avoid rater bias in cases of uncertainty with regards to the presence or absence of a CV in a lesion. It is possible that a rater would be more likely to confirm the presence of a CV if most of the previously analysed lesions in that brain had had a CV. All of the blocks were subsequently randomised.

3.3.5 Lesion analysis

Prior to the analysis, the raters assessed a small number of scans together to ensure high consensus agreement on the presence of the CV sign in the lesions. Regarding my intra-rater agreement, it has been previously been reported using Cohen's Kappa as κ=0.95 (95% CI, 0.85-1) (Clarke et al., 2016).

3D Slicer version 4.6 (Fedorov et al., 2012) was then used to view an overlay of the IR and GRE images to help detect the presence of individual lesions and veins in each of the blocks by blinded raters (see Figure 3-2). Each WM lesion and vein was

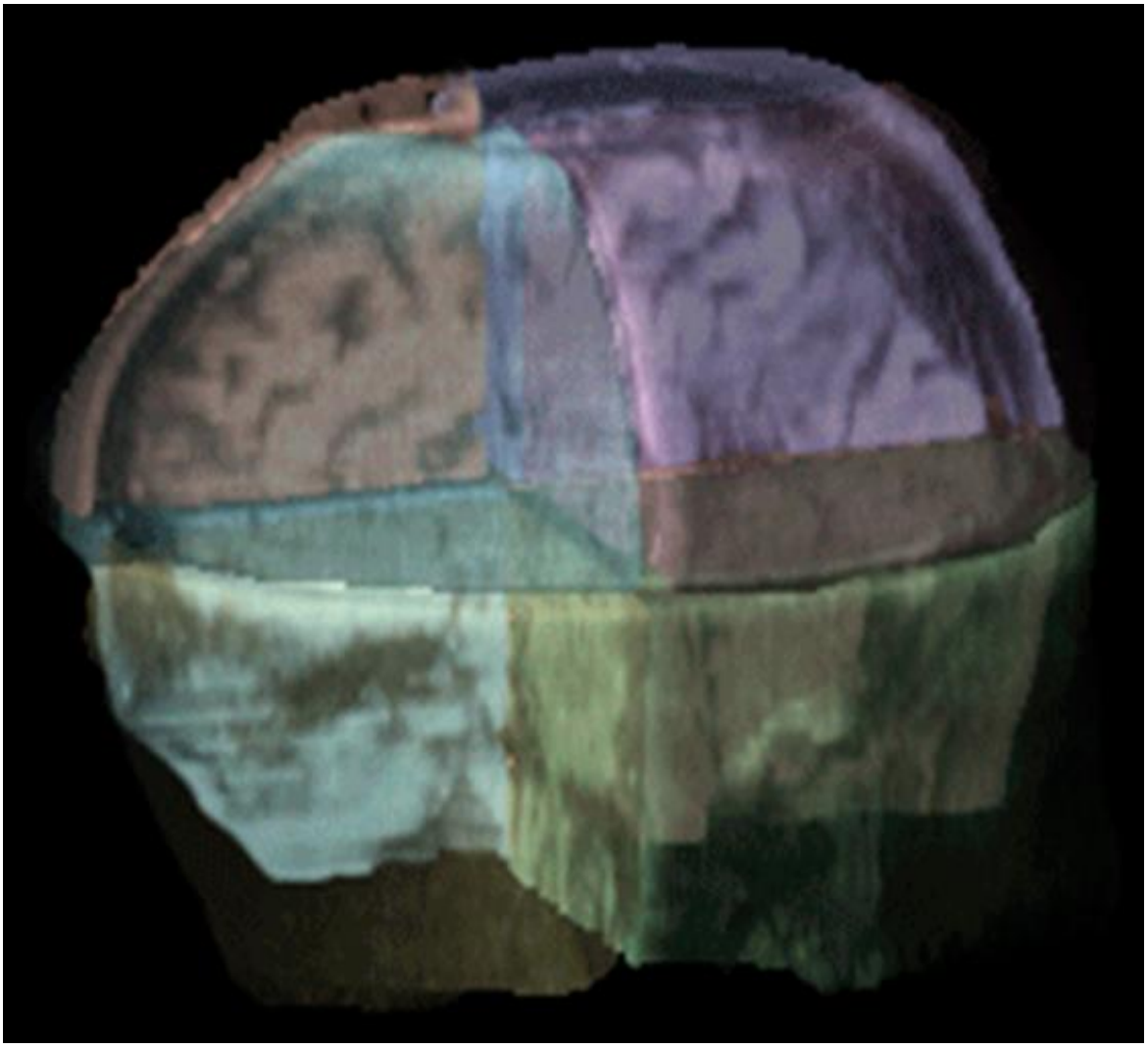


Figure 3-1 The parcelation of the brain into eight equally-sized blocks.

Each coloured part represents an individual block (eight in total), which was used to blind the raters to the distribution and characteristics of lesions in the other parts of the brain.

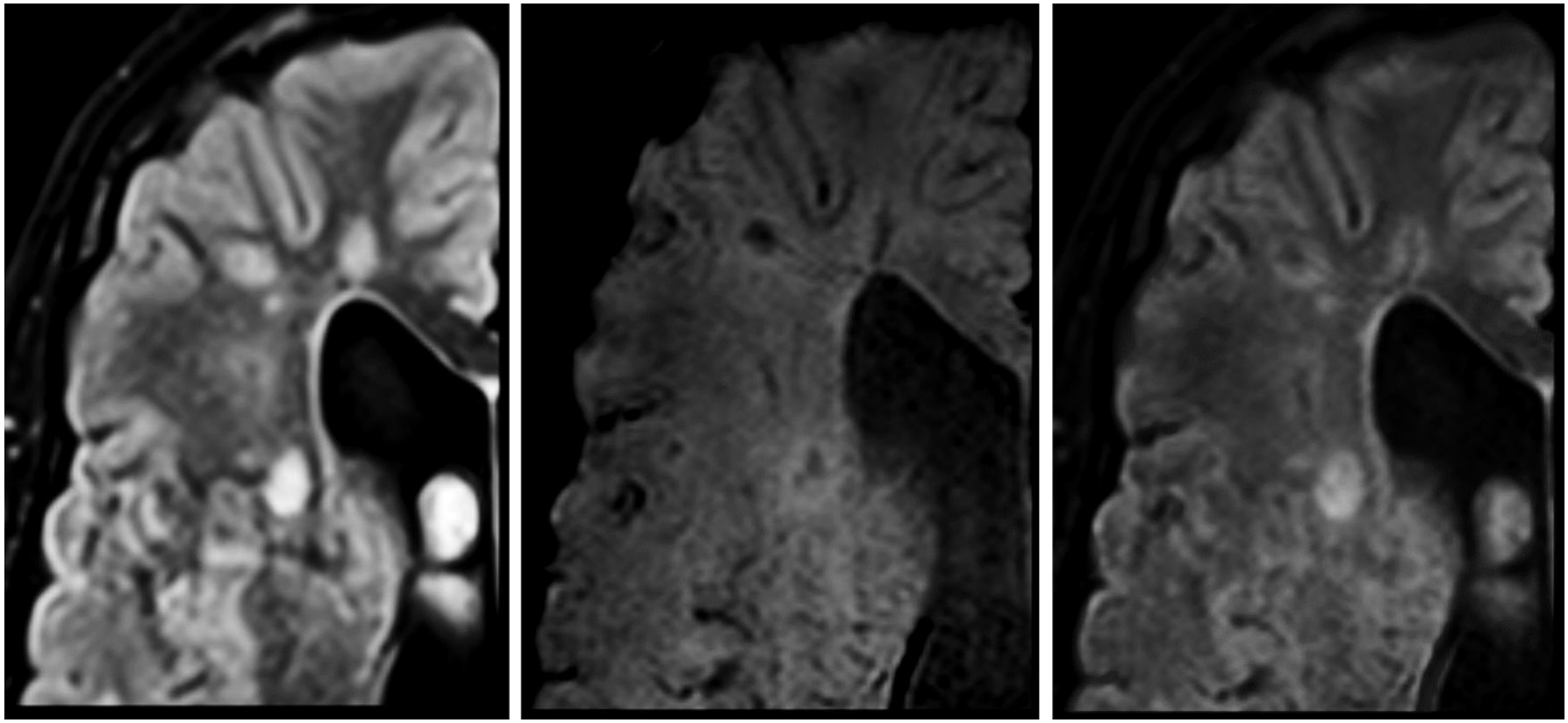


Figure 3-2 Images used to identify WM lesions and CVs.

(A) An example block from a FLAIR image showing multiple WM lesions. (B) The corresponding SWI image showing veins. (C) An overlay of FLAIR and SWI images which allows the visualisation of both lesions and veins. The raters were able to adjust the transparency level between the two types of images from 100% IR to 100% GRE image and any value in-between.

manually identified if they satisfied the pre- specified criteria defined in accordance to the guidelines released by the North American Imaging in Multiple Sclerosis (NAIMS) group (Sati et al., 2016) (see Appendix A*).* Each WM lesion was measured (in mm) alongside the long and short axes using the Annotations option, and classified based on their location as: juxtacortical (in direct contact with the cerebral cortex), periventricular (in direct contact with the ventricles) or subcortical WM lesion. Lesions, which were not fully visible on the GRE sequence (due to artefacts) or which were split between the different blocks, were not analysed.

3.3.6 Statistics

The demographic and clinical characteristics of all the patients were described using means and SD, unless otherwise stated. Normal distribution was assessed visually and an ANOVA was selected to test for inter-group differences with Tukey as a post-hoc test.

Agreement between the central vein sign with clinical diagnoses

I used MedCalc (https://www.medcalc.org/calc/diagnostic_test.php) to plot a receiver operating characteristic (ROC) curve, which was plotted as a function of sensitivity against the false positive rate (100 – Specificity) for different threshold values for the CV sign. Based on that curve, Youden's index J was used to find the ideal cut-off threshold for an MS diagnosis in this cohort (Youden, 1950). Briefly, Youden's index calculates the optimal threshold for a diagnosis, giving equal weight to sensitivity and specificity (sensitivity + specificity – 1). Youden's J statistic ranges between zero and one with a value of zero meaning the test cannot differentiate between the two groups at all whilst a value of one indicates the test can perfectly differentiate

between two disease groups. Based on that threshold, I calculated the optimal sensitivity, specificity, with their confidence intervals (CIs) and the positive and negative likelihood ratios (LRs). The sensitivity values were calculated as the number of true positives (patients with a positive CV sign and clinically diagnosed with MS) among all patients with a clinical diagnosis of MS, including true positives and false negatives (patients clinically diagnosed with MS who had a negative CV sign), expressed as a percentage. The specificity values were calculated as the number of true negatives (patients clinically diagnosed with a condition other than MS who also had a negative CV sign) among all patients subsequently clinically classed as non-MS, including true negatives and false positives (patients clinically diagnosed with a condition other than MS with a positive CV sign), expressed as a percentage. The CIs were calculated using exact Clopper-Pearson CIs (Clopper and Pearson, 1934). Positive LR was defined as the ratio between the probability of a positive test result given the presence of the disease and the probability of a positive test result given the absence of the disease i.e. true positive rate / false positive rate. Negative LR was defined as the ratio between the probability of a negative test result given the presence of the disease and the probability of a negative test result given the absence of the disease i.e. false negative rate / true negative rate.

Simplifying the central vein sign diagnostic rule

A simulated analysis of two simplified diagnostic rules of the CV sign was carried out using WM lesion information collected for each patient. The first one involved a simplified proportion approach whereby a fixed number of WM lesions, ranging from two to five was randomly picked, and a diagnosis of MS was proposed if a pre-defined number of lesions with CVs could be detected in that subset of lesions (e.g.

finding at least two lesions with CVs out of three randomly picked lesions). The second rule involved finding a fixed number of WM lesions with CVs, ranging from one to five, for each patient and establishing a diagnosis of MS if the minimum number of lesions with CVs could be found.

The R code used to test the two simplified rules, developed by Dr Chris Tench, can be found in Appendix B. The analysis for each rule was based on 100 random iterations to determine their mean sensitivity and specificity (and their standard deviation [SD]), taking into account the sampling error.

Inter-rater agreement

The inter-rater reproducibility, based on a subset of 76 WM lesions from 30 randomly selected patients, was expressed as % agreement between the raters.

Effects of clinical characteristics and scanning parameters on the CVS

A least absolute shrinkage and selection operator (LASSO) regression analysis was used to determine (1) the effect of clinical and demographic characteristics on the proportion of WM lesions with CVs in MS patients and (2) the effect of scanning protocol parameters on the agreement between the clinical diagnoses and those made according to the CV sign (using the ideal threshold calculated earlier). Briefly, LASSO analysis is used to select variables, through an L1 regulation method, that contribute significantly to the model whilst rejecting the ones that do not by shrinking their coefficient values to zero (Tibshirani, 1996). The remaining, significant variables can then be selected and used in, for example, a regression model. To carry out the LASSO regressions, the data were split into training (60%, 123 observations) and test (40%, 84 observations) datasets. A repeated cross-validation approach was

used whereby the training data were broken into 10 parts: nine parts for model validation and one for error estimation. This process was repeated five times.
The selected predictor variables to determine the effect of patient clinical characteristics were: gender, age, EDSS, and time since diagnosis. Time since symptom development was excluded as too many values were missing from the dataset. The effect of scanning parameters on the agreement between the diagnoses made using the central vein sign and the clinical were tested using the following predictor variables: GRE sequence used (categorical variables: SWAN, SWI, T2*; SWI and T2* were compared to the SWAN sequence), voxel volume of the GRE scan, slice thickness of the GRE scan, GRE sequence TR (repetition time), GRE sequence TE (echo time), voxel volume of the IR scan, IR sequence TR, IR sequence TE and IR sequence TI (inversion time).

All statistical analyses were carried out in SPSS version 24 apart from the LASSO regression (Tibshirani, 1996) which was done using the glmnet package (Friedman et al., 2010) in R (R Core Team, 2013).

3.4 Results

3.4.1 Patient sample

606 subjects, of whom 413 were female, with a mean age of 43 years (SD=16 years) were included in this study. In the MS group, median time since diagnosis (known for 227 MS patients) was equal to 5.4 years (range: 0-33.2 years) whilst median time since first symptom development (known for 166 MS patients) was equal to 7.0 years (range: 0-31.5 years). 108 MS patients were classed as ‘early MS’ with a disease duration of less than five years. 44 CIS patients had information about time

since first symptom developments, which was equal to 0.4 years (range: 0.1-4 years). Median EDSS score was 2 (range: 0-6.5) in the MS group and 1.5 (range: 0-4) in the CIS group. Diagnoses were unknown for 14 of the patients and their scans were subsequently excluded from the analysis.

3.4.2 WM lesion and CV analysis

119 subjects had no analysable lesions; specifically, 109 had no detectable lesions >3mm, eight had only confluent lesions and in two patients all the analysable lesions were split between different blocks during the blinding procedure. As a result, we analysed scans from 487 patients with a total of 4,447 WM lesions (see *Table* 3-2). In total, there were 1,857 lesions with CVs, 1,335 (71.9%) of which belonged to patients with a confirmed MS diagnosis and 374 (20.1%) to CIS patients. 91.3% of all the CIS and MS patients had at least one WM lesion with a CV whilst only 45.1% of the non-MS patients had one or more CV. Location-specific analysis of the proportion of WM lesions with CVs in the CIS/MS and non-MS groups can be seen in Figure 3-3. A χ_2 goodness-of-fit test showed that there was a significant difference in the distribution of lesions with CVs in CIS/MS patients, with most lesions found subcortically (n=897 [52.5%]) and periventricularly (615 [36.0%]) rather than juxtacortically (n=197 [11.5%]), $\chi_2(2)= 435.49$, $p<0.001$. The difference in the proportion of CV positive lesions was highest for periventricular WM lesions between the two patient groups.

A patientwise analysis of the data showed that the median percentage of WM lesions with CVs in MS and CIS patients was equal to 50% (range 0-100) and 60% (range 0-100), respectively (see *Figure 3-4*). The highest median proportion of CVs in non-MS

Table 3-2 Lesion-wise summary classified by disease group

Disease	Number of patients (n)	Median WML number (range) per patient	Total WMLs (n)	Total WMLs with CVs (n)	% WMLs with CVs (95% CI)
RRMS	225	10 (1-59)	2815	1335	47.4 (45.6-49.3)
CIS	98	5 (1-29)	690	374	54.2 (50.4-58.0)
Migraine*	29	3 (1-14)	121	21	17.4 (11.1-25.3)
Small vessel disease	88	2 (1-38)	473	75	15.9 (12.7-19.5)
SLE	14	2 (1-17)	54	11	20.4 (10.6-33.5)
NMOSD	13	2 (1-18)	54	9	16.7 (7.9-29.3)
Diabetes mellitus	20	8 (2-61)	240	32	13.3 (9.3-18.3)
Total	487	-	4447	1857	-

*Including 4 cluster headache patients

WML – white matter lesion; CV – central vein; RRMS – relapsing-remitting multiple sclerosis; CIS – clinically isolated syndrome; SLE – systemic lupus erythematosus NMOSD – neuromyelitis optica spectrum disorder; CI – confidence intervals

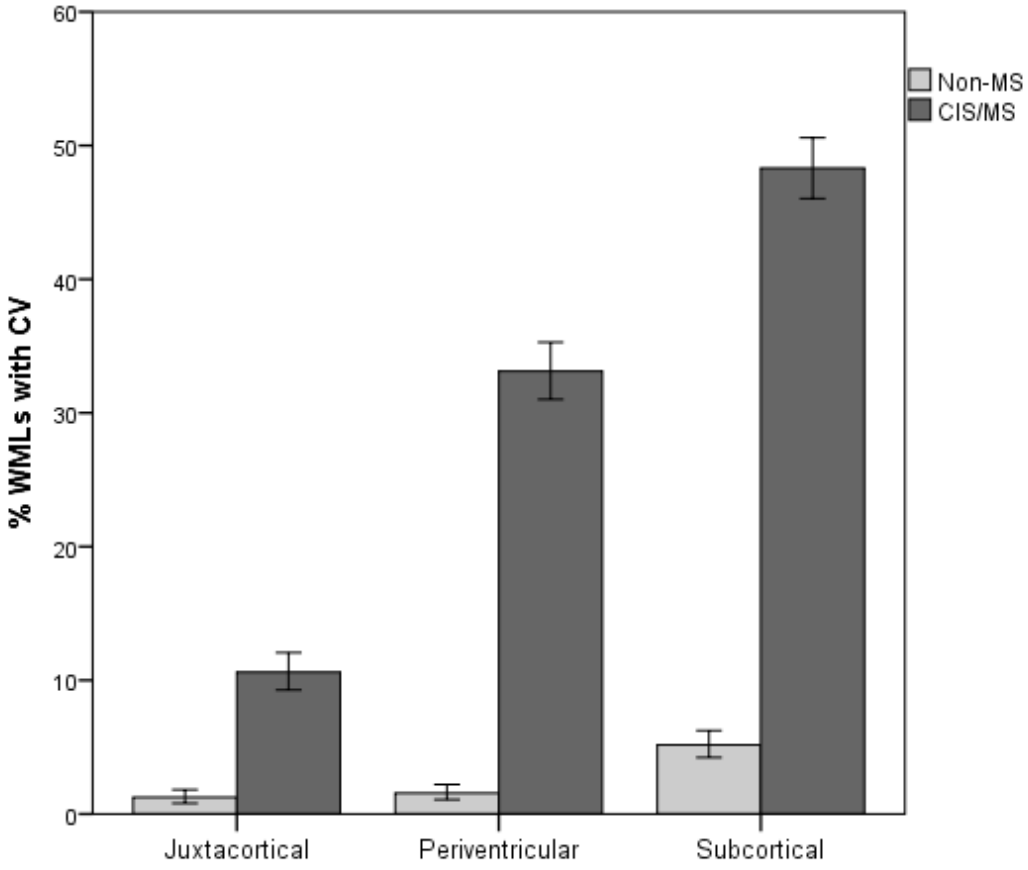


Figure 3-3 Location-specific proportion of WM lesions with CVs in non-MS and CIS/MS individuals expressed as percentage of al the lesions included in the study.

Error bars represent standard error. WMLs – white matter lesions; CV – central vein

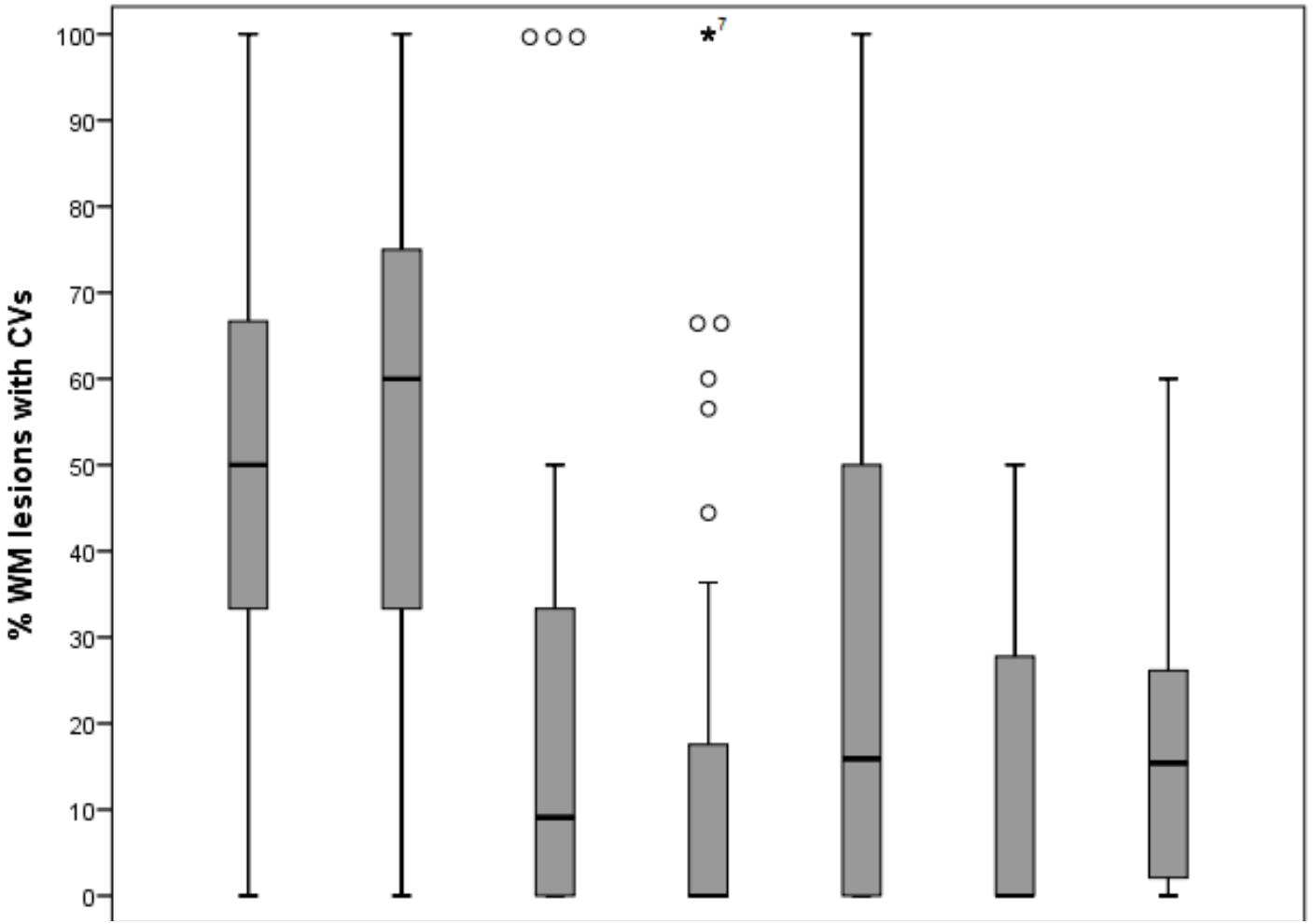


Figure 3-4 Patientwise proportion of WM lesions with CVs for each disease group.

Outliers with values more than 1.5 times the box length from the nearest hinge are marked as dots, whereas outliers with values more than 3 times the box length from a hinge are marked as asterisks. Numerical values have been assigned to >4 outliers with the same values.

WM – white matter; CVs – central veins; MS – multiple sclerosis; CIS – clinically isolated syndrome; SLE – systemic lupus erythematosus; NMOSD – neuromyelitis optica spectrum disorder

was observed in SLE (15.9%, range: 0-100%), closely followed by diabetes mellitus (15.4%, range: 0-60%). A one-way ANOVA showed there was a statistically significant difference in the proportion of WM lesions with CVs between the different disease groups $F(6,487)=25.568$, $p<0.001$.

3.4.3 Comparison between WM lesion and CV numbers in different disorders

The clinical diagnoses of patients were compared to the CV sign diagnoses made using a variety of proportion thresholds, ranging from 0 to 60%, which can be seen in *Table* 3-3. For each of the proposed threshold, the sensitivity, specificity and their 95% CIs, as well as the positive and negative LRs were calculated. An ROC curve was also plotted (see *Figure 3-5*) and the area under the curve (AUC) was calculated as 0.81 (SE=0.02; 95% CI=0.77-0.86, $p<0.0001$). Youden index J was equal to 0.58 when associated with a threshold of >25%. This resulted in sensitivity of 80.38% (95% CI, 75.6-84.6) and specificity of 77.91% (95% CI, 70.80-84.00). When the previously proposed threshold of 40% lesions with CVs was used, the resulting sensitivity and specificity were equal to 61.71% (95% CI, 56.1-67.1) and 84.05% (95%CI, 77.5-89.3).

3.4.4 Simplifying the central vein diagnostic rule

The sensitivity and specificity values for the simplified analysis of the CV sign using a simplified proportion approach can be seen in Table 3-4. The results of the second simplified analysis, which involved finding a fixed number of WM lesions with CVs to propose a diagnosis of MS, can be found in Table 3-5. Additionally, based on the results of the location analysis of lesions with CVs in CIS/MS, I also tested the performance of using a fixed number of subcortical WM lesions only (see Table 3-6).

Table 3-3 Sensitivity, specificity, positive and negative LR for different thresholds of WM lesions with CVs

Threshold (%)	Sensitivity (95% CIs)	Specificity (95% CIs)	Positive LR	Negative LR
0	100 (98.8-100)	0 (0-2.2)	1.00	-
>5	90.82 (87.1-93.8)	55.83 (47.9-63.6)	2.06	0.16
>10	88.92 (84.9-92.2)	57.67 (49.7-65.4)	2.10	0.19
>15	87.97 (83.9-91.3)	66.87 (59.1-74.0)	2.66	0.18
>20	84.81 (80.4-88.6)	72.39 (64.9/79.1)	3.07	0.21
>25	80.38 (75.6-84.6)	77.91 (70.8-84.0)	3.64	0.25
>30	79.11 (74.2-83.5)	79.14 (72.1-85.1)	3.79	0.26
>35	68.04 (62.6-73.1)	82.82 (76.1-88.3)	3.96	0.39
>40	61.71 (56.1-67.1)	84.05 (77.5-89.3)	3.87	0.46
>45	59.18 (53.5-64.6)	85.89 (79.6-90.8)	4.19	0.48
>50	46.52 (40.9-52.2)	89.57 (83.8-93.8)	4.46	0.60
>60	34.18 (29.0-39.7)	91.41 (86.0-95.2)	3.98	0.72

CIs - confidence intervals; LR – likelihood ratio

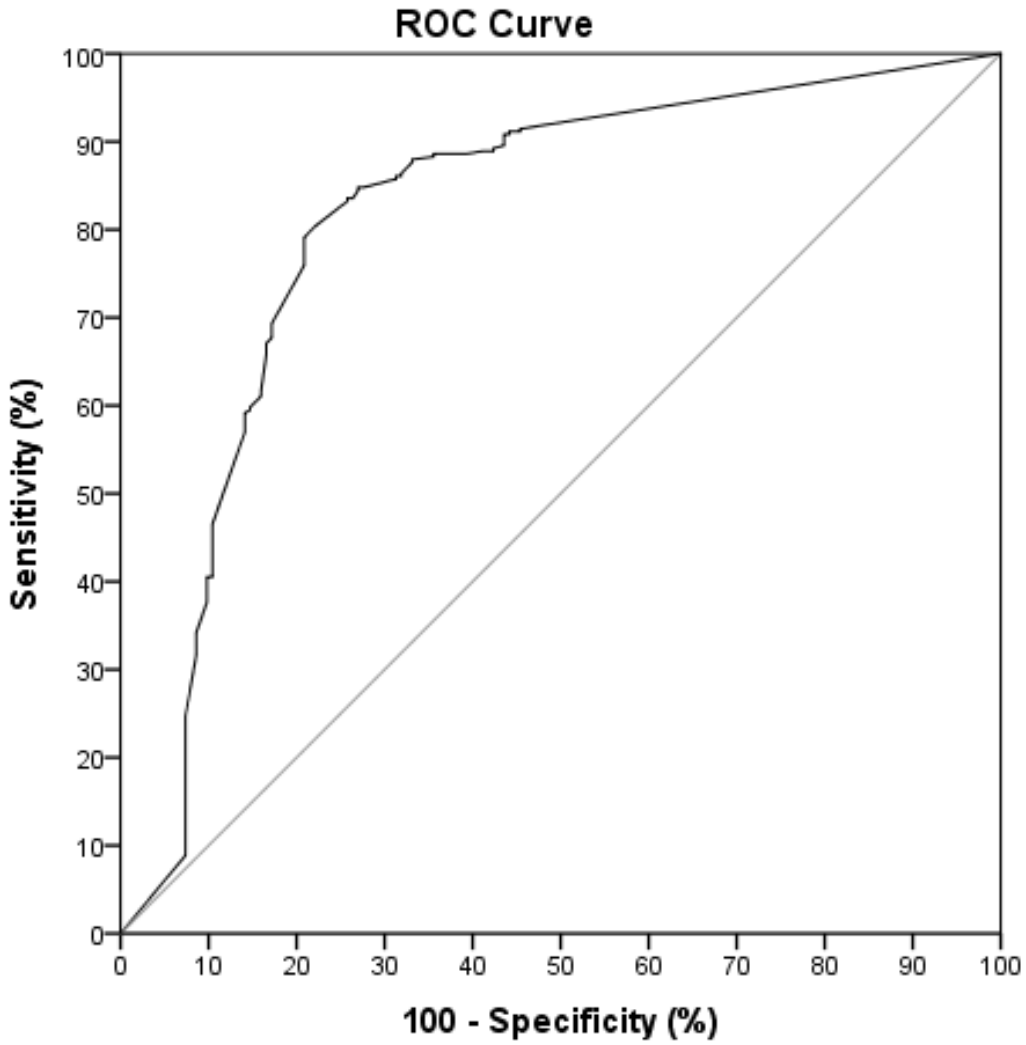


Figure 3-5 The receiver operator characteristics (ROC) curve (black) showing the sensitivity plotted against 100-specificity for the CVS to differentiate between MS and non-MS

Table 3-4 Sensitivity and specificity values for a range of simplified central vein sign diagnostic criteria using a subset of randomly picked WM lesions

No. of WMLs	No. of CVs	Mean sensitivity (%)	SD (%)	Mean specificity (%)	SD (%)
2	1	67.27	1.81	75.35	2.34
	2	33.02	2.20	88.86	1.49
3	1	76.34	1.51	70.88	2.03
	2	49.94	2.17	84.53	1.74
	3	23.67	1.89	90.77	1.05
4	1	80.95	1.61	68.20	2.14
	2	61.55	2.00	81.15	1.93
	3	38.28	2.02	88.43	1.29
	4	18.94	1.42	91.71	0.66
5	1	83.85	1.22	65.75	1.75
	2	69.13	1.98	77.75	1.97
	3	50.28	1.83	85.62	1.56
	4	30.65	1.88	90.16	1.00
	5	15.84	1.22	92.23	0.44

WMLs - white matter lesions; CVs – central veins; SD – standard deviation

Table 3-5 Sensitivity and specificity values for a range of simplified CV criteria using a fixed number of lesions with CVs

No of WMLs with CVs	Mean Sensitivity (%)	SD (%)	Mean Specificity (%)	SD (%)
1	85.40	7.72	61.73	6.57
2	75.51	18.26	69.90	10.73
3	64.59	27.15	76.29	12.05
4	54.89	31.89	81.40	12.24

Table 3-6 Sensitivity and specificity values for a range of simplified CV criteria using a fixed number of subcortical WM lesions

No. of subcortical WMLs with CVs	Mean Sensitivity (%)	SD (%)	Mean Specificity (%)	SD (%)
1	61.27	11.15	74.19	5.73
2	47.65	19.20	81.02	7.90
3	36.49	21.78	85.73	8.52
4	28.19	21.63	89.15	7.95

WMLs - white matter lesions; CVs – central veins; SD – standard deviation

3.4.5 Agreement between the raters

The two main raters (myself and Dr Sinnecker) agreed on the presence or absence of a CV in 64 out of 76 lesions resulting in 84.2% agreement. Rater1 identified 48 CVs whereas Rater2 identified 44 CVs. Patientwise comparison of diagnoses made by the two raters using the proposed threshold of over 25% WMLs with CVs resulted in two patients out of 30 being given a different diagnosis, which translates as 7% of the sample population. In both cases it was a difference of one lesion which led to a different overall proportion of WM lesions with CVs and hence a different diagnosis.

3.4.6 Relationship between MS patient clinical characteristics and the proportion of WM lesions with CVs

A LASSO regression was carried out on data from 207 MS cases (patients with missing values in any of the predictor variables were excluded) to select variables that contributed significantly to the proportion of WM lesions with CVs. The analysis established the optimal lambda value at 0.03 resulting in root mean square error (RMSE) of 0.25 and mean absolute error (MAE) of 0.21. R_2=0.12 and the pooled variables could explain approximately 6% of the deviance in the data with one of the variables, gender, accounting for 5 % of the variability Supplementary Figure 1 in Appendix C.

3.4.7 Effect of scanning parameters on diagnostic accuracy

A LASSO regression was carried out on data from 479 patients to predict the agreement between the clinical and CV sign diagnoses. The optimal lambda

value was equal to 1, resulting in accuracy of 77.7% and Kappa=0. The pooled variables could explain around 7% of the deviance in the model. The choice of the GRE sequence was the most important predictor, with the Nottingham T2*-weighted protocol showing increased agreement between the clinical and CV sign diagnoses compared to the other GRE sequences, although none of the model variables were deemed significant (Supplementary Figure 2 in Appendix C).

3.5 Discussion

This is the first multi-centre study of a large cohort of subjects with CIS, MS and mimicking disorders using MRI data from European MS clinics, to test the diagnostic utility of the central vein sign. Here I demonstrate that the CV sign alone can successfully discriminate between MS and a range of other conditions, which also result in the appearance WM lesions, and consequently can be a useful biomarker of MS lesions. In this large cohort, CV positive lesions were present in half of all the MS WM lesions whereas none of the other disorders exceeded a median proportion of 20% thus confirming previous findings at 3T.

3.5.1 The central vein sign as a diagnostic marker of MS

Overall, we achieved 80% sensitivity and over 78% specificity when applying the CV sign to differentiate between CIS/MS and non-MS using a threshold of over 25% WM lesions with CVs. A recent study of over 200 CIS patients showed that the 2017 McDonald diagnostic criteria for MS had a sensitivity and specificity of 68% and 61%, respectively (van der Vuurst de Vries et al., 2018).

The 2017 criteria resulted in more diagnoses made at baseline, compared to the 2010 criteria, with two in three extra diagnoses made thanks to the addition of OCBs testing. Compared to the 2010 criteria, the 2017 criteria had a higher sensitivity (68% vs. 36%) but their specificity was lower (61% vs. 85%), an effect attributed to the addition of OCBs. The results, reported here, indicate that the CV sign outperforms the current diagnostic criteria resulting in increased sensitivity and specificity and as such has the potential to be applied clinically to help differentiate between MS and non-MS.

A recent study found that the inappropriate application of the McDonald criteria was one of the main factors contributing to MS misdiagnosis (Solomon et al., 2016). The current diagnostic criteria require a level of expertise in correctly identifying clinical and radiological findings that are typical of MS and rejecting any symptoms or MRI abnormalities inconsistent with MS. In this study, the CV sign diagnoses were based solely on the information about the proportion and/or number of lesions with CVs suggesting the CV sign is robust in differentiating between MS and non-MS irrespective of the clinical and radiological presentation of a demyelinating disorder and requires less expertise to correctly diagnose patients.

3.5.2 Simplifying the rule

However, if the CV sign is to be used in clinical practice, its application should be simplified. Counting all the WM lesions and then calculating the proportion of lesions with CVs would be impractical in a clinical setting, hence I tested two simplified diagnostic rules, which involved only assessing a small number of

WM lesions for each patient. When testing a simplified proportion approach of ‘1 in 3’ lesions, whereby one out of three analysed lesions had to have a CV in order to propose a diagnosis of MS, the sensitivity and specificity were 76% and 71%, respectively. However, in clinical practice, selecting a random number of lesions might not be practical as doctors might be predisposed towards picking lesions with particular characteristics or in particular locations. As a result, the second simplified analysis involved finding a minimum number of lesions with CVs to make a diagnosis. Finding two WM lesions with CVs resulted in similar sensitivity and specificity values, 76% and 70% respectively, demonstrating the robustness of the CV sign for diagnostic purposes. Focusing solely on subcortical lesions (WM lesions which were not juxtacortical or periventricular) increased the specificity to 81%, at the expense of sensitivity which was equal to 48%. As the existing diagnostic criteria are more sensitive than they are specific, it may be clinically useful to use a threshold, which favours specificity over sensitivity, instead of giving equal weighting to both. This could prove useful if the CV sign were to be used as an additional test to help diagnose MS, similarly to OCBs. As such, finding four or more WM lesions with CVs would result in very high specificity values of over 80%, which increases to over 90% if four subcortical lesions with CVs can be identified. This could help reduce the number of patients misdiagnosed with MS and reducing the risk of commencing them on expensive MS DTMs with potentially significant side effects. Ultimately, the decision about the minimum number of lesions with CVs should be left to the neurologist who, taking into account both the clinical symptoms and radiological findings of each individual patient, can decide how many CVs are ‘enough’ for a diagnosis.

3.5.3 The central vein sign in MS and other disorders

In this study, CIS and MS patients had the highest proportions of WM lesions with CVs. The CIS cohort consisted of patients with a single demyelinating event (excluding optic neuritis), indicative of CNS inflammatory, demyelinating disease (Polman et al., 2011). According to a large, prospective study, nine in 10 patients CIS patients with abnormal baseline MRI will develop MS within 15 years (Brownlee et al., 2019). The high median proportion of CVs in that group suggests that the CV sign could be useful in confirming the demyelinating nature of CIS; however, only long-term follow-up of this cohort of patients will be able to determine the value of the CV sign in those patients. As a small proportion of the CIS patients will, undoubtedly, be subsequently diagnosed with other conditions, it would be interesting to see whether those patients were also the patients with the lowest proportions of CVs. The value of the CV sign in diagnosing patients with CIS, as detected on 3T MRI, remains to be determined.

SLE had the highest proportion of CVs, outside of CIS and MS, with a quarter of all SLE patients having CVs in more than half their lesions. SLE can lead to the formation of vascular lesions in the brain, whereby immunomolecules, such as immunoglobulins, can aggregate in the walls of small and medium veins leading to their narrowing and blockage, leading to micro-haemorrhages, which appear dark on SWI (Cieslik et al., 2008). Prior findings suggest that the use of MRI remains limited in distinguishing between MS and SLE (Kim et al., 2014), however, previous studies of the CV sign in SLE did not report a high proportion

of WM lesions with CVs in that disorder. Further studies using a larger number of SLE patients are needed to establish whether the CV sign is a common feature of lupus.

In total, 12 non-MS patients had 100% of WM lesions with CVs. A retrospective review of these patients showed that nine had only one WM lesion and the remaining three patients had two WM lesions suggesting that caution should be exercised when analysing patients with a small lesion load and other clinical and radiological information should be taken into account when proposing a diagnosis in those patients.

3.5.4 Inter-rater reliability

In this study, the two main raters used for the assessment of the lesions demonstrated high inter-rater agreement of 84% indicating that the CV sign can be successfully used by various raters (i.e. radiologists in clinical practice) to help differentiate between MS and non-MS lesions. However, as the study protocol dictated that the lesions and veins be classified according to the guidelines proposed by the NAIMS initiate (Sati et al. 2016) some training would be required to ensure radiologists follow the same practice. Care should be exercises in patients with a small lesion load as any disagreement about the number of CVs could translate into disagreement regarding the diagnosis proposed to the patient.

3.5.5 Clinical and MRI effects on the detectability of CVs

Clinical and demographic characteristics did not explain the differences in the percentage of CVs in individual MS patients, accounting for as little as 5% of the variability detected between patients. Regarding the MR protocol effects, the T2*-weighted protocol from the Nottingham centre seemed to be the most optimal sequence for the detection of the CV sign adding to the previous findings whereby the same T2*-weighted protocol was found to differentiate between MS and non-MS subjects with the highest accuracy (Samaraweera et al. 2016). Nevertheless, there were only 30 scans contributed from that centre, which is likely why the results were non-significant despite an upwards trend. More patients scanned using that particular sequence would be needed to establish its superiority compared to the other GRE sequences such as the SWI and SWAN. Even so, I found that the different MR protocols added little to being able to predict the agreement between the clinical and CV sign diagnoses, accounting for around 7% of the deviance in the model. This suggests the central vein sign is readily detectable across a range of sequences which facilitate the detection of cerebral WM lesions and veins. This, in turn, would make the potential addition of the CV sign into clinical practice efficient as all hospitals already use FLAIR and some use a GRE sequence as part of a possible-MS MR scan. Thus, these findings further suggested that the CV sign is detectable on clinical quality scans in a multi-centre setting where different scanners and protocols are used and is not largely influenced by individual patient characteristics.

3.5.6 Comparison of the results

The results of this study were analysed independently by researchers from Basel, namely Dr. Tim Sinnecker and Dr. Dominik Meier (Sinnecker et al., 2019), and Nottingham (myself). Our results and conclusions are very similar yet there are a few notable differences. Firstly, the Basel team picked a 35% threshold of WM lesions with CVs to differentiate between MS and non-MS as they wished to favour specificity over sensitivity (82.9% and 68.1%, respectively vs. 77.9% and 80.4% using the >25% threshold). Their suggested simplified 3 lesion criteria had a sensitivity of 61.9% and specificity of 89.0%, which are higher than the values reported in this analysis (finding 3 lesions with CVs was associated with sensitivity and specificity values of 64.6% and 76.3%, respectively). This is likely due to the differences in methodology as the Basel team reported their results based on a single assessment of the reported results, whereas I used a simulation based on 100 iterations and thus I also took into account the sampling error associated with the simplified rules. Furthermore, the results of the LASSO analysis predicting agreement between the clinical and the CV sign diagnoses showed differences too with exactly the same variables explaining up to 12% of the deviance in the model when the 35% threshold was applied, whereas only up to 7% could be explained using the 25% threshold. Similar predictors were identified by both models with the T2* sequence consistently contributing to higher agreement between the clinical and central vein sign diagnoses. Overall, the two independent analysis allow us to conclude that the CV sign is a robust diagnostic biomarker of MS lesions, largely irrespective of the GRE or IR sequence used, and that simplified lesion

criteria can be successfully applied in a clinical setting to classify patients as MS and non-MS.

3.5.7 Study strengths

One of the main strengths of this study is its multi-centre setting which allowed the collection of MR scans from hundreds of individuals with a variety of conditions using MR images from variable scanners, thus providing a real-life setting reminiscent of that in large MS clinics. The blinding approach used, which involved parcelation of the brains and their subsequent randomisation, reduced potential bias resulting from the lesion distribution patters and/or the raters' previous decisions concerning the presence or absence of CVs. Nevertheless, most of the patients used in this study had already established diagnoses and as such, a prospective, multi-centre study is needed to further validate the role of the CV sign as a diagnostic biomarker in MS.

3.5.8 Study limitations

This study also faced limitations. Firstly, although useful, the CV sign did not lead to the correct diagnosis of all the patients. Previous smaller studies, at both 7T and 3T, have resulted in 100% sensitivity and specificity (Tallantyre et al., 2009, Tallantyre et al., 2011, Wuerfel et al., 2012, Mistry et al., 2013, Mistry et al., 2015, Maggi et al., 2018). However, all of the previous studies have included small samples of patients and the raters were not blinded to the scans as rigorously as was the case in this study. Despite the imperfect distinction between MS and non-MS in our study, the overall diagnostic accuracy was very good and resulted in higher sensitivity and specificity values than the current

diagnostic criteria. Another limitation was due to our parcelation approach. Some lesions had to be excluded from the analysis, as they were positioned between different blocks, which is the likely reason why some of our patients had no detectable lesions. Finally, although the MRI data consisted of variable GRE and IR protocols, some protocols were clearly under-represented. As a result, future, prospective studies should aim to study the effects of different imaging protocols in a more balanced sample. It would also be interesting to compare standard protocols, such as those used in our study, to optimised sequences such as the previously reported FLAIR* sequence (Campion et al., 2017).

3.5.9 Conclusion

The CV sign appears to be a highly sensitive and specific marker of MS lesions, which works well when using a proportion approach and with a simplified approach when only two WM lesions with CVs were used to make a diagnosis of MS. The results from this chapter demonstrate that the CV sign is detectable across variable, real-life, clinical-quality scans, with high inter-rater agreement, and could therefore be used as an additional test for the diagnosis of MS. Patient clinical characteristics did not have an effect on the expression of the CV sign and MRI protocol parameters did not have a significant effect on the detectability of the CV sign demonstrating the robustness of the CV sign as an MS imaging biomarker. As the current findings are encouraging, a large, multi-centre, prospective study is needed to further evaluate the diagnostic accuracy of the CV sign in newly presenting patients.

CHAPTER 4: STAR-MS: SINGLE TEST TO ARRIVE AT MS DIAGNOSIS; A PROSPECTIVE PILOT STUDY ASSESSING THE ACCURACY OF THE CENTRAL VEIN SIGN IN PREDICTING MULTIPLE SCLEROSIS IN CASES OF DIAGNOSTIC UNCERTAINTY

4.1 Introduction

There is no single diagnostic test for MS. The McDonald criteria were specifically developed in order to facilitate a diagnosis in patients with a typical CIS and high probability of MS (Thompson et al., 2018), not as a tool to differentiate MS from other disorders. The 2017 criteria (Thompson et al., 2018) continue to incorporate the principle of 'no better explanation' for neurological symptoms experienced by patients that might have a demyelinating disease; when patients fulfil the proposed criteria and no other disease can explain their symptoms, the diagnosis of MS is offered. These criteria, and their most recent predecessors (2010 McDonald diagnostic criteria (Polman et al., 2011)) work well in patients with a typical presentation of demyelination often allowing a quick diagnosis using a single MRI scan. This is important as early treatment is beneficial by reducing long-term disability (Goodin et al., 2011, Butzkueven et al., 2014). Nevertheless, despite, and possibly due to early application of those criteria, the rate of misdiagnosis in MS remains high (Solomon and Corboy, 2017). Solomon and colleagues showed that 95% of 122 American neurologists had seen at least one misdiagnosis, with 40% of the neurologists reporting between three and five patients misdiagnosed in a year (Solomon et al., 2012). Sometimes misdiagnoses arise when patients experience symptoms consistent with MS, however, their MRI findings are not typical of MS. On

the contrary, patients with atypical symptoms, can have MRI findings consistent with MS (Shetty et al., 2014). As a result the specificity of MRI in MS, despite the improvement offered by the application of specific MRI criteria (Filippi et al., 2018), is still low and patients often have to undergo a lumbar puncture, an invasive test that does not always result in a diagnosis. This highlights a real need to explore additional MS-specific biomarkers which could help differentiate MS from other mimicking conditions.

The CV sign is frequently seen in MS lesions on MRI, whereby demyelinating plaques form perivascularly resulting in what has been described as a ‘doughnut’ or ‘coffee bean’ appearance on GRE MRI images, due to the bright appearance of the lesion and a dark CV (Mistry et al., 2013). Its presence has been demonstrated at UHF, where over 80% of all WM lesions have consistently been shown to have CVs, (Tallantyre et al., 2008, Sinnecker et al., 2012b, Wuerfel et al., 2012, Kuchling et al., 2014) and at clinically available field-strengths (1.5T and 3T), where it is typically detected in over 40% of all WM lesions in MS (Tan et al., 2000, Tallantyre et al., 2009, Mistry et al., 2015, Samaraweera et al., 2016). However, so far most studies have mainly involved patients with already established diagnoses. Mistry et al. prospectively studied the usefulness of the CV sign in patients with diagnostic uncertainty at presentation (Mistry et al., 2013). Using a single UHF field 7T scan, all patients were subsequently diagnosed as MS and non-MS using the previously mentioned 40% threshold achieving 100% positive and negative predictive values (PPV and NPV, respectively) with respect to the diagnosis from follow-up clinical scans. Kau and colleagues (Kau et al., 2013) reported the analysis of 28 lesions in

14 young patients who presented for MRI review and found that the CV sign was useful in discriminating between MS and non-MS lesions.

Here, I report the first prospective, investigator-blinded pilot study of the CV sign in patients with diagnostic uncertainty presenting to Neurology and/or MS clinics. I report the accuracy of using the CV sign at the time of disease presentation, in predicting a final diagnosis of MS, using a single, clinically available, 3T T2*-weighted MRI scan.

4.2 Aims

1. To prospectively test the usefulness of the CV sign in patients with diagnostic uncertainty at presentation using a single 3T T2*-weighted scan
2. To calculate the optimal proportion of WM lesions with CV to make a diagnosis of MS in patients presenting with atypical CIS
3. To calculate the percentage agreement between blinded-raters when using the CV sign to diagnose patients
4. To calculate the rate of patient recruitment for this pilot study to inform future multi-centre studies

4.3 Methods

The study received ethical approval from the local research ethics committee (Research Ethics Committee reference number: 15/NW/0286). The full study protocol can be found on https://www.clinicaltrials.gov using the study reference: NCT02485223. Written informed consent was obtained from all patients prior to

participation in the study. I followed the STARD protocol (Cohen et al., 2016) for reporting the results of a diagnostic test.

4.3.1 Patient sample

In this prospective study, patients were recruited from the MS or general neurology clinics at the Nottingham University Hospitals Trust, UK by their MS specialist neurologist between May 2015 and February 2018. The inclusion criteria comprised patients aged over 18 years old with WM lesions and possible MS where a) the neuroradiologist's report was not in line with the clinical impression of a typical CIS presentation (such as unilateral optic neuritis, focal brainstem or cerebellar syndrome, myelitis) or b) where the symptoms experienced by the patient were unusual for an MS presentation (atypical headaches, cognitive concerns, fleeting sensory symptoms etc.), however, their MRI was consistent with the disease. Patients fulfilling the above criteria, identified by the study team, were recruited sequentially.

4.3.2 MRI acquisition

MRI scans were obtained using a 3T Philips Achieva scanner (Philips Healthcare, Best, The Netherlands) with a 32 channel head coil. The protocols included a 3D FLAIR (TR=4800ms, TE=291.87ms, TI=1650ms, 256x256x180 matrix, 1x1x1mm voxels, flip angle 90°, pixel bandwidth=1221, 1mm thick slices, acquisition time round 2.25 minutes) and a T2*-weighted scan with high EPI factor of 15, similar to Sati and colleagues (Sati et al., 2014): TR= 54ms, TE=29ms, 448x448x336 matrix, 0.54x0.54x0.55mm voxel size flip angle 10°, acquisition time around 4.23 minutes.

4.3.3 Lesion analysis

All scans were analysed using in-house image analysis software NeuROI (www.nottingham.ac.uk/research/groups/clinicalneurology/neuroi.aspx). The T2*-weighted scan was used to assess the presence of WM lesions with CVs and the 3D FLAIR scan was used to aid lesion identification. An experienced rater (myself) analysed the lesions by manually placing ROIs over individual lesions on T2* scans and calculating the proportion of WM lesions with CVs for each patient. A second experienced rater (Dr Yasser Falah) also analysed the lesions in order to calculate inter-rater reproducibility. Each WM lesion was assessed on the axial plane and the presence of a CV was then confirmed by establishing its presence on one of the other two orthogonal planes, as recommended by the NAIMS cooperative (Sati et al., 2016). Lesions smaller than 3mm in their shorter axis, confluent and infratentorial lesions were not analysed (Sati et al., 2016). Both the raters were blinded to all the clinical information.

4.3.4 Establishing the diagnoses

Once the proportion of WM lesions with CVs was calculated for each patient (index test), I compared the results to the known clinical diagnoses (reference standard) and calculated the optimal threshold of WM lesions with CVs to propose a diagnosis of MS using the CV sign. The clinical diagnoses were reached by the patient's neurologist following any additional tests and/or clinical follow-up and were made according to the latest available McDonald criteria. Neurologists did not have access to the MRI research scans or the CV sign results when making the diagnoses. Additionally, I also tested the diagnostic performance of the previously proposed ´rule of six´ (Mistry et al., 2015).

4.3.5 Statistics

The statistical analysis was planned prior to the start of the study. I followed the same methods as described in Chapter 3 to calculate the sensitivity and specificity. Additionally, I also calculated PPV and NPV; the PPV were calculated as the number of true positives among true positives and false positives, expressed as a percentage, and the NPV were calculated as the number of true negatives among the true negatives and false negatives, expressed as a percentage. The CIs for the predictive values were calculated using the standard logit method (Mercaldo et al., 2007). Inter-rater reproducibility was calculated as percentage agreement between the raters with regards to their proposed diagnoses using the CV sign (MS vs. non-MS), based on the ideal threshold calculated in this study. The rate of recruitment was calculated for the months of active recruitment during the first year of the study.

4.4 Results

4.4.1 Demographics

Thirty eight patients (27 female) with a mean age of 50 years (SD=15 years) were recruited. Data from one patient was excluded from the analysis due to excessive movement artefacts on MRI. At the time of the analysis, 35 patients had been given a diagnosis by their neurologist. Median time between the index test (CV assessment) and the reference standard (clinical diagnosis) was equal to 6.5 months (range: 0-41 months). Twelve patients were subsequently clinically diagnosed with MS and 23 as non-MS. Figure 4-1 shows the patient flow diagram. Confirmed MS diagnoses were usually based on new typical relapses, radiological progression or

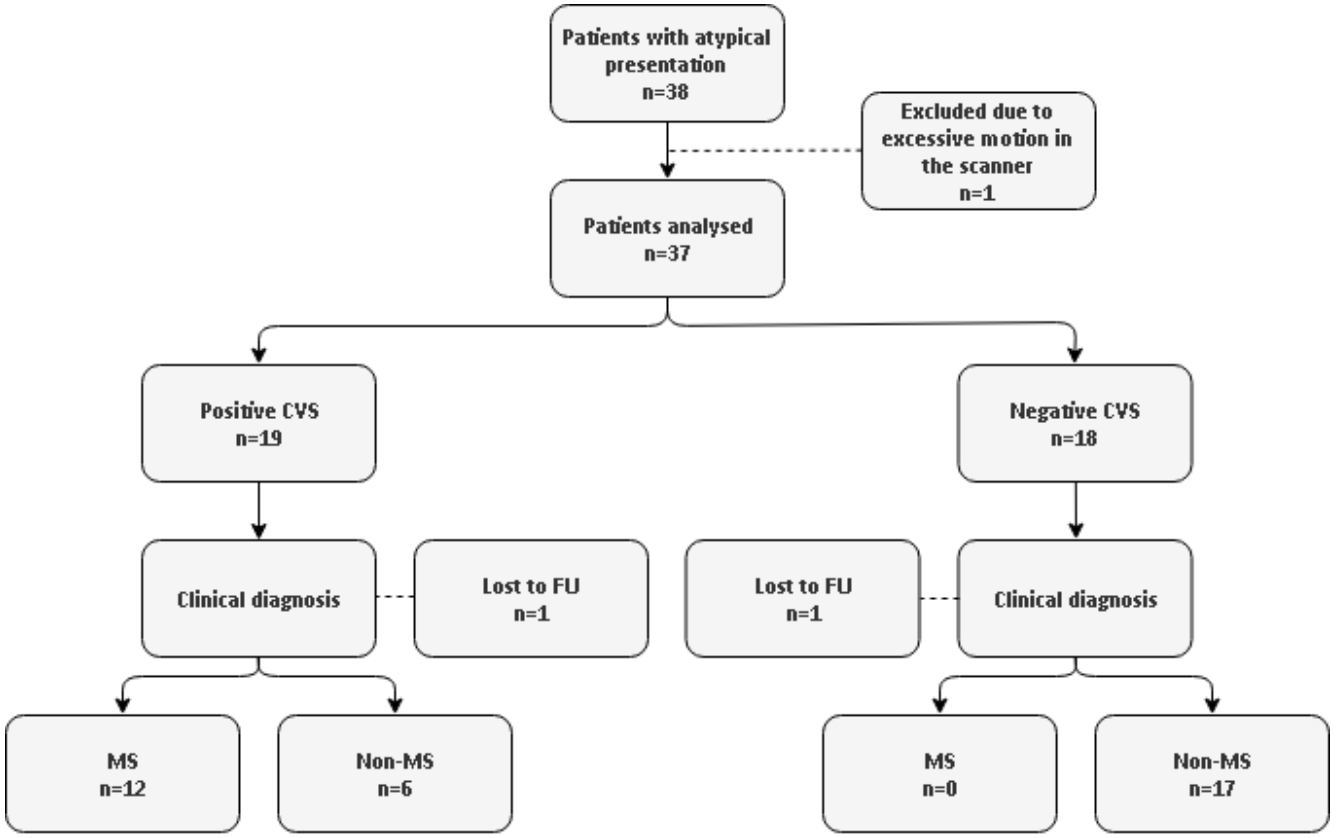


Figure 4-1 Patient flow diagram comparing the diagnoses made using the central vein sign (index test) and McDonald criteria (reference standard).

All patients had a 3T T2* MRI, which was used to calculate the proportion of lesions with CVs. A threshold of >40.74% was used to identify MS patients (positive CV sign). Patients who did not reach that threshold (negative CVS) were classed as non-MS. One patient's scan was excluded due to excess motion in the scanner. The remaining patients' diagnoses were made using the CV sign and then compared to the eventual clinical diagnoses. 12 patients, who were eventually diagnosed with MS by their neurologist, had MS according to the diagnosis made using the CV sign and 6 were diagnosed with alternative conditions. All 17 patients classed as non-MS using the CV sign were subsequently diagnosed as such clinically. Clinical diagnoses remain unknown for 2 patients, who were lost to follow-up and subsequently excluded from analyses.

CVS - central vein sign; MS - multiple sclerosis; FU – follow-up

supportive lumbar puncture results (see Table 4-1). The most common non-MS diagnosis was small vessel disease (13 patients). Median time between the index test and the last follow-up visit was 20 months (range: 1-49 months). Patients experienced no adverse events as a result of the index test or the reference standard. Patients eventually diagnosed as MS had minimal to no disability at the time of their index test.

4.4.2 Lesion analysis

The typical appearance of MS and non-MS lesions can be seen in Figure 4-2. Rater 1 (myself) identified a total of 785 WM lesions; patients subsequently diagnosed with MS had a median of 20 WM lesions (range: 3-71) whilst non-MS individuals had a median of 21 lesions (range: 2-63). The proportion of CVs (out of all the WM lesions) in patients subsequently clinically diagnosed as having MS was equal to 51% (SE=3%) whilst in non-MS patients it was equal to 28% (SE=2%). Figure 4-3 presents the proportions of WM lesions with CVs for the MS and non-MS patients.

Rater 2 identified a total of 715 WM lesions. The proportion of lesions with CVs in patients subsequently diagnosed with MS was equal to 56% (SE=3%) and 30% (SE=2%) in the non-MS group.

4.4.3 Agreement between the diagnoses

The clinical diagnoses were compared to the CV sign diagnoses made by Rater 1 and the ideal threshold for a diagnosis using the CV sign was calculated. Youden's J statistic was equal to 0.74 when associated with a threshold of >40.7% WM lesions

Table 4-1 Cross-tabulation of results from the CV assessment and clinical diagnoses (sorted by %WMLs with CVs) with the results of OCBs for comparison.

Patient	Age (y)	Sex	% WMLs with CVs	Rule of 6 diagnosis	Baseline				Follow-up					Clinical diagnosis
					Typical clinical presentation	DIS	DIT	OCBs	Clinical progression	Radiological progression	Further paraclinical testing	Time to clinical diagnosis (m)	Total duration (m)	
1	29	F	0.0	Non-MS	Y	N	N	Negative	Y	N	Y	8	8	Functional neurological disorder
2	66	M	0.0	Non-MS	N	N	N	Negative	N	N	Y	17	18	Neurosyphilis
3	53	F	10.5	Non-MS	Y	Y	N	Negative	Y	Y	N	12	12	Small vessel disease
4	35	M	12.5	Non-MS	Y	Y	N	Negative	Y	N	N	1	20	Behçet disease
5	77	F	12.9	Non-MS	Y	Y	N	Negative	N	N	Y	1	1	Small vessel disease
6	58	M	15.2	Non-MS	Y	Y	N	Negative	N	N	Y	1	20	Headache
7	51	F	17.2	Non-MS	Y	Y	N	Negative	N	N	Y	2	13	Cervical cystic lesion
8	50	F	23.8	MS	N	Y	N	Negative	N	N	Y	1	15	Epilepsy and migraine
9	60	F	24.1	MS	Y	Y	N	Not tested	Y	N	Y	24	37	Small vessel disease
10	55	M	24.2	MS	N	Y	Y	Not tested	N	N	Y	7	7	Small vessel disease, hypertensive micro-haemorrhages
11	34	F	25.0	Non-MS	Y	Y	N	Not tested	N	N	Y	3	3	Migraine
12	56	F	26.5	MS	Y	Y	N	Result uninterpretable	Lost to follow-up					Currently Unknown
13	56	F	28.6	Non-MS	Y	Y	Y	Not tested	Y	Y	Y	1	37	Small vessel disease

Patient	Age (y)	Sex	% WMLs with CVs	Rule of 6 diagnosis	Baseline				Follow-up					Clinical diagnosis
					Typical clinical presentation	DIS	DIT	OCBs	Clinical progression	Radiological progression	Further paraclinical testing	Time to clinical diagnosis (m)	Total duration (m)	
14	58	F	29.4	Non-MS	Y	Y	N	Positive	Y	N	Y	11	40	Functional neurological disorder
15	61	F	34.4	MS	N	Y	N	Negative	Y	Y	Y	2	20	Small vessel disease
16	55	M	35.7	Non-MS	Y	Y	N	Negative	N	N	Y	1	1	Small vessel disease
17	50	F	40.7	MS	Y	Y	N	Negative	N	N	Y	9	9	Small vessel disease
18	74	F	40.7	MS	Y	Y	N	Negative	Y	N	Y	32	44	Small vessel disease
19	65	F	40.9	MS	Y	Y	N	Result uninterpretable	N	N	N	1	19	MS
20	62	F	41.7	MS	N	Y	N	Positive	Y	Y	Y	1	45	MS
21	47	F	46.2	MS	N	Y	N	Negative	Y	N	Y	3	20	Small vessel disease
22	49	M	47.1	MS	Y	Y	N	Negative	Y	N	Y	9	25	MS
23	51	M	50.0	Non-MS	Y	Y	N	Negative	Y	N	Y	1	24	Functional sensory symptoms
24	49	M	50.0	MS	N	Y	N	Positive	Y	N	Y	2	16	MS
25	53	F	53.6	MS	Y	Y	N	Negative	N	N	Y	3	3	Small vessel disease
26	48	F	53.9	MS	Y	Y	N	Positive	Y	N	Y	17	49	MS
27	72	F	53.9	MS	Y	Y	N	Negative	Y	N	Y	37	37	MS
28	69	M	57.1	Non-MS	N	Y	Y	Negative	Y	N	N	14	14	MS
29	45	M	58.3	MS	Y	Y	N	Not tested	Y	Y	N	1	46	MS

Patient	Age (y)	Sex	% WMLs with CVs	Rule of 6 diagnosis	Baseline				Follow-up					Clinical diagnosis
					Typical clinical presentation	DIS	DIT	OCBs	Clinical progression	Radiological progression	Further paraclinical testing	Time to clinical diagnosis (m)	Total duration (m)	
30	44	F	63.6	MS	N	Y	Y	Not tested	Y	N	N	13	36	MS
31	42	F	66.7	MS	N	Y	N	Negative	Y	N	Y	41	41	Small vessel disease
32	38	F	66.7	MS	Y	N	N	Not tested	N	N	Y	15	15	Small vessel disease
33	45	F	71.4	Non-MS	Unknown			Negative	Lost to follow-up					Currently unknown
34	26	F	100.0	MS	N	Y	Y	Not tested	Y	Y	N	1	39	MS, nephrotic syndrome
35	22	F	100.0	MS	Y	Y	N	Positive	Y	Y	Y	9	45	MS
36	53	F	100.0	MS	N	Y	Y	Not tested	Y	N	Y	6	6	Neurosarcoidosis
37	30	F	100.0	MS	N	Y	Y	Not tested	N	Y	N	8	10	MS

%WMLs with CVs - percentage white matter lesions with central veins; OCBs - oligoclonal bands; MS - multiple sclerosis; DIS - Dissemination in space; DIT - Dissemination in time.

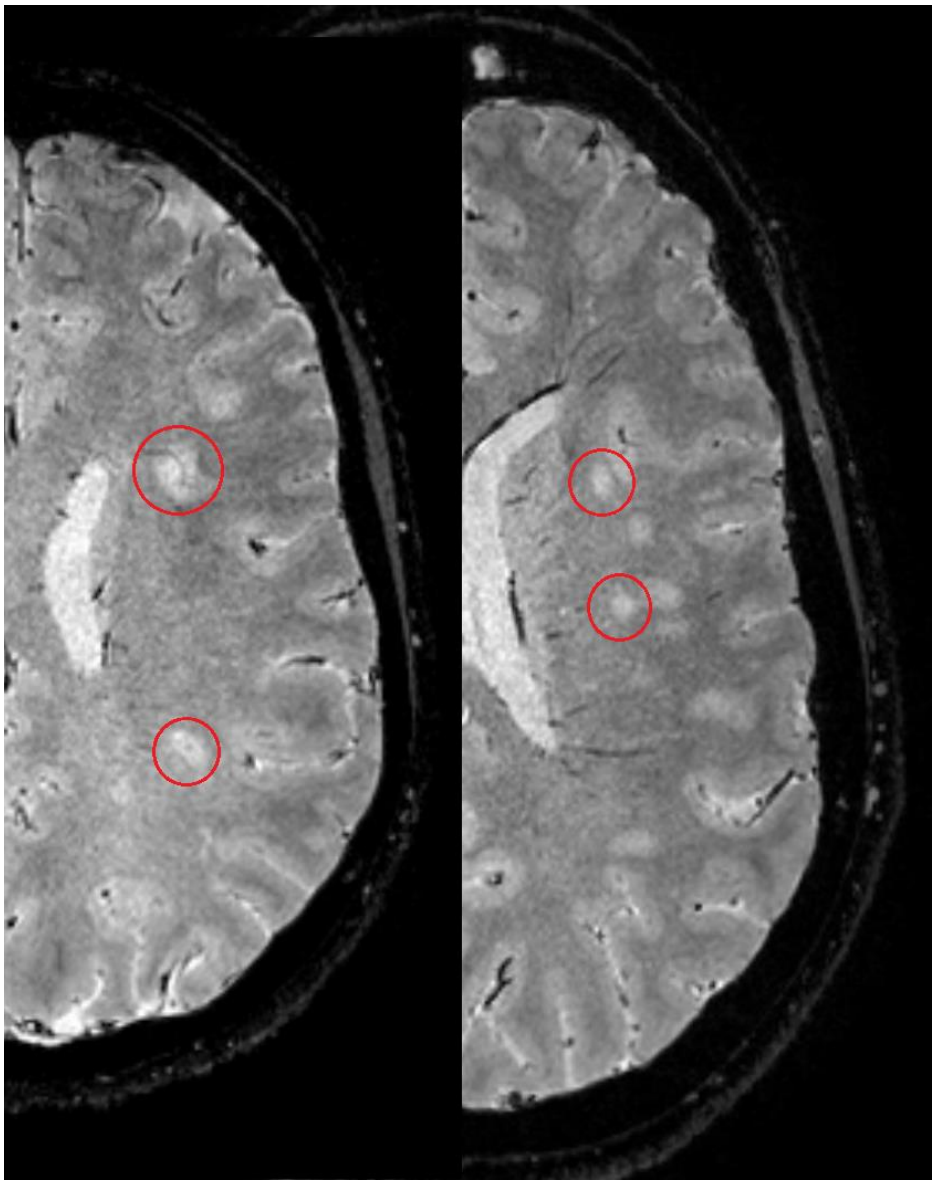


Figure 4-2 Typical appearance of lesions in MS (left) and non-MS (right).

The WM lesions in MS have clearly visible dark veins, which stand out against the bright appearance of the WM lesions.

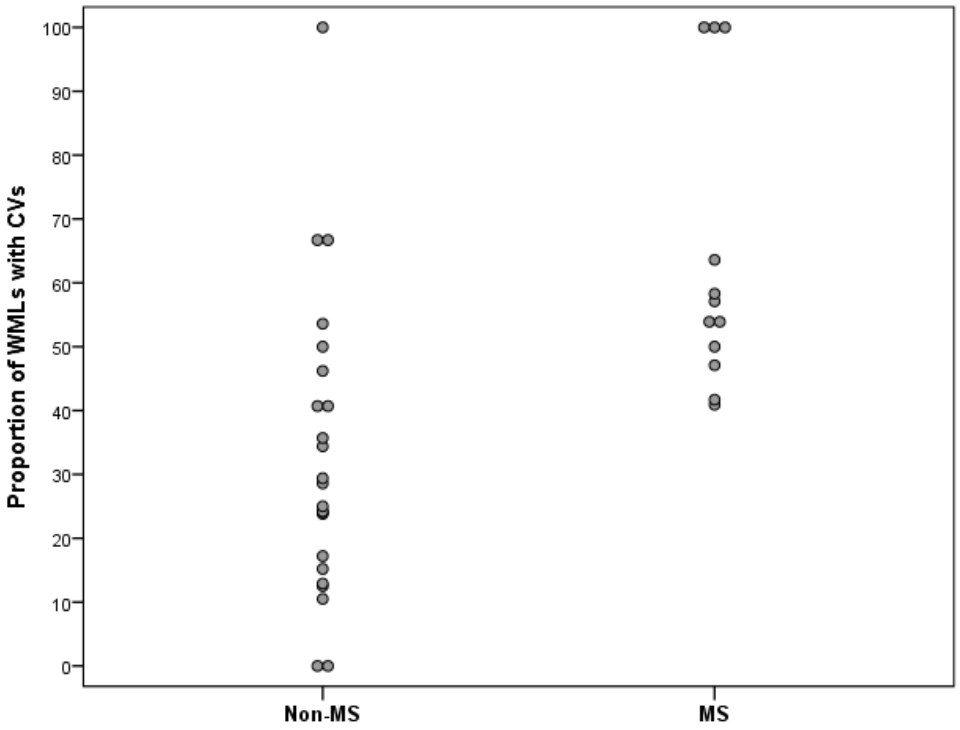


Figure 4-3 Distribution of the index test results (percentage of WM lesions with CVs), assessed by the two raters, as compared to the reference standard (clinical diagnosis).

with CVs. This resulted in a sensitivity of 100% (95% CI, 73.5%-100%), specificity of 73.9% (95%CI, 51.6%-89.8%) and PPV and NPV were equal to 66.7% (95%CI, 50.1%-79.9%) and 100%, respectively. The AUC was 0.86, p<0.001 (see Figure 4-4*).* Rater 1 achieved 82.9% agreement with the clinical diagnoses. The sensitivity and specificity of the diagnoses according to the rule of six were 91.7% (95%CI, 61.5-99.8) and 52.2% (95% CI, 30.6-73.2), respectively. A patient-wise cross-tabulation of Rater 1's results using the index and the reference standard, including paraclinical data, are found in Table 4-1. There were no statistically significant differences in age (p=0.139) and sex (p=0.92) between patients with and without an MS diagnosis according to the CV sign.

In comparison, Rater 2 achieved 71.43 % agreement with the clinical diagnoses with 91.67% sensitivity (95% CI, 61.5%-99.8%) and 60.9% specificity (95% CI, 38.5%-80.3%). The PPV and NPV were equal to 55.0% (95% CI, 41.7-67.7%) and 93.3% (95% CI, 67.6%-99.0%), respectively. The inter-rater agreement was 73.0%.

4.4.4 Rate of patient recruitment

Since the start of the study in May 2015, 57 patients have been approached by their neurologist about the study. Out of those, 38 patients agreed to take part and were subsequently scanned. Reasons for not participating included no contact after an initially positive response (6 patients), unwillingness/inability to have another MRI scan (4 patients), being too busy (4 patients), incompatibility

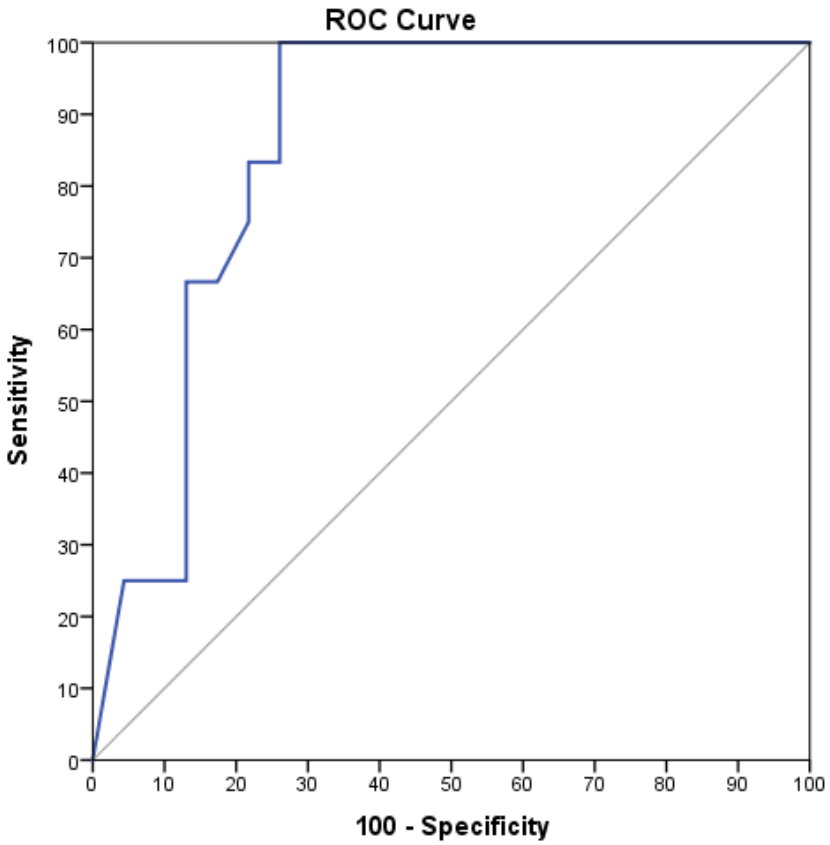


Figure 4-4 The ROC curve (blue) showing the diagnostic accuracy of the central vein sign in our study.

with study dates (3 patients) and a change of mind (3 patients). Rate of recruitment for the first year of the study (May 2015 to April 2016) was 2 patients a month with a screen failure rate of 0.25, based on 24 patients out of a total of 32 whom we approached about the study that year.

4.5 Discussion

4.5.1 Usefulness of the CV sign in predicting MS in patients with atypical presentation

To our knowledge, this is the first prospective study testing the usefulness of the CV sign, assessed using the NAIMS criteria (Sati et al., 2016), in diagnosing MS using a single 3T T2* scan in a population of patients with diagnostic uncertainty at presentation to an MS clinic. I demonstrate that the CV sign can successfully differentiate MS from non-MS in patients who remained undiagnosed by their MS specialist neurologists after initial clinical assessment, MRI review, and any additional investigation such as lumbar puncture. Even in this challenging, older cohort with an atypical presentation, the CV sign proved to have high sensitivity, specificity, and good agreement between raters when compared with the final clinical diagnoses.

The McDonald criteria were designed to be applied in patients with a high probability of MS at presentation, presenting with a CIS. However, a significant proportion of patients referred to MS clinics have an atypical history, examination or MR findings, leading to delays in diagnosis or misdiagnosis (Shetty et al., 2014). Offering the diagnosis of MS depends on the clinical presentation and MRI findings, clinician’s interpretation, experience, and

diagnostic confidence. For example, many of the patients who were ultimately clinically diagnosed with non-MS conditions in this study had evidence of DIS at presentation, increasing their chances of a misdiagnosis. As a result, the CV sign could be used in all atypical cases, not only to predict a diagnosis of MS earlier but also to exclude a diagnosis of MS in patients with a low percentage of CVs.

WM lesions on MRI most frequently misdiagnosed as MS are often due to small vessel disease, migraine and NMOSD seen in patients with non-MS specific symptoms (Solomon et al., 2012). The utility of the CV sign in diagnosing MS has previously been tested in a range of MS-mimicking disorders including small vessel disease (Tallantyre et al., 2011, Kilsdonk et al., 2014b, Mistry et al., 2015, Samaraweera et al., 2016, Sparacia et al., 2018), migraine (Solomon et al., 2015), NMOSD (Sinnecker et al., 2012a, Kister et al., 2013), Susac syndrome (Wuerfel et al., 2012) and SADs (Maggi et al., 2018), allowing for a successful differentiation between MS and non-MS. In fact, many previous studies have reported no overlap between MS and other groups when blinded raters used the CV sign in already established cases (Tallantyre et al., 2011, Wuerfel et al., 2012, Mistry et al., 2015, Maggi et al., 2018). In this prospective study, the typically small number of lesions in many of the CIS patients meant that there is an overlap between the two groups. Moreover, there was a trend whereby the oldest patients included in this study generally had a lower proportion of lesions with CVs, compared to younger patients. Although accurate differentiation between MS and non-MS was still possible in the older population, it is likely that age- and comorbidity-related ischaemic lesions might

affect the proportion of CV positive lesions. This shows that although useful, the CV sign might not be diagnostically accurate in all cases. Indeed, whilst the proportion approach performed very well in predicting a final diagnosis of MS and confirmed previous findings at 3T, the ´rule of six´ provided low specificity and high sensitivity (52 and 92%, respectively). This suggests that caution needs to be exercised when applying simplified rules to exclude a diagnosis of MS in patients with atypical CIS. Long-term follow-up of this group of patients is needed in order to ensure the correct clinical diagnoses have been given.

4.5.2 Comparison of the CV sign to OCBs testing

Comparison of the CV sign and OCBs was not the aim of this study; nevertheless, they are the most commonly used ancillary test in the diagnosis of MS. The use of lumbar punctures varies enormously from centre to centre reflecting the variable perception of neurologists and patients on its usefulness and discomfort. It is expected that the recent modifications to the McDonald criteria (Thompson et al., 2018), which allow positive OCBs to replace the need to fulfil DIT, will lead to an increased rate of lumbar punctures used. The sensitivity and specificity of OCBs in a population of CIS patients with a typical presentation were found to be 81% and 43%, respectively (Tintore et al., 2001) and another study found that 67% of patients who fulfilled the 2017 McDonald criteria on the basis of OCBs testing went on to develop CDMS (Beesley et al., 2018). The diagnostic values of OCBs are lower than the findings reported in this study and involved patients with a typical presentation. Previous research suggests that patients not fulfilling the criteria for an MS diagnosis had a low proportion of WM lesions with CVs, which also coincided with negative OCBs

(Grammatico et al., 2017). In this study, all but one patient diagnosed as non-MS, both clinically and using the CV sign, had negative OCBs. The incidence of positive OCBs in patients subsequently diagnosed as MS was more variable, thus suggesting that in the future CVs could be used in combination with lumbar punctures if required, although of course, MRI is much more tolerable than lumbar punctures for the majority of patients.

4.5.3 Inter-rater agreement

In order to be clinically useful, different raters identifying the CV sign need to have a high level of agreement between themselves and with the clinical diagnoses. Identification of WM lesions on MR scans is known to often result in poor inter-rater agreement (Ashton et al., 2003); nevertheless, when identifying the CV sign, it is the proportion of lesions with CVs that determines the diagnosis, regardless of lesion number. In this study, myself and Dr Falah identified slightly different numbers of WM lesions (785 & 715, respectively), however, despite this difference, the overall proportions of lesions with CVs were similar in MS (51% and 56%, respectively) and non-MS (28% and 30%, respectively). We also achieved good, but no perfect, agreements with the clinical diagnoses (82.7% and 71.4%) and with each other (73%). If the CV sign is to be used clinically, training should be provided to ensure uniform practices across clinicians. The existing guidelines can be found in a consensus statement on the evaluation of the CV sign (Sati et al. 2016).

4.5.4 Rate of patient recruitment

The recruitment rate during the first year of the study was 0.75 suggesting a high enrolment rate. As this study was designed as a pilot, the results suggest that a larger, multi-centre study testing the diagnostic value of the CV sign in populations with diagnostic uncertainty at presentation is feasible.

4.5.5 Study strengths

A key strength of this study is its prospective, blinded design. Most previous studies of the CV sign involved retrospective classification of patients based on the percentage of lesions with CVs and as a result are more susceptible to bias. For example, even when a blinded-rater is blinded to the diagnosis, they may still be influenced by their knowledge of the typical lesion appearance in the two disorders that they are comparing. In this study, both the clinicians and the raters were blinded at the time of the scan and analysis. Furthermore, the cohort of patients included in this chapter better reflects the true clinical scenario when the CV sign would be used in clinical practice as I examined patients presenting with a first clinical or radiological attack, rather than assessing patients with already established diagnoses.

4.5.6 Study limitations

However, the study also faced certain limitations; the cohort of patients with diagnostic uncertainty included in the analysis was relatively small and all the patients were recruited from a single centre. Moreover, I assessed the presence of CVs using only one sequence (T2* with high EPI). Whilst the results from the previous chapter of this thesis suggest that the detectability of the CV sign is not

significantly influenced by the GRE protocol used, it would be informative to perform a large-scale, multi-centre, prospective study to address the effects of scanners and sequences on the detectability of the CV sign in patients scanned at the earliest stages of the disease. This is because CIS patients generally have fewer lesions than patients with an established disease and, as such, there is an increased risk of a misdiagnosis according to the CV sign. Finally, all of the lesion assessment in this study was performed manually, which is time-consuming and subject to inter-rater variability. Recent reports on the automated detection of the CV sign, based on lesion masks from T_1- and T_2-weighted scans and vessel maps, based on T2*-weighted images, are promising although further work is required as the automated pipeline was more susceptible to motion artefacts and performed significantly worse than manual analysis of the lesions (Dworkin et al., 2018).

4.5.7 Conclusion

High resolution GRE protocols are becoming part of clinical MR assessment of patients suspected of having MS in many centres as clinicians use all the available information to support the correct diagnosis of their patients. This study reinforces the value of the CV sign as an additional tool to help diagnose MS in patients with diagnostic uncertainty at presentation. The reported results demonstrate high agreement with clinical diagnoses and between blinded raters suggesting the CV sign could be useful clinically in patients with diagnostic uncertainty at presentation. The findings confirm previous studies advocating the 40% threshold of WM lesions with CVs to make a diagnosis of MS suggesting that the CV sign is a robust biomarker of early MS lesions

irrespective of the type of presentation. Considering the different advantages and disadvantages of the CV sign and OCB a multi-centre study assessing the individual and combined value of those two tests in the diagnosis of MS is needed.

CHAPTER 5: DIAGNOSING MULTIPLE SCLEROSIS USING A SINGLE, 3T SUSCEPTIBILITY-WEIGHTED MRI SCAN

5.1 Introduction

SWI is an MRI technique sensitive to compounds that distort the local magnetic field, such as deoxyhemoglobin found in veins, and iron (Reichenbach et al., 1997, Haacke et al., 2004). Recently, the role of SWI in MS has gained attention as it offers additional information about MS WM lesions, which cannot be appreciated on conventional T_1- and T_2-weighted images, commonly used to diagnose and monitor patients (Haacke et al., 2009, Dal-Bianco et al., 2017).

Evidence from 7T, 3T and 1.5T studies shows that MS lesions form around small veins, a phenomenon termed the CV sign (Tallantyre et al., 2008, Tallantyre et al., 2009, Mistry et al., 2016, Samaraweera et al., 2016, Sati et al., 2016, Maggi et al., 2018). 3T studies of patients with established disease have proposed a 40% threshold of WM lesions with CVs to differentiate between MS and other disorders, which can mimic it on MRI (Tallantyre et al., 2011, Mistry et al., 2013, Campion et al., 2015).

More recently, the study of lesions with paramagnetic, iron rims (PIRs), which appear as dark, ring-like features on the edges of some MS lesions, has gained attention thanks to the increasing use of susceptibility weighted MRI sequences. Other patterns of iron accumulation, including scattered deposits, which appear as hypointense ´dots´ have also been described and they could indicate the initial

stages of iron deposition to form PIRs (Hagemeier et al., 2012). The significance of iron presence in MS lesions is still unclear, however, lesions with PIRs are believed to be chronic active lesions, also known as smoldering or slowly expanding lesions, which fail to remyelinate and have subsequently been proposed to be a marker of tissue destruction, continuing inflammation and failure to repair (Absinta et al., 2015, Absinta et al., 2016). Indeed, evidence from pathology and 7T MRI has shown that rimmed MS lesions were found to be significantly larger and underwent expansion by almost a third, unlike non-rimmed lesions which reduced their size by 10% (Dal-Bianco et al., 2017). Another study found that PIRs were more frequently found in patients with progressive MS and higher disability (EDSS score ≥5) (Harrison et al., 2016). Interestingly, PIRs have been reported in all subtypes of MS (Harrison et al., 2016, Dal-Bianco et al., 2017) but appear absent in NMOSD (Chawla et al., 2016), Susac´s syndrome (Wuerfel et al., 2012) and ischaemic lesions (Hosseini et al., 2018). This suggests that PIRs might be specific features of MS lesions, which could be used not only as biomarkers of disease progression but also diagnostically.

However, almost all studies of the PIRs so far have been carried out on small cohorts, using UHF scanners and/or non-conventional protocols. Whilst a recent study found that almost all of the IR visible on 7T MRI could be seen on 3T phase images (Absinta et al., 2018), assessment of PIRs for clinical purposes remains unreported. Furthermore, the diagnostic performance of the CV sign and PIRs in isolation and in combination has never before been tested in a large longitudinally studied cohort of patients with typical CIS. In this chapter, I describe the

prevalence and diagnostic value of both the CV sign and PIRs, detected on SWI images acquired as part of routine, clinical assessment in patients scanned at the time of typical CIS and in patients with MS-mimicking disorders. I hypothesized that patients, who went on to develop MS, would have a higher number of lesions with CVs and PIR compared to patients who did not.

5.2 Aims

1. To assess the prevalence of lesions with CVs and PIRs, detected on clinical 3T SWI, in patients with CIS and MS-mimicking disorders.
2. To calculate the diagnostic value of lesions with CVs and PIRs to predict conversion to MS in a typical CIS cohort
3. To test the diagnostic performance of CVs and PIRs when combined with evidence of DIS, against the performance of DIS and simultaneous presence of gadolinium-enhancing and non-enhancing lesions, and OCBs, currently used in the diagnosis of MS.

5.3 Methods

5.3.1 Ethics

This study received approval from the local ethics committee (reference no: PR(AG)302/2018). All patients signed written informed consent.

5.3.2 Patient sample

Patients were recruited from the Vall d'Hebron University Hospital in Barcelona, Spain between October 2010 and February 2019. All the scans were acquired as part of routine, clinical assessment.

Two groups of patients were recruited using consecutive sampling. The first one has been described previously elsewhere (Tintore et al., 2015, Arrambide et al., 2018) and included a prospective cohort of patients, aged <50 years with a typical CIS suggestive of CNS demyelination, scanned within 3-5 months of the first attack (CIS group). The second group comprised patients with multi-focal WM abnormalities, which were attributed to disorders considered to be MS imaging mimics, including the following groups of diseases: autoimmune (anti-MOG associated disease, Susac syndrome, anti-AQP4+ NMOSD, primary antiphospholipid syndrome, CNS vasculitis, Sjögren, neurosarcoidosis), small vessel vascular disease, infectious (HTLV-associated myelopathy), headache (including migraine), incidental findings and other diseases (non-specific leukoencephalopathy, non-specific paresthesia, neuralgia, non-specific visual symptoms, cranial nerve palsy) (non-MS group). Patients included in this group included a mix of prospective and retrospective cases scanned before or after being formally given a non-MS diagnosis.

The neurologists/radiologists who provided the clinical diagnoses for these two cohorts were not involved in the lesion analysis reported in this chapter. MS patients were diagnosed according to the 2017 McDonald criteria (Thompson et al. 2018). Patients were excluded from the analysis if they did not have the required SWI or T_2-weighted FLAIR protocols, had scans of insufficient quality for analysis or who had diagnostic doubt.

5.3.3 IgG oligoclonal bands

Intrathecal immunoglobulin OCB testing was performed in most CIS patients within 3 months of the first clinical attack, and in some of the non-MS group. OCBs were determined using agarose isoelectric focusing combined with immunoblotting in the CSF and serum (Freedman et al., 2005).

5.3.4 MRI acquisition

MR images were acquired on a 3T Siemens Trio MRI system (Siemens, Germany) with a 12-channel phased-array head coil and a whole-body transmit coil. The following sequences were obtained in all the patients: transverse 2D SWI (TR=33 ms, TE1=6.08 ms, TE2=24.6 ms, flip angle=15°, matrix size=288x384x104, voxel size=0.65×0.65×3.0 mm) and transverse 2D T2-FLAIR (TR=9000ms, TE=87ms, TI=2500ms, flip angle=119°, matrix size=412x512x46, voxel size=0.49×0.49×3.0 mm) or sagittal 3D FLAIR (TR=5000 ms, TE=394 ms, TI=1800 ms, flip angle=120°, matrix size =240x256x176 , voxel size=1.0x1.0x1.0 mm).

5.3.5 Image registration

Each patient´s FLAIR and SWI images were co-registered using Statistical Parametric Mapping (SPM12), and 3D slicer version 4.10.0 (Fedorov et al., 2012) was used to review the images.

5.3.6 Lesion analysis

WM lesions were identified on the T2-FLAIR images and analyzed for the presence of both CVs and PIRs on the axial plane of the SWI image. For the

analysis of CVs, I followed the guidelines described by the NAIMS Cooperative (Sati et al., 2016); however, confluent lesions were included in this analysis. If one of the ´fingers´ of a confluent lesion had CVs, I classed that lesion as CV positive. PIRs were identified as areas of hypointense ring-like signal, which corresponded to the lesion´s edge, encircling it fully or partially and visible on at least two consecutive slices. Information was also recorded about lesions with scattered iron deposition in the form hypointense iron ´dots´ on SWI (Hagemeier et al., 2012). Lesions that were <3mm in their shortest axis, infratentorial lesions, and lesions not fully visible on the SWI were not analysed. For each lesion, I recorded their topography: juxtacortical, periventricular, subcortical (WM not directly in contact with the cortex or the ventricles) or deep GM.

Image analysis was carried out by myself and a subset of 25 randomly selected scans was analysed by a second rater (Dr Francesco Crescenzo) to calculate inter-rater agreement and determine the effect of rater experience on identifying the lesion features of interest. The second rater (a neurology resident with four years' experience) had no previous experience in identifying CVs and PIRs, and underwent training prior to analysis. Both the raters were blinded to all of the clinical information at the time of the analysis.

This study included three main analyses. Data from both the CIS and non-MS patients was used in the first part of the analysis and I subsequently focused on the CIS population only to assess the diagnostic performance of the CV sign and PIRs in predicting conversion from CIS to MS in a population of patients with a typical first demyelinating event (part 2 & 3).

5.3.6.1 Part 1: Incidence of CVs and PIRs in CIS and non-MS patients

Firstly, I compared the incidence of lesions with PIRs, iron dots and CVs between the CIS and non-MS groups.

5.3.7 Diagnostic performance of central veins and iron rims

5.3.7.1 Part 2: Diagnostic value of CVs and PIRs in predicting conversion to MS according to the 2017 McDonald criteria

I used information about the incidence of PIRs and CVs in the CIS group to carry out an exploratory analysis of the diagnostic accuracy of using absolute number of lesions with PIRs and CVs to predict conversion to MS. For this part of the analysis, I only included patients with a minimum of 3 years of follow-up or a confirmed diagnosis of MS according to the 2017 McDonald criteria.

5.3.7.2 Part 3: DIS+ analysis

I also assessed the diagnostic value of PIRs, iron dots and CVs in combination with evidence of DIS by testing the following proposed criteria:

DIS + PIR: evidence of DIS and simultaneous presence of rim+ and rim-lesions

DIS + Iron: evidence of DIS, and rim+ and rim- lesions or iron dots

DIS+CVs: evidence of DIS and lesions with central veins

DIS+ any: evidence of PIR, Iron, 2CVs, OCBs and/or Gad.

I compared these results to the clinically used combinations of baseline DIS and: positive OCBs (DIS+OCB) and simultaneous presence of gadolinium-enhancing and non-enhancing lesions (DIS+GAD). I assessed the diagnostic value of the

DIS+ criteria compared to the following outcomes: radiological conversion to MS (at baseline or during the follow-up period) and the Poser criteria (clinically defined MS with evidence of two separate clinical attacks). Only patients with a minimum of 3 years of follow-up or a positive outcome (DIS+DIT or Poser) were included.

5.3.8 Statistics

I selected the minimum number of lesions with CVs and PIRs needed to establish an MS diagnosis to ensure high specificity. The sensitivity, specificity, accuracy, positive predictive value (PPV) and negative predictive value (NPV), with 95% confidence intervals (CIs), were calculated the same way as in Chapter 4. Inter-rater agreement was calculated separately for lesions with CVs and PIRs using intra-class correlation (ICC) (two-way mix model, single measures, absolute agreement).

I also tested for location preference of lesions with CVs and PIRs using a chi squared test. A Cox regression was used to calculate the risk of conversion to MS for each of the proposed DIS+ criteria. CIS patients not fulfilling the given criteria (radiological or clinical conversion to MS) were used as the reference group.

Statistical analysis was carried out using SPSS version 24 and the diagnostic performance was assessed using MedCalc for Windows, version 15.0. A p value of <0.05 was used to indicate statistically significant results.

5.4 Results

5.4.1 Patient demographics

112 CIS and 35 non-MS patients were included in this analysis. Patient flow diagram can be seen in Figure 5-1. Demographic and clinical data are shown in Table 5-1. In the CIS group, the mean follow-up time was 4.6 years (SD= 2 years). During the follow-up period, 94 patients (84%) were given a diagnosis of MS according to the 2017 McDonald criteria, including 42 (37.5%) who had experienced a second clinical attack and were classed as CDMS (Poser et al., 1983).

The breakdown of the non-MS diagnoses can be seen in Table 5-1. In the non-MS group, 19 individuals (54%) did not have an established diagnosis at the time of the scan and the mean follow-up period of that group of patients was 4.7 years (SD=2 years). For the non-MS patients with an already established diagnosis prior to the scan (n=16), median disease duration was 2 years (range: 0-49 years).

5.4.2 Lesion analysis

In total, 956 focal WM lesions were analyzed; 636 in the CIS group and 319 in the non-MS group. Figure 5-2 shows example lesions with CVs and PIRs and without them.

Regarding the inter-rater agreement, the ICC based on the number of lesions

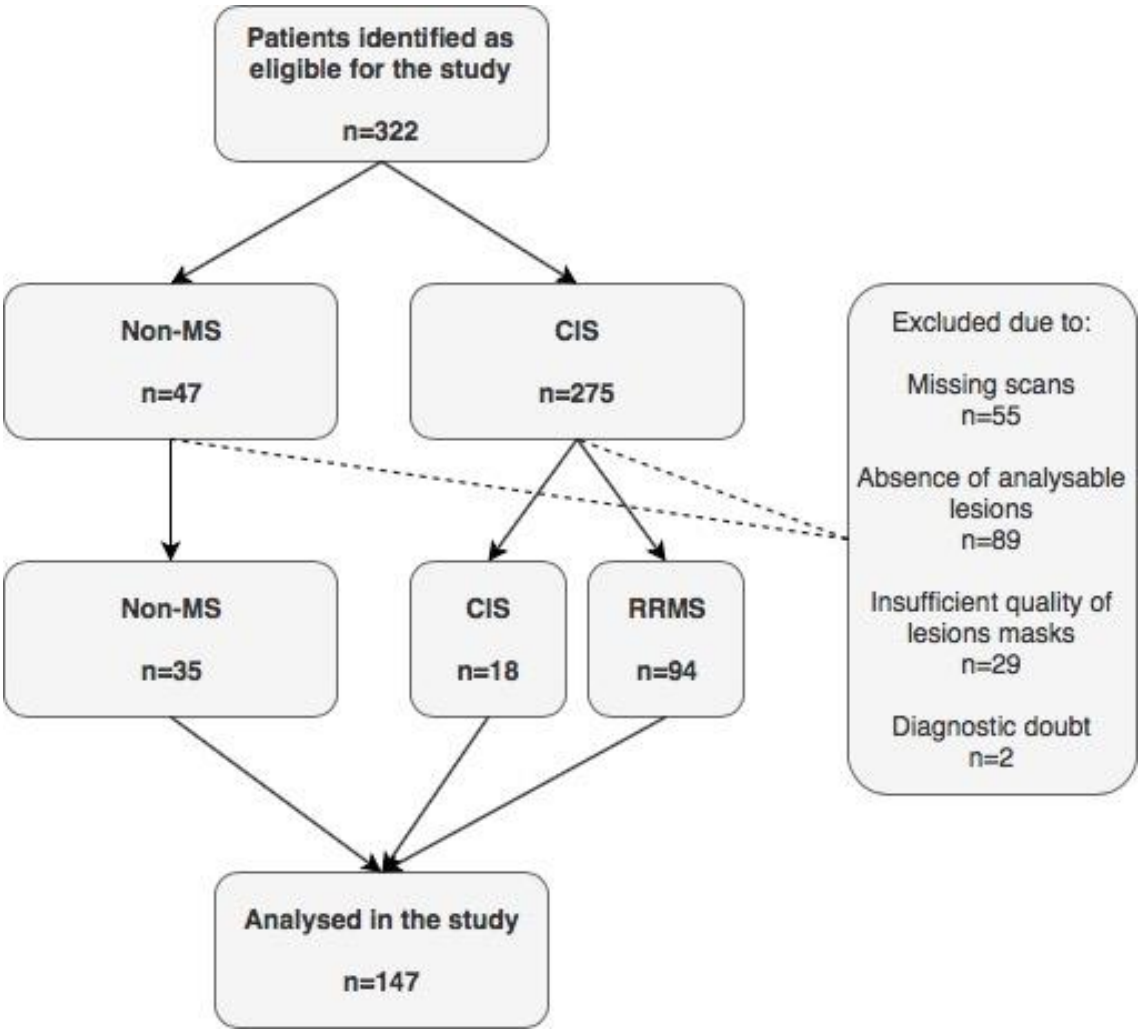


Figure 5-1 Patient flow diagram showing how patients were selected for the study and reasons for exclusion.

Non-MS - non-multiple sclerosis; CIS - clinically isolated syndrome; RRMS – relapsing-remitting multiple sclerosis

Table 5-1 Demographic and clinical data of the patients included in the study

	CIS Group (n=112)	Non-MS group (n=35)
Mean age, years (SD, range)	35.4 (7.9, 19-49)	41.7 (11.5, 20-67)
Female gender (%)	70 (70.5%)	23 (65.7%)
Clinical diagnosis at the end of the study (%)	MS=94 (83.9%) (including CDMS=42 [37.5%]) CIS=18 (16.1%)	Autoimmune disease=13 (37.1%) Vascular disease =8 (22.9%) Incidental findings =1 (2.9%) Infectious disease =1 (2.9%) Headache=3 (8.6%) Other=9 (25.7%)
+/- OCBs, n (not performed)	80/21 (11)	2/15 (18)
Median EDSS (range)	1.5 (0-4.5)	N/A
Median FLAIR WM lesion number (range)	4 (1-31)	7 (1-31)

CIS - clinically isolated syndrome; MS - multiple sclerosis; SD - standard deviation; OCBs - oligoclonal bands; EDSS - expanded disability status scale; WM - white matter; CDMS - clinically definitive multiple sclerosis; N/A - not available; n -number

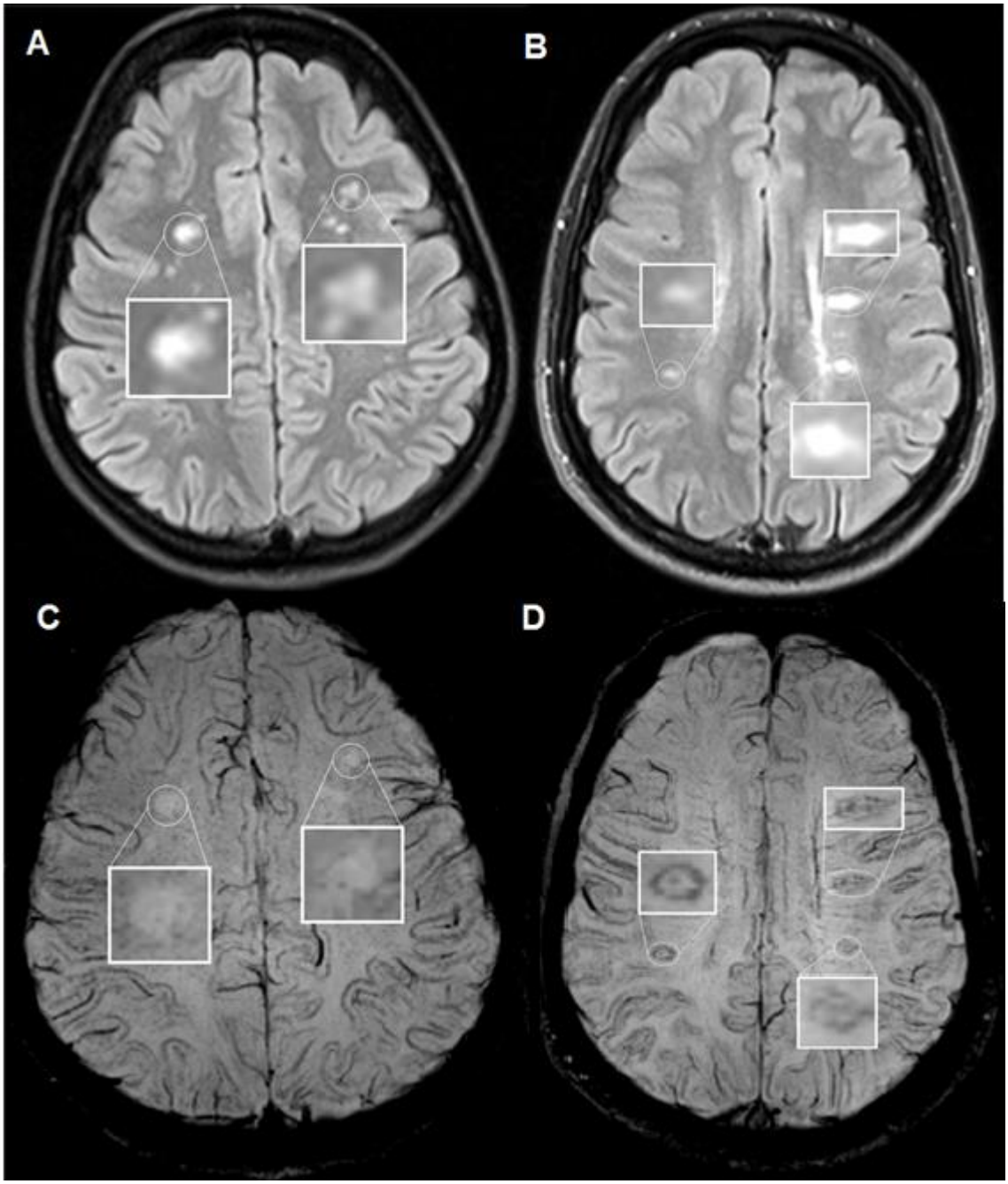


Figure 5-2 Lesion appearance on axial FLAIR (top) and the corresponding SWI (bottom) images in a patient who was subsequently diagnosed with Sjögren disease (A and C) and MS (B and D).

The Sjögren patient has no visible CVs or PIRs on the SWI. The MS patient has clearly visible PIRs, which correspond to the lesion edges visible on the FLAIR. CVs are also visible inside the lesions as dots and lines.

with CVs was equal to 0.84 (95%CI, 0.67-0.93) and 0.84 (95%CI, 0.64-0.93) for lesions with PIRs.

5.4.2.1 Part 1: Incidence of CVs and PIRs in CIS and non-MS patients

In the CIS group, 410 lesions (64.5%) had a CV, while only 53 lesions (16.6%) had a CV in the non-MS group. Figure 5-3 shows the location of lesions with CVs in the CIS and non-MS groups. The difference in the distribution of the lesions with CVs in the CIS group was significant ($X_2(3)=169.805$; $p<0.001$) showing preference for periventricular and subcortical locations. Figure 5-4 shows the differences in the proportions and number of lesions with CVs between the two groups. None of the non-MS patients reached the 40% threshold of WM lesions with CVs (range 0-37.5%).

127 lesions (19.9%) in the CIS group had a PIR, while none of the lesions in the non-MS group had them. None of the patients who remained CIS at the end of the study, both according to the 2017 McDonald criteria and the Poser criteria, had any lesions with PIRs. Over half of all the lesions with PIRs were located periventricularly, demonstrating a significant location preference ($X_2(3)=86.4$; $p<0.001$) and three quarters of all the patients with PIRs had at least one periventricular lesion with a PIR. 13.2% of all CIS lesions had both a CV and a PIR. See Figure 5-5 for a summary of the incidence and location of lesions with PIRs in this study.

46 CIS lesions (7.2%) and 16 non-MS lesions (5%) had an iron dot. For both the CIS and non-MS groups, the median number of lesions with iron dots was 0

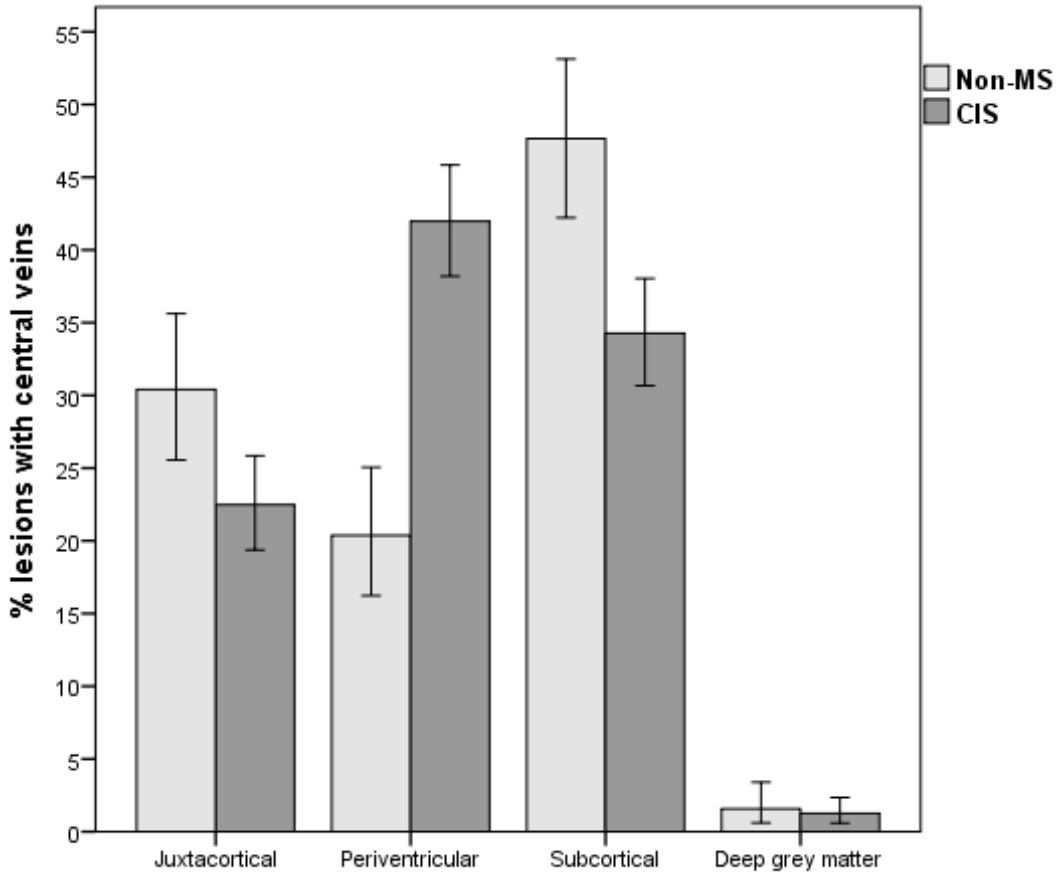


Figure 5-3 The location of lesions with central veins reported separately for the CIS and non-MS groups.

Error bars represent standard error.

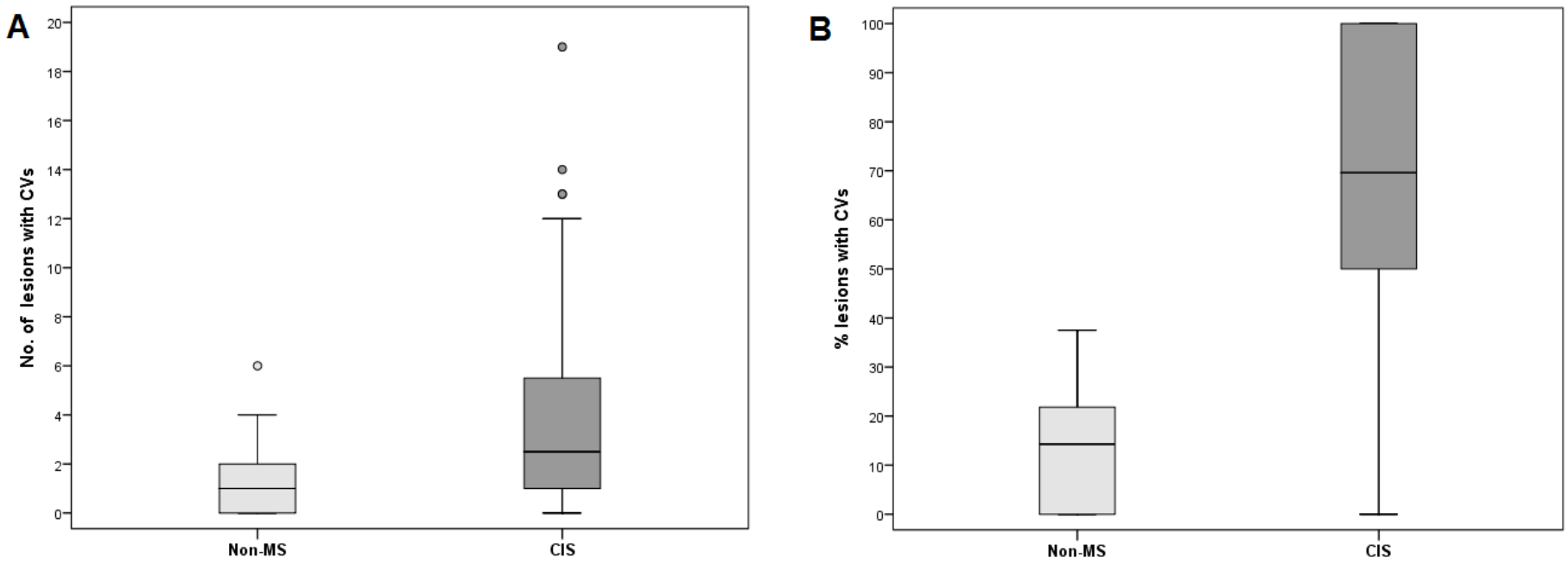


Figure 5-4 Summary of the incidence of lesions with central veins in the CIS and non-MS groups.

(A) Number of lesions with CVs (per patient) in the two groups. (B) Percentage of lesions with central veins (per patient) in the two groups.

CVs – central veins; non-MS – non-multiple sclerosis; CIS – clinically isolated syndrome

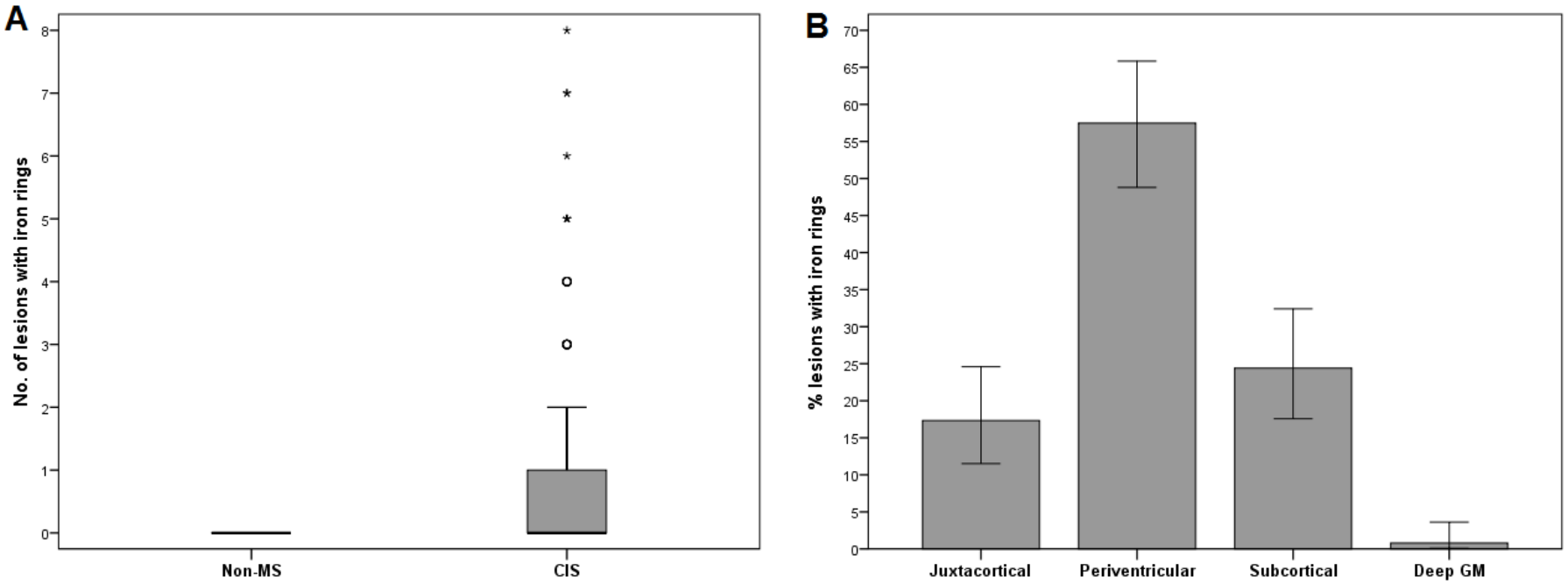


Figure 5-5 Summary of the lesions with iron rings in the CIS and non-MS groups.

(A) Number of lesions with iron rings (per patient) in the two groups. (B) Location of lesions with iron rings in the CIS group.

Non-MS – non-multiple sclerosis; CIS – clinically isolated syndrome; GM – grey matter

(range: 0-3). 24 CIS and 8 non-MS patients had at least one lesion with an iron dot.

5.4.3 Diagnostic performance of central veins and iron rims

5.4.3.1 Part 2: Diagnostic value of CVs and PIRs in predicting conversion to MS according to the 2017 McDonald criteria

Finding 3 lesions with CVs and/or 1 lesion with a PIR on the baseline SWI scan resulted in 70.21% sensitivity (95% CIs, 59.90-79.21) and 87.50% specificity (95% CI, 61.65-98.45) when tested against the 2017 McDonald criteria. The PPV and NPV were 97.06% (95% CI, 89.97-99.18) and 33.33 (95% CI, 25.83-41.78), respectively.

5.4.3.2 Part 3: DIS+ analysis

Table 5-2 shows the sensitivity, specificity, PPV and NPV and hazard ratios for the different DIS+ criteria. For DIS+CVs, I selected 2 CVs (rather than 3) to classify patients as MS as this increased the number of patients who could be diagnosed at baseline, whilst still ensuring high sensitivity and specificity. Each of the proposed tests resulted in the following number of patients receiving a diagnosis at baseline: DIS+OCB=69, DIS+Gad=55, DIS+PIR=48, DIS+Iron=55, DIS+2CVs=64,DIS+Any=83.

Table 5-2 The diagnostic properties and hazard ratios for radiological and clinical conversion to MS over time

Criteria at baseline	Outcome	Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)	Agreement (%)	HR(95%CI)	P value
DIS + OCB	DIS + DIT	**84.06** (73-26-91.76)	**50.00** (28.22-71.78)	**84.06** (77.42-89.02)	**50.00** (33.53-66.47)	75.82	**2.64** (1.38-5.06)	<0.01
	CDMS	**82.50** (67.22-92.66)	**32.50** (18.57-49.13)	**55.00** (48.56-61.27)	**65.00** (45.30-80.64)	57.50	**1.81** (0.80-4.12)	0.16
DIS + GAD	DIS + DIT	**73.33** (61.86-82.89)	**100** (85.75-100)	**100** (-)	**54.55** (45.19-63.59)	79.80	**38.92** (16.53-91.67)	<0.001
	CDMS	**73.81** (57.96-86.14)	**65.91** (50.08-79.51)	**67.39** (56.89-76.40)	**72.50** (60.32-82.05)	69.77	**3.52** (1.77-7.02)	<0.001
DIS + PIR	DIS + DIT	**58.67** (46.70-69.92)	**83.33** (62.62-95.26)	**91.67** (81.51-96.49)	**39.22** (31.82-47.14)	64.65	**2.73** (1.70-4.37)	<0.001
	CDMS	**54.76** (38.67-70.15)	**59.09** (43.25-73.66)	**56.10** (44.92-66.69)	57.78 (47.50-67.42)	56.98	**1.73** (0.94-3.18)	0.08
DIS + Iron	DIS + DIT	**66.67** (54.83-77.14)	**79.17** (57.85-92.87)	**90.91** (81.85-95.68)	**43.18** (34.19-52.64)	69.70	**2.76** (1.69-4.52)	<0.001
	CDMS	**64.29** (48.03-78.45)	**59.09** (43.25-73.66)	**60.00** (49.62-69.55)	**63.41** (51.89-71.92)	61.63	**2.21** (1.17-4.16)	0.014
DIS + 2CVs	DIS + DIT	**76.00** (65.75-85.11)	**70.83** (48.91-87.38)	**89.06** (81.17-93.90)	**48.57** (36.94-60.36)	74.75	**3.16** (1.83-5.44)	<0.001
	CDMS	**73.81** (57.96-86.14)	**52.27** (36.69-67.54)	**59.62** (50.79-67.86)	**67.65** (53.91-78.89)	62.79	**2.43** (1.21-4.87)	0.013
DIS + any	DIS + DIT	**94.67** (86.90-98.53)	**50.00** (29.12 –70.88)	**85.54** (79.80-89.86)	**75.00** (51.61-89.40)	83.84	**7.45** (2.70-20.58)	<0.001
	CDMS	**95.24** (83.84-99.42)	**29.55** (16.76-45.20)	**56.34** (51.30-61.25)	**86.67** (60.93-96.44)	61.63	**5.86** (1.42-24.43)	0.015

CI - confidence intervals, PPV- positive predictive values; NPV- negative predictive values; HZ - hazard ratio; DIS- dissemination in space; OCBs- oligoclonal bands; PIR- paramagnetic iron ring positive and negative lesions; CVs -central veins; DIT- dissemination in time; CDMS - clinically definite multiple sclerosis

5.5 Discussion

5.5.1 Diagnostic utility of the central vein sign and iron rims in CIS

This is the first study of the diagnostic value of the CV sign and PIRs, detected on 3T SWI, in patients presenting with a typical CIS and MS-mimicking disorders. I report that clinically acquired SWI can successfully detect CVs and PIRs with high inter-rater agreement.

In this study, finding at least one lesion with a PIR resulted in 100% specificity when used to differentiate between patients with a typical CIS and MS-mimicking disorders. In patients with a typical CIS, finding at least three lesions with CVs or one lesion with a PIR on a baseline SWI scan demonstrated high sensitivity and specificity (70% and 88%, respectively) against the 2017 McDonald criteria. Moreover, combining evidence of DIS and simultaneous presence of iron+ and iron- lesions or presence of at least two lesions with CVs predicted a 2 to 3-fold increased risk of a subsequent MS diagnosis (irrespective of whether using the radiological or clinical criteria) and demonstrated increased diagnostic specificity compared to DIS+OCBs.

5.5.2 Usefulness of SWI in MS diagnosis

The International Panel on Diagnosis of MS has identified the study of the CV sign and PIRs as a high-priority research area (Thompson et al., 2018). The CV sign has previously been studied using a variety of MRI protocols. Some of them, such as FLAIR* (Campion et al., 2017), require extra post-processing steps whilst others have used a contrast agent (Maggi et al., 2018). In practice, financial restraints of hospitals, as well as health concerns over gadolinium accumulation

in the brain (Guo et al., 2018) would limit the use of these protocols. The only previous, prospective study of the CV sign using 3T SWI involved 14 patients and concluded that the CV sign was useful in differentiating between MS and non-MS lesions (Kau et al., 2013). In this study of over 100 CIS patients, I confirm that a single SWI scan can offer valuable diagnostic information whilst being relatively easy to implement in hospital radiology departments.

Similarly, most of the studies of PIRs have been performed using 7T quantitative susceptibility mapping (QSM) and phase imaging (Hammond et al., 2008, Bagnato et al., 2011, Yao et al., 2012, Absinta et al., 2013, Bian et al., 2013, Chawla et al., 2016, Harrison et al., 2016). Recently, Absinta and colleagues compared 7T and 3T phase images and found that almost all 7T rings were also visible at 3T (Absinta et al., 2018). However, QSM and phase imaging are not commonly used, or even available, in clinical practice. In this study, using routinely acquired SWI images, PIRs were completely absent in patients who did not have an MS diagnosis whereas 56% of patients who fulfilled the 2017 McDonald criteria had at least one lesion with a PIR.

5.5.3 Alternatives to DIT

As lesions with PIRs are thought to represent the chronic, active stage of lesion evolution (Dal-Bianco et al., 2017, Absinta et al., 2018), we could speculate that simultaneous presence of iron+ and iron- lesions on a single scan provides objective evidence of DIT, similarly to the way gadolinium is currently used to support a diagnosis of MS at the time of the baseline scan. Whilst presence of CVs cannot be considered as evidence of DIT, it could be used as an alternative

to DIT, similarly to the way OCBs have been incorporated into the 2017 diagnostic criteria.

In this study, presence of iron+ and iron- lesions or finding two lesions with CVs in combination with evidence of DIS had higher specificity compared to OCBs, which had the lowest specificity of all the tests. Although useful diagnostically as an alternative to DIT, OCB testing is invasive and can result in side effects including back pain and headache (Engelborghs et al., 2017). As a result, some patients refuse to have a lumbar puncture. Specific, non-invasive MRI markers have the potential to help diagnose MS without exposing patients to unnecessary invasive procedures. Moreover, from the patients' perspective, a short (in our study <5min), additional MRI would be far more convenient than a hospital day case admission for OCB testing.

5.5.4 Study strengths

One of the strengths of this study is that all the CIS patients were scanned before a diagnosis was reached, which reflects the way the CV sign and PIRs would be applied in clinical practice. So far, only one study of the CV sign at 3T has tested its diagnostic value in CIS patients (Kau et al., 2013). I am not aware of any studies of the diagnostic value of PIRs using 3T MRI. Moreover, the proposed diagnostic scheme would allow for easy implementation in a clinical setting. Finding a fixed number of lesions with CVs and/or PIRs is more practical than using a proportion-based approach, which requires the analysis of all the lesions. Moreover, the low number of lesions required for a diagnosis (3 lesions with CVs

or one with a PIR) means it can be applied even in patients with a small lesion load.

5.5.5 Study limitations

This study also had limitations. The non-MS cohort was relatively small and I only analyzed scans from one center using a single scanner and protocol. This makes the results promising but currently cannot be generalized to other centers or scanners. Future studies should assess the visibility of PIRs on different susceptibility-weighted images, including T2*-weighted images used in some MS centres. Moreover, even with a mean follow-up period of four and a half years, some patients did not develop CDMS and only long-term follow-up of these patients will demonstrate the real usefulness of SWI in MS diagnosis. With longer follow-up, it would also be interesting to determine whether patients who had PIR at the time of the baseline scan developed disability or progression earlier than those without.

5.5.6 Conclusions

Clinically available, 3T SWI MRI can successfully visualize CVs and PIRs, which appear to be highly-specific features of early MS lesions. Whilst the diagnostic value of the CV sign has previously been validated in cohorts of patients with established disease, this is the first report of its diagnostic utility in patients scanned within a few months of CIS, who were followed-up for a period of almost five years. Moreover, this study was also the first to assess the prevalence and diagnostic value of PIRs in that population. The proposed easy-to-implement diagnostic criteria of finding 3 lesions with CVs and/or 1 lesion with a PIR could

be applied in a clinical setting without the need for post-processing. Moreover, in combination with DIS, finding at least 2 CVs, or evidence of iron positive and negative lesions offered good, non-invasive alternatives to gadolinium for demonstrating DIT or to OCB testing. Future, prospective, multi-center studies are needed to confirm the findings of the diagnostic role of the CV sign and PIRs in MS.

CHAPTER 6: BRAIN ATROPHY ESTIMATION FROM ROUTINE, NON-STANDARDISED, CLINICAL SCANS WITH INCOMPLETE BRAIN COVERAGE – A METHOD VALIDATION

6.1 Introduction

Currently most hospitals use Patient Archiving and Communication Systems (PACS) to electronically store medical images, including MRI. With appropriate ethical approval, these images could be explored to address important research questions in MS and offer a unique opportunity to carry out longitudinal studies combining clinical and MRI information at a fraction of the cost compared to acquiring the data prospectively. Other potential applications include safety monitoring of treatment responses in diverse MS populations.

Brain atrophy is accelerated in MS compared to normal ageing and has been shown to predict disability (Di Filippo et al., 2010, Popescu et al., 2013b) cognitive performance (Zivadinov et al., 2001) and conversion to CDMS (Perez-Miralles et al., 2013). Moreover, atrophy is a well-established secondary outcome in clinical trials of all new DMTs (De Stefano et al., 2014). Studies attempting to use atrophy values obtained from routine, non-standardised, clinical MR scans were first reported in the last couple of years. A report of the MS-MRIUS cohort, using data from 1,800 individuals from 33 US clinics conducted longitudinally over an average follow-up period of five years, demonstrated that the beneficial effects of Fingolimod, detected on routine 2D and 3D MR scans performed on 1.5T and 3T scanners, were comparable to those reported in phase III clinical trials (Zivadinov

et al., 2019). Raji and colleagues conducted a cross-sectional study to test the added value of total intracranial volume, calculated from 3D, 3T, clinical MR images obtained on two scanners as part of routine, clinical assessment (Raji et al., 2018). They reported that patients with the highest levels of whole brain and thalamic atrophy had more disability, higher FLAIR lesion volume and were more likely to have entered the SP course. Similarly, another study of the MS-MRIUS cohort found that brain atrophy, detected on 3D, T1 scans, was significantly associated with longer disease duration, more disability and higher MSSS in MS patients, with patients with confirmed disease progression having over 20% higher PBVC than those who did not progress (Ghione et al., 2018).

However, atrophy assessment has not, so far, been incorporated into routine, clinical practice due to the limitations of current methodologies and their effect on the accuracy of measurements. As mentioned in De Stefano et al. (2014), significant differences in terms of atrophy across two groups can only be reliably detected based on power calculations that would allow for significant detection of atrophy. Reliable results in small sample studies come from images obtained from the same scanner manufacturer and similar image acquisition parameters (Beadnall et al., 2019). However, this approach is problematic in real-life MS clinics where time and financial constraints mean most patients are not scanned on the same scanner during their follow-up MRI visits. In fact, a recent study found that around half of a group of 250 patients had their 9-24 month follow-up clinical scan taken following major hardware, software or coil upgrades suggesting true image standardisation is difficult in clinical practice even across relatively short periods of time (Zivadinov et al., 2017). Indeed, atrophy

estimations from clinical scans reportedly have a failure rate of 37-82% with the majority of errors attributed to scanner, protocol or orientation changes between the two time points (Ghione et al., 2018, Zivadinov et al., 2018). This suggests that, whilst possible, reliable atrophy estimation in clinical practice is difficult to achieve. I this chapter, I aimed to compare atrophy estimates obtained from clinical and research scans, performed in the same patients, and determine if, and under what circumstances, routine, non-standardised clinical scans could be used to successfully estimate brain atrophy.

6.2 Aims

1. To audit routine, clinical MRI protocols used at the Nottingham University Hospitals NHS Trusts in order to assess the intra-site MR scan variability
2. To determine if, and under what circumstances, routine, non-standardised clinical scans can be used to reliably detect atrophy in MS patients

6.3 Methods

6.3.1 Ethical consent

For this study I selected patients who had previously agreed for their medical data to be used for research purposes. This was either done through the UK MS Register informed consent or having already given consent when patients participated in previous research studies at the University of Nottingham. The MS Register database has been reviewed and given approval by an independent research NHS Research Ethics Committee NRES Committee South West-Central Bristol (11/SW/0160). A copy of the consent form can be found in Appendix D. Patients included in this study have participated in one (or more) of

the following studies carried out at the University of Nottingham: *Investigation of grey matter pathology using ultra high field (7T) Magnetic Resonance Imaging scanner* (ethics number: 12/EM/0452), *STAR-MS Study; Single Test to ARrive at MS diagnosis. Investigating the use of a single MRI brain scan to help diagnose Multiple Sclerosis in instances where there is uncertainty* (ethics number: 15/NW/0286) and *High Resolution MRI Scanning of the brain* (ethics number: 05/Q2404/44). All patients signed written informed consent.

6.3.2 Patient sample

Patients from the Nottingham University Hospitals NHS Trust with the appropriate ethical consent, over 18 years old, with a diagnosis of MS (although the diagnosis did not have to be confirmed at the time of the first scan) and who had at least two clinical and two research MRIs were included. Failure to meet any of the above criteria resulted in exclusion from the study.

6.3.3 Audit of clinical MRI protocols

I carried out an audit of the clinical protocols used at the Nottingham University Hospitals NHS Trust. Information about the scanner, field strength, slice thickness, orientation, TE and TR was recorded for all scans. The collected MRI data represented all major scanner manufacturers, field strengths, types of scans and sequence parameters used in most UK clinics. I subsequently focused only on the protocols, which can be used for atrophy estimation (T_1-weighted). A sub-analysis of 76 randomly selected T_1-weighted scans demonstrated that 76% of the images provided incomplete coverage of the brain, which resulted in missing slices in the anterior-posterior or superior-inferior planes (see Figure 6-1).

6.3.4 Research MRI protocols

The research T_1-weighted scans were obtained from previous research studies carried out at the University of Nottingham. Table 6-1 shows information about the T_1-weighted research protocols used in this study.

6.3.5 Atrophy estimation from clinical scans with missing slices and research scans

I selected SIENA version 2.6 (Smith et al., 2002), part of FSL (Smith et al., 2004), to perform atrophy estimations on the clinical and the research scans as it is a widely-adopted and validated longitudinal approach for atrophy estimation, which has previously been used in both clinical and research studies as well as clinical trials in MS (Smeets et al., 2016). Each output was visually assessed and re-run using appropriate optimisation parameters, if necessary, to ensure optimal brain extraction, tissue segmentation and image registration. I developed a simple numerical scoring system to assess the quality of the outputs from SIENA:

1 - The brain has been extracted precisely with no obvious errors. Tissue around the eyes and neck has been classified as brain and non-brain correctly.

2 - Small areas of the brain, especially around the eyes or neck are underestimated or non-brain tissue is classified as brain tissue.

3 - Large portions of the brain are missed or non-brain tissue is classed as brain

4 - Complete failure with very little brain tissue classified correctly

Scores of 3 and 4 were deemed insufficient for inclusion in the study and such scans were rejected from subsequent analysis analysis.

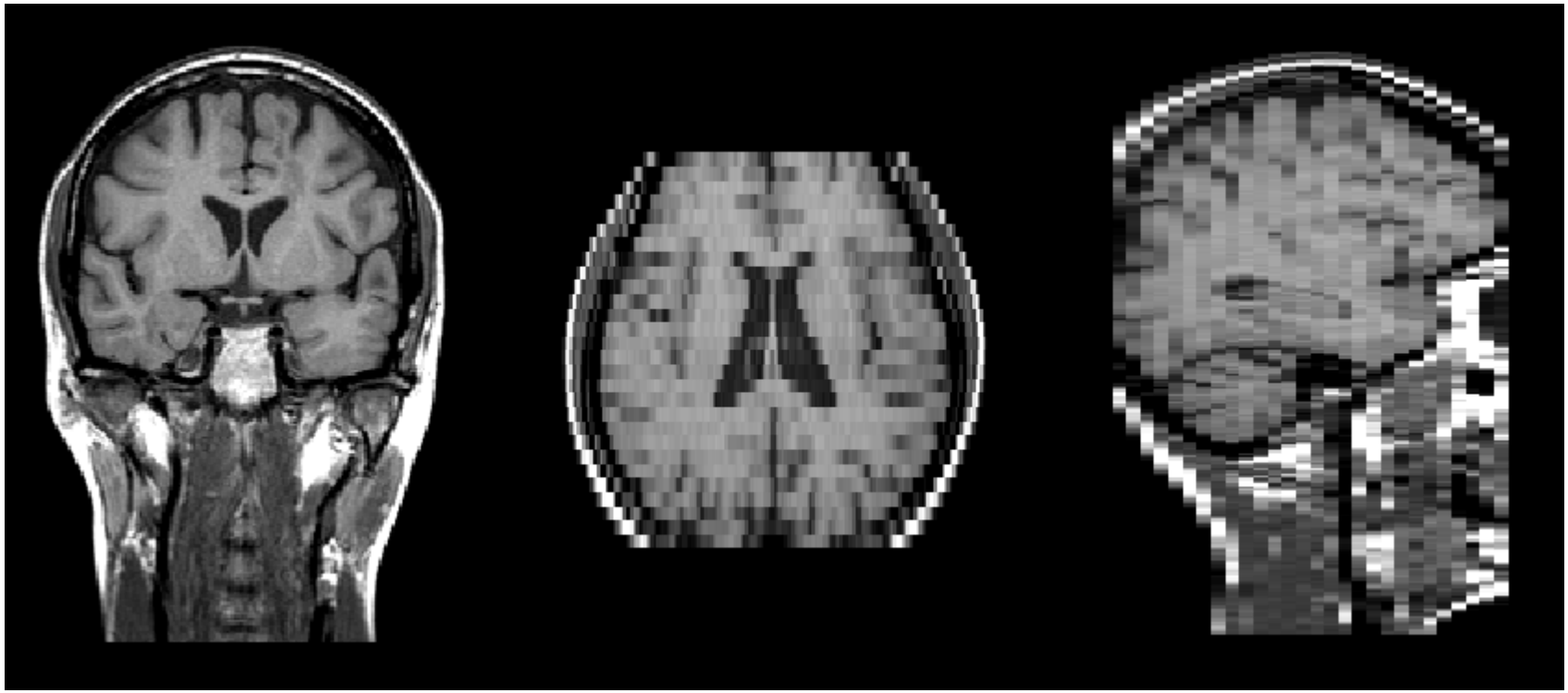


Figure 6-1 An example 2D hospital scan taken in the coronal plane.

The axial and sagittal planes demonstrate the incomplete coverage of the brain, where the anterior and posterior parts are not included in the field of view.

Table 6-1 Protocol information for the T1-weighted research scans

Sequence	Scanner	2D/3D	Thickness (mm)	TR (ms)	TE (ms)
T_1-weighted	Philips 7T	3D	0.5	15	7
	Philips 3T	3D	1	7	3
	GE 3T	3D	1	8	3

TR – repetition time; TE – echo time

6.3.6 The Frankenstein method

SIENA was unable to process any of the clinical scans with missing slices (more details are presented in the results section). In order to overcome atrophy estimation failures in those scans, Dr. Alain Pitiot developed an approach, termed the *Frankenstein method*. This approach combines incomplete, clinical scans with slices from another scan that provide coverage of the whole brain to create a complete scan without missing slices. To achieve that, each original clinical scan with missing slices was registered to a fully-covered, 3D MPRAGE scan (template) using the ROBEX package (Iglesias et al., 2011). Ideally, the template came from the same patient (from a different time-point) and had similar characteristics in terms of image resolution and noise. For patients who did not have such a scan available, we used fully a covered, 3D MPRAGE template from another patient. We selected scans exhibiting varied levels of brain atrophy to ensure the scan with missing slices and the template were as similar as possible (see Figure 6-2). Once the original scan and the template were registered, we harmonised their intensity distribution and blended them at the borders to create a softer transition between the two images. This process was repeated for both time points before the two images were analysed in SIENA. Figure 6-3 shows an example *Frankenstein* image.

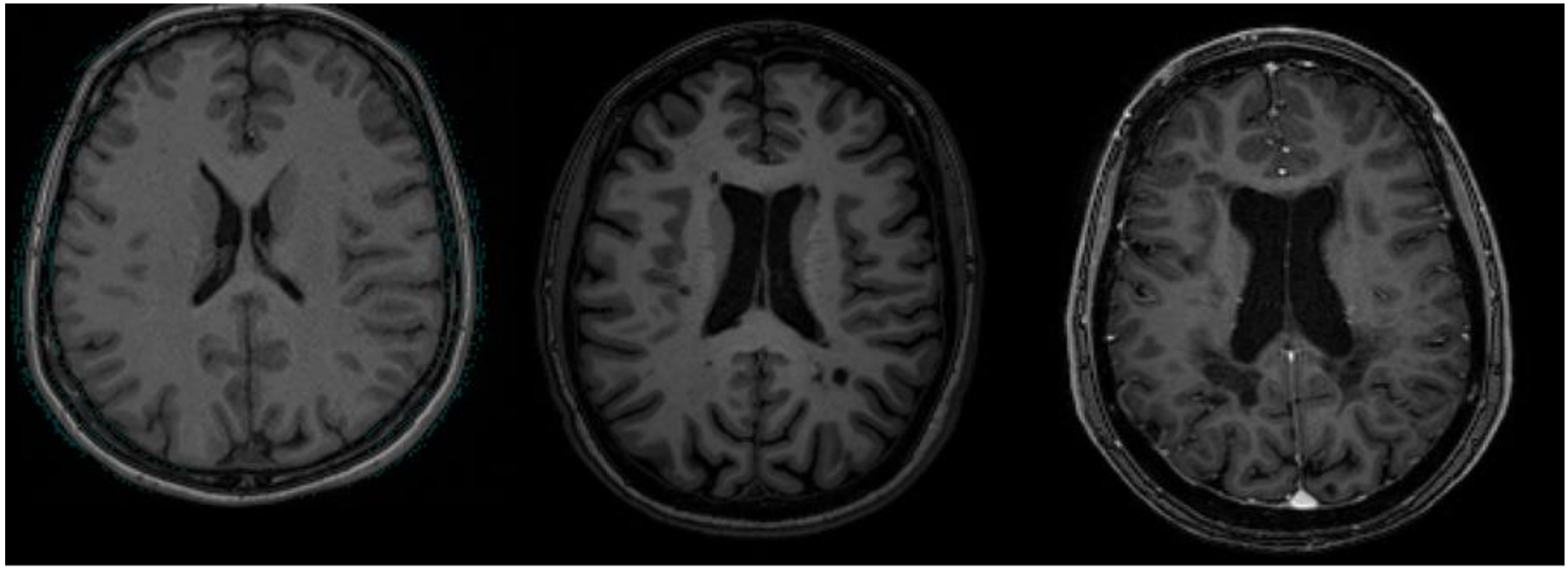


Figure 6-2 3D MPRAGE MRI scans with low (left), medium (middle) and high (right) atrophy levels.

Clinical scan with missing slices were visually assessed for atrophy levels based on the appearance and size of the gyri, sulci and lateral ventricles and its atrophy levels deemed low, medium or high. The images were then registered with one of the 3D, fully covered scans (shown above).

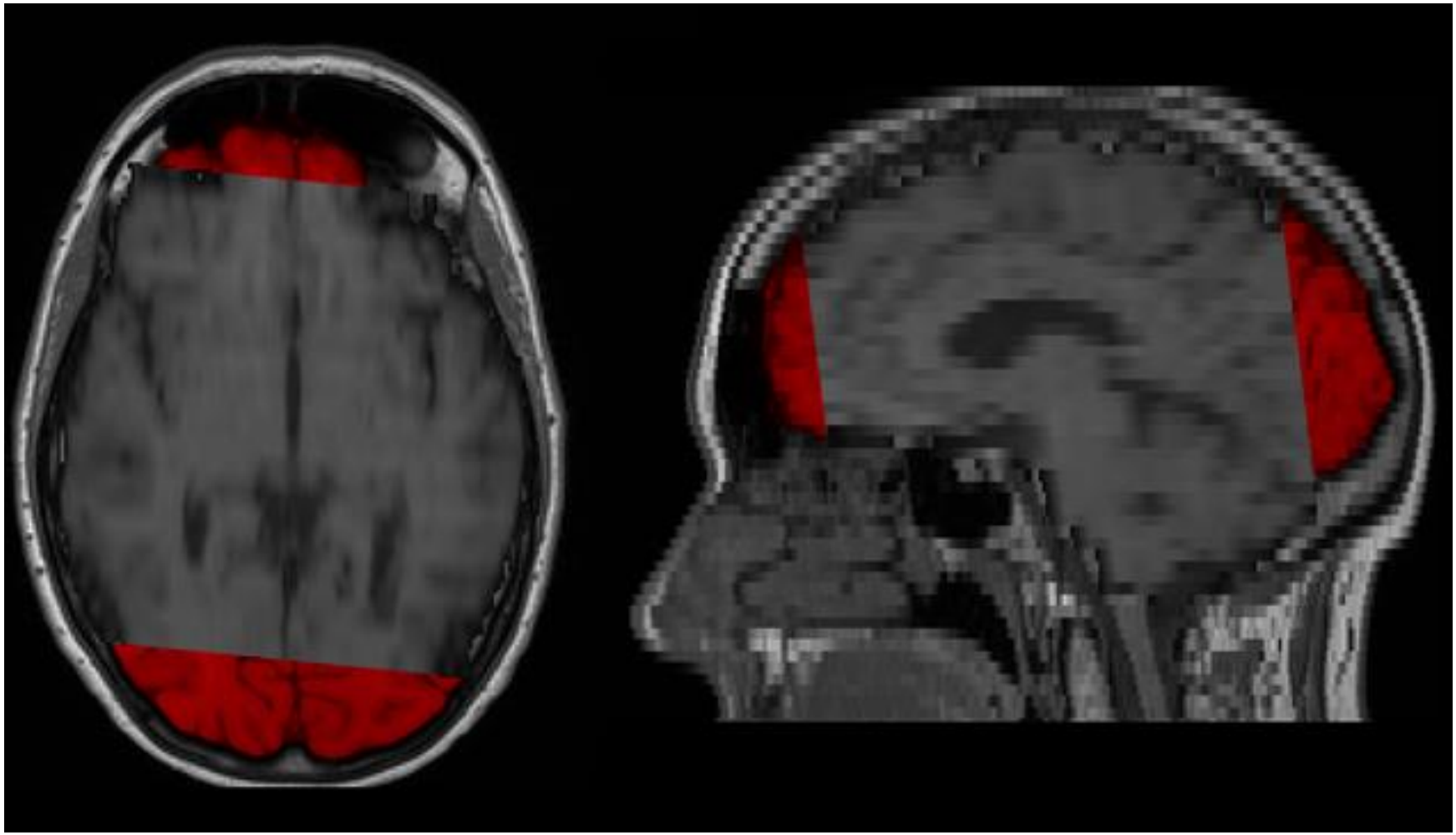


Figure 6-3 A “Frankenstein” brain MR

The images consist of the original MRI with missing slices (grey) and the added anterior and posterior slices (red) from the 3D images with similar atrophy levels.

For each *Frankenstein* brain, SIENA estimated local atrophy maps that gave, for each voxel, the estimated atrophy between the time points. We then discarded those voxels that did not belong to the original scan and estimated a global atrophy rate. For example, if the original baseline scan had slices missing on top of the brain and the follow-up scan had slices missing in the frontal and occipital lobes, the local atrophy maps from all of those areas would be discarded. Following a successful visual inspection for quality control, a global estimate of the percentage volume change was then recalculated based on the tissue present in the original brain images only.

6.3.7 Frankenstein validation

In order to estimate the effect of varying levels of missing slices on atrophy estimation, we used data from a healthy control male, aged 44 at the time of the baseline scan, who had two 3T MPRAGE scans (Philips Achieva, Philips Medical Systems) taken 7 years apart. The scans provided 100% coverage of the head and we artificially removed varying numbers of slices and applied the *Frankenstein* approach. The number of missing slices in either or both scans ranged from 0% to 35%.

6.3.8 Statistics

The interquartile range was used to assess the variability of the results reported in the validation of the Frankenstein approach. A paired T test was used to test for differences in atrophy estimates obtained from the clinical and research scans. Additionally, a one sample T Test was used to test whether the atrophy values obtained using the Frankenstein approach matched atrophy values previously reported in the literature. I selected an atrophy value of 0.4% per year to indicate

pathological brain volume loss, based on previous literature (De Stefano et al., 2016). All analyses were done using SPSS version 24.

6.4 Results

6.4.1 Variability of clinical protocols

An initial assessment of the intra-site variability of T_1-, T_2-weighted and FLAIR MR protocols used in the scanning of patients reviewed at the Nottingham University Hospitals NHS Trust, based on scans from 115 patients, is presented in Table 6-2. A visual inspection of the 477 individual T_1-, T_2-weighted and FLAIR hospital scans allowed to conclude that the vast majority of the routine, non-standardised scans were of high quality; only 60 scans (12.7%) of all the clinical scans had significant artefacts, which included Gibb's rings, excess movement and low contrast.

6.4.2 Patient demographics

Figure 6-4 outlines the patient selection process for this study. Out of the 115 patients with appropriate ethical consent, 27 patients (14 females) had at least two clinical and two research scans and were subsequently included in the analysis. Mean age, at the time of the first MRI, was 35 years (SD=9 years) and at the time of their first scan, 2 patients had CIS, 23 RR, 1 SP and 1 PP MS. The median time between the baseline and the latest clinical MRI scan was equal to 4 years 3 months (range: 4 months to 10 years 8 months) and 9 months (range: 1 month to 4 years 10 months) for the research scans.

Table 6-2 Acquisition parameters for selected common sequences in a sample of 78 patients from the Nottingham MS Clinic.

Sequence	Scanner	2D/3D	Orientation	Thickness (mm)	TR (ms)	TE (ms)
T_1-weighted	GE 1.5T Philips 1.5T Siemens 1.5T Philips 3T	2D	Sagittal, axial, coronal	3-5	230-850	7-25
	Philips 1.5T Siemens 1.5T Philips 3T	3D	-	1-3	8-22	5-9
T_2-weighted	GE 1.5T Philips 1.5T Siemens 1.5T Philips 3T	2D	Sagittal, axial, coronal	3-5	1,900-12.500	80-120
	Philips 1.5T Philips 3T	3D	-	1-3	1,600-2,000	100-200
FLAIR	GE 1.5T Philips 1.5T Siemens 1.5T GE 3T Philips 3T	2D	Sagittal, axial, coronal	2.5-5	5,000-11,000	80-150

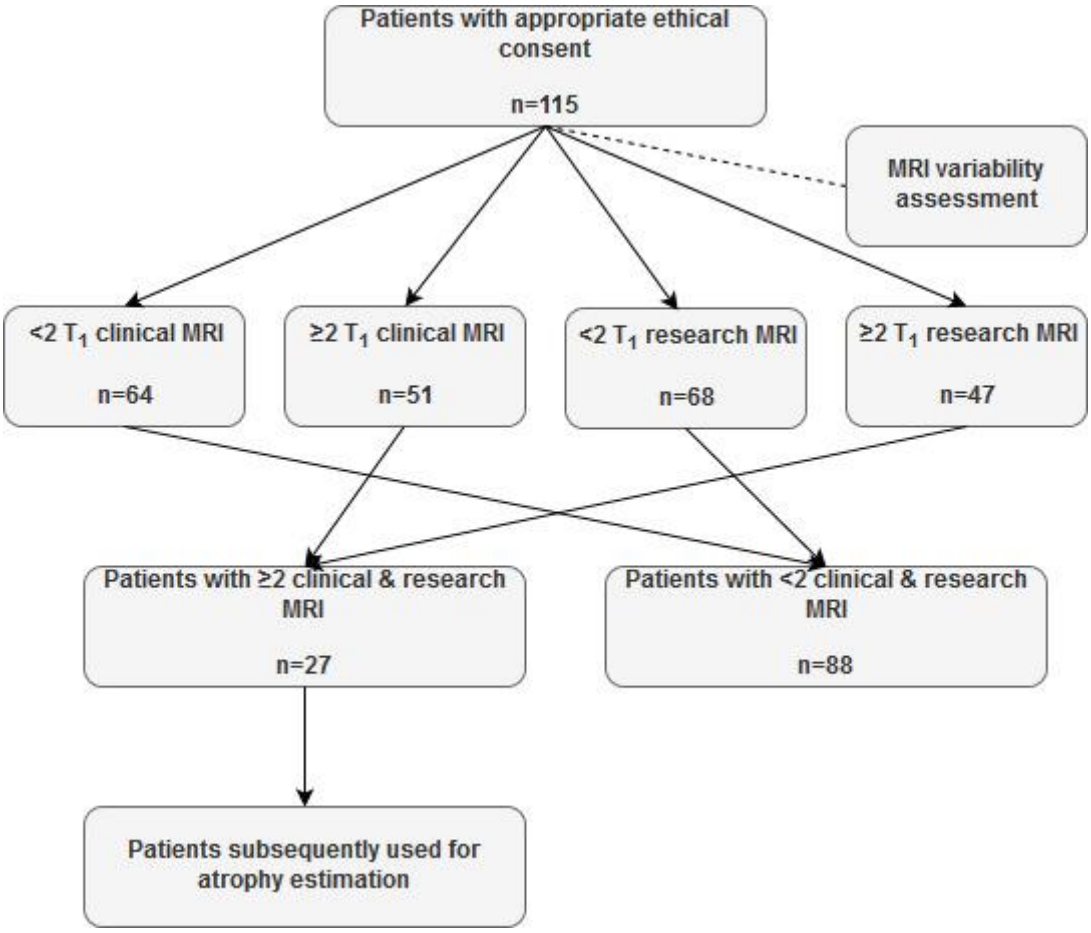


Figure 6-4 Patient flow diagram

6.4.3 Atrophy estimation in incompletely covered clinical scans

None of the 27 pairs of clinical images produced outputs which passed visual inspection despite undertaking multiple possible adjustments of the various SIENA parameters to improve them, including the previously proposed most-optimal parameters of "-B" (bias field correction and neck clean-up) and "-f" (fractional intensity threshold) (Popescu et al., 2012). A detailed outline of the reasons for failure and the parameters used to improve the outcomes is listed in Supplementary Table 1 in Appendix E.

We hypothesised that the reason for failure was that each patient had at least one scan with missing slices. As SIENA relies on the skull to perform image registration between two scans (Smith et al. 2002), the lack of a continuous signal from that structure ultimately led to image processing failures. An example mask, which should only contain brain tissue (orange) shows a large portion of the neck tissue included as part of the mask (see Figure 6-5).

6.4.4 Frankenstein validation

SIENA was able to provide atrophy results in 100% of the artificially simulated Frankenstein scans from the healthy control (see Table 6-3) even after removing 35% of the brain. The average annual PBVC of -0.19% (SE= -0.01%) for the healthy control was in line with the values reported in literature (Hedman et al., 2012). The interquartile range was -0.04%.

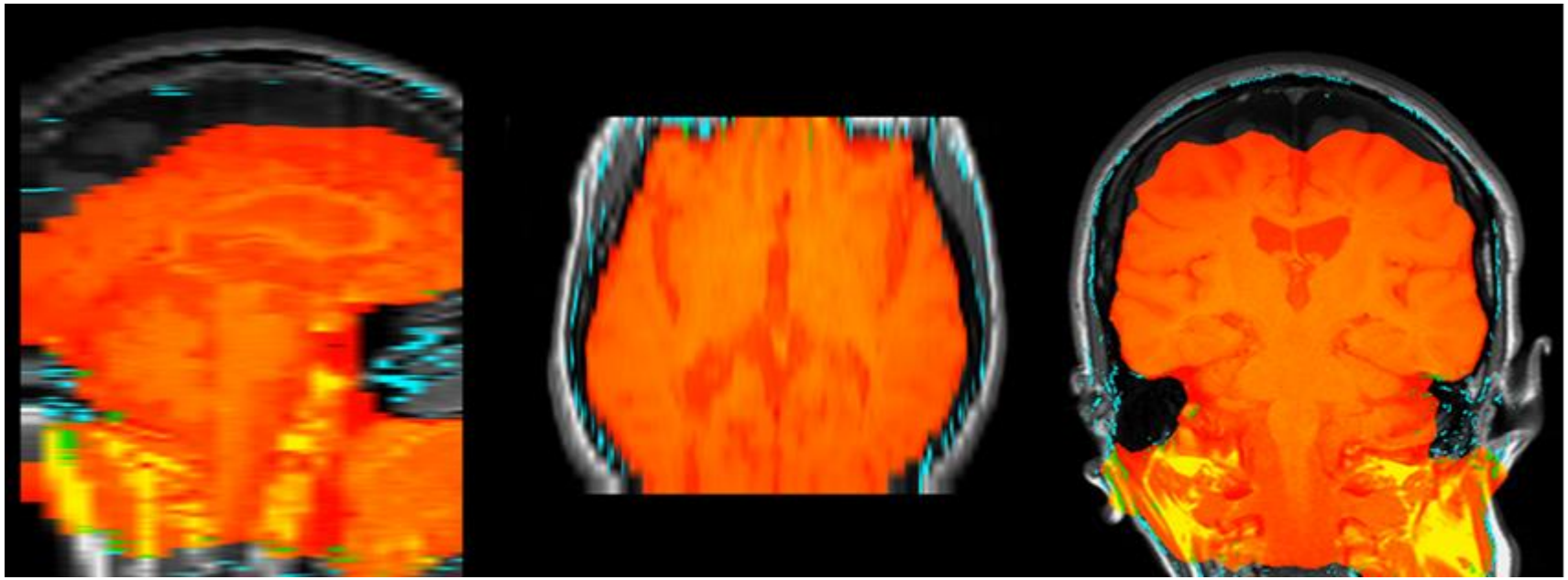


Figure 6-5 Extracted brain tissue mask (orange) from an incompletely acquired clinical MRI scan.

SIENA largely misclassified non-brain tissue, particularly sections of the jaw and neck, as brain tissue leading to atrophy estimation failures.

Table 6-3 Results of the Frankenstein validation

Original slices present (%)		7 year PBVC	Annual PBVC
Baseline scan	**Follow-up scan**		
100	100	-1.54	-0.22
100	80	-1.37	-0.20
65	100	-1.12	-0.16
65	80	-1.23	-0.18
80	80	-1.35	-0.19

PBVC – percentage brain volume change

6.4.5 Atrophy estimation from “Frankenstein” scans

In the clinical cohort, 16 out of 27 Frankenstein scan pairs were processed by Siena, resulting in a 59% success rate. Figure 6-6 demonstrates the marked visual improvement in brain extraction in *Frankenstein* scans, compared to scans with missing slices. Table 6-4 shows the individual results from each patient and lists reasons for failure in the unsuccessful scans.

The mean annual PBVC was equal to -0.005% (SD= 0.44). Annual atrophy values obtained using the *Frankenstein* approach were significantly lower by 0.40% than a typical value of -0.40% cited as *pathological* in the literature, $t(15)=3.592$, $p=0.003$.

6.4.6 Atrophy estimation from research scans

Similarly to the original clinical scans, none of the research scans could be processed by SIENA, which resulted in multiple issues with brain extraction and tissue misclassification. No reliable atrophy estimates could be obtained from 3T or 7T scans despite adjusting the various SIENA parameters, as recommended by Popescu et al. 2012. Examples of multiple failed attempts at brain extraction from a standardised, 7T T_1-weighted research scan can be seen in Figure 6-7. As a result, it was impossible to directly compare the results between atrophy values obtained from the Frankenstein scans and the research scans.

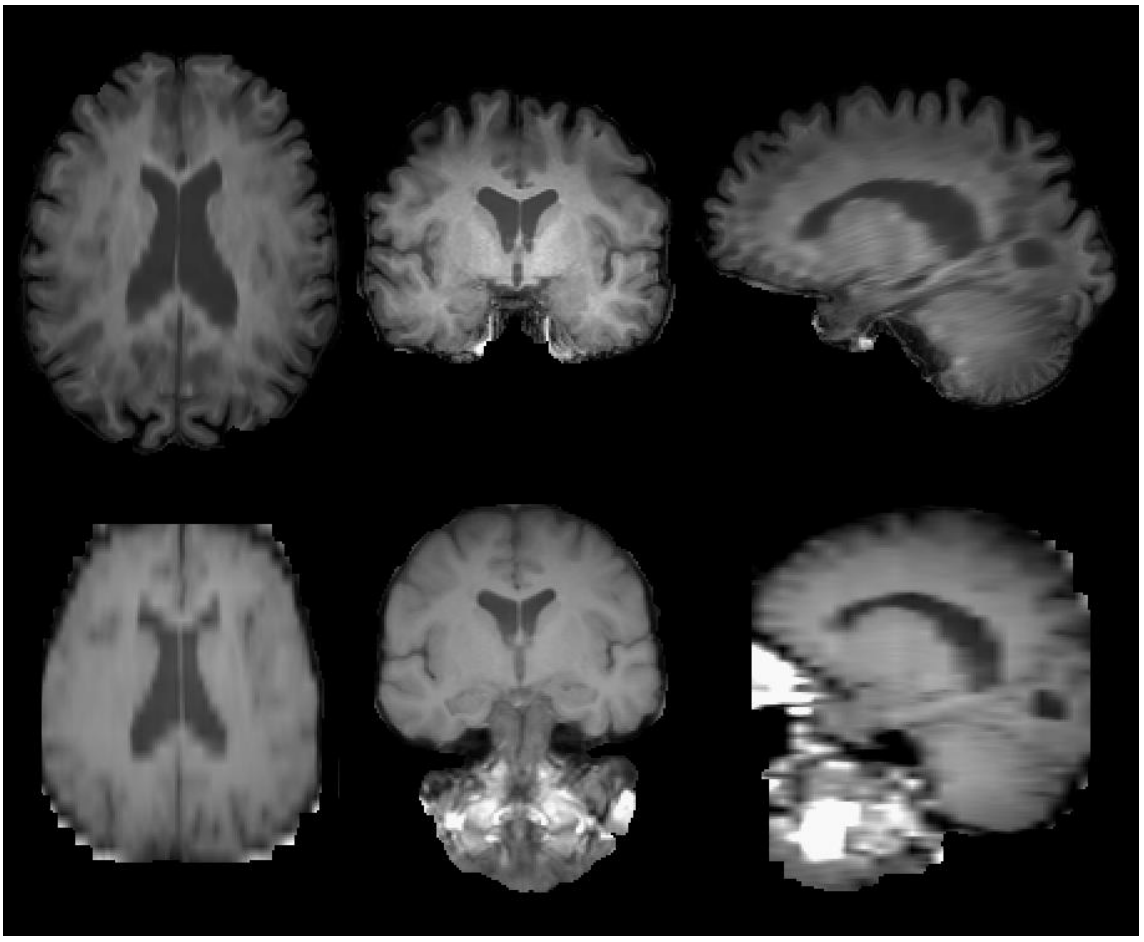


Figure 6-6 Improvement in brain extraction using SIENA, following the application of the Frankenstein approach (top row), compared to to pre-Frankenstein results (bottom row).

Table 6-4 Results from atrophy estimation using SIENA following the application of the Frankenstein approach

ID	Most optimal parameters	Quality rating of baseline scan	Quality rating of the follow-up scan	Atrophy (raw PBVC)	Atrophy (annual PBVC)	Reason for failure
1	-B –f 0.3	3	2	-	-	A large portion of non-brain tissue classed as brain around midbrain (-B -f 0.3 best results)
2	-B –f 0.4	3	2	-	-	Slight overestimation of tissue around the temporal lobes in places and small chunks of tissue missing in some areas (-B -f 0.4) best results)
3	-B -f 0.3	2	2	-0.37	-0.09	
4	-B -f 0.3	1	1	-3.32	-0.46	
5	-B -f 0.3	2	2	-0.99	-0.35	
6	-B -f 0.3	2	2	-0.92	-0.26	
7	-	3	3	-	-	Chunks of the brain are missing in some parts of the brain
8	-B -f 0.2	1	1	-0.37	-0.07	
9	-B f -0.4	2	2	1.79	0.51	
10	-B -f 0.4	2	2	-1.06	-0.30	
11	-B -f 0.3	1	2	-1.19	-0.13	
12	-B -f 0.4	2	2	-0.15	-0.03	
13	-B -f 0.3	2	2	-0.25	-0.07	
14	-	3	3	-	-	Inclusion of optic nerve as brain
15	-B -f 0.4	2	2	-0.30	-0.40	
16	-	2	3	-	-	Time point 2 has bits of skull included and there is general overestimation of what is brain tissue whilst the cerebellum is slightly underestimated.

17	-B -f 0.3	1	2	0.80	0.26	
18	-	3	3	-	-	Brain tissue overestimated around midbrain. Small chunks of brain tissue are missing in the middle of the brain
19	-B			-	-	
20	-	3	3	-	-	Brain tissue underestimated in many places
21	-	3	3	-	-	Brain tissue overestimated close to midbrain and some gyri are underestimated.
22	-B -B 0.4	2	2	0.25	0.06	
23	-B -f 0.5	2	2	-0.78	-0.07	
24	-B -f 0.2	2	2	2.51	1.37	
25	-	3	4	-	-	Brain tissue largely underestimated. Time point 2 lacks brain tissue in the middle of the brain.
26	-B -f 0.4	2	2	-0.39	-0.05	A little chunk missing from time point 2 in the frontal lobe.
27	-	3	3	-	-	Brain tissue around temporal lobes largely overestimated.

PBVC – percentage brain volume change

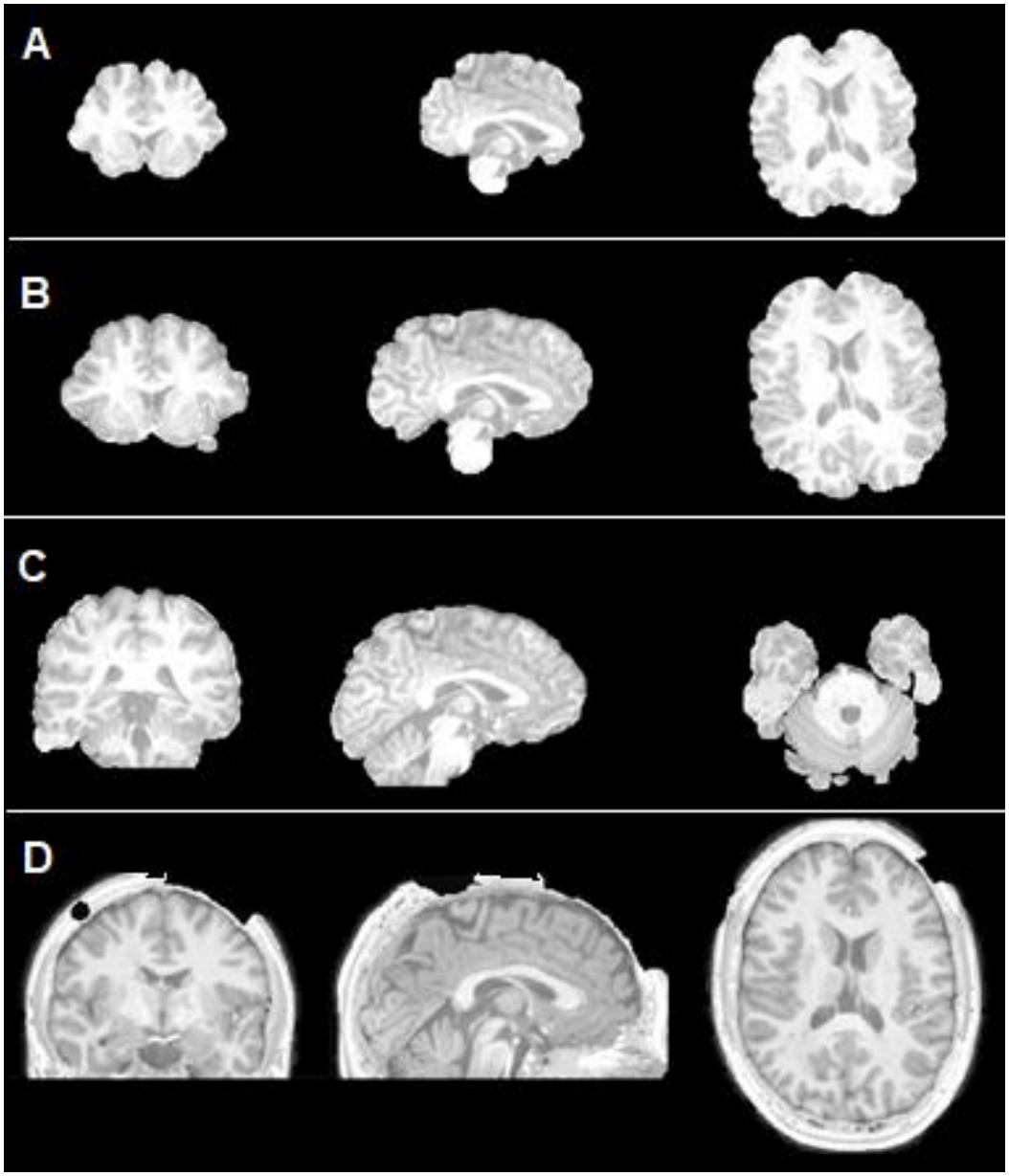


Figure 6-7 Issues with brain extraction from standardised, 7T T1-weighted scans.

A – Initial attempt at extracted brain tissue using BET option –R (robust brain centre estimation);
B – Extracted brain tissue using BET options –R and –f 0.3 (fractional intensity threshold which, in this case, aims to increase the brain outline estimate);
C – Extracted brain tissue using BET options –R –f 0.4. Despite an overall improvement in brain outline estimation, tissue in the cerebellum is still largely underestimated;
D – Extracted brain tissue using BET options –R –f 0.1, which resulted in overestimation of brain tissue and inclusion of the skull.

6.5 Discussion

Cerebral atrophy quantification has been hailed as a promising surrogate marker in MS, which could provide important information about disease progression, disability accumulation, cognitive performance and therapeutic response. In this study, I attempted to use routinely acquired, non-standardised scans from MS patients in order to determine if atrophy could be reliably measured in a clinical setting and introduced a novel *Frankenstein* approach to deal with atrophy estimation failures in clinical scans with missing slices.

6.5.1 Assessment of the variability of clinical protocols in the East Midlands

An audit of the T_1-, T_2-weighted and FLAIR scans used in the MS clinic at Nottingham University Hospitals NHS Trust showed that the variability of the MR protocols is reflective of that in other MS centres across the United Kingdom. The scans were generally of high quality with only around one in ten having more serious artefacts, including severe movement artefacts. Nevertheless, over 75% of all the T_1-weighted images had missing slices resulting in an incomplete coverage of the head, which resulted in image processing failures by SIENA. Discussions with radiology staff suggest that whilst many modern imaging systems acquire 3D data in a relatively short period of time and image the whole head, that was not available in this centre. Incomplete brain coverage during scanning helps facilitate shorter scanning times, which is necessary in the NHS setting, However, this is largely problematic in studies of atrophy, which rely on information from the entire head to calculate brain tissue loss across time.

6.5.2 The *Frankenstein* approach

Our proposed *Frankenstein* approach combines an incompletely imaged scan with a different scan that covers the entire head to allow successful processing through the pipeline, before the non-original slices are discarded and atrophy rate is calculated for the brain tissues present in the original scan only. This resulted in a marked improvement, achieving successful processing of images in 59% of cases (compared to 0% when using the original, pre-Frankenstein scans). Whilst SIENA has the option to provide atrophy estimates in scans with missing slices by ignoring tissue in the top or bottom parts of the brain (SIENA options '-t' and '-b', respectively), there are no options to deal with missing slices in the anterior-posterior directions. A recent study comparing atrophy estimates between whole brain and 'central' brain atrophy (which uses the middle portion of the brain with the top and bottom slices removed) with SIENA showed that the latter resulted in a mean increase of 14% in annual PBVC reported (Ruberte et al., 2018). The validation of our approach, which accommodates for missing slices in any direction, demonstrated that even scans simulated to have over a third of their slices missing, obtained very similar atrophy values to those from the fully covered, original scans. The interquartile range of 0.04% confirmed that our approach provides highly consistent, robust atrophy estimates.

However, when applying this approach to MS patients, the results were more variable and some produced positive atrophy values whilst others reported virtually no change in brain volume across time. This is not unexpected as studies of atrophy generally report group-level results that demonstrate accelerated tissue loss in MS compared to healthy controls but individual-level

measurements do not always result in atrophy values that within the pathological range or negative (De Stefano et al., 2015). As in this study, we have only examined 16 patients using the Frankenstein approach, a larger number of subjects is needed to validate our results. Indeed, a recent study using non-standardised, clinical images from over 500 MS patients showed that the median annualised PBVC was equal to -0.3% suggesting that perhaps atrophy is underestimated when scans acquired on different scanners and protocols are used (Zivadinov et al., 2018)

The higher rate of atrophy estimation failures when applying the *Frankenstein* approach to scans of MS patients, compared to scans from the healthy control, could be attributed to WM pathologies present in the scans. Previously, a study of atrophy measurements using standardised, clinical scans reported highest atrophy estimation failure rates for MS patients, followed by CIS patients, whilst less than 2% of scans from healthy controls failed the analysis (Ghione et al., 2018). Conversely, it is highly plausible that differences in image contrast, gradient distortions or issues with tissue segmentation and/or registration led to misclassification errors, thus masking any real loss of tissue when using non-standardised scans. Future analyses using a larger population of patients are needed to determine whether pathological rates of atrophy can be detected with the *Frankenstein* approach.

6.5.3 Previous studies of atrophy assessment in the clinic

Ghione and colleagues studied atrophy assessment using 3D T_1-weighted clinical images with 1mm thick slices without gaps, which had higher resolution

compared to the scans used in this study (slice thickness ranged from 1 to 5mm). Despite that, longitudinal assessment of atrophy failed in over a third of all the subjects from that study, with the highest error rates reported for images acquired on different scanners or with different protocols, orientation or slice thickness. Another study from that group reported failure rates in estimating PBVC from non-standardised, clinical MRI scans in 56% of patients with 2D T_1-weighted scans and 82% of 3D T_1-weighted (Zivadinov et al., 2018). This seems surprising as generally 3D MRI is preferable to 2D thanks to its higher spatial resolution and reduction of partial volume effects (De Stefano et al., 2014, Zivadinov et al., 2016a). In fact, both the studies concluded that whilst it is possible to measure atrophy in a clinical setting using SIENA, measurements of the percentage lateral ventricular volume change resulted in fewer failures presumably as a result of more accelerated atrophy rates in the ventricles and better contrast at the ventricle-brain tissue border. Whilst ventricular measurements could be a useful proxy of cerebral atrophy in MS, whole brain atrophy provides a global measure of neurodegeneration and has been found to be one of the best predictors of disease progression, therapeutic response and long-term disability in MS patients (Barkhof et al., 2009, Zivadinov et al., 2013, Maghzi et al., 2014).

6.5.4 Atrophy estimation failures using high resolution research scans

In this study, high resolution 3D 3T and 7T T_1-weighted research scans could not be successfully processed by SIENA. Previous studies of atrophy have showed that SIENA resulted in a higher failure rate than SIENAX, a segmentation-based approach, when using standardised, clinical scans (37% vs. 19% failure rate, respectively) (Ghione et al., 2018) and clinical trial data (values not specified)

(Radue et al., 2017). Ghione suggests that this is likely due to the fact that segmentation-based approaches rely on the analysis of one scan at a time, whereas registration-based approaches analyse two scans simultaneously, making it more susceptible to encountering errors. This is most likely due to errors encountered during the brain extraction step, which was the main reason behind the failed attempts in atrophy estimation in this study. Attempts at improving brain extraction have been reported only recently (Battaglini et al., 2018, Nakamura et al., 2018) and we expect that future attempts using these proposed approaches might improve the results, at least in a proportion of scans.

6.5.5 Study limitations

This study faced limitations. First and foremost, the sample included was limited to 27 patients. Whilst I initially identified over one hundred patients as potentially eligible to take part, ultimately less than a quarter had at least two clinical and two research scans of high-quality. This restricted the analysis to a small population of patients making our conclusions limited until the *Frankenstein* approach can be tested on a bigger population of MS patients.

Moreover, previous literature suggests that atrophy does not affect all brain areas at the same rate and, as such, the *Frankenstein* approach could lead to under- or over-estimation of atrophy in individuals. It is important to validate the measurements with clinical outcomes. For example, both the frontal and temporal lobes have been previously reported as areas with accelerated atrophy in healthy individuals during ageing (Raz et al., 1997, Scahill et al., 2003). In MS, thalamic atrophy and ventricular enlargement have been proposed as the most prominent markers of neurodegeneration (Fox et al., 2000, Azevedo et al., 2018)

suggesting that the atrophy estimates obtained from Frankenstein images might still provide a good estimate of the brain tissue loss in MS.

Finally, due to time constraints of this project, I was unable to test other brain extraction approaches, which can be combined with SIENA. A future analysis of this dataset will help determine whether the proposed approaches improve atrophy estimates from SIENA obtained using the clinical and research scans from our centre.

6.5.6 Future analyses

Whilst this study allowed us to validate our proposed approach using data from a healthy volunteer and apply our technique to a small sample of MS patients, future analyses on a larger cohort of patients are still needed. More work aimed at trying to better understand why the method did not work with high resolution scans is also needed. If the *Frankenstein* approach proves to be successful in providing reliable and robust atrophy estimates, large scale studies could be undertaken. Moreover, recent advances in image quality transfer, which allows information from high quality MRI data to be transferred onto low quality images to enhance them (Alexander et al., 2014) or volumetric image reconstruction of 2D data into 3D (Ebner et al., 2018), in combination with the Frankenstein approach, might further improve the scans with missing slices and successful assessment of atrophy .

6.5.7 Conclusions

Atrophy measurements obtained from clinical scans could be used to answer research questions in MS at a fraction of the cost compared to prospectively acquiring the data. However, before that can happen, the quality of the scans and accuracy of atrophy measurements scans needs to be verified. An audit of the clinical protocols from the Nottingham University Hospitals NHS Trust revealed that the T_1-, T_2-weighted and FLAIR protocols were representative of MR scans used across the UK. Whilst our attempts at comparing atrophy using the registration-based, longitudinal approach SIENA on both the non-standardised, routine, clinical MRI images and high-resolution, 3D research-quality images failed, the proposed method provided atrophy values in line with those cited in the literature in a healthy control and significantly decreased the rate of failure in SIENA in a small population of MS patients. Future analyses using a larger cohort of patients and applying the recently proposed, improved brain extraction techniques with SIENA, will demonstrate whether atrophy measurements obtained using the *Frankenstein* method are reliable when compared to high-quality research scans images using 3T and 7T. Moreover, future developments in image processing, including image quality transfer or aritificial intelligence could help further optimise the scans for reliable atrophy estimation.

CHAPTER 7: THESIS CONCLUSIONS

7.1 The usefulness of the central vein sign in the diagnosis of multiple sclerosis

Previous studies of the CV sign using ultra-high field MRI have demonstrated its presence in nearly all MS lesions (Tallantyre et al., 2008, Mistry et al., 2013). Despite the reduced visibility of deoxygenated haemoglobin in veins at 1.5T and 3T, compared to 7T (Tallantyre et al., 2009, Maggi et al., 2018), successful differentiation between MS and non-MS is still possible with many previous studies reporting up to 100% success in differentiating between MS and non-MS at 3T (Campion et al., 2017, Cortese et al., 2018, Maggi et al., 2018). In this thesis, I tested the detectability of the CV sign in a large, multi-centre study of over 600 patients using variable, clinically-used GRE sequences. Whilst 3D T2*-weighted MRI with high EPI was the most optimal sequence, the CV sign proved to be a useful diagnostic marker of MS, irrespective of the protocol and scanner used and resulted in 80% sensitivity and 78% specificity, when a threshold of 25% of lesions with CVs was used to differentiate between CIS/MS and non-MS. This is an important finding as the lack of scanner and protocol standardization is one of the major obstacles in translating research findings into clinical practice. Furthermore, some of the previously reported studies of the CV sign have used contrast enhancement or post-processing of images (combining FLAIR and T2* images to produce FLAIR* (Campion et al., 2017)), which could potentially limit the application of the CV sign in a clinical setting with limited financial and time resources. Findings from this study reinforce of the value of CV assessment using conventional GRE protocols utilised by major European MS centres and

demonstrate that the CV sign is not significantly affected by patient clinical characteristics, including patient age, and as a result, can be considered a robust biomarker of MS lesions.

It is encouraging that the CV sign can be used to differentiate between patients with MS and non-MS, however, its diagnostic usefulness in a ´real´ clinical setting can only be determined when tested in a longitudinal cohort of patients presenting with a first demyelinating attack and a suspicion of MS. So far, only one previous study had prospectively tested the value of the CV sign in differentiating between MS and non-MS in a prospective study of 28 WM lesions (Kau et al., 2013). The results from the two prospective studies presented in this thesis demonstrate that the CV sign can successfully predict a final diagnosis of MS in patients presenting with a typical CIS (unilateral optic neuritis, focal brainstem or cerebellar syndrome, or myelitis) and atypical CIS or RIS (where the clinical and MRI presentation of the disease were not in line with each other). In the STAR MS study, all the patients presenting with atypical clinical or MRI findings, who went on to develop MS, had CVs in over 40% of all of their WM lesions suggesting the CV sign is a highly sensitive marker of MS lesions. Conversely, when applied to differentiate between typical CIS and non-MS, the CV sign proved 100% specific as none of the non-MS patients had CVs in over 40% of their WM lesions.

In this thesis, I tested a variety of ´rules´ which aim to facilitate differentiation between MS and non-MS patients, based on a proposed threshold or number of lesions with CVs. Whilst using a 25% threshold (in the MAGNIMS study) and a

41% threshold (in the STAR MS study) led to high diagnostic sensitivity and specificity, the integration of CV assessment into clinical practice would be most facilitated by the use of a fixed number of lesions, rather than a proportion approach. Results from this thesis support the idea that finding as few as two WM lesions with CVs can achieve over 80% specificity in differentiating between MS and non-MS, based on results using subcortical lesions in the MAGNIMS study. Moreover, when combined with evidence of DIS, 2 CVs showed better specificity for the diagnosis of MS than OCBs, which are already incorporated into the current diagnostic criteria for MS. Based on the finding from this thesis, the proposed threshold of two lesions appears as a very promising candidate to replace the 40% threshold thanks to its easy implementation in a clinical setting in patients with a small lesion load, characteristic of CIS patients and should be further tested in future prospective studies, also when incorporated with evidence of DIS.

Nevertheless, whilst the CV sign can be used to correctly differentiate between MS and non-MS and predict a diagnosis of MS at the time of CIS in the majority of cases, it did not lead to the correct classification of all the patients. One of the drawbacks of diagnostic studies in MS is the lack of a gold standard. The current diagnostic criteria allow for a quick diagnosis on the basis of radiological progression, however, a diagnosis of CDMS can only be confirmed in patients with a second clinical attack and many publications still report significant problems with misdiagnosis of MS even after many years of follow-up. Confidence for a clinical diagnosis increases with time so, only long term follow-

up of the patients included in the two prospective studies in this thesis will be able to confirm the true diagnostic sensitivity and specificity of the CV sign.

Whilst most of the work in this thesis focused on the detection of the CV sign on conventional GRE sequences, the role of iron, principally in the form of PIRs surrounding some MS lesions, showed equally promising results in the study of typical CIS and non-MS patients. Iron accumulation within MS lesions has previously been visualised primarily using non-conventional sequences at 7T (Dal-Bianco et al., 2017, Absinta et al., 2018, Yao et al., 2018). The role of iron accumulation in MS lesions has been linked to chronic lesion activity and disease progression (Absinta et al., 2015, Absinta et al., 2018) and in this thesis, I demonstrated that it can be useful diagnostically to differentiate between MS and non-MS, and to predict conversion to MS. Moreover, combining PIRs and CVs with evidence of DIS, demonstrated they are a promising alternative to DIT or OCBs at the time of the baseline scan. Concerns over Gadolinium accumulation in the brain and side-effects from lumbar punctures make the integration of these two MS lesion features into clinical practice, promising.

7.2 Atrophy estimation from routine, non-standardised, clinical MRI

Brain atrophy is an attractive measure of global neurodegeneration in MS, which has been shown to predict conversion to CDMS, future disability, disease progression as well as cognitive decline (Zivadinov et al., 2001, Di Filippo et al., 2010, Uher et al., 2014). In this thesis, I aimed to verify whether brain atrophy could be reliably measured in a clinical setting using variable, non-standardised

MRI data. If proven successful, routinely collected scans could be explored to answer research questions in MS in large, observational studies.

The audit of clinical MRI scans performed at the Nottingham University Hospitals NHS Trust demonstrated that there was a vast variability of scanners and protocols used. This suggests that the findings presented in this thesis, could be applied to MRI data from other hospitals in the country. Nevertheless, the incomplete coverage of the head in the clinical MRIs led to atrophy estimation errors using SIENA, the most widely used longitudinal approach for brain volume loss estimation.

In this thesis, I presented an approach, which combines scans with missing slices with a fully-covered template scan to allow successful processing through SIENA, before atrophy is estimated in the original scans only. The *Frankenstein* approach provided robust and reasonable atrophy estimations in a healthy control, in line with those cited in the literature, even in scans with over a third of their slices missing. It also made it possible to successfully process with SIENA, 59% of the clinical scans from MS patients, compared to none using the original, pre-Frankenstein scans. These findings are exciting as they open up a way to conduct large-scale studies combining clinical and MRI data from patients who have consented for their data to be used as part of the UK MS Register. Whilst randomised clinical trials are currently the gold standard for evidence-based practise, wider use of registries could prove invaluable in filling the information gaps in MS that currently cannot be explored due to logistic, time and financial constraints (Frieden, 2017).

Whilst promising, the *Frankenstein* approach needs to be validated against high-resolution, standardised research scans obtained a large cohort of patients with MS. With new improvements in brain extraction, routinely collected MRI data could vastly complement current research.

7.3 Future research

Based on the findings presented in this thesis, the next step in determining whether the CV sign can be incorporated as a diagnostic marker of MS lesions and used in the diagnostic criteria of MS, would require a multi-centre study of patients presenting with a suspicion of MS. A comparison of the diagnostic performance of the CV sign against the use of positive OCBs, in patients with typical and atypical presentations of a demyelinating attack, would help determine which of the tests is the most accurate, well-tolerated by patients and financially viable. Such a study is already under way in the United Kingdom.

Whilst the data presented in this thesis showed that findings as few as two lesions with CVs can provide high sensitivity and specificity in predicting MS, future studies should determine the ideal threshold for a diagnosis, using a fixed number of lesions. This is especially important in CIS patients, who might present with very few lesions on their baseline scan.

The CV sign remains unreported in the spinal cord, due to methodological issues. As presence of spinal cord lesions at the time of the baseline significantly increases the risk of a subsequent MS diagnosis (Brownlee et al.,

2019), it would be interesting to determine if there is a way to reliably visualise it outside of the brain.

Whilst brain atrophy assessment is unlikely to benefit of individual patients, reliable estimation of atrophy values obtained from clinical scans needs further work as can facilitate the use of wealth of scans collected every year for clinical purposes. The atrophy values obtained using the *Frankenstein* approach need further validation in a larger cohort of patients against clinical outcomes. With increasing use of patient registries and interest in ´big data´, combining the Nottingham data from scans from other centres could help validate the Frankenstein approach and, if successful, help explore the vast swaths of MRI data, which remain currently unexplored in the NHS system.

With more developments in the field of image processing and artificial intelligence, it is likely that soon both the CV sign and atrophy assessment in scans with missing slices will be able to be fully automated. If shown to be accurate and reliable when assessed without manual processing, these two biomarkers have the potential to help improve the lives of people with MS.

APPENDIX A

Guidelines for the identification of lesions and central veins in the MAGNIMS study:

1. Image quality

MR sessions that are not amenable to the detection of veins need to be excluded; e.g. in such cases where the normal appearing periventricular venous vasculature is not clearly visible on the SWI images due to motion artifacts, to much noise or a poor resolution.

2. Lesion selection

white matter lesions only; grey matter lesions will be excluded

confluent lesions without clearly distinguishable „lesion fingers“ will be excluded

all lesions need to be larger than 3mm or/and 5 voxel (short axis diameter)

each lesion must be clearly distinguishable from normal appearing white matter

the signal-to-noise ratio and the overall image quality need to be sufficient in the lesion area

3. Lesion characterization

individual lesion ID for each lesion

lesion location (juxtacortical (direct contact to cortex), periventricular (direct contact to ventricle), other white matter lesions); periventricular lesions that additionally fulfil other lesion criteria will be named periventricular lesion

lesion size: both short and long axis diameter

for confluent lesions: how many fingers are there (if possible, break confluent into fingers)

existence of an intra-lesional vein?

existence of a central vein?

4. Definition of an intra-lesional vein

1.) the vein appears hypointense on T2*w or SWI images

2.) the vein is detectable in at least 2 perpendicular planes (only applicable for 3D dataset)

3.) the vein passes through the edges of the lesion on at least one place. Thus the vein is visible outside of the lesion too

5. Definition of “centrality“

1# - in general: equidistance of the vein to the lesions edges + the vein passes through the edges of the lesions on one or two places

2# - a more specific definition for ovoid lesions: the vein needs to meet criteria 1# and need to run along the longitudinal axis of an ovoid lesion or a single „lesion finger“. The lesion must follow the course of the vessel.

3# - a more specific definition for confluent lesions with distinct lesion fingers: the vein needs to meet criteria 1# and need to run along the longitudinal axis of a single „lesion finger“ in at least 2 slices. The lesion must follow the course of the vessel(s).

6. 3D-Slicer/Annotations/Description

- Layout: ruler-ID,lesion localization, existence of an intralesional vein, existence of a central vein,number of lesion fingers (if applicable)

- ruler-ID
- lesion localization: A= juxtacortical, B= periventricular, C= other white matter lesion
- existence of an intralesional vein: 1=yes, 0=no
- existence of a central vein: 1=yes, 0=no
- number of lesion fingers (if applicable)
- ex.: M_1,B,1,1,3 = M_1 + periventricular + intralesional vein + central vein + 3 lesion fingers
- ex.: M_4,A,0,0 = M_4 + juxtacortical + no intralesional vein + no central vein

APPENDIX B

The following R code, developed by Dr Chris Tench, was used to calculate the sensitivity and specificity of the CV sign when only counting a subset of lesions.

Simplified analysis no. 1 based on selecting a fixed number of WM lesions and applying various diagnostic thresholds to make a diagnosis of MS. The variables used were as follows: Ncv (number of central veins reported per patient), Nl (number of lesions reported per patient), MS (clinical diagnosis coded as a binary variable).:

```
CV<-read.table("c:\\temp\\CVs.txt",header=T)

p<-CV$Ncv/CV$Nl

pick= [number]      #ADJUST THIS FOR THE RULE
threshold= [threshold]   #ADJUST THIS FOR THE THRESHOLD
iterations=100

sensitivity=rep(0,iterations)
specificity=rep(0,iterations)

q<-p
for (i in 1:iterations)
{
     for (j in 1:length(n)) q[j]=rbinom(1,pick,p[j])/pick

     Tp=0
     Fp=0
     Tn=0
     Fn=0
     for (j in 1:length(n))
     {
          if (q[j]>=threshold)
          {
               if (CV$MS[j]==1) Tp=Tp+1
               else Fp=Fp+1
          }
          else
          {
               if (CV$MS[j]==1) Fn=Fn+1
               else Tn=Tn+1
```

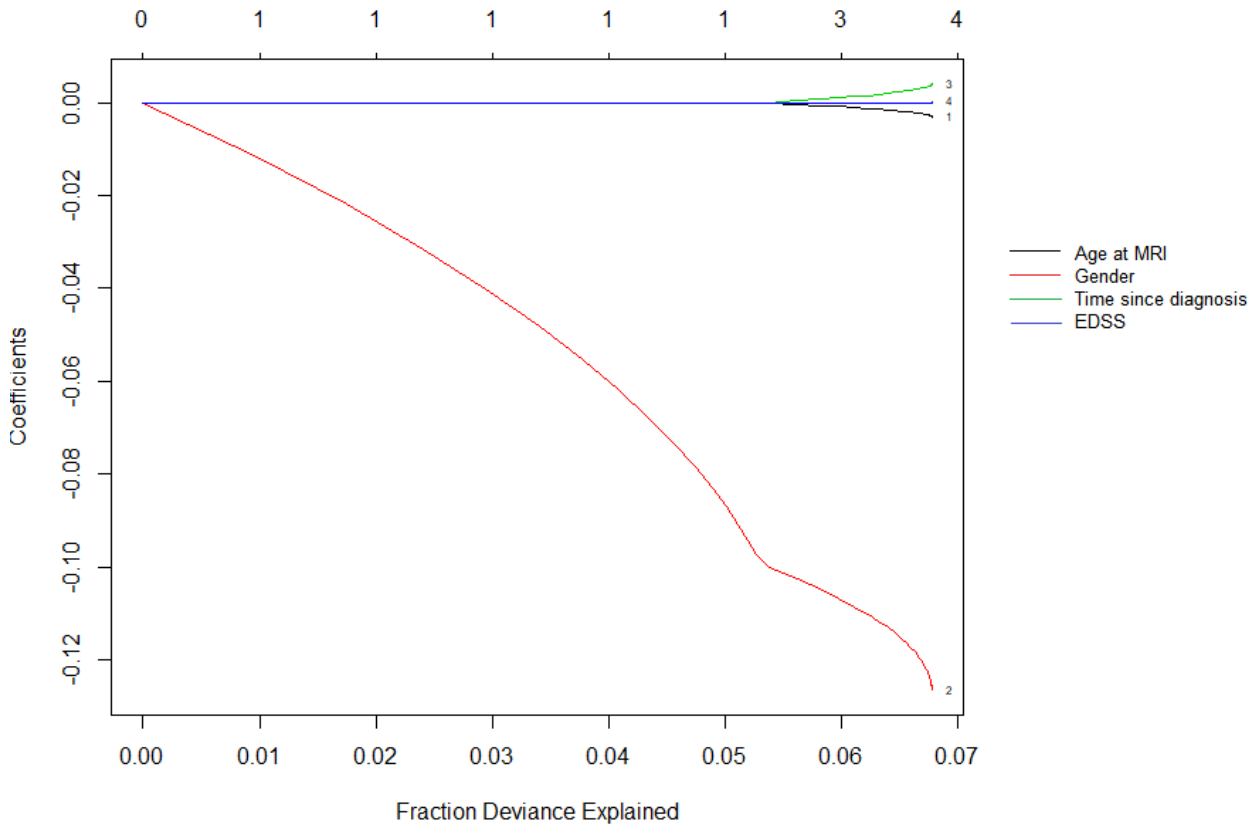
```
          }
     }
     sensitivity[i]=Tp/(Tp+Fn)
     specificity[i]=Tn/(Tn+Fp)
}
hist(sensitivity)
hist(specificity)

print(mean(sensitivity))
print(sd(sensitivity))


print(mean(specificity))
print(sd(specificity))
```

Simplified analysis no. 2 based on selecting a fixed number of lesions with CVs to make a diagnosis of MS:

```
CV<-
read.table("c:\\Users\\lpxmac\\Desktop\\CVs.txt",header=T)

p<-CV$Ncv/CV$Nl

find=2#ADJUST THIS FOR THE RULE
iterations=100

sensitivity=rep(0,iterations)
specificity=rep(0,iterations)

nl<-CV$Nl
for (i in 1:iterations)
{
     for (j in 1:length(nl)) nl[j]=rbinom(1,CV$Nl[i],p[j])

     Tp=0
     Fp=0
     Tn=0
     Fn=0
     for (j in 1:length(nl))
     {
          if (nl[j]>=find)
          {
               if (CV$MS[j]==1) Tp=Tp+1
               else Fp=Fp+1
          }
          else
          {
               if (CV$MS[j]==1) Fn=Fn+1
               else Tn=Tn+1
          }
     }
     sensitivity[i]=Tp/(Tp+Fn)
     specificity[i]=Tn/(Tn+Fp)
}
```

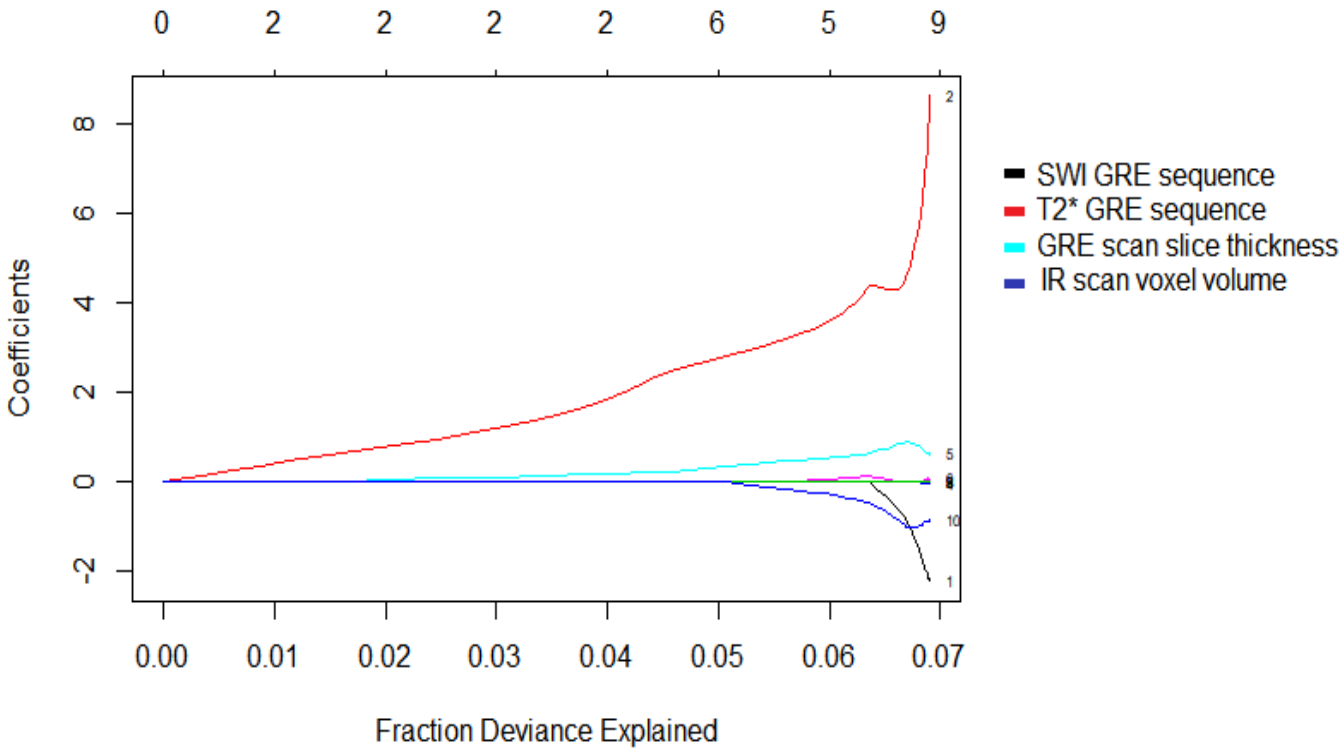
```
hist(sensitivity)
hist(specificity)

print(mean(sensitivity))
print(sd(sensitivity))


print(mean(specificity))
print(sd(specificity))
```

APPENDIX C



Supplementary Figure 1 The LASSO analysis of the effect of clinical variables on the proportion of WM lesions with CVs

The clinical variables included in the model (age, gender, time since diagnosis and EDSS) could explain up to 6% of the deviance in the data with the variable gender explaining 5% of that deviance.



Supplementary Figure 2 The LASSO analysis of the effect of MRI protocol variables on the agreement between the clinical and CV diagnoses using the 25% threshold

The Fraction Deviance Explained by the model was 7%. None of the variables contributed significantly to the model. The variables with the highest contributions are depicted in the graph.

APPENDIX D

MS REGISTER

Swansea University
Prifysgol Abertawe

UK MS Register Consent Form

	Please initial box
1. I confirm that I have read and understood the participant information sheet v3.1 (28/5/2014) for this study. I have had the opportunity to consider this information and to ask questions.	☐
2. I understand that my ongoing participation is voluntary and that I am free to withdraw at any time without giving a reason.	☐
3. I understand that identifiable information from my medical notes and from health related records will be collected and data linkage will be carried out by the UK MS Register.	☐
4. I understand that the UK MS Register will not publish identifiable information from any data that has been collected.	☐
5. I understand that my data will be anonymously linked to other health care related data stored by the UK MS Register.	☐
6. I understand that this anonymous data may be shared with other MS researchers subject to appropriate governance.	☐
7. I agree to be contacted again by the UK MS Register team if I may be eligible to take part in further research.	☐

Participant details

Your name: _____

Your signature: _____

Your email: _____

Your mobile phone: _____

Date signed: _____

Person taking consent details

Name of person taking consent:

_____

Signature of person taking consent:

_____

Date signed: _____

Useful information

Identifiable Information: Name, Gender, Postcode, Date of Birth & NHS Number (Participant Information Leaflet p6)

Medical Notes such as: Doctors notes, nursing records, referral letters (Participant Information Leaflet p6)

Health Related Records such as: Clinical System, GP Systems, MRI, Biochemistry (Participant Information Leaflet p7)

Data Linkage: Joining health care related data together for research purposes (Participant Information Leaflet p7)

Anonymised Data: Identifiable information is used to link your data. Once linked all identifiers are removed and replaced with an anonymous code.

V3.1 28/5/2014

Study ID: | 1 | 7 | 1 | 4 | 2 | 7 |

APPENDIX E

Supplementary Table 1 Atrophy estimation failures in clinical scans with missing slices

Patient	Baseline scan date	Follow-up scan date	Parameters used	Baseline scan comments	Follow-up scan comments	Sufficient output quality
1	30/11/13	14/04/14	Default	Whole neck included	Whole neck included	N
			-B	Cerebellum is significantly underestimated. Some parts of the meninges included	Parts of the cerebellum and superior parts of the brain are missing	N
			-B -f0.4	Bits of cerebellum cut. Large portions of the meninges included	Overestimation around the pons. Parts of the meninges included.	N
2	12/04/13	29/06/16	Default	Large portion of the neck included	Large portion of the neck included	N
			-B	Superior skull large classed as brain. Optic nerves partially included	Some parts of the meninges and a bit of skull included	N
			-R	Large parts of the skull included as brain	Large part of the neck classed as brain	N
3	11/05/12	17/05/16	Default	Large portion of the eyes and neck included	Neck included	N
			-R	Large portion of the neck included	Very small overestimation around the pons, small chunks of brain missing in places	N
			-B	Overestimation around the pons	Very small overestimation in the temporal lobe	N
4	25/06/09	09/09/16	Default	Eyes and neck included	Neck included	N
			-B	Overestimation around the eyes and neck	Overestimation around the pons and cerebellum underestimated	N

			-R	Neck largely overestimated	Area around the pons largely overestimated	N
5	15/08/13	26/06/16	Default	Slight underestimation of tissue in places. Inclusion of neck region.	Inclusion of meninges and a large chunk of the neck	N
			-B -f 0.6	Gyrus rectus (frontal lobe, eye level) is slightly cut, minimal overestimation around temporal lobes.	Large portion of non-brain tissue around pons classed as brain	N
			-R	Tissue overestimation around the pons. Frontal lobe slightly underestimated	Large portion of the neck included	N
6	07/09/13	02/04/17	Default	Overestimation around pons and eyes	Large portion of the neck included	N
			-B	Overestimation around the eyes and small chunks of the superior brain are missing	Small chunks of the brain missing in places	N
			-R	Overestimation around the eyes and small chunks of the superior brain are missing	Small parts of the eyes included	N
			-S	Overestimation around the pons, parts of the eyes included and small chunks of the brain missing in places	Some parts of the neck included	N
7	05/01/11	01/09/16	Default	Parts of the neck included	Small parts of the neck included	N
			-B	Generally quite good. Very small parts of the brain missing, others have very slight overestimation.	Small parts of the meninges included	N

				Edgepoints do not align well with brain.		
			-B -f 0.6	Posterior brain cut in some places	Tips of the temporal lobes are cut	N
8	12/08/11	09/11/16	Default	Large part of the neck included	Large part of the neck included	N
			-B	Overestimation around the pons	Cerebellum slightly cut in places	N
			-R	Overestimation around the pons	Slight overestimation around the pons, cerebellum slightly cut in places	N
9	04/02/12	12/08/15	Default	Inclusion of eyes and meninges	Overestimation around the pons and inclusion of meninges	N
			-B	Large portion of the meninges included and eyes	Sections of the meninges included, small parts of cerebellum missing	N
			-S	Parts of non-brain tissue included and eyes	A large chunk of the neck included	N
10	22/04/13	27/11/16	Default	Large overestimation of the neck	Overestimation of tissue around the pons	N
			-R	Overestimation around the neck and temporal lobes	Inclusion of the meninges in some parts. Superior brain slightly underestimated	N
			-B	Parts of the brain under and overestimated	Inclusion of the meninges in many places	N
11	03/12/07	11/12/16	Default	Overestimation around the pons	Overestimation around the pons	N
			-B	Small parts of the frontal lobes are missing, eyes included	Parts of the frontal lobes are missing	N
			-S -f 0.4	Large portion of the eyes and neck included	Large portion of the neck included	N

12	02/11/11	26/02/17	Default	The whole neck and eyes included	Large portion of the neck included	N
			-R	The whole neck and eyes included	Overestimation around the pons, large part of the meninges included	N
			-B	Some parts of the meninges included and overestimation of tissue around the pons	Parts of the meninges included, slight overestimation around the pons and small chunks missing in the posterior brain	N
13	21/03/12	25/08/15	Default	Large potion of the neck and eyes included	Large portion of the neck included	N
			-R	Large portion of the neck included	Small overestimation around the pons and occipital brain	N
			-B	Tissue overestimation around the pons	Small chunks of non-brain tissue classed as brain	N
14	12/07/09	29/01/13	Default	Almost the whole neck included	Eyes and neck included	N
			-B	Parts of cerebellum missing	Eyes included & overestimation by the temporal lobe	N
			-R	Eyes and some meninges included	Eyes included & overestimation by the pons	N
15	14/01/10	29/10/10	Default	Eyes included	Small chunk of overestimated tissue around pons	N
			-S	Cerebellum has bits cut in places	Small portions of the neck included and small parts of cerebellum cut	N
			-R	Cerebellum has bits cut in places	Small portions of the neck included and small parts of cerebellum cut	N
16	23/12/13	02/06/17	Default	Large parts of the neck, skull and eye included	Large part of the neck included, tissue underestimated in places	N

			-B	Parts of meninges and eyes included	Tissue significantly underestimated	N
			-R	Overestimation around the pons and inclusion of parts of the meninges	Large portion of the neck included	N
17	16/12/13	20/01/17	Default	Right eye mostly included	The whole neck included	N
			-B	Eyes mostly included	Very small chunks of the brain missing here and there	N
			-S	Eyes mostly included	Neck included	N
			-R	Eyes and neck included	Neck included	N
18	20/06/10	17/02/17	Default	Overestimation around the pons	A large portion of the neck included	N
			-B	Overestimation around the pons	Small overestimation around the pons	N
			-R	Overestimation around the pons	Very small overestimation in places	N
19	16/07/09	11/05/16	Default	Overestimation around the pons and eyes	Brain largely underestimated, large portion of the neck included	N
			-R	Overestimation around the pons and eyes included.	Brain largely underestimated	N
			-B	Eyes included	Brain largely underestimated	N
20	08/12/08	20/02/15	Default	Large overestimation around eyes/midbrain	Large overestimation around the neck	N
			-R	Large overestimation around the eyes	Some overestimation around the neck	N
			-S	Inclusion of eyes and overestimation around midbrain/pons	Inclusion of the neck	N
			-B	Frontal lobe underestimated in places. Overestimation around the eyes	Some overestimation around the neck	N

21	01/03/11	02/02/17	Default	Large overestimation around the pons	Large part of the neck classed as brain	N
			-B	Overestimation around midbrain and eyes	Small chunks of brain missing in places	N
			-S	Overestimation around the pons	Large part of the neck classed as brain	N
			-R	Overestimation around midbrain and eyes	Small chunks of brain missing in places	N
22	12/02/12	25/02/16	Default	Large portion of the neck included	Large portion of the neck included	N
			-B	Eyes slightly overestimated	Meninges included in some place but overall good.	N
			-R	Small parts of the meninges and neck overestimated	Small parts of the meninges and neck overestimated	N
			-S	Large portion of the neck included	Large portion of the neck included	N
23	14/06/06	24/02/17	Default	Large portion of the neck included	Large portion of the neck included	N
			-B	Parts of the eyes included	Small chunks of the brain missing in places	N
			-S	Large portion of the neck included	Large portion of the neck included	N
			-R	Overestimation around the eyes and neck	Overestimation around the eyes and neck	N
24	25/07/12	19/06/14	Default	Large overestimation around pons/neck	Large overestimation around pons/neck	N
			-B	Tissue overestimation around the pons	Small chunks missing from the cerebellum	N
			-R	Tissue overestimation around the pons	Tissue overestimation around the pons	N
25	14/01/10	09/11/16	Default	Overestimation of the pons and meninges	Large portion of the neck and some meninges included	N
			-R	Parts of cerebellum missing, inclusion of meninges	Parts of cerebellum missing, inclusion of meninges	N

			-B	Parts of cerebellum missing, inclusion of meninges	Parts of cerebellum missing, inclusion of meninges	N
26	10/06/08	13/02/16	Default	Large part of the neck included	Large part of the neck included	N
			-B	Parts of the frontal lobe underestimated	Meninges included in some place but overall good.	N
			-R	Overestimation around pons and eyes included	Small overestimation around the pons	N
			-S	Large part of the neck included	Large part of the neck included	N
27	29/01/13	13/07/17	Default	Overestimation around the pons and eyes included	Large portion of the neck and eyes classed as brain	N
			-R	Eyes included	Overestimation around the eyes and pons	N
			-S	Overestimation around the pons	Large portion of the neck included	N
			-B	Eyes included	Small chunks of the brain missing in places	N

REFERENCES

ABSINTA, M., SATI, P., FECHNER, A., SCHINDLER, M. K., NAIR, G. & REICH, D. S. 2018. Identification of Chronic Active Multiple Sclerosis Lesions on 3T MRI. *American Journal of Neuroradiology*.

ABSINTA, M., SATI, P., FILIPPI, M., OHAYON, J., CORTESE, I. & REICH, D. 2015. Failure Of Early Lesion Repair In Multiple Sclerosis Lesions With Persistent Phase RIM (P4.006). *Neurology,* 84.

ABSINTA, M., SATI, P., GAITAN, M. I., MAGGI, P., CORTESE, I. C., FILIPPI, M. & REICH, D. S. 2013. Seven-tesla phase imaging of acute multiple sclerosis lesions: a new window into the inflammatory process. *Ann Neurol,* 74**,** 669-78.

ABSINTA, M., SATI, P., SCHINDLER, M., LEIBOVITCH, E. C., OHAYON, J., WU, T., MEANI, A., FILIPPI, M., JACOBSON, S., CORTESE, I. C. & REICH, D. S. 2016. Persistent 7-tesla phase rim predicts poor outcome in new multiple sclerosis patient lesions. *J Clin Invest,* 126**,** 2597-609.

AGGARWAL, S., KUMAR, S. & TOPALOGLU, H. 2016. PND67 - Trends In Design, Conduct, And Outcomes Measures For Patient Registries For Multiple Sclerosis. *Value in Health,* 19**,** A70.

AL-ABDULLA, N. A., PORTERA-CAILLIAU, C. & MARTIN, L. J. 1998. Occipital cortex ablation in adult rat causes retrograde neuronal death in the lateral geniculate nucleus that resembles apoptosis. *Neuroscience,* 86**,** 191-209.

ALEXANDER, D. C., ZIKIC, D., ZHANG, J., ZHANG, H. & CRIMINISI, A. 2014. Image Quality Transfer via Random Forest Regression: Applications in Diffusion MRI. *Cham. Springer International Publishing*, 225-232.

ANDERSON, V. M., BARTLETT, J. W., FOX, N. C., FISNIKU, L. & MILLER, D. H. 2007a. Detecting treatment effects on brain atrophy in relapsing remitting multiple sclerosis: Sample size estimates. *Journal of Neurology,* 254**,** 1588.

ANDERSON, V. M., FERNANDO, K. T., DAVIES, G. R., RASHID, W., FROST, C., FOX, N. C. & MILLER, D. H. 2007b. Cerebral atrophy measurement in clinically isolated syndromes and relapsing remitting multiple sclerosis: a comparison of registration-based methods. *J Neuroimaging,* 17**,** 61-8.

ARRAMBIDE, G., TINTORE, M., ESPEJO, C., AUGER, C., CASTILLO, M., RIO, J., CASTILLO, J., VIDAL-JORDANA, A., GALAN, I., NOS, C., MITJANA, R., MULERO, P., DE BARROS, A., RODRIGUEZ-ACEVEDO, B., MIDAGLIA, L., SASTRE-GARRIGA, J., ROVIRA, A., COMABELLA, M. & MONTALBAN, X. 2018. The value of oligoclonal bands in the multiple sclerosis diagnostic criteria. *Brain,* 141**,** 1075-1084.

ASHBURNER, J. & FRISTON, K. J. 2005. Unified segmentation. *NeuroImage,* 26**,** 839-851.

ASHTON, E. A., TAKAHASHI, C., BERG, M. J., GOODMAN, A., TOTTERMAN, S. & EKHOLM, S. 2003. Accuracy and reproducibility of manual and semiautomated quantification of MS lesions by MRI. *J Magn Reson Imaging,* 17**,** 300-8.

AZEVEDO, C. J., CEN, S. Y., KHADKA, S., LIU, S., KORNAK, J., SHI, Y., ZHENG, L., HAUSER, S. L. & PELLETIER, D. 2018. Thalamic atrophy in

multiple sclerosis: A magnetic resonance imaging marker of neurodegeneration throughout disease. *Ann Neurol,* 83**,** 223-234.
BAGNATO, F., HAMETNER, S., YAO, B., VAN GELDEREN, P., MERKLE, H., CANTOR, F. K., LASSMANN, H. & DUYN, J. H. 2011. Tracking iron in multiple sclerosis: a combined imaging and histopathological study at 7 Tesla. *Brain,* 134**,** 3602-15.
BALCHANDANI, P. & NAIDICH, T. P. 2015. Ultra-High-Field MR Neuroimaging. *AJNR Am J Neuroradiol,* 36**,** 1204-15.
BARKHOF, F., BRÜCK, W., DE GROOT, C. A. & ET AL. 2003. Remyelinated lesions in multiple sclerosis: Magnetic resonance image appearance. *Archives of Neurology,* 60**,** 1073-1081.
BARKHOF, F., CALABRESI, P. A., MILLER, D. H. & REINGOLD, S. C. 2009. Imaging outcomes for neuroprotection and repair in multiple sclerosis trials. *Nat Rev Neurol,* 5.
BARKHOF, F., SCHELTENS, P., FREQUIN, S. T., NAUTA, J. J., TAS, M. W., VALK, J. & HOMMES, O. R. 1992. Relapsing-remitting multiple sclerosis: sequential enhanced MR imaging vs clinical findings in determining disease activity. *AJR Am J Roentgenol,* 159**,** 1041-7.
BARTZOKIS, G., CUMMINGS, J. L., SULTZER, D., HENDERSON, V. W., NUECHTERLEIN, K. H. & MINTZ, J. 2003. White matter structural integrity in healthy aging adults and patients with Alzheimer disease: a magnetic resonance imaging study. *Arch Neurol,* 60**,** 393-8.
BARTZOKIS, G., SULTZER, D., LU, P. H., NUECHTERLEIN, K. H., MINTZ, J. & CUMMINGS, J. L. 2004. Heterogeneous age-related breakdown of white matter structural integrity: implications for cortical "disconnection" in aging and Alzheimer's disease. *Neurobiol Aging,* 25**,** 843-51.
BATTAGLINI, M., JENKINSON, M., DE STEFANO, N. & INITIATIVE, F. T. A. S. D. N. 2018. SIENA-XL for improving the assessment of gray and white matter volume changes on brain MRI. *Human Brain Mapping,* 39**,** 1063-1077.
BEADNALL, H. N., WANG, C., VAN HECKE, W., RIBBENS, A., BILLIET, T. & BARNETT, M. H. 2019. Comparing longitudinal brain atrophy measurement techniques in a real-world multiple sclerosis clinical practice cohort: towards clinical integration? *Therapeutic Advances in Neurological Disorders,* 12**,** 1756286418823462.
BEESLEY, R., ANDERSON, V., HARDING, K. E., JOSEPH, F., TOMASSINI, V., PICKERSGILL, T. P., ROBERTSON, N. P. & TALLANTYRE, E. C. 2018. Impact of the 2017 revisions to McDonald criteria on the diagnosis of multiple sclerosis. *Mult Scler***,** 1352458518778007.
BERG, D., MÄURER, M., WARMUTH-METZ, M., RIECKMANN, P. & BECKER, G. 2000. The correlation between ventricular diameter measured by transcranial sonography and clinical disability and cognitive dysfunction in patients with multiple sclerosis. *Archives of Neurology,* 57**,** 1289-1292.
BERGAMASCHI, R., TONIETTI, S., FRANCIOTTA, D., CANDELORO, E., TAVAZZI, E., PICCOLO, G., ROMANI, A. & COSI, V. 2004. Oligoclonal bands in Devic's neuromyelitis optica and multiple sclerosis: differences in repeated cerebrospinal fluid examinations. *Mult Scler,* 10**,** 2-4.
BIAN, W., HARTER, K., HAMMOND-ROSENBLUTH, K. E., LUPO, J. M., XU, D., KELLEY, D. A., VIGNERON, D. B., NELSON, S. J. & PELLETIER, D.

2013. A serial in vivo 7T magnetic resonance phase imaging study of white matter lesions in multiple sclerosis. *Mult Scler,* 19**,** 69-75.
BIBERACHER, V., SCHMIDT, P., KESHAVAN, A., BOUCARD, C. C., RIGHART, R., SÄMANN, P., PREIBISCH, C., FRÖBEL, D., ALY, L., HEMMER, B., ZIMMER, C., HENRY, R. G. & MÜHLAU, M. 2016. Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. *NeuroImage,* 142**,** 188-197.
BILLER, A., CHOLI, M., BLAIMER, M., BREUER, F. A., JAKOB, P. M. & BARTSCH, A. J. 2014. Combined Acquisition Technique (CAT) for Neuroimaging of Multiple Sclerosis at Low Specific Absorption Rates (SAR). *PLOS ONE,* 9**,** e91030.
BITSCH, A., SCHUCHARDT, J., BUNKOWSKI, S., KUHLMANN, T. & BRÜCK, W. 2000. Acute axonal injury in multiple sclerosisCorrelation with demyelination and inflammation. *Brain,* 123**,** 1174-1183.
BJARTMAR, C., KINKEL, R. P., KIDD, G., RUDICK, R. A. & TRAPP, B. D. 2001. Axonal loss in normal-appearing white matter in a patient with acute MS. *Neurology,* 57**,** 1248-52.
BLINZINGER, K. & KREUTZBERG, G. 1968. Displacement of synaptic terminals from regenerating motoneurons by microglial cells. *Z Zellforsch Mikrosk Anat,* 85**,** 145-57.
BO, L., VEDELER, C. A., NYLAND, H. I., TRAPP, B. D. & MORK, S. J. 2003. Subpial demyelination in the cerebral cortex of multiple sclerosis patients. *J Neuropathol Exp Neurol,* 62**,** 723-32.
BREWER, J. B. 2009. Fully-automated volumetric MRI with normative ranges: translation to clinical practice. *Behav Neurol,* 21**,** 21-8.
BREX, P. A., JENKINS, R., FOX, N. C., CRUM, W. R., O'RIORDAN, J. I., PLANT, G. T. & MILLER, D. H. 2000. Detection of ventricular enlargement in patients at the earliest clinical stage of MS. *Neurology,* 54.
BRONNUM-HANSEN, H., KOCH-HENRIKSEN, N. & STENAGER, E. 2011. The Danish Multiple Sclerosis Registry. *Scand J Public Health,* 39**,** 62-4.
BROSNAN, C. F. & RAINE, C. S. 2013. The astrocyte in multiple sclerosis revisited. *Glia,* 61**,** 453-465.
BROWNLEE, W. J., ALTMANN, D. R., PRADOS, F., MISZKIEL, K. A., ESHAGHI, A., GANDINI WHEELER-KINGSHOTT, C. A. M., BARKHOF, F. & CICCARELLI, O. 2019. Early imaging predictors of long-term outcomes in relapse-onset multiple sclerosis. *Brain,* 142**,** 2276-2287.
BROWNLEE, W. J. & MILLER, D. H. 2014. Clinically isolated syndromes and the relationship to multiple sclerosis. *Journal of Clinical Neuroscience,* 21**,** 2065-2071.
BRUCK, W., PORADA, P., POSER, S., RIECKMANN, P., HANEFELD, F., KRETZSCHMAR, H. A. & LASSMANN, H. 1995. Monocyte/macrophage differentiation in early multiple sclerosis lesions. *Ann Neurol,* 38**,** 788-96.
BRUCK, W., SOMMERMEIER, N., BERGMANN, M., ZETTL, U., GOEBEL, H. H., KRETZSCHMAR, H. A. & LASSMANN, H. 1996. Macrophages in multiple sclerosis. *Immunobiology,* 195**,** 588-600.
BUTZKUEVEN, H., KAPPOS, L., PELLEGRINI, F., TROJANO, M., WIENDL, H., PATEL, R. N., ZHANG, A., HOTERMANS, C. & BELACHEW, S. 2014. Efficacy and safety of natalizumab in multiple sclerosis: interim

observational programme results. *J Neurol Neurosurg Psychiatry,* 85**,** 1190-7.
CALABRESE, M., AGOSTA, F., RINALDI, F., MATTISI, I., GROSSI, P., FAVARETTO, A., ATZORI, M., BERNARDI, V., BARACHINO, L. & RINALDI, L. 2009. Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis. *Arch Neurol,* 66.
CAMARA-REY, O., SNELLER, B. I., RIDGWAY, G. R., GARDE, E., FOX, N. C. & HILL, D. L. G. Simulation of Acquisition Artefacts in MR Scans: Effects on Automatic Measures of Brain Atrophy. 2006 Berlin, Heidelberg. Springer Berlin Heidelberg, 272-280.
CAMPION, T., SMITH, P., TURNER, B., ALTMANN, D., SATI, P., EVANSON, J., GEORGE, I., MIQUEL, M., REICH, D. & SCHMIERER, K. 2015. FLAIR* for the non-invasive histological diagnosis of multiple sclerosis (S29.003). *Neurology,* 84**,** S29.003.
CAMPION, T., SMITH, R. J., ALTMANN, D. R., BRITO, G. C., TURNER, B. P., EVANSON, J., GEORGE, I. C., SATI, P., REICH, D. S., MIQUEL, M. E. & SCHMIERER, K. 2017. FLAIR* to visualize veins in white matter lesions: A new tool for the diagnosis of multiple sclerosis? *Eur Radiol*.
CARMICHAEL, O. T., KULLER, L. H., LOPEZ, O. L., THOMPSON, P. M., DUTTON, R. A., LU, A., LEE, S. E., LEE, J. Y., AIZENSTEIN, H. J., MELTZER, C. C., LIU, Y., TOGA, A. W. & BECKER, J. T. 2007. Cerebral ventricular changes associated with transitions between normal cognitive function, mild cognitive impairment, and dementia. *Alzheimer Dis Assoc Disord,* 21**,** 14-24.
CERQUEIRA, J. J., COMPSTON, D. A. S., GERALDES, R., ROSA, M. M., SCHMIERER, K., THOMPSON, A., TINELLI, M. & PALACE, J. 2018. Time matters in multiple sclerosis: can early treatment and long-term follow-up ensure everyone benefits from the latest advances in multiple sclerosis? *Journal of Neurology, Neurosurgery & Psychiatry,* 89**,** 844-850.
CHARCOT, J. M. & MARIE, P. 1886. *Sur une forme particulière d'atrophie musculaire progressive: souvent familiale débutant par les pieds et les jambes et atteignant plus tard les mains*, Félix Alcan.
CHARCOT, M. 1868. Histologie de la sclerose en plaque. *Gaz. Hôsp.,* 41**,** 554-556.
CHAWLA, S., KISTER, I., WUERFEL, J., BRISSET, J. C., LIU, S., SINNECKER, T., DUSEK, P., HAACKE, E. M., PAUL, F. & GE, Y. 2016. Iron and Non-Iron-Related Characteristics of Multiple Sclerosis and Neuromyelitis Optica Lesions at 7T MRI. *AJNR Am J Neuroradiol,* 37**,** 1223-30.
CHERIYAN, J., KIM, S., WOLANSKY, L. J., COOK, S. D. & CADAVID, D. 2012. Impact of inflammation on brain volume in multiple sclerosis. *Archives of Neurology,* 69**,** 82-88.
CICCARELLI, O., WERRING, D. J., BARKER, G. J., GRIFFIN, C. M., WHEELER-KINGSHOTT, C. A., MILLER, D. H. & 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. *J Neurol,* 250**,** 287-92.
CIESLIK, P., HRYCEK, A. & KLUCINSKI, P. 2008. Vasculopathy and vasculitis in systemic lupus erythematosus. *Pol Arch Med Wewn,* 118**,** 57-63.

CLARKE, M., SAMARAWEERA, A., MOUGIN, O., SOUTH, M., DINEEN, R., MORGAN, P. & EVANGELOU, N. Utility of the central vein sign to diagnose multiple sclerosis; A comparison of 3T T2* and FLAIR-SWI. MULTIPLE SCLEROSIS JOURNAL, 2016. SAGE PUBLICATIONS LTD 1 OLIVERS YARD, 55 CITY ROAD, LONDON EC1Y 1SP, ENGLAND, 536-536.

CLOPPER, C. J. & PEARSON, E. S. 1934. THE USE OF CONFIDENCE OR FIDUCIAL LIMITS ILLUSTRATED IN THE CASE OF THE BINOMIAL. *Biometrika,* 26**,** 404-413.

COHEN, J. F., KOREVAAR, D. A., ALTMAN, D. G., BRUNS, D. E., GATSONIS, C. A., HOOFT, L., IRWIG, L., LEVINE, D., REITSMA, J. B., DE VET, H. C. W. & BOSSUYT, P. M. M. 2016. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open,* 6.

COMPSTON, A. & COLES, A. 2002. Multiple sclerosis. *Lancet,* 359**,** 1221-31.

CONFAVREUX, C., AIMARD, G. & DEVIC, M. 1980. Course and prognosis of multiple sclerosis assessed by the computerized data processing of 349 patients. *Brain,* 103**,** 281-300.

CORTESE, R., COLLORONE, S., CICCARELLI, O. & TOOSY, A. T. 2019. Advances in brain imaging in multiple sclerosis. *Therapeutic advances in neurological disorders,* 12**,** 1756286419859722-1756286419859722.

CORTESE, R., MAGNOLLAY, L., TUR, C., ABDEL-AZIZ, K., JACOB, A., DE ANGELIS, F., YIANNAKAS, M. C., PRADOS, F., OURSELIN, S., YOUSRY, T. A., BARKHOF, F. & CICCARELLI, O. 2018. Value of the central vein sign at 3T to differentiate MS from seropositive NMOSD. *Neurology*.

CRUVEILHIER, J. 1829-1842. Pathological anatomy of the human body. Billiére, Paris, France.

DAL-BIANCO, A., GRABNER, G., KRONNERWETTER, C., WEBER, M., HÖFTBERGER, R., BERGER, T., AUFF, E., LEUTMEZER, F., TRATTNIG, S., LASSMANN, H., BAGNATO, F. & HAMETNER, S. 2017. Slow expansion of multiple sclerosis iron rim lesions: pathology and 7 T magnetic resonance imaging. *Acta Neuropathologica,* 133**,** 25-42.

DALE, A. M., FISCHL, B. & SERENO, M. I. 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage,* 9**,** 179-194.

DALTON, C. M., CHARD, D. T., DAVIES, G. R., MISZKIEL, K. A., ALTMANN, D. R., FERNANDO, K., PLANT, G. T., THOMPSON, A. J. & MILLER, D. H. 2004. Early development of multiple sclerosis is associated with progressive grey matter atrophy in patients presenting with clinically isolated syndromes. *Brain,* 127.

DAWSON, J. 1916. The histology of multiple sclerosis. *Trans R Soc Edinburgh,* 50**,** 517-740.

DE GROOT, C., BERGERS, E., KAMPHORST, W., RAVID, R., POLMAN, C., BARKHOF, F. & VAN DER VALK, P. 2001. Post-mortem MRI-guided sampling of multiple sclerosis brain lesions: increased yield of active demyelinating and (p) reactive lesions. *Brain,* 124**,** 1635-1645.

DE STEFANO, N., AIRAS, L., GRIGORIADIS, N., MATTLE, H. P., O'RIORDAN, J., OREJA-GUEVARA, C., SELLEBJERG, F., STANKOFF, B., WALCZAK, A., WIENDL, H. & KIESEIER, B. C. 2014. Clinical Relevance

of Brain Volume Measures in Multiple Sclerosis. *CNS Drugs,* 28**,** 147-156.
DE STEFANO, N., GIORGIO, A., BATTAGLINI, M., ROVARIS, M., SORMANI, M. P., BARKHOF, F., KORTEWEG, T., ENZINGER, C., FAZEKAS, F., CALABRESE, M., DINACCI, D., TEDESCHI, G., GASS, A., MONTALBAN, X., ROVIRA, A., THOMPSON, A., COMI, G., MILLER, D. H. & FILIPPI, M. 2010. Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. *Neurology,* 74**,** 1868-76.
DE STEFANO, N., NARAYANAN, S., FRANCIS, S. J., SMITH, S., MORTILLA, M., TARTAGLIA, M. C., BARTOLOZZI, M. L., GUIDI, L., FEDERICO, A. & ARNOLD, D. L. 2002. Diffuse axonal and tissue injury in patients with multiple sclerosis with low cerebral lesion load and no disability. *Arch Neurol,* 59**,** 1565-71.
DE STEFANO, N., STROMILLO, M. L., GIORGIO, A., BARTOLOZZI, M. L., BATTAGLINI, M., BALDINI, M., PORTACCIO, E., AMATO, M. P. & SORMANI, M. P. 2015. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*.
DE STEFANO, N., STROMILLO, M. L., GIORGIO, A., BARTOLOZZI, M. L., BATTAGLINI, M., BALDINI, M., PORTACCIO, E., AMATO, M. P. & SORMANI, M. P. 2016. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. *J Neurol Neurosurg Psychiatry,* 87**,** 93-9.
DE STEFANO, N., STROMILLO, M. L., ROSSI, F., BATTAGLINI, M., GIORGIO, A., PORTACCIO, E., HAKIKI, B., MALENTACCHI, G., GASPERINI, C., SANTANGELO, M., BARTOLOZZI, M. L., SORMANI, M. P., FEDERICO, A. & AMATO, M. P. 2011. Improving the characterization of radiologically isolated syndrome suggestive of multiple sclerosis. *PLoS One,* 6**,** e19452.
DENNEY, D. R., LYNCH, S. G. & PARMENTER, B. A. 2008. A 3-year longitudinal study of cognitive impairment in patients with primary progressive multiple sclerosis: Speed matters. *Journal of the Neurological Sciences,* 267**,** 129-136.
DI FILIPPO, M., ANDERSON, V. M., ALTMANN, D. R., SWANTON, J. K., PLANT, G. T., THOMPSON, A. J. & MILLER, D. H. 2010. Brain atrophy and lesion load measures over 1 year relate to clinical status after 6 years in patients with clinically isolated syndromes. *J Neurol Neurosurg Psychiatry,* 81**,** 204-8.
DOW, R. S. & BERGLUND, G. 1942. Vascular pattern of lesions of multiple sclerosis. *Archives of Neurology & Psychiatry,* 47**,** 1-18.
DUNING, T., KLOSKA, S., STEINSTRÄTER, O., KUGEL, H., HEINDEL, W. & KNECHT, S. 2005. Dehydration confounds the assessment of brain atrophy. *Neurology,* 64**,** 548-550.
DURAND-DUBIEF, F., BELAROUSSI, B., ARMSPACH, J. P., DUFOUR, M., ROGGERONE, S., VUKUSIC, S., HANNOUN, S., SAPPEY-MARINIER, D., CONFAVREUX, C. & COTTON, F. 2012. Reliability of Longitudinal Brain Volume Loss Measurements between 2 Sites in Patients with Multiple Sclerosis: Comparison of 7 Quantification Techniques. *American Journal of Neuroradiology*.

DUTTA, R. & TRAPP, B. D. 2014. Relapsing and progressive forms of multiple sclerosis: insights from pathology. *Curr Opin Neurol,* 27**,** 271-8.
DWORKIN, J. D., SATI, P., SOLOMON, A., PHAM, D. L., WATTS, R., MARTIN, M. L., ONTANEDA, D., SCHINDLER, M. K., REICH, D. S. & SHINOHARA, R. T. 2018. Automated Integration of Multimodal MRI for the Probabilistic Detection of the Central Vein Sign in White Matter Lesions. *American Journal of Neuroradiology*.
DWYER, M., RAMASAMY, D., DURFEE, J. & ZIVADINOV, R. Accurate and reliable atrophy measurement on clinical quality FLAIR scans. MULTIPLE SCLEROSIS JOURNAL, 2015a. SAGE PUBLICATIONS LTD 1 OLIVERS YARD, 55 CITY ROAD, LONDON EC1Y 1SP, ENGLAND, 185-186.
DWYER, M. G., ZIVADINOV, R., TAO, Y., ZHANG, X., KENNEDY, C., BERGSLAND, N., RAMASAMY, D. P., DURFEE, J., HOJNACKI, D., WEINSTOCK-GUTTMAN, B., HAYWARD, B., DANGOND, F. & MARKOVIC-PLESE, S. 2015b. Immunological and short-term brain volume changes in relapsing forms of multiple sclerosis treated with interferon beta-1a subcutaneously three times weekly: an open-label two-arm trial. *BMC Neurol,* 15**,** 232.
DZIEDZIC, T., METZ, I., DALLENGA, T., KÖNIG, F. B., MÜLLER, S., STADELMANN, C. & BRÜCK, W. 2010. Wallerian Degeneration: A Major Component of Early Axonal Pathology in Multiple Sclerosis. *Brain Pathology,* 20**,** 976-985.
EBNER, M., CHUNG, K. K., PRADOS, F., CARDOSO, M. J., CHARD, D. T., VERCAUTEREN, T. & OURSELIN, S. 2018. Volumetric reconstruction from printed films: Enabling 30 year longitudinal analysis in MR neuroimaging. *NeuroImage,* 165**,** 238-250.
ENGELBORGHS, S., NIEMANTSVERDRIET, E., STRUYFS, H., BLENNOW, K., BROUNS, R., COMABELLA, M., DUJMOVIC, I., VAN DER FLIER, W., FRÖLICH, L., GALIMBERTI, D., GNANAPAVAN, S., HEMMER, B., HOFF, E., HORT, J., IACOBAEUS, E., INGELSSON, M., JAN DE JONG, F., JONSSON, M., KHALIL, M., KUHLE, J., LLEÓ, A., DE MENDONÇA, A., MOLINUEVO, J. L., NAGELS, G., PAQUET, C., PARNETTI, L., ROKS, G., ROSA-NETO, P., SCHELTENS, P., SKÅRSGARD, C., STOMRUD, E., TUMANI, H., VISSER, P. J., WALLIN, A., WINBLAD, B., ZETTERBERG, H., DUITS, F. & TEUNISSEN, C. E. 2017. Consensus guidelines for lumbar puncture in patients with neurological diseases. *Alzheimer's & dementia (Amsterdam, Netherlands),* 8**,** 111-126.
ERIKSSON, M., ANDERSEN, O. & RUNMARKER, B. 2003. Long-term follow up of patients with clinically isolated syndromes, relapsing-remitting and secondary progressive multiple sclerosis. *Multiple Sclerosis Journal,* 9**,** 260-274.
ERNERUDH, J., OLSSON, T., LINDSTRÖM, F. & SKOGH, T. 1985. Cerebrospinal fluid immunoglobulin abnormalities in systemic lupus erythematosus. *Journal of Neurology, Neurosurgery, and Psychiatry,* 48**,** 807-813.
EVANGELOU, N., ESIRI, M. M., SMITH, S., PALACE, J. & MATTHEWS, P. M. 2000a. Quantitative pathological evidence for axonal loss in normal appearing white matter in multiple sclerosis. *Ann Neurol,* 47**,** 391-5.

EVANGELOU, N., KONZ, D., ESIRI, M. M., SMITH, S., PALACE, J. & MATTHEWS, P. M. 2000b. Regional axonal loss in the corpus callosum correlates with cerebral white matter lesion volume and distribution in multiple sclerosis. *Brain,* 123 (Pt 9)**,** 1845-9.

EVANGELOU, N., KONZ, D., ESIRI, M. M., SMITH, S., PALACE, J. & MATTHEWS, P. M. 2001. Size-selective neuronal changes in the anterior optic pathways suggest a differential susceptibility to injury in multiple sclerosis. *Brain,* 124**,** 1813-20.

FEDOROV, A., BEICHEL, R., KALPATHY-CRAMER, J., FINET, J., FILLION-ROBIN, J. C., PUJOL, S., BAUER, C., JENNINGS, D., FENNESSY, F., SONKA, M., BUATTI, J., AYLWARD, S., MILLER, J. V., PIEPER, S. & KIKINIS, R. 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging,* 30**,** 1323-41.

FERGUSON, B., MATYSZAK, M. K., ESIRI, M. M. & PERRY, V. H. 1997. Axonal damage in acute multiple sclerosis lesions. *Brain: a journal of neurology,* 120**,** 393-399.

FILIPPI, M., PREZIOSA, P., MEANI, A., CICCARELLI, O., MESAROS, S., ROVIRA, A., FREDERIKSEN, J., ENZINGER, C., BARKHOF, F., GASPERINI, C., BROWNLEE, W., DRULOVIC, J., MONTALBAN, X., CRAMER, S. P., PICHLER, A., HAGENS, M., RUGGIERI, S., MARTINELLI, V., MISZKIEL, K., TINTORE, M., COMI, G., DEKKER, I., UITDEHAAG, B., DUJMOVIC-BASUROSKI, I. & ROCCA, M. A. 2018. Prediction of a multiple sclerosis diagnosis in patients with clinically isolated syndrome using the 2016 MAGNIMS and 2010 McDonald criteria: a retrospective study. *Lancet Neurol,* 17**,** 133-142.

FILIPPI, M., ROCCA, M. A., CICCARELLI, O., DE STEFANO, N., EVANGELOU, N., KAPPOS, L., ROVIRA, A., SASTRE-GARRIGA, J., TINTORE, M., FREDERIKSEN, J. L., GASPERINI, C., PALACE, J., REICH, D. S., BANWELL, B., MONTALBAN, X. & BARKHOF, F. 2016. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *Lancet Neurol,* 15**,** 292-303.

FILIPPI, M., ROCCA, M. A., GASPERINI, C., SORMANI, M. P., BASTIANELLO, S., HORSFIELD, M. A., POZZILLI, C. & COMI, G. 1999. Interscanner variation in brain MR lesion load measurements in multiple sclerosis using conventional spin-echo, rapid relaxation-enhanced, and fast-FLAIR sequences. *AJNR Am J Neuroradiol,* 20**,** 133-7.

FILIPPI, M., ROCCA, M. A., MARTINO, G., HORSFIELD, M. A. & COMI, G. 1998. Magnetization transfer changes in the normal appearing white matter precede the appearance of enhancing lesions in patients with multiple sclerosis. *Ann Neurol,* 43**,** 809-14.

FISCHL, B., SALAT, D. H., BUSA, E., ALBERT, M., DIETERICH, M., HASELGROVE, C., VAN DER KOUWE, A., KILLIANY, R., KENNEDY, D., KLAVENESS, S., MONTILLO, A., MAKRIS, N., ROSEN, B. & DALE, A. M. 2002. Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron,* 33**,** 341-355.

FISCHL, B., SALAT, D. H., VAN DER KOUWE, A. J., MAKRIS, N., SEGONNE, F., QUINN, B. T. & DALE, A. M. 2004. Sequence-independent segmentation of magnetic resonance images. *Neuroimage,* 23 Suppl 1**,** S69-84.

FISHER, E., RUDICK, R. A., SIMON, J. H., CUTTER, G., BAIER, M., LEE, J. C., MILLER, D., WEINSTOCK-GUTTMAN, B., MASS, M. K., DOUGHERTY, D. S. & SIMONIAN, N. A. 2002. Eight-year follow-up study of brain atrophy in patients with MS. *Neurology,* 59**,** 1412-20.

FISNIKU, L. K., BREX, P. A., ALTMANN, D. R., MISZKIEL, K. A., BENTON, C. E., LANYON, R., THOMPSON, A. J. & MILLER, D. H. 2008. Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain,* 131**,** 808-17.

FLACHENECKER, P., BUCKOW, K., PUGLIATTI, M., KES, V. B., BATTAGLIA, M. A., BOYKO, A., CONFAVREUX, C., ELLENBERGER, D., ESKIC, D., FORD, D., FRIEDE, T., FUGE, J., GLASER, A., HILLERT, J., HOLLOWAY, E., IOANNIDOU, E., KAPPOS, L., KASILINGAM, E., KOCH-HENRIKSEN, N., KUHLE, J., LEPORE, V., MIDDLETON, R., MYHR, K.-M., OROLOGAS, A., OTERO, S., PITSCHNAU-MICHEL, D., RIENHOFF, O., SASTRE-GARRIGA, J., SCHYNS-LIHARSKA, T., SUTOVIC, D., THALHEIM, C., TROJANO, M., VLASOV, Y. V. & YALDIZLI, Ö. 2014. Multiple sclerosis registries in Europe – results of a systematic survey. *Multiple Sclerosis Journal,* 20**,** 1523-1532.

FORD, D. V., JONES, K. H., MIDDLETON, R. M., LOCKHART-JONES, H., MARAMBA, I. D., NOBLE, G. J., OSBORNE, L. A. & LYONS, R. A. 2012. The feasibility of collecting information from people with Multiple Sclerosis for the UK MS Register via a web portal: characterising a cohort of people with MS. *BMC Med Inform Decis Mak,* 12**,** 73.

FOX, N. C., JENKINS, R., LEARY, S. M., STEVENSON, V. L., LOSSEFF, N. A., CRUM, W. R., HARVEY, R. J., ROSSOR, M. N., MILLER, D. H. & THOMPSON, A. J. 2000. Progressive cerebral atrophy in MS: a serial study using registered, volumetric MRI. *Neurology,* 54.

FOX, R. J., FISHER, E., TKACH, J., LEE, J. C., COHEN, J. A. & RUDICK, R. A. 2005. Brain atrophy and magnetization transfer ratio following methylprednisolone in multiple sclerosis: short-term changes and long-term implications. *Mult Scler,* 11**,** 140-5.

FRANK, J. A., RICHERT, N., BASH, C., STONE, L., CALABRESI, P. A., LEWIS, B., STONE, R., HOWARD, T. & MCFARLAND, H. F. 2004. Interferon-β-1b slows progression of atrophy in RRMS. *Three-year follow-up in NAb− and NAb+ patients,* 62**,** 719-725.

FREEDMAN, M. S., THOMPSON, E. J., DEISENHAMMER, F., GIOVANNONI, G., GRIMSLEY, G., KEIR, G., OHMAN, S., RACKE, M. K., SHARIEF, M., SINDIC, C. J., SELLEBJERG, F. & TOURTELLOTTE, W. W. 2005. Recommended standard of cerebrospinal fluid analysis in the diagnosis of multiple sclerosis: a consensus statement. *Arch Neurol,* 62**,** 865-70.

FRIEDEN, T. R. 2017. Evidence for Health Decision Making — Beyond Randomized, Controlled Trials. *New England Journal of Medicine,* 377**,** 465-475.

FRIEDMAN, J., HASTIE, T. & TIBSHIRANI, R. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw,* 33**,** 1-22.

FRISCHER, J. M., BRAMOW, S., DAL-BIANCO, A., LUCCHINETTI, C. F., RAUSCHKA, H., SCHMIDBAUER, M., LAURSEN, H., SORENSEN, P. S. & LASSMANN, H. 2009. The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain,* 132.

FRISONI, G. B., GANZOLA, R., CANU, E., RUB, U., PIZZINI, F. B., ALESSANDRINI, F., ZOCCATELLI, G., BELTRAMELLO, A., CALTAGIRONE, C. & THOMPSON, P. M. 2008. Mapping local hippocampal changes in Alzheimer's disease and normal ageing with MRI at 3 Tesla. *Brain,* 131**,** 3266-76.

FROHMAN, E. M., RACKE, M. K. & RAINE, C. S. 2006. Multiple sclerosis--the plaque and its pathogenesis. *N Engl J Med,* 354**,** 942-55.

FU, L., MATTHEWS, P., DE STEFANO, N., WORSLEY, K., NARAYANAN, S., FRANCIS, G., ANTEL, J., WOLFSON, C. & ARNOLD, D. 1998. Imaging axonal damage of normal-appearing white matter in multiple sclerosis. *Brain: a journal of neurology,* 121**,** 103-113.

GANTER, P., PRINCE, C. & ESIRI, M. M. 1999. Spinal cord axonal loss in multiple sclerosis: a post-mortem study. *Neuropathology and applied neurobiology,* 25**,** 459-467.

GE, Y. 2006. Multiple Sclerosis: The Role of MR Imaging. *American Journal of Neuroradiology,* 27**,** 1165-1176.

GE, Y., GROSSMAN, R. I., BABB, J. S., RABIN, M. L., MANNON, L. J. & KOLSON, D. L. 2002. Age-related total gray matter and white matter changes in normal adult brain. Part I: volumetric MR imaging analysis. *AJNR Am J Neuroradiol,* 23**,** 1327-33.

GE, Y., GROSSMAN, R. I., UDUPA, J. K., BABB, J. S., NYUL, L. G. & KOLSON, D. L. 2001. Brain atrophy in relapsing-remitting multiple sclerosis: fractional volumetric analysis of gray matter and white matter. *Radiology,* 220.

GERALDES, R., CICCARELLI, O., BARKHOF, F., DE STEFANO, N., ENZINGER, C., FILIPPI, M., HOFER, M., PAUL, F., PREZIOSA, P., ROVIRA, A., DELUCA, G. C., KAPPOS, L., YOUSRY, T., FAZEKAS, F., FREDERIKSEN, J., GASPERINI, C., SASTRE-GARRIGA, J., EVANGELOU, N. & PALACE, J. 2018. The current role of MRI in differentiating multiple sclerosis from its imaging mimics. *Nat Rev Neurol*.

GHIONE, E., BERGSLAND, N., DWYER, M. G., HAGEMEIER, J., JAKIMOVSKI, D., PAUNKOSKI, I., RAMASAMY, D. P., SILVA, D., CARL, E., HOJNACKI, D., KOLB, C., WEINSTOCK-GUTTMAN, B. & ZIVADINOV, R. 2018. Brain Atrophy Is Associated with Disability Progression in Patients with MS followed in a Clinical Routine. *American Journal of Neuroradiology,* 39**,** 2237-2242.

GIORGIO, A., PORTACCIO, E., STROMILLO, M. L., MARINO, S., ZIPOLI, V., BATTAGLINI, M., BLANDINO, A., BARTOLOZZI, M. L., SIRACUSA, G. & AMATO, M. P. 2010. Cortical functional reorganization and its relationship with brain structural damage in patients with benign multiple sclerosis. *Mult Scler,* 16.

GIOVANNONI, G., BERMEL, R., PHILLIPS, T. & RUDICK, R. 2018. A brief history of NEDA. *Mult Scler Relat Disord,* 20**,** 228-230.

GOLDENBERG, M. M. 2012. Multiple sclerosis review. *P & T : a peer-reviewed journal for formulary management,* 37**,** 175-184.

GOODIN, D. S., TRABOULSEE, A., KNAPPERTZ, V., REDER, A. T., LI, D., LANGDON, D., WOLF, C., BECKMANN, K., KONIECZNY, A. & EBERS, G. C. 2011. Relationship between early clinical characteristics and long term disability outcomes: 16 year cohort study (follow-up) of the pivotal

interferon β-1b trial in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*.
GRAMMATICO, M., CARLUCCI, G., DALLAGIACOMA, S., VUOLO, L., REPICE, A. M., MECHI, C., MAGNANI, G., BARILARO, A., FAINARDI, E. & MASSACESI, L. 2017. Accuracy of the central vein sign evaluation by brain MRi for the pathologic characterization of multiple sclerosis cases with markers of “better explanation” of the diagnosis. *Multiple Sclerosis Journal,* 23**,** 432.
GUO, B. J., YANG, Z. L. & ZHANG, L. J. 2018. Gadolinium Deposition in Brain: Current Scientific Evidence and Future Perspectives. *Frontiers in molecular neuroscience,* 11**,** 335-335.
GUTTMANN, C. R., ROUSSET, M., ROCH, J. A., HANNOUN, S., DURAND-DUBIEF, F., BELAROUSSI, B., CAVALLARI, M., RABILLOUD, M., SAPPEY-MARINIER, D., VUKUSIC, S. & COTTON, F. 2016. Multiple sclerosis lesion formation and early evolution revisited: A weekly high-resolution magnetic resonance imaging study. *Mult Scler,* 22**,** 761-9.
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. & GROSSMAN, R. I. 2009. Characterizing iron deposition in multiple sclerosis lesions using susceptibility weighted imaging. *Journal of magnetic resonance imaging : JMRI,* 29**,** 537-544.
HAACKE, E. M., XU, Y., CHENG, Y. C. & REICHENBACH, J. R. 2004. Susceptibility weighted imaging (SWI). *Magn Reson Med,* 52**,** 612-8.
HAGEMEIER, J., HEININEN-BROWN, M., POLONI, G. U., BERGSLAND, N., MAGNANO, C. R., DURFEE, J., KENNEDY, C., CARL, E., WEINSTOCK-GUTTMAN, B., DWYER, M. G. & ZIVADINOV, R. 2012. Iron deposition in multiple sclerosis lesions measured by susceptibility-weighted imaging filtered phase: a case control study. *J Magn Reson Imaging,* 36**,** 73-83.
HAMMOND, K. E., METCALF, M., CARVAJAL, L., OKUDA, D. T., SRINIVASAN, R., VIGNERON, D., NELSON, S. J. & PELLETIER, D. 2008. Quantitative in vivo magnetic resonance imaging of multiple sclerosis at 7 Tesla with sensitivity to iron. *Ann Neurol,* 64**,** 707-13.
HANSEN, K., CRUZ, M. & LINK, H. 1990. Oligoclonal Borrelia burgdorferi-specific IgG antibodies in cerebrospinal fluid in Lyme neuroborreliosis. *J Infect Dis,* 161**,** 1194-202.
HARDMEIER, M., WAGENPFEIL, S., FREITAG, P., FISHER, E., RUDICK, R. A., KOOIJMANS, M., CLANET, M., RADUE, E. W. & KAPPOS, L. 2005. Rate of brain atrophy in relapsing MS decreases during treatment with IFNβ-1a. *Neurology,* 64**,** 236-240.
HARRISON, D. M., LI, X., LIU, H., JONES, C. K., CAFFO, B., CALABRESI, P. A. & VAN ZIJL, P. 2016. Lesion Heterogeneity on High-Field Susceptibility MRI Is Associated with Multiple Sclerosis Severity. *AJNR Am J Neuroradiol,* 37**,** 1447-53.
HAVRDOVA, E., GALETTA, S., HUTCHINSON, M., STEFOSKI, D., BATES, D., POLMAN, C. H., O'CONNOR, P. W., GIOVANNONI, G., PHILLIPS, J. T., LUBLIN, F. D., PACE, A., KIM, R. & HYDE, R. 2009. Effect of natalizumab on clinical and radiological disease activity in multiple sclerosis: a retrospective analysis of the Natalizumab Safety and Efficacy

in Relapsing-Remitting Multiple Sclerosis (AFFIRM) study. *The Lancet Neurology,* 8**,** 254-260.
HEDMAN, A. M., VAN HAREN, N. E., SCHNACK, H. G., KAHN, R. S. & HULSHOFF POL, H. E. 2012. Human brain changes across the life span: a review of 56 longitudinal magnetic resonance imaging studies. *Hum Brain Mapp,* 33**,** 1987-2002.
HEINEN, R., BOUVY, W. H., MENDRIK, A. M., VIERGEVER, M. A., BIESSELS, G. J. & DE BRESSER, J. 2016. Robustness of Automated Methods for Brain Volume Measurements across Different MRI Field Strengths. *PLoS One,* 11**,** e0165719.
HILLERT, J. & STAWIARZ, L. 2015. The Swedish MS registry – clinical support tool and scientific resource. *Acta Neurologica Scandinavica,* 132**,** 11-19.
HOSSEINI, Z., MATUSINEC, J., RUDKO, D. A., LIU, J., KWAN, B. Y. M., SALEHI, F., SHARMA, M., KREMENCHUTZKY, M., MENON, R. S. & DRANGOVA, M. 2018. Morphology-Specific Discrimination between MS White Matter Lesions and Benign White Matter Hyperintensities Using Ultra-High-Field MRI. *American Journal of Neuroradiology*.
HUH, S. Y., MIN, J. H., KIM, W., KIM, S. H., KIM, H. J., KIM, B. J., KIM, B. J. & LEE, K. H. 2014. The usefulness of brain MRI at onset in the differentiation of multiple sclerosis and seropositive neuromyelitis optica spectrum disorders. *Mult Scler,* 20**,** 695-704.
HUPPERTZ, H. J., KROLL-SEGER, J., KLOPPEL, S., GANZ, R. E. & KASSUBEK, J. 2010. Intra- and interscanner variability of automated voxel-based volumetry based on a 3D probabilistic atlas of human cerebral structures. *Neuroimage,* 49**,** 2216-24.
IGLESIAS, J. E., LIU, C. Y., THOMPSON, P. M. & TU, Z. 2011. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Trans Med Imaging,* 30**,** 1617-34.
JACOBS, L. D., BECK, R. W., SIMON, J. H., KINKEL, R. P., BROWNSCHEIDLE, C. M., MURRAY, T. J., SIMONIAN, N. A., SLASOR, P. J. & SANDROCK, A. W. 2000. Intramuscular interferon beta-1a therapy initiated during a first demyelinating event in multiple sclerosis. CHAMPS Study Group. *N Engl J Med,* 343**,** 898-904.
JAIN, S., SIMA, D. M., RIBBENS, A., CAMBRON, M., MAERTENS, A., VAN HECKE, W., DE MEY, J., BARKHOF, F., STEENWIJK, M. D., DAAMS, M., MAES, F., VAN HUFFEL, S., VRENKEN, H. & SMEETS, D. 2015. Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. *Neuroimage Clin,* 8**,** 367-75.
JENKINSON, M., BANNISTER, P., BRADY, M. & SMITH, S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage,* 17**,** 825-41.
JONES, B. C., NAIR, G., SHEA, C. D., CRAINICEANU, C. M., CORTESE, I. C. M. & REICH, D. S. 2013. Quantification of multiple-sclerosis-related brain atrophy in two heterogeneous MRI datasets using mixed-effects modeling. *NeuroImage: Clinical,* 3**,** 171-179.
JULIAN, L. J., VELLA, L., VOLLMER, T., HADJIMICHAEL, O. & MOHR, D. C. 2008. Employment in multiple sclerosis: Exiting and re-entering the work force. *Journal of neurology,* 255**,** 1354-1360.

KAISEY, M., SOLOMON, A. J., LUU, M., GIESSER, B. S. & SICOTTE, N. L. 2019. Incidence of multiple sclerosis misdiagnosis in referrals to two academic centers. *Mult Scler Relat Disord,* 30**,** 51-56.

KALKERS, N. F., AMEZIANE, N., BOT, J. C., MINNEBOO, A., POLMAN, C. H. & BARKHOF, F. 2002. Longitudinal brain volume measurement in multiple sclerosis: rate of brain atrophy is independent of the disease subtype. *Arch Neurol,* 59**,** 1572-6.

KAPPOS, L., DE STEFANO, N., FREEDMAN, M. S., CREE, B. A., RADUE, E. W., SPRENGER, T., SORMANI, M. P., SMITH, T., HARING, D. A., PIANI MEIER, D. & TOMIC, D. 2015. Inclusion of brain volume loss in a revised measure of 'no evidence of disease activity' (NEDA-4) in relapsing-remitting multiple sclerosis. *Mult Scler*.

KAPPOS, L., MOERI, D., RADUE, E. W., SCHOETZAU, A., SCHWEIKERT, K., BARKHOF, F., MILLER, D., GUTTMANN, C. R., WEINER, H. L., GASPERINI, C. & FILIPPI, M. 1999. Predictive value of gadolinium-enhanced magnetic resonance imaging for relapse rate and changes in disability or impairment in multiple sclerosis: a meta-analysis. Gadolinium MRI Meta-analysis Group. *Lancet,* 353**,** 964-9.

KATZ, D., TAUBENBERGER, J. K., CANNELLA, B., MCFARLIN, D. E., RAINE, C. S. & MCFARLAND, H. F. 1993. Correlation between magnetic resonance imaging findings and lesion development in chronic, active multiple sclerosis. *Ann Neurol,* 34**,** 661-9.

KAU, T., TASCHWER, M., DEUTSCHMANN, H., SCHONFELDER, M., WEBER, J. R. & HAUSEGGER, K. A. 2013. The "central vein sign": is there a place for susceptibility weighted imaging in possible multiple sclerosis? *Eur Radiol,* 23**,** 1956-62.

KEMPER, T. L. 1994. Neuroanatomical and neuropathological changes during aging and dementia.

KEMPTON, M. J., ETTINGER, U., SCHMECHTIG, A., WINTER, E. M., SMITH, L., MCMORRIS, T., WILKINSON, I. D., WILLIAMS, S. C. R. & SMITH, M. S. 2007. Effects of acute dehydration on brain morphology in healthy humans. *Human Brain Mapping,* 30**,** 291-298.

KILSDONK, I. D., JONKMAN, L. E., KLAVER, R., VAN VELUW, S. J., ZWANENBURG, J. J. M., KUIJER, J. P. A., POUWELS, P. J. W., TWISK, J. W. R., WATTJES, M. P., LUIJTEN, P. R., BARKHOF, F. & GEURTS, J. J. G. 2016. Increased cortical grey matter lesion detection in multiple sclerosis with 7 T MRI: a post-mortem verification study. *Brain,* 139**,** 1472-1481.

KILSDONK, I. D., LOPEZ-SORIANO, A., KUIJER, J. P., DE GRAAF, W. L., CASTELIJNS, J. A., POLMAN, C. H., LUIJTEN, P. R., GEURTS, J. J., BARKHOF, F. & WATTJES, M. P. 2014a. Morphological features of MS lesions on FLAIR* at 7 T and their relation to patient characteristics. *J Neurol,* 261**,** 1356-64.

KILSDONK, I. D., WATTJES, M. P., LOPEZ-SORIANO, A., KUIJER, J. P., DE JONG, M. C., DE GRAAF, W. L., CONIJN, M. M., POLMAN, C. H., LUIJTEN, P. R., GEURTS, J. J., GEERLINGS, M. I. & BARKHOF, F. 2014b. Improved differentiation between MS and vascular brain lesions using FLAIR* at 7 Tesla. *Eur Radiol,* 24**,** 841-9.

KIM, S. S., RICHMAN, D. P., JOHNSON, W. O., HALD, J. K. & AGIUS, M. A. 2014. Limited utility of current MRI criteria for distinguishing multiple

sclerosis from common mimickers: primary and secondary CNS vasculitis, lupus and Sjogren's syndrome. *Mult Scler,* 20**,** 57-63.
KINGWELL, E., MARRIOTT, J. J., JETTÉ, N., PRINGSHEIM, T., MAKHANI, N., MORROW, S. A., FISK, J. D., EVANS, C., BÉLAND, S. G., KULAGA, S., DYKEMAN, J., WOLFSON, C., KOCH, M. W. & MARRIE, R. A. 2013. Incidence and prevalence of multiple sclerosis in Europe: a systematic review. *BMC Neurology,* 13**,** 1-13.
KISTER, I., HERBERT, J., ZHOU, Y. & GE, Y. 2013. Ultrahigh-Field MR (7 T) Imaging of Brain Lesions in Neuromyelitis Optica. *Multiple Sclerosis International,* 2013**,** 1-7.
KOCH-HENRIKSEN, N., MAGYARI, M. & LAURSEN, B. 2015. Registers of multiple sclerosis in Denmark. *Acta Neurologica Scandinavica,* 132**,** 4-10.
KOLLIA, K., MADERWALD, S., PUTZKI, N., SCHLAMANN, M., THEYSOHN, J. M., KRAFF, O., LADD, M. E., FORSTING, M. & WANKE, I. 2009. First Clinical Study on Ultra-High-Field MR Imaging in Patients with Multiple Sclerosis: Comparison of 1.5T and 7T. *American Journal of Neuroradiology,* 30**,** 699-702.
KUCHLING, J., RAMIEN, C., BOZIN, I., DÖRR, J., HARMS, L., ROSCHE, B., NIENDORF, T., PAUL, F., SINNECKER, T. & WUERFEL, J. 2014. Identical lesion morphology in primary progressive and relapsing–remitting MS–an ultrahigh field MRI study. *Multiple Sclerosis Journal,* 20**,** 1866-1871.
KUTZELNIGG, A., LUCCHINETTI, C. F., STADELMANN, C., BRUCK, W., RAUSCHKA, H., BERGMANN, M., SCHMIDBAUER, M., PARISI, J. E. & LASSMANN, H. 2005. Cortical demyelination and diffuse white matter injury in multiple sclerosis. *Brain,* 128**,** 2705-12.
LAI, M., HODGSON, T., GAWNE-CAIN, M., WEBB, S., MACMANUS, D., MCDONALD, W. I., THOMPSON, A. J. & MILLER, D. H. 1996. A preliminary study into the sensitivity of disease activity detection by serial weekly magnetic resonance imaging in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry,* 60**,** 339-341.
LANE, J. I., BOLSTER, B., CAMPEAU, N. G., WELKER, K. M. & GILBERTSON, J. R. 2015. Characterization of multiple sclerosis plaques using susceptibility-weighted imaging at 1.5 T: can perivenular localization improve specificity of imaging criteria? *J Comput Assist Tomogr,* 39**,** 317-20.
LASSMANN, H. 1998. Neuropathology in multiple sclerosis: new concepts. *Multiple Sclerosis Journal,* 4**,** 93-98.
LASSMANN, H. From neuropathology to new patho-physiological concepts and clinical perspectives. MULTIPLE SCLEROSIS JOURNAL, 2017. SAGE PUBLICATIONS LTD 1 OLIVERS YARD, 55 CITY ROAD, LONDON EC1Y 1SP, ENGLAND, 19-19.
LECHNER-SCOTT, J., SPENCER, B., DE MALMANCHE, T., ATTIA, J., FITZGERALD, M., TROJANO, M., GRAND'MAISON, F., GOMEZ, J. A., IZQUIERDO, G., DUQUETTE, P., GIRARD, M., GRAMMOND, P., OREJA-GUEVARA, C., HUPPERTS, R., BERGAMASCHI, R., BOZ, C., GIULIANI, G., VAN PESCH, V., IULIANO, G., FIOL, M., CRISTIANO, E., VERHEUL, F., SALADINO, M. L., SLEE, M., BARNETT, M., DERI, N., FLECHTER, S., VELLA, N., SHAW, C., HERBERT, J., MOORE, F.,

PETKOVSKA-BOSKOVA, T., JOKUBAITIS, V. & BUTZKUEVEN, H. 2012. The frequency of CSF oligoclonal banding in multiple sclerosis increases with latitude. *Mult Scler,* 18**,** 974-82.
LOVAS, G., SZILÁGYI, N., MAJTÉNYI, K., PALKOVITS, M. & KOMOLY, S. 2000. Axonal changes in chronic demyelinated cervical spinal cord plaques. *Brain,* 123**,** 308-317.
LUBLIN, F. D. 2005. Clinical features and diagnosis of multiple sclerosis. *Neurol Clin,* 23**,** 1-15, v.
LUBLIN, F. D., REINGOLD, S. C., COHEN, J. A., CUTTER, G. R., SØRENSEN, P. S., THOMPSON, A. J., WOLINSKY, J. S., BALCER, L. J., BANWELL, B., BARKHOF, F., BEBO, B., CALABRESI, P. A., CLANET, M., COMI, G., FOX, R. J., FREEDMAN, M. S., GOODMAN, A. D., INGLESE, M., KAPPOS, L., KIESEIER, B. C., LINCOLN, J. A., LUBETZKI, C., MILLER, A. E., MONTALBAN, X., O'CONNOR, P. W., PETKAU, J., POZZILLI, C., RUDICK, R. A., SORMANI, M. P., STÜVE, O., WAUBANT, E. & POLMAN, C. H. 2014. Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology,* 83**,** 278-286.
LUCCHINETTI, C., BRÜCK, W., PARISI, J., SCHEITHAUER, B., RODRIGUEZ, M. & LASSMANN, H. 1999. A quantitative analysis of oligodendrocytes in multiple sclerosis lesionsA study of 113 cases. *Brain,* 122**,** 2279-2295.
LUDWIN, S. K. 2000. Understanding multiple sclerosis: Lessons from pathology. *Annals of Neurology,* 47**,** 691-693.
LUDWIN, S. K., RAO, V. T., MOORE, C. S. & ANTEL, J. P. 2016. Astrocytes in multiple sclerosis. *Multiple Sclerosis Journal,* 22**,** 1114-1124.
LUO, W., C., A. & ALBRIGHT, J. 2015. The NeuroQuant normative database: Comparing individual brain structures. . CorTechs Labs Inc.
MACKENZIE, I. S., MORANT, S. V., BLOOMFIELD, G. A., MACDONALD, T. M. & O'RIORDAN, J. 2014. Incidence and prevalence of multiple sclerosis in the UK 1990–2010: a descriptive study in the General Practice Research Database. *Journal of Neurology, Neurosurgery & Psychiatry,* 85**,** 76-84.
MADDEN, D. J., BENNETT, I. J., BURZYNSKA, A., POTTER, G. G., CHEN, N. K. & SONG, A. W. 2012. Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochim Biophys Acta,* 1822**,** 386-400.
MAGGI, P., ABSINTA, M., GRAMMATICO, M., VUOLO, L., EMMI, G., CARLUCCI, G., SPAGNI, G., BARILARO, A., REPICE, A. M., EMMI, L., PRISCO, D., MARTINELLI, V., SCOTTI, R., SADEGHI, N., PERROTTA, G., SATI, P., DACHY, B., REICH, D. S., FILIPPI, M. & MASSACESI, L. 2018. The central vein sign differentiates MS from CNS inflammatory vasculopathies. *Ann Neurol*.
MAGHZI, A. H., REVIRAJAN, N., JULIAN, L. J., SPAIN, R., MOWRY, E. M., LIU, S., JIN, C., GREEN, A. J., MCCULLOCH, C. E., PELLETIER, D. & WAUBANT, E. 2014. Magnetic resonance imaging correlates of clinical outcomes in early multiple sclerosis. *Mult Scler Relat Disord,* 3**,** 720-7.
MAGRANER, M., CORET, F. & CASANOVA, B. 2012. The relationship between inflammatory activity and brain atrophy in natalizumab treated patients. *Eur J Radiol,* 81**,** 3485-90.
MAGYARI, M., KOCH-HENRIKSEN, N. & SØRENSEN, P. S. 2016. The Danish Multiple Sclerosis Treatment Register. *Clinical Epidemiology,* 8**,** 549-552.

MASSACESI, L. 2016. Evaluation by brain Mri of white matter perivenular lesions in inflammatory micro-angiopatic ischemia and in demyelinating multiple sclerosis lesions. *European Journal of Neurology,* 23**,** 86.
MCDONALD, W. I., COMPSTON, A., EDAN, G., GOODKIN, D., HARTUNG, H. P., LUBLIN, F. D., MCFARLAND, H. F., PATY, D. W., POLMAN, C. H. & REINGOLD, S. C. 2001. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. *Ann Neurol,* 50.
MEIER, D. S., WEINER, H. L. & GUTTMANN, C. R. 2007. MR imaging intensity modeling of damage and repair in multiple sclerosis: relationship of short-term lesion recovery to progression and disability. *AJNR Am J Neuroradiol,* 28**,** 1956-63.
MERCALDO, N. D., LAU, K. F. & ZHOU, X. H. 2007. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat Med,* 26**,** 2170-83.
MEWS, I., BERGMANN, M., BUNKOWSKI, S., GULLOTTA, F. & BRÜCK, W. 1998. Oligodendrocyte and axon pathology in clinically silent multiple sclerosis lesions. *Multiple Sclerosis Journal,* 4**,** 55-62.
MILLER, D. H., BARKHOF, F., FRANK, J. A., PARKER, G. J. M. & THOMPSON, A. J. 2002. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. *Brain,* 125**,** 1676-1695.
MILLER, D. H., SOON, D., FERNANDO, K. T., MACMANUS, D. G., BARKER, G. J., YOUSRY, T. A., FISHER, E., O'CONNOR, P. W., PHILLIPS, J. T., POLMAN, C. H., KAPPOS, L., HUTCHINSON, M., HAVRDOVA, E., LUBLIN, F. D., GIOVANNONI, G., WAJGT, A., RUDICK, R., LYNN, F., PANZARA, M. A. & SANDROCK, A. W. 2007. MRI outcomes in a placebo-controlled trial of natalizumab in relapsing MS. *Neurology,* 68**,** 1390-1401.
MISTRY, N., ABDEL-FAHIM, R., SAMARAWEERA, A., MOUGIN, O., TALLANTYRE, E., TENCH, C., JASPAN, T., MORRIS, P., MORGAN, P. S. & EVANGELOU, N. 2015. Imaging central veins in brain lesions with 3-T T2*-weighted magnetic resonance imaging differentiates multiple sclerosis from microangiopathic brain lesions. *Mult Scler*.
MISTRY, N., ABDEL-FAHIM, R., SAMARAWEERA, A., MOUGIN, O., TALLANTYRE, E., TENCH, C., JASPAN, T., MORRIS, P., MORGAN, P. S. & EVANGELOU, N. 2016. Imaging central veins in brain lesions with 3-T T2*-weighted magnetic resonance imaging differentiates multiple sclerosis from microangiopathic brain lesions. *Multiple Sclerosis Journal,* 22**,** 1289-1296.
MISTRY, N., DIXON, J., TALLANTYRE, E., TENCH, C., ABDEL-FAHIM, R., JASPAN, T., MORGAN, P. S., MORRIS, P. & EVANGELOU, N. 2013. Central veins in brain lesions visualized with high-field magnetic resonance imaging: a pathologically specific diagnostic biomarker for inflammatory demyelination in the brain. *JAMA Neurol,* 70**,** 623-8.
MORAAL, B., POHL, C., UITDEHAAG, B. M. J., POLMAN, C. H., EDAN, G., FREEDMAN, M. S., HARTUNG, H.-P., KAPPOS, L., MILLER, D. H., MONTALBAN, X., LANIUS, V., SANDBRINK, R. & BARKHOF, F. 2009. Magnetic Resonance Imaging Predictors of Conversion to Multiple

Sclerosis in the BENEFIT StudyMRI Predictors of Conversion to MS. *JAMA Neurology,* 66**,** 1345-1352.
MS-SOCIETY. 2012. *Our research priorities* [Online]. Available: https://www.mssociety.org.uk/research/explore-our-research/our-research-priorities# [Accessed 07/08/2019 2019].
MSBASE 2018a. MSBase Patient enrolment.
MSBASE. 2018b. *Peer-reviewed journal aricles* [Online]. Available: https://www.msbase.org/data-and-findings/peer-reviewed-journal-articles/ [Accessed 27/07/2018].
MULERO, P., MIDAGLIA, L. & MONTALBAN, X. 2018. Ocrelizumab: a new milestone in multiple sclerosis therapy. *Therapeutic advances in neurological disorders,* 11**,** 1756286418773025-1756286418773025.
MÜLLER, M., ESSER, R., KÖTTER, K., VOSS, J., MÜLLER, A. & STELLMES, P. 2013. Third ventricular enlargement in early stages of multiple sclerosis is a predictor of motor and neuropsychological deficits: a cross-sectional study. *BMJ Open,* 3.
MYHR, K. M., RIISE, T., VEDELER, C., NORTVEDT, M. W., GRØNNING, M., MIDGARD, R. & NYLAND, H. I. 2001. Disability and prognosis in multiple sclerosis: demographic and clinical variables important for the ability to walk and awarding of disability pension. *Multiple Sclerosis Journal,* 7**,** 59-65.
NAKAMURA, K., BROWN, R. A., ARAUJO, D., NARAYANAN, S. & ARNOLD, D. L. 2014. Correlation between brain volume change and T2 relaxation time induced by dehydration and rehydration: Implications for monitoring atrophy in clinical studies. *NeuroImage: Clinical,* 6**,** 166-170.
NAKAMURA, K., BROWN, R. A., NARAYANAN, S., COLLINS, D. L. & ARNOLD, D. L. 2015. Diurnal fluctuations in brain volume: Statistical analyses of MRI from large populations. *Neuroimage,* 118**,** 126-32.
NAKAMURA, K., ESKILDSEN, S. F., NARAYANAN, S., ARNOLD, D. L. & COLLINS, D. L. 2018. Improving the SIENA performance using BEaST brain extraction. *PLoS One,* 13**,** e0196945.
NARAYANA, P. A., DOYLE, T. J., LAI, D. & WOLINSKY, J. S. 1998. Serial proton magnetic resonance spectroscopic imaging, contrast-enhanced magnetic resonance imaging, and quantitative lesion volumetry in multiple sclerosis. *Ann Neurol,* 43**,** 56-71.
OCHS, A. L., ROSS, D. E., ZANNONI, M. D., ABILDSKOV, T. J. & BIGLER, E. D. 2015. Comparison of Automated Brain Volume Measures obtained with NeuroQuant and FreeSurfer. *J Neuroimaging,* 25**,** 721-7.
OFFICE FOR NATIONAL STATISTICS 2018. UK labour market: June 2018. Estimates of employment, unemployment, economic inactivity and other employment-related statistics for the UK.
OGAWA, S., LEE, T.-M., KAY, A. R. & TANK, D. W. 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences,* 87**,** 9868-9872.
OLEK, M. J. 2005. Differential Diagnosis, Clinical Features, and Prognosis of Multiple Sclerosis. *In:* OLEK, M. J. (ed.) *Multiple Sclerosis: Etiology, Diagnosis, and New Treatment Strategies.* Totowa, NJ: Humana Press.
PANTONI, L. & GARCIA, J. H. 1995. The Significance of Cerebral White Matter Abnormalities 100 Years After Binswanger's Report. *Stroke,* 26**,** 1293.

PATRUCCO, L., ROJAS, J., MIGUEZ, J. & CRISTIANO, E. 2013. Application of the McDonald 2010 criteria for the diagnosis of multiple sclerosis in an Argentinean cohort of patients with clinically isolated syndromes. *Multiple Sclerosis Journal,* 19**,** 1297-1301.

PENNY, W. D., FRISTON, K. J., ASHBURNER, J. T., KIEBEL, S. J. & NICHOLS, T. E. 2011. *Statistical parametric mapping: the analysis of functional brain images*, Elsevier.

PEREZ-MIRALLES, F., SASTRE-GARRIGA, J., TINTORE, M., ARRAMBIDE, G., NOS, C., PERKAL, H., RIO, J., EDO, M. C., HORGA, A., CASTILLO, J., AUGER, C., HUERGA, E., ROVIRA, A. & MONTALBAN, X. 2013. Clinical impact of early brain atrophy in clinically isolated syndromes. *Mult Scler,* 19**,** 1878-86.

PETERSON, J. W., BO, L., MORK, S., CHANG, A. & TRAPP, B. D. 2001. Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. *Ann Neurol,* 50.

POLMAN, C. H., REINGOLD, S. C., BANWELL, B., CLANET, M., COHEN, J. A., FILIPPI, M., FUJIHARA, K., HAVRDOVA, E., HUTCHINSON, M., KAPPOS, L., LUBLIN, F. D., MONTALBAN, X., O'CONNOR, P., SANDBERG-WOLLHEIM, M., THOMPSON, A. J., WAUBANT, E., WEINSHENKER, B. & WOLINSKY, J. S. 2011. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol,* 69**,** 292-302.

POPESCU, B. F. & LUCCHINETTI, C. F. 2012a. Pathology of demyelinating diseases. *Annu Rev Pathol,* 7**,** 185-217.

POPESCU, B. F. G. & LUCCHINETTI, C. F. 2012b. Pathology of Demyelinating Diseases. *Annual Review of Pathology: Mechanisms of Disease,* 7**,** 185-217.

POPESCU, B. F. G., PIRKO, I. & LUCCHINETTI, C. F. 2013a. Pathology of multiple sclerosis: where do we stand? *CONTINUUM: Lifelong Learning in Neurology,* 19**,** 901.

POPESCU, V., AGOSTA, F., HULST, H. E., SLUIMER, I. C., KNOL, D. L., SORMANI, M. P., ENZINGER, C., ROPELE, S., ALONSO, J., SASTRE-GARRIGA, J., ROVIRA, A., MONTALBAN, X., BODINI, B., CICCARELLI, O., KHALEELI, Z., CHARD, D. T., MATTHEWS, L., PALACE, J., GIORGIO, A., DE STEFANO, N., EISELE, P., GASS, A., POLMAN, C. H., UITDEHAAG, B. M., MESSINA, M. J., COMI, G., FILIPPI, M., BARKHOF, F. & VRENKEN, H. 2013b. Brain atrophy and lesion load predict long term disability in multiple sclerosis. *J Neurol Neurosurg Psychiatry,* 84**,** 1082-91.

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**,** 1484-1494.

POSER, C. M., PATY, D. W., SCHEINBERG, L., MCDONALD, W. I., DAVIS, F. A., EBERS, G. C., JOHNSON, K. P., SIBLEY, W. A., SILBERBERG, D. H. & TOURTELLOTTE, W. W. 1983. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol,* 13.

PREBOSKE, G. M., GUNTER, J. L., WARD, C. P. & JACK, C. R., JR. 2006. Common MRI acquisition non-idealities significantly impact the output of the boundary shift integral method of measuring brain atrophy on serial MRI. *Neuroimage,* 30**,** 1196-202.
PRINEAS, J., KWON, E., CHO, E. S. & SHARER, L. 1984. Continual Breakdown and Regeneration of Myelin in Progressive Multiple Sclerosis Plaques a. *Annals of the New York Academy of Sciences,* 436**,** 11-32.
PROSPERINI, L., GALLO, V., PETSAS, N., BORRIELLO, G. & POZZILLI, C. 2009. One-year MRI scan predicts clinical response to interferon beta in multiple sclerosis. *Eur J Neurol,* 16**,** 1202-9.
PUGLIATTI, M., SOTGIU, S. & ROSATI, G. 2002. The worldwide prevalence of multiple sclerosis. *Clinical Neurology and Neurosurgery,* 104**,** 182-191.
R CORE TEAM 2013. R: A language and environment for statistical computing.
RADUE, E.-W., SPRENGER, T., GAETANO, L., MUELLER-LENKE, N., CAVALIER, S., THANGAVELU, K., PANZARA, M. A., DONALDSON, J. E., WOODWARD, F. M., WUERFEL, J., WOLINSKY, J. S. & KAPPOS, L. 2017. Teriflunomide slows BVL in relapsing MS: A reanalysis of the TEMSO MRI data set using SIENA. *Neurology(R) neuroimmunology & neuroinflammation,* 4**,** e390-e390.
RAJI, A., OSTWALDT, A.-C., OPFER, R., SUPPA, P., SPIES, L. & WINKLER, G. 2018. MRI-Based Brain Volumetry at a Single Time Point Complements Clinical Evaluation of Patients With Multiple Sclerosis in an Outpatient Setting. *Frontiers in neurology,* 9**,** 545-545.
RAZ, N., GUNNING, F. M., HEAD, D., DUPUIS, J. H., MCQUAIN, J., BRIGGS, S. D., LOKEN, W. J., THORNTON, A. E. & ACKER, J. D. 1997. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cereb Cortex,* 7**,** 268-82.
RAZ, N., RODRIGUE, K. M., KENNEDY, K. M. & ACKER, J. D. 2004. Hormone replacement therapy and age-related brain shrinkage: regional effects. *Neuroreport,* 15**,** 2531-2534.
REICHENBACH, J. R., VENKATESAN, R., SCHILLINGER, D. J., KIDO, D. K. & HAACKE, E. M. 1997. Small vessels in the human brain: MR venography with deoxyhemoglobin as an intrinsic contrast agent. *Radiology,* 204**,** 272-7.
REUTER, M., TISDALL, M. D., QURESHI, A., BUCKNER, R. L., VAN DER KOUWE, A. J. W. & FISCHL, B. 2015. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *Neuroimage,* 107**,** 107-115.
ROCCA, M. A., MESAROS, S., PAGANI, E., SORMANI, M. P., COMI, G. & FILIPPI, M. 2010. Thalamic damage and long-term progression of disability in multiple sclerosis. *Radiology,* 257**,** 463-9.
ROSATI, G. 2001. The prevalence of multiple sclerosis in the world: an update. *Neurol Sci,* 22**,** 117-39.
ROSS, D. E., OCHS, A. L., TATE, D. F., TOKAC, U., SEABAUGH, J., ABILDSKOV, T. J. & BIGLER, E. D. 2018. High correlations between MRI brain volume measurements based on NeuroQuant((R)) and FreeSurfer. *Psychiatry Res Neuroimaging*.
ROVIRA, A., WATTJES, M. P., TINTORE, M., TUR, C., YOUSRY, T. A., SORMANI, M. P., DE STEFANO, N., FILIPPI, M., AUGER, C., ROCCA, M. A., BARKHOF, F., FAZEKAS, F., KAPPOS, L., POLMAN, C.,

MILLER, D. & MONTALBAN, X. 2015. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis-clinical implementation in the diagnostic process. *Nat Rev Neurol,* 11**,** 471-82.

RUBERTE, E., SINNECKER, T., AMANN, M., GAETANO, L., NAEGELIN, Y., PENNER, I. K., KUHLE, J., DERFUSS, T., KAPPOS, L., GRANZIERA, C., WUERFEL, J. & YALDIZLI, O. 2018. Central Slab versus Whole Brain to Measure Brain Atrophy in Multiple Sclerosis. *Eur Neurol,* 80**,** 207-214.

SACHDEV, P., WEN, W., CHEN, X. & BRODATY, H. 2007. Progression of white matter hyperintensities in elderly individuals over 3 years. *Neurology,* 68**,** 214-222.

SAHRAIAN, M. A., RADUE, E. W., HALLER, S. & KAPPOS, L. 2010. Black holes in multiple sclerosis: definition, evolution, and clinical correlations. *Acta Neurologica Scandinavica,* 122**,** 1-8.

SAMARAWEERA, A. P., CLARKE, M. A., WHITEHEAD, A., FALAH, Y., DRIVER, I. D., DINEEN, R. A., MORGAN, P. S. & EVANGELOU, N. 2016. The Central Vein Sign in Multiple Sclerosis Lesions Is Present Irrespective of the T2* Sequence at 3 T. *J Neuroimaging*.

SARTORI, A., ABDOLI, M. & FREEDMAN, M. S. 2017. Can we predict benign multiple sclerosis? Results of a 20-year long-term follow-up study. *J Neurol,* 264**,** 1068-1075.

SASTRE-GARRIGA, J., PARETO, D. & ROVIRA, A. 2017. Brain Atrophy in Multiple Sclerosis: Clinical Relevance and Technical Aspects. *Neuroimaging Clin N Am,* 27**,** 289-300.

SASTRE-GARRIGA, J., TUR, C., PARETO, D., VIDAL-JORDANA, A., AUGER, C., RIO, J., HUERGA, E., TINTORE, M., ROVIRA, A. & MONTALBAN, X. 2015. Brain atrophy in natalizumab-treated patients: A 3-year follow-up. *Mult Scler,* 21**,** 749-56.

SATI, P., GEORGE, I. C., SHEA, C. D., GAITÁN, M. I. & REICH, D. S. 2012. FLAIR*: a combined MR contrast technique for visualizing white matter lesions and parenchymal veins. *Radiology,* 265**,** 926-932.

SATI, P., OH, J., CONSTABLE, R. T., EVANGELOU, N., GUTTMANN, C. R. G., HENRY, R. G., KLAWITER, E. C., MAINERO, C., MASSACESI, L., MCFARLAND, H., NELSON, F., ONTANEDA, D., RAUSCHER, A., ROONEY, W. D., SAMARAWEERA, A. P. R., SHINOHARA, R. T., SOBEL, R. A., SOLOMON, A. J., TREABA, C. A., WUERFEL, J., ZIVADINOV, R., SICOTTE, N. L., PELLETIER, D., REICH, D. S. & ON BEHALF OF THE, N. C. 2016. The central vein sign and its clinical evaluation for the diagnosis of multiple sclerosis: a consensus statement from the North American Imaging in Multiple Sclerosis Cooperative. *Nat Rev Neurol,* 12**,** 714-722.

SCAHILL, R. I., FROST, C., JENKINS, R., WHITWELL, J. L., ROSSOR, M. N. & FOX, N. C. 2003. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch Neurol,* 60**,** 989-94.

SCALFARI, A., NEUHAUS, A., DAUMER, M., MURARO, P. A. & EBERS, G. C. 2014. Onset of secondary progressive phase and long-term evolution of multiple sclerosis. *J Neurol Neurosurg Psychiatry,* 85**,** 67-75.

SCALFARI, A., NEUHAUS, A., DEGENHARDT, A., RICE, G. P., MURARO, P. A., DAUMER, M. & EBERS, G. C. 2010. The natural history of multiple sclerosis, a geographically based study 10: relapses and long-term disability. *Brain,* 133**,** 1914-1929.

SCHILD, H. H. 1992. *MRI Made Easy:(--well Almost)*, Berlex Laboratories.

SCHOONHEIM, M. M. & CICCARELLI, O. 2018. The value of including thalamic atrophy as a clinical trial endpoint in multiple sclerosis. *Neurology*.

SCHUMACHER, G. A., BEEBE, G., KIBLER, R. F., KURLAND, L. T., KURTZKE, J. F., MCDOWELL, F., NAGLER, B., SIBLEY, W. A., TOURTELLOTTE, W. W. & WILLMON, T. L. 1965. PROBLEMS OF EXPERIMENTAL TRIALS OF THERAPY IN MULTIPLE SCLEROSIS: REPORT BY THE PANEL ON THE EVALUATION OF EXPERIMENTAL TRIALS OF THERAPY IN MULTIPLE SCLEROSIS. *Ann N Y Acad Sci,* 122**,** 552-68.

SHETTY, A. A., MUKHERJI, A., JOSHI, M., FULARA, S. & PATEL, N. 2014. An Atypical Presentation of Multiple Sclerosis. *Open Journal of Clinical Diagnostics,* Vol.04No.03**,** 4.

SIMON, J. H., KINKEL, R. P., JACOBS, L., BUB, L. & SIMONIAN, N. 2000. A Wallerian degeneration pattern in patients at risk for MS. *Neurology,* 54**,** 1155-60.

SINNECKER, T., CLARKE, M. A., MEIER, D., ENZINGER, C., CALABRESE, M., DE STEFANO, N., PITIOT, A., GIORGIO, A., SCHOONHEIM, M. M., PAUL, F., PAWLAK, M. A., SCHMIDT, R., KAPPOS, L., MONTALBAN, X., ROVIRA, A., EVANGELOU, N. & WUERFEL, J. 2019. Evaluation of the Central Vein Sign as a Diagnostic Imaging Biomarker in Multiple Sclerosis. *JAMA Neurol*.

SINNECKER, T., DORR, J., PFUELLER, C. F., HARMS, L., RUPRECHT, K., JARIUS, S., BRUCK, W., NIENDORF, T., WUERFEL, J. & PAUL, F. 2012a. Distinct lesion morphology at 7-T MRI differentiates neuromyelitis optica from multiple sclerosis. *Neurology,* 79**,** 708-714.

SINNECKER, T., DORR, J., PFUELLER, C. F., HARMS, L., RUPRECHT, K., JARIUS, S., BRUCK, W., NIENDORF, T., WUERFEL, J. & PAUL, F. 2012b. Distinct lesion morphology at 7-T MRI differentiates neuromyelitis optica from multiple sclerosis. *Neurology,* 79**,** 708-14.

SMEETS, D., RIBBENS, A., SIMA, D. M., CAMBRON, M., HORAKOVA, D., JAIN, S., MAERTENS, A., VAN VLIERBERGHE, E., TERZOPOULOS, V., VAN BINST, A. M., VANECKOVA, M., KRASENSKY, J., UHER, T., SEIDL, Z., DE KEYSER, J., NAGELS, G., DE MEY, J., HAVRDOVA, E. & VAN HECKE, W. 2016. Reliable measurements of brain atrophy in individual patients with multiple sclerosis. *Brain and Behavior,* 6.

SMITH, S. M., JENKINSON, M., WOOLRICH, M. W., BECKMANN, C. F., BEHRENS, T. E., JOHANSEN-BERG, H., BANNISTER, P. R., DE LUCA, M., DROBNJAK, I., FLITNEY, D. E., NIAZY, R. K., SAUNDERS, J., VICKERS, J., ZHANG, Y., DE STEFANO, N., BRADY, J. M. & MATTHEWS, P. M. 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage,* 23 Suppl 1**,** S208-19.

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.
SOLOMON, A. J., BOURDETTE, D. N., CROSS, A. H., APPLEBEE, A., SKIDD, P. M., HOWARD, D. B., SPAIN, R. I., CAMERON, M. H., KIM, E., MASS, M. K., YADAV, V., WHITHAM, R. H., LONGBRAKE, E. E., NAISMITH, R. T., WU, G. F., PARKS, B. J., WINGERCHUK, D. M., RABIN, B. L., TOLEDANO, M., TOBIN, W. O., KANTARCI, O. H., CARTER, J. L., KEEGAN, B. M. & WEINSHENKER, B. G. 2016. The contemporary spectrum of multiple sclerosis misdiagnosis: A multicenter study. *Neurology,* 87**,** 1393-9.
SOLOMON, A. J. & CORBOY, J. R. 2017. The tension between early diagnosis and misdiagnosis of multiple sclerosis. *Nat Rev Neurol,* 13**,** 567-572.
SOLOMON, A. J., KLEIN, E. P. & BOURDETTE, D. 2012. "Undiagnosing" multiple sclerosis: the challenge of misdiagnosis in MS. *Neurology,* 78**,** 1986-91.
SOLOMON, A. J., SCHINDLER, M. K., HOWARD, D. B., WATTS, R., SATI, P., NICKERSON, J. P. & REICH, D. S. 2015. “Central vessel sign” on 3T FLAIR* MRI for the differentiation of multiple sclerosis from migraine. *Annals of Clinical and Translational Neurology,* 3**,** 82-87.
SOLOMON, A. J., WATTS, R., ONTANEDA, D., ABSINTA, M., SATI, P. & REICH, D. S. 2018. Diagnostic performance of central vein sign for multiple sclerosis with a simplified three-lesion algorithm. *Mult Scler,* 24**,** 750-757.
SORMANI, M. P., ARNOLD, D. L. & DE STEFANO, N. 2014. Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. *Ann Neurol,* 75**,** 43-9.
SORMANI, M. P. & BRUZZI, P. 2013. MRI lesions as a surrogate for relapses in multiple sclerosis: a meta-analysis of randomised trials. *The Lancet Neurology,* 12**,** 669-676.
SPARACIA, G., AGNELLO, F., GAMBINO, A., SCIORTINO, M. & MIDIRI, M. 2018. Multiple sclerosis: High prevalence of the 'central vein' sign in white matter lesions on susceptibility-weighted images. *Neuroradiol J***,** 1971400918763577.
STONNINGTON, C. M., TAN, G., KLOPPEL, S., CHU, C., DRAGANSKI, B., JACK, C. R., JR., CHEN, K., ASHBURNER, J. & FRACKOWIAK, R. S. 2008. Interpreting scan data acquired from multiple scanners: a study with Alzheimer's disease. *Neuroimage,* 39**,** 1180-5.
STREITBÜRGER, D.-P., MÖLLER, H. E., TITTGEMEYER, M., HUND-GEORGIADIS, M., SCHROETER, M. L. & MUELLER, K. 2012. Investigating Structural Brain Changes of Dehydration Using Voxel-Based Morphometry. *PLOS ONE,* 7**,** e44195.
SVENNERHOLM, L., BOSTROM, K. & JUNGBJER, B. 1997. Changes in weight and compositions of major membrane components of human brain during the span of adult human life of Swedes. *Acta Neuropathol,* 94**,** 345-52.
TAKAO, H., HAYASHI, N. & OHTOMO, K. 2011. Effect of scanner in longitudinal studies of brain volume changes. *J Magn Reson Imaging,* 34**,** 438-44.

TALLANTYRE, E., BROOKES, M., DIXON, J., MORGAN, P., EVANGELOU, N. & MORRIS, P. 2008. Demonstrating the perivascular distribution of MS lesions in vivo with 7-Tesla MRI. *Neurology,* 70**,** 2076-2078.
TALLANTYRE, E. C., DIXON, J. E., DONALDSON, I., OWENS, T., MORGAN, P. S., MORRIS, P. G. & EVANGELOU, N. 2011. Ultra-high-field imaging distinguishes MS lesions from asymptomatic white matter lesions. *Neurology,* 76**,** 534-9.
TALLANTYRE, E. C., MORGAN, P. S., DIXON, J. E., AL-RADAIDEH, A., BROOKES, M. J., EVANGELOU, N. & MORRIS, P. G. 2009. A comparison of 3T and 7T in the detection of small parenchymal veins within MS lesions. *Invest Radiol,* 44**,** 491-4.
TAN, I. L., VAN SCHIJNDEL, R. A., POUWELS, P. J., VAN WALDERVEEN, M. A., REICHENBACH, J. R., MANOLIU, R. A. & BARKHOF, F. 2000. MR venography of multiple sclerosis. *AJNR Am J Neuroradiol,* 21**,** 1039-42.
TEDESCHI, G., LAVORGNA, L., RUSSO, P., PRINSTER, A., DINACCI, D., SAVETTIERI, G., QUATTRONE, A., LIVREA, P., MESSINA, C. & REGGIO, A. 2005. Brain atrophy and lesion load in a large population of patients with multiple sclerosis. *Neurology,* 65.
THOMPSON, A., KOBELT, G., BERG, J., CAPSA, D., ERIKSSON, J. & MILLER, D. 2017. New insights into the burden and costs of multiple sclerosis in Europe: Results for the United Kingdom. *Mult Scler,* 23**,** 204-216.
THOMPSON, A. J., BANWELL, B. L., BARKHOF, F., CARROLL, W. M., COETZEE, T., COMI, G., CORREALE, J., FAZEKAS, F., FILIPPI, M., FREEDMAN, M. S., FUJIHARA, K., GALETTA, S. L., HARTUNG, H. P., KAPPOS, L., LUBLIN, F. D., MARRIE, R. A., MILLER, A. E., MILLER, D. H., MONTALBAN, X., MOWRY, E. M., SORENSEN, P. S., TINTORÉ, M., TRABOULSEE, A. L., TROJANO, M., UITDEHAAG, B. M. J., VUKUSIC, S., WAUBANT, E., WEINSHENKER, B. G., REINGOLD, S. C. & COHEN, J. A. 2018. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology,* 17**,** 162-173.
TIBSHIRANI, R. 1996. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological),* 58**,** 267-288.
TINTORE, M., ROVIRA, A., BRIEVA, L., GRIVE, E., JARDI, R., BORRAS, C. & MONTALBAN, X. 2001. Isolated demyelinating syndromes: comparison of CSF oligoclonal bands and different MR imaging criteria to predict conversion to CDMS. *Mult Scler,* 7**,** 359-63.
TINTORE, M., ROVIRA, A., RIO, J., OTERO-ROMERO, S., ARRAMBIDE, G., TUR, C., COMABELLA, M., NOS, C., AREVALO, M. J., NEGROTTO, L., GALAN, I., VIDAL-JORDANA, A., CASTILLO, J., PALAVRA, F., SIMON, E., MITJANA, R., AUGER, C., SASTRE-GARRIGA, J. & MONTALBAN, X. 2015. Defining high, medium and low impact prognostic factors for developing multiple sclerosis. *Brain,* 138**,** 1863-74.
TOMIC, D., KAPPOS, L., MEIER, D. P., HÄRING, D., MEINERT, R., GIOVANNONI, G. & CHITNIS, T. 2018. Predictors of Conversion to Secondary Progressive Multiple Sclerosis in Patients With Relapsing–Remitting Multiple Sclerosis (P2.393). *Neurology,* 90**,** P2.393.
TRABOULSEE, A., SIMON, J., STONE, L., FISHER, E., JONES, D., MALHOTRA, A., NEWSOME, S., OH, J., REICH, D. & RICHERT, N.

2016. Revised recommendations of the consortium of MS centers task force for a standardized MRI protocol and clinical guidelines for the diagnosis and follow-up of multiple sclerosis. *American Journal of Neuroradiology,* 37**,** 394-401.
TRABOULSEE, A. L. & LI, D. K. 2006. The role of MRI in the diagnosis of multiple sclerosis. *Adv Neurol,* 98**,** 125-46.
TRAPP, B. D., PETERSON, J., RANSOHOFF, R. M., RUDICK, R., MORK, S. & BO, L. 1998. Axonal transection in the lesions of multiple sclerosis. *N Engl J Med,* 338**,** 278-85.
TRATTNIG, S., BOGNER, W., GRUBER, S., SZOMOLANYI, P., JURAS, V., ROBINSON, S., ZBYN, S. & HANEDER, S. 2016. Clinical applications at ultrahigh field (7 T). Where does it make the difference? *NMR Biomed,* 29**,** 1316-34.
TREFLER, A., SADEGHI, N., THOMAS, A. G., PIERPAOLI, C., BAKER, C. I. & THOMAS, C. 2016. Impact of time-of-day on brain morphometric measures derived from T 1-weighted magnetic resonance imaging. *NeuroImage,* 133**,** 41-52.
TRIP, S. A. & MILLER, D. H. 2005. Imaging in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry,* 76**,** iii11-iii18.
TROJANO, M., LUCCHESE, G., GRAZIANO, G., TAYLOR, B. V., SIMPSON, S., JR., LEPORE, V., GRAND'MAISON, F., DUQUETTE, P., IZQUIERDO, G., GRAMMOND, P., AMATO, M. P., BERGAMASCHI, R., GIULIANI, G., BOZ, C., HUPPERTS, R., VAN PESCH, V., LECHNER-SCOTT, J., CRISTIANO, E., FIOL, M., OREJA-GUEVARA, C., SALADINO, M. L., VERHEUL, F., SLEE, M., PAOLICELLI, D., TORTORELLA, C., D'ONGHIA, M., IAFFALDANO, P., DIRENZO, V. & BUTZKUEVEN, H. 2012. Geographical variations in sex ratio trends over time in multiple sclerosis. *PLoS One,* 7**,** e48078.
TROJANO, M., PAOLICELLI, D., BELLACOSA, A. & CATALDO, S. 2003. The transition from relapsing-remitting MS to irreversible disability: clinical evaluation. *Neurol Sci,* 24 Suppl 5**,** S268-70.
TUR, C., MOCCIA, M., BARKHOF, F., CHATAWAY, J., SASTRE-GARRIGA, J., THOMPSON, A. J. & CICCARELLI, O. 2018. Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. *Nat Rev Neurol,* 14**,** 75-93.
TURNER, B., RAMLI, N., BLUMHARDT, L. D. & JASPAN, T. 2001. Ventricular enlargement in multiple sclerosis: a comparison of three-dimensional and linear MRI estimates. *Neuroradiology,* 43**,** 608-14.
UHER, T., HORAKOVA, D., BERGSLAND, N., TYBLOVA, M., RAMASAMY, D. P., SEIDL, Z., VANECKOVA, M., KRASENSKY, J., HAVRDOVA, E. & ZIVADINOV, R. 2014. MRI correlates of disability progression in patients with CIS over 48 months. *Neuroimage Clin,* 6**,** 312-9.
UK MS REGISTER. 2018. *UK MS Register* [Online]. Available: http://www.ukmsregister.org [Accessed 10/07/2018].
VAN DER VALK, P. & AMOR, S. 2009. Preactive lesions in multiple sclerosis. *Current Opinion in Neurology,* 22**,** 207-213.
VAN DER VUURST DE VRIES, R. M., MESCHERIAKOVA, J. Y., WONG, Y. M. & ET AL. 2018. Application of the 2017 revised mcdonald criteria for multiple sclerosis to patients with a typical clinically isolated syndrome. *JAMA Neurology*.

VAN MUNSTER, C. E. P. & UITDEHAAG, B. M. J. 2017. Outcome Measures in Clinical Trials for Multiple Sclerosis. *CNS Drugs,* 31**,** 217-236.
VAN OSCH, M. J. P. & WEBB, A. G. 2014. Safety of Ultra-High Field MRI: What are the Specific Risks? *Current Radiology Reports,* 2**,** 61.
VAN WAESBERGHE, J. H., KAMPHORST, W., DE GROOT, C. J., VAN WALDERVEEN, M. A., CASTELIJNS, J. A., RAVID, R., LYCKLAMA A NIJEHOLT, G. J., VAN DER VALK, P., POLMAN, C. H., THOMPSON, A. J. & BARKHOF, F. 1999. Axonal loss in multiple sclerosis lesions: magnetic resonance imaging insights into substrates of disability. *Ann Neurol,* 46**,** 747-54.
VIDAL-JORDANA, A., SASTRE-GARRIGA, J., PEREZ-MIRALLES, F., TUR, C., TINTORE, M., HORGA, A., AUGER, C., RIO, J., NOS, C., EDO, M. C., AREVALO, M. J., CASTILLO, J., ROVIRA, A. & MONTALBAN, X. 2013. Early brain pseudoatrophy while on natalizumab therapy is due to white matter volume changes. *Mult Scler,* 19**,** 1175-81.
VOLLMER, T., SIGNOROVITCH, J., HUYNH, L., GALEBACH, P., KELLEY, C., DIBERNARDO, A. & SASANE, R. 2015. The natural history of brain volume loss among patients with multiple sclerosis: a systematic literature review and meta-analysis. *J Neurol Sci,* 357**,** 8-18.
VRENKEN, H., JENKINSON, M., HORSFIELD, M. A., BATTAGLINI, M., VAN SCHIJNDEL, R. A., ROSTRUP, E., GEURTS, J. J. G., FISHER, E., ZIJDENBOS, A., ASHBURNER, J., MILLER, D. H., FILIPPI, M., FAZEKAS, F., ROVARIS, M., ROVIRA, A., BARKHOF, F., DE STEFANO, N. & GRP, M. S. 2013. Recommendations to improve imaging and analysis of brain lesion load and atrophy in longitudinal studies of multiple sclerosis. *Journal of Neurology,* 260**,** 2458-2471.
VRENKEN, H., VOS, E. K., VAN DER FLIER, W. M., SLUIMER, I. C., COVER, K. S., KNOL, D. L. & BARKHOF, F. 2014. Validation of the automated method VIENA: an accurate, precise, and robust measure of ventricular enlargement. *Hum Brain Mapp,* 35**,** 1101-10.
WALHOVD, K. B., FJELL, A. M., REINVANG, I., LUNDERVOLD, A., DALE, A. M., EILERTSEN, D. E., QUINN, B. T., SALAT, D., MAKRIS, N. & FISCHL, B. 2005. Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiol Aging,* 26**,** 1261-70; discussion 1275-8.
WANG, C., BEADNALL, H. N., HATTON, S. N., BADER, G., TOMIC, D., SILVA, D. G. & BARNETT, M. H. 2016. Automated brain volumetrics in multiple sclerosis: a step closer to clinical application. *J Neurol Neurosurg Psychiatry,* 87**,** 754-7.
WEINSHENKER, B. G., BASS, B., RICE, G. P. A., NOSEWORTHY, J., CARRIERE, W., BASKERVILLE, J. & EBERS, G. C. 1989. THE NATURAL HISTORY OF MULTIPLE SCLEROSIS: A GEOGRAPHICALLY BASED STUDYI. CLINICAL COURSE AND DISABILITY. *Brain,* 112**,** 133-146.
WERRING, D. J., BRASSAT, D., DROOGAN, A. G., CLARK, C. A., SYMMS, M. R., BARKER, G. J., MACMANUS, D. G., THOMPSON, A. J. & MILLER, D. H. 2000. The pathogenesis of lesions and normal-appearing white matter changes in multiple sclerosis: a serial diffusion MRI study. *Brain,* 123 (Pt 8)**,** 1667-76.
WINGERCHUK, D. M., BANWELL, B., BENNETT, J. L., CABRE, P., CARROLL, W., CHITNIS, T., DE SEZE, J., FUJIHARA, K.,

GREENBERG, B., JACOB, A., JARIUS, S., LANA-PEIXOTO, M., LEVY, M., SIMON, J. H., TENEMBAUM, S., TRABOULSEE, A. L., WATERS, P., WELLIK, K. E. & WEINSHENKER, B. G. 2015. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. *Neurology,* 85**,** 177-189.

WU, E. & RAINE, C. S. 1992. Multiple sclerosis. Interactions between oligodendrocytes and hypertrophic astrocytes and their occurrence in other, nondemyelinating conditions. *Lab Invest,* 67**,** 88-99.

WU, G. F. & ALVAREZ, E. 2011. The immunopathophysiology of multiple sclerosis. *Neurologic clinics,* 29**,** 257-278.

WUERFEL, J., BELLMANN-STROBL, J., BRUNECKER, P., AKTAS, O., MCFARLAND, H., VILLRINGER, A. & ZIPP, F. 2004. Changes in cerebral perfusion precede plaque formation in multiple sclerosis: a longitudinal perfusion MRI study. *Brain,* 127**,** 111-9.

WUERFEL, J., SINNECKER, T., RINGELSTEIN, E. B., JARIUS, S., SCHWINDT, W., NIENDORF, T., PAUL, F., KLEFFNER, I. & DORR, J. 2012. Lesion morphology at 7 Tesla MRI differentiates Susac syndrome from multiple sclerosis. *Multiple Sclerosis Journal,* 18**,** 1592-1599.

YAO, B., BAGNATO, F., MATSUURA, E., MERKLE, H., VAN GELDEREN, P., CANTOR, F. K. & DUYN, J. H. 2012. Chronic multiple sclerosis lesions: characterization with high-field-strength MR imaging. *Radiology,* 262**,** 206-15.

YAO, Y., NGUYEN, T. D., PANDYA, S., ZHANG, Y., HURTADO RÚA, S., KOVANLIKAYA, I., KUCEYESKI, A., LIU, Z., WANG, Y. & GAUTHIER, S. A. 2018. Combining Quantitative Susceptibility Mapping with Automatic Zero Reference (QSM0) and Myelin Water Fraction Imaging to Quantify Iron-Related Myelin Damage in Chronic Active MS Lesions. *American Journal of Neuroradiology,* 39**,** 303-310.

YOUDEN, W. J. 1950. Index for rating diagnostic tests. *Cancer,* 3**,** 32-5.

ZALC, B. 2018. One hundred and fifty years ago Charcot reported multiple sclerosis as a new neurological disease. *Brain : a journal of neurology,* 141**,** 3482-3488.

ZAREI, M., CHANDRAN, S., COMPSTON, A. & HODGES, J. 2003. Cognitive presentation of multiple sclerosis: evidence for a cortical variant. *J Neurol Neurosurg Psychiatry,* 74.

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 Trans Med Imaging,* 20**,** 45-57.

ZIVADINOV, R., BERGSLAND, N., DOLEZAL, O., HUSSEIN, S., SEIDL, Z., DWYER, M. G., VANECKOVA, M., KRASENSKY, J., POTTS, J. A., KALINCIK, T., HAVRDOVA, E. & HORAKOVA, D. 2013. Evolution of cortical and thalamus atrophy and disability progression in early relapsing-remitting MS during 5 years. *AJNR Am J Neuroradiol,* 34**,** 1931-9.

ZIVADINOV, R., BERGSLAND, N., KORN, J. R., DWYER, M. G., KHAN, N., MEDIN, J., PRICE, J. C., WEINSTOCK-GUTTMAN, B. & SILVA, D. 2018. Feasibility of Brain Atrophy Measurement in Clinical Routine without Prior Standardization of the MRI Protocol: Results from MS-MRIUS, a Longitudinal Observational, Multicenter Real-World Outcome

Study in Patients with Relapsing-Remitting MS. *American Journal of Neuroradiology,* 39**,** 289-295.
ZIVADINOV, R., JAKIMOVSKI, D., GANDHI, S., AHMED, R., DWYER, M. G., HORAKOVA, D., WEINSTOCK-GUTTMAN, B., BENEDICT, R. R., VANECKOVA, M., BARNETT, M. & BERGSLAND, N. 2016a. Clinical relevance of brain atrophy assessment in multiple sclerosis. Implications for its use in a clinical routine. *Expert Rev Neurother,* 16**,** 777-93.
ZIVADINOV, R., KHAN, N., MEDIN, J., CHRISTOFFERSEN, P., PRICE, J., KORN, J. R., BONZANI, I., DWYER, M. G., BERGSLAND, N., CARL, E., SILVA, D. & WEINSTOCK-GUTTMAN, B. 2017. An Observational Study to Assess Brain MRI Change and Disease Progression in Multiple Sclerosis Clinical Practice—The MS-MRIUS Study. *Journal of Neuroimaging,* 27**,** 339-347.
ZIVADINOV, R., MEDIN, J., KHAN, N., KORN, J. R., CHITNIS, T., NAISMITH, R. T., ALVAREZ, E., DWYER, M. G., BERGSLAND, N., CARL, E., SILVA, D. & WEINSTOCK-GUTTMAN, B. 2019. Impact of fingolimod on clinical and magnetic resonance imaging outcomes in routine clinical practice: A retrospective analysis of the multiple sclerosis, clinical and MRI outcomes in the USA (MS-MRIUS) study. *Multiple Sclerosis and Related Disorders,* 27**,** 65-73.
ZIVADINOV, R., REDER, A. T., FILIPPI, M., MINAGAR, A., STUVE, 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.
ZIVADINOV, R., SEPCIC, J., NASUELLI, D., DE MASI, R., BRAGADIN, L. M., TOMMASI, M. A., ZAMBITO-MARSALA, S., MORETTI, R., BRATINA, A., UKMAR, M., POZZI-MUCELLI, R. S., GROP, A., CAZZATO, G. & ZORZON, M. 2001. A longitudinal study of brain atrophy and cognitive disturbances in the early phase of relapsing-remitting multiple sclerosis. *Journal of neurology, neurosurgery, and psychiatry,* 70**,** 773-780.
ZIVADINOV, R., UHER, T., HAGEMEIER, J., VANECKOVA, M., RAMASAMY, D. P., TYBLOVA, M., BERGSLAND, N., SEIDL, Z., DWYER, M. G., KRASENSKY, J., HAVRDOVA, E. & HORAKOVA, D. 2016b. A serial 10-year follow-up study of brain atrophy and disability progression in RRMS patients. *Mult Scler,* 22**,** 1709-1718.