Clinically isolated syndromes and multiple sclerosis: prospective clinical and MRI follow up after 30 years and features at earlier time-points

Dr Karen K Chung MBBS BSc (hons) MRCP

Queen Square MS Centre, NMR Research Unit,

Institute of Neurology,

University College London, London, UK

Thesis submitted to University College London for the degree of

Doctor of Medicine Research

September 2019

Declaration

I, Karen Ka Yan Chung, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Dr Michael Ebner and Dr Ferran Prados designed the methodology and created the volumetric reconstruction of old film prints, described in Chapter 2.

Dr Declan Chard, Prof Frederik Barkhof and Dr Katherine Miszkiel contributed to image analysis of lesions described in Chapter 2. Dr Lukas Haider, Ms Giselle Birch and Dr Declan Chard contributed to brain atrophy measurements described in Chapter 2.

Dr Dan Altmann assisted with the statistical analyses described in Chapter 2, and presented in Chapters 3 and 4. Dr Lukas Haider assisted with the statistical analyses relating to brain atrophy measurements, described in Chapter 2, and presented in Chapter 4.

Abstract

This thesis is based on a 30-year follow-up study of a cohort of people who initially presented, in the 1980's, with clinically isolated syndromes (CIS), suggestive of relapse-onset multiple sclerosis (MS). The main aims were: 1) to study the very long-term outcome of the cohort, with particular attention on those who have fared well over time, 2) review the idea and definition of 'benign' MS, a controversial entity, and 3) to identify any potential early clinical and radiological features, of 30-year outcome.

MS is a very heterogenous condition and biomarkers of long-term prognostication remain limited. With the increasing range of disease modifying therapies available, it is important that treatment decisions should, as far as possible, involve a personalized risk-benefit analysis.

At 30 years, I found that the clinical outcomes of the cohort were diverse. Approximately a third remained CIS, and two thirds developed MS. Within the MS group, who were largely untreated, ~40% remained ambulatory, ~35% had developed significant disability, and 20% had died related to their MS. Comparisons between the ambulatory MS group and CIS group, showed that the groups were not significantly different across several clinical measures.

In this cohort, the strongest early predictors of 30-year outcomes identified were radiological features. The presence of MRI white matter lesions in

specific locations in the brain, within one year of presentation, were able to predict 30-year clinical outcomes with accuracies in the 70-75% range. These results could potentially be applied in a clinical setting and help inform treatment decisions.

Acknowledgements

I would like to thank the United Kingdom MS Society for their sustained funding of this work; this research would not have been possible without their financial support.

I am indebted to Professor David Miller and the late Professor W Ian McDonald for the inception and early follow-up of this cohort. Their early vision enabled the establishment of this historical cohort, from which major insights on CIS and its conversion to MS have been derived.

A sincere and heartfelt thank you to all the participants in this study who gave up their time, in particular those who travelled from afar, those living with disabilities, and those who have attended every one of the seven time-points over the course of the three decades.

I would like to thank my supervisors: Dr Declan Chard, Professor David Miller and Professor Frederik Barkhof. I am grateful for their support, guidance and motivation. I thank Declan Chard for his in-depth knowledge and precision, which have helped designed this study and shaped this thesis and my overall work; his overall guidance, not only on my research, but also on my broader career. I thank Frederik Barkhof for his extensive wisdom on all things radiological and for reviewing endless scans for the project. I thank Professor Miller for his encouragement and for giving me the opportunity to undertake

this research; I would like to wish him a happy retirement.

I thank my predecessors, Drs Sean Morrissey, Jonathan O'Riordan, Peter Brex, and Leonora Fisniku, for their perseverance and follow-up of this cohort, and data collection at earlier time-points.

I would like to thank the wonderful team and colleagues at Queen Square MS Centre, without whom this work would not be possible: Tina Holmes, Marie Braisher, and Charlotte Burt, who provided essential administrative support; Jon Steel for his IT support, who has been particularly patient to my many queries; Dr Marios Yiannakis and Luke Hoy for undertaking the MRI scanning; Michael Ebner from the Translational Imaging Group (TIG), and Dr Ferran Prados from the Physics team, for their reconstruction of images from old films, and general support with image-related matters; Dr Katherine Miszkiel for her expert input on image-analysis; Dr Lukas Haider for image and statistical analyses relating to brain atrophy measurements; I am indebted to Dr Dan Altmann for the complex statistical analyses and overall guidance and patience with all my questions relating to statistics; Professor Olga Ciccarelli, for her overall guidance and encouragement. It has been a privilege to work alongside people of such expertise and knowledge.

I have also made friends with many of my colleagues over the past four years. I would like to thank the support and friendship of: Dr Wallace Brownlee, Dr Will Brown, Dr Steven van de Pavert, Dr Rosanna Cortese, Dr Sara Collorone, Dr Anisha Doshi, Dr Floriana de Angelis, Dr Arman Eshaghi,

Tiggy Beyene, Vanessa Bassan, Ifrah Iidow, Sarah Alexander, Dr Nevin John, and Dr Tom Williams.

Finally, I must thank my family. My parents, who gave me so many opportunities, and believed and supported in my decisions. I would like to thank my husband Nish, who has been a continuous source of encouragement, despite my absence while I worked on my thesis. I would like to thank my children Maya and Seth, and apologise for my time away from them: I hope this makes you proud.

Publications associated with this thesis

Ebner M, Chung KK, Prados F, Cardoso MJ, Chard DT, Vercauteren T, Ourselin S. Volumetric Reconstruction from Printed Films: Enabling 30 Year Longitudinal Analysis in MR Neuroimaging. *Neuroimage* 2017; 165:238-250.

Chung KK, Altmann D, Barkhof F, Miszkiel K, Brex PA, O'Riordan J, Ebner M, Prados F, Cardoso MJ, Vercauteren T, Ourselin S, Thompson AJ, Ciccarelli O, Chard DT. A thirty-year clinical and MRI observational study of multiple sclerosis and clinically isolated syndromes. *Annals Neurol* 2019; 00:1-12.

Platform presentations associated with this thesis

Chung KK, Altmann DR, Miller DH, Chard D. Opportunities and challenges of a long-term clinically isolated syndrome study. Presented at Multiple Sclerosis Frontiers conference, June 2017, Edinburgh, UK

Chung KK, Barkhof F, Altmann DR, Miller DH, Chard D. Does 'benign' multiple sclerosis exist? A 30-year follow-up study of people presenting with clinically isolated syndrome. Presented at 7th Joint European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS) and Americas Committee for Treatment and Research in Multiple Sclerosis (ACTRIMS), October 2017, Paris, France

Chung KK, Barkhof F, Altmann DR, Miszkiel KA, Fisniku LK, Brex PA,

O'Riordan JI, Miller DH, Chard D. Multiple sclerosis outcomes 30-years after symptom onset: a longitudinal study. Presented at the Association of British Neurologists (ABN) annual meeting, May 2018, Birmingham, UK

Chung KK, Altmann D, Barkhof F, Ciccarelli O, Chard D. Early MRI predictors of long-term multiple sclerosis outcomes: a 30-year follow-up study of people with clinically isolated syndromes. Presented at 34th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS), October 2018, Berlin, Germany

Table of Contents

List of abbreviations

ANOVA	Analysis of variance
BICAMS	Brief international cognitive assessment for multiple sclerosis
BMS	Benign multiple sclerosis
BVMTR	Brief visuospatial memory test
CDMS	Clinically definite multiple sclerosis
CI	Confidence interval
CIS	Clinically isolated syndrome
CNS	Central nervous system
CSF	Cerebral spinal fluid
CVLT	California verbal learning test
DICOM	Digital imaging and communications in medicine
DIS	Dissemination in space
DIT	Dissemination in time
DMT	Disease modifying therapy
DWM	Deep white matter
EDSS	Expanded Disability Status Scale
FLAIR	Fluid attenuation inversion recovery
GdE	Gadolinium enhancing
GM	Grey matter
HIV	Human immunodeficiency virus
HR	Hazard ratio

IT	Infra-tentorial
JC	Juxta-cortical
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MSE	Mean square error
MSFC	Multiple sclerosis functional composite
MTR	Magnetization transfer ratio
NART	National Adult Reading Test
NHS	National Health Service
NPV	Negative predictive value
ON	Optic neuritis
OR	Odds ratio
PASAT	Paced auditory serial addition test
PD	Proton density
PPV	Positive predictive value
PSIR	Phase sensitive inversion recovery
PV	Peri-ventricular
RIS	Radiologically isolated syndrome
RRMS	Relapsing-remitting multiple sclerosis
SD	Standard deviation
SDMT	Symbol digit modalities test
SPMS	Secondary progressive multiple sclerosis
T	Tesla (field strength)
TM	Transverse myelitis

TVW	Third ventricular width
WM	White matter

List of figures

3.8 Number of total intracranial lesions at 30-years, by number of known vascular risk factors in participants, n= 63

3.9 Third ventricular width measurements over 30 years, by 30-year status

3.10 Employment and retirement status in people with MS, by EDSS at 30 years

List of tables

3.5 Lesion count by anatomical location, based on classification at 30 years

3.6 Number of participants with each CIS presentation, at baseline and at 30-years

3.7 Number of BICAMS assessments with validated z-scores (where available), performed on people with MS, categorized by EDSS at 30 years

3.8 Comparison of age-adjusted cognitive measurements between MS EDSS ≤3.5 and EDSS >3.5 groups, with the CIS group

4.1 Early univariable predictors of 30-year outcomes: demographic and clinical features

4.2 Contingency tables of early EDSS scores as predictors of 30-year EDSS ≤3.5 versus EDSS >3.5

4.3 Contingency tables of early EDSS scores as predictors of SPMS, and of MS-related death, by 30 years

4.4 Early EDSS values as univariable predictors of 30-year outcomes

4.5 Early WM lesion variables as univariable predictors of 30-year outcomes

4.6 Contingency tables of best early MRI predictors of 30-year EDSS ≤3.5 versus EDSS >3.5

4.7 Contingency tables of best early MRI predictors of SPMS by 30 years

Chapter 1: Introduction to multiple sclerosis

1.1 Introduction and Aims

This thesis is based on work that aimed to assess the very long-term outcomes, and early predictors of these outcomes, in people presenting with clinically isolated syndrome (CIS) suggestive of multiple sclerosis (MS). It also specifically considered 'benign' MS, a controversial phenotype. There are addressed by asking these main questions:

1. Does 'benign' MS exist? If so, what proportion of people with MS run a 'benign' clinical course over their lifetime?
2. How should we define 'benign' MS (BMS)?
3. Can we predict a very long-term (30-years) outcome using early and intermediate clinical and radiological features?

1.2 Multiple Sclerosis

1.2.1 Introduction

Multiple Sclerosis is a chronic, immune-mediated and neurodegenerative condition affecting the central nervous system (CNS). More than 100,000 people in the United Kingdom have MS, and it is the leading non-traumatic cause of disability in young adults in developed countries. Globally it is estimated to affect more than 2 million people (Thompson *et al.*, 2018b; Wallin *et al.*, 2019). The aetiology of MS is unknown, and is likely to be

multifactorial. The current commonly held view is that MS develops in genetically susceptible people, who have been exposed to one or more environmental triggers (Ascherio *et al.*, 2008; Thompson *et al.*, 2018b).

1.2.2 Pathology

Focal demyelinating plaques, found within the CNS white matter of people with MS, have always been considered to be the pathological hallmark of the disease. These lesions have a predilection for the periventricular, juxtacortical, and infratentorial regions within the brain, as well as the spinal cord (Swanton *et al.*, 2007; Filippi *et al.*, 2016). Their presence in these characteristic anatomical sites within the CNS, disseminated in time and in space, is the foundation of the contemporary MS diagnostic criteria (see section 1.2.4)(McDonald *et al.*, 2001). It is thought that plaques evolve over time from being new, active, inflammatory lesions to chronic lesions, with varying degrees of remyelination (Frohman *et al.*, 2006). It is widely believed that early symptoms of MS result from axonal demyelination, with the regression of symptoms attributed to the resolution of inflammation and partial remyelination (Noseworthy *et al.*, 2000). Based on this, drugs developed to treat MS all have the principal aim to reduce or suppress inflammation within the CNS.

In recent years, there has been recognition that the pathological spectrum of MS is not exclusive to the WM, and that the grey matter (GM) is also involved. Pathological processes including axonal loss and atrophy are also

important, and neurodegeneration is increasingly considered to be the main cause of irreversible long-term disability in MS (Jacobsen *et al*., 2014; Eshaghi *et al.*, 2018). However, the temporal relationship(s) between inflammation, demyelination, remyelination, and neurodegeneration, remain uncertain. While brain atrophy occurs very early in the course of MS (Brex *et al.*, 2000; Dalton *et al.*, 2002b; Bergsland *et al*., 2012), it remains unknown at which point brain atrophy becomes clinically relevant. In particular, it is unclear if early brain atrophy has a significant link with long-term clinical outcomes, and so whether or not it represents an early treatment target.

1.2.3 Use of magnetic resonance imaging

Since its development and widespread use in clinical medicine, magnetic resonance imaging (MRI) has been the most important paraclinical tool in the diagnosis of MS, and MRI features have been incorporated in the MS diagnostic criteria since 2001 (described in the next section). In addition to its role in the diagnosis of MS, MRI is used in the monitoring of disease course, clinical trials, and have provided insights into the pathophysiology of MS.

White matter (WM) lesions, seen as hyperintense on proton density (PD)- and T2-weighted MRI, are used as the primary radiological biomarker in clinical practice. While their presence, or lack of, have an important role to play in diagnosis, the correlation between overall lesion quantity at the time of CIS and future disability, is only moderate (Barkhof, 1999; Fisniku *et al.*, 2008a). The use of fluid attenuation inversion recovery (FLAIR) sequences is

increasingly adopted in routine clinical setting, as they can improve the sensitivity of detecting WM lesions.

A proportion of hyperintense lesions on T2-weighted images can be seen as hypointense on a T1-weighted image. These have been shown to represent areas of axonal loss on histopathology (Bitsch *et al.*, 2001), and are commonly termed ‘T1 black holes’. It is thought that ~40% of acute T1 black holes may evolve and become persistent (Bagnato, 2003; Bermel and Bakshi, 2006). T1-weighted images can also be a helpful tool, following the administration of gadolinium contrast, to compare with a corresponding T2-weight image, in order to identify any acutely active MS lesion(s).

Unlike WM lesions, GM lesions are not visible on conventional MRI sequences, and as such, the measurement and quantification of GM pathology are currently not available in routine practice, but limited to research imaging modalities. As brain atrophy is viewed as the principal *in vivo* measurement of neurodegeneration (Roosendaal *et al.*, 2011; Eshaghi *et al.*, 2018), it is increasingly used as a potential biomarker in clinical practice and in MS clinical trials.

There are numerous methods for measuring brain atrophy and change in atrophy over time. A high-resolution three-dimensional T1-weighted image is the most preferred, and using segmentation-based techniques, quantitative three-dimensional atrophy measurements are obtained (Rocca *et al.*, 2017). Quantitative atrophy measurements can also be obtained using two-

dimensional images, for example, linear measurements of third ventricular width, bi-caudate ratio, frontal horn width, and inter-caudate distance (Butzkueven *et al.*, 2008). In one study looking at three-dimensional versus linear estimates of atrophy in a small group of MS patients, measurements of the third ventricle width showed the strongest correlation (Turner *et al.*, 2001). While these linear methods have the advantage that they can be easily applied to any brain image, they are limited by a lack of reproducibility, compared with three-dimensional measures (Bermel and Bakshi, 2006).

Imaging of both WM lesions and atrophy can be similarly applied to the spinal cord.

The role of MRI features as prognostic indicators is discussed in section 1.3.2.

1.2.4 MS Diagnostic criteria

This section concerns the evolving diagnostic criteria for relapse-onset MS. The underpinning feature of MS is the presence of clinical neurological symptoms, caused by demyelination, which are disseminated in time (DIT) and in space (DIS).

The Schumacher criteria, proposed in 1965, were the first internationally recognized criteria for diagnosis. Developed before the availability of diagnostic investigations, they relied on clinical information only. The criteria

required clinical symptoms or relapses, involving two or more areas of the CNS, in a person aged between 10 and 50 years (Schumacher *et al.*, 1965).

In 1983, the Poser criteria was published: this required two or more clinical relapses, with evidence of DIT and DIS. The Poser criteria introduced the concept of clinically definite MS (CDMS), and that paraclinical information, derived from investigative tests, could provide supportive evidence of DIS (Poser *et al.*, 1983). The Poser criteria has been used widely in clinical settings, as well as in research.

As MRI became available in clinical practice, it became a useful tool in the diagnosis of MS. Based on this, the first McDonald diagnostic criteria was proposed in 2001 (McDonald *et al.*, 2001). This has since been revised in 2005, 2010, and 2017 (Polman *et al.*, 2005, 2011; Thompson *et al.*, 2018a), with the requirements for DIT and DIS modified in each revision. Following the 2010 revision, a diagnosis of MS could be made in patients following their first clinical event, i.e. CIS, if there was radiological evidence of DIS and DIT. The most recent revision, in 2017, further accepted the substitution of DIT by the presence of positive oligoclonal bands in the cerebral spinal fluid (CSF).

Table 1.1 outlines the McDonald criteria and its evolution with time. In those with clinical features suggestive of MS, the McDonald criteria are most commonly used by clinicians. They should, however, not be applied unequivocally, and other diagnoses with a potential to mimic MS should be

considered, particularly if atypical clinical and/or paraclinical features are present.

Table 1.1 Evolution and comparison of McDonald criteria for dissemination in space and time, following onset of clinically isolated syndrome

	Dissemination in space	**Dissemination in time**
McDonald 2001 criteria (McDonald *et al.*, 2001)	Three of four of the following: - One GdE or nine T2 lesions - At least one infratentorial lesion - At least one juxtacortical lesion - At least three periventricular lesions One spinal cord lesion can be substituted for one brain lesion	One of the following: - Presence of new GdE lesion, ≥ 3 months after onset of clinical event - Presence of new T2 and/ or GdE lesion at any time, compared to a scan done ≥ 3 months after the clinical event
McDonald 2005 criteria (Polman *et al.*, 2005)	Three of four of the following: - One GdE lesion or nine T2 brain and/or spinal cord lesions - At least one infratentorial or spinal cord lesion - At least one juxtacortical lesion - At least three periventricular lesions	One of the following: - Presence of new GdE lesion, ≥ 3 months after onset of clinical event - Presence of new T2 lesion compared with a reference scan done at least 30 days after the initial clinical event

McDonald 2010 criteria (Polman *et al.*, 2011)	One or more T2 lesion in two of four sites typically affected in demyelination*: - Periventricular - Juxtacortical - Infratentorial - Spinal cord * Lesions in the symptomatic region excluded	One of the following: - Simultaneous presence of GdE and non-enhancing lesions at any time - New T2 and/or GdE lesion on follow-up MRI done at any time point
McDonald 2017 criteria (Thompson *et al.*, 2018a)	One or more T2 lesion in two of four sites typically affected in demyelination*: - Periventricular - Cortical or juxtacortical, - Infratentorial - Spinal cord *No distinction between symptomatic and asymptomatic lesion is required	One of the following: - Demonstration of oligoclonal bands in CSF - Simultaneous presence of GdE and non-enhancing lesions at any time - New T2 and/or GdE lesion on follow-up MRI done at any time point

GdE= gadolinium enhacing; CSF= cerebral spinal fluid

1.3 Clinical course and prognosis

1.3.1 Clinical course

In about 85% of people with MS, their initial course is characterised by relapses and remissions, known as relapsing-remitting MS (RRMS). A relapse is an episode of neurological symptoms, that typically evolves over hours or days, lasts for days to weeks, and is, at least partially, reversible. The initial presenting relapse is termed CIS. More common CIS presentations include optic neuritis, spinal cord or brainstem syndromes. Approximately two thirds of people who start with RRMS develop secondary progressive (SP) MS, when disability gradually worsens over time, in the absence of relapses. Natural history studies have demonstrated that disability from MS mainly accrues in the secondary progressive phase (Confavreux, 2003).

Clinically, MS is a highly variable condition and an individual's disease course is largely unpredictable. While some people have aggressive active disease and accumulate disability rapidly, termed malignant MS, others accrue little or no neurological disability over decades (Lublin, F, Reingold, 1996). Overall, natural history studies have identified that approximately 50% of patients will require assistance with walking within 15 years after disease onset, and approximately 70% will have secondary progression (Weinshenker *et al.*, 1989).

1.3.2 Prognostic factors

Clinical features

Thus far, the most consistent clinical and demographic features that have been associated with a favourable clinical outcome include: an early age at symptom onset, female sex, an initially RR course, optic neuritis or sensory relapse as the presenting CIS, complete remission after the first episode, and a longer internal between the first and subsequent relapse. Worse prognostic factors identified include male sex, older age of disease onset, multifocal onset, and high relapse rate within the first five years (Hawkins and Mcdonnell, 1999; Eriksson *et al.*, 2003; Confavreux and Vukusic, 2006; Degenhardt *et al.*, 2009; Dobson *et al.*, 2012; Scalfari A, Knappertz V, Cutter G, Goodin DS, Ashton R, 2013).

Radiological features

In terms of radiological prognostic markers, both intracranial lesions and atrophy are relevant. Following a CIS, the MRI features associated with an increased likelihood of converting to MS and subsequent development of disability include a higher initial brain lesion load, and a greater accrual of lesions over the initial five years (Fisniku *et al.*, 2008a). In the same study, lesion accrual over 20 years following a CIS was shown to be lower in people with benign compared with non-benign RRMS, though the relationship is a modest one. Other studies, however, have not demonstrated this, and

instead showed that MRI disease progression is similar between benign and non-benign patient groups (Thompson *et al*., 1990; Filippi *et al*., 1995). These studies, however, had small (<20) numbers per patient-group, and classified benign MS as those with minimal disability up to a disease duration of 10 years. As discussed in sections 1.3.1 and 1.3.3, people who have minimal disability at 10 years can develop significant disability later on in the disease course. Studies looking at cognitive impairment and MRI correlation have similarly shown that early accrual of brain lesion load following a CIS is associated with cognitive impairment seven years later (Summers *et al.*, 2008). It is possible that this is because WM lesions seen on MRI are pathologically non-specific, and likely to reflect a broad spectrum of pathological processes, including inflammation, oedema, demyelination and axonal loss (Filippi *et al*., 1999; Noseworthy *et al.*, 2000).

It is likely that overall lesion burden is dependent on quantity, lesion location, as well as lesion characteristics. Brainstem presentations, and corresponding lesions within the brainstem, appear to carry a greater risk of conversion to MS and subsequent disability (Swanton *et al.*, 2009; Tintore *et al.*, 2010b). It has been described that patients with less disability have lower infratentorial lesion volumes, compared to SPMS patients (Filippi *et al.*, 1996). Further down in the neuro-axis, there has been increasing interest in spinal cord lesion load in MS and its relationship with disability. In one study the presence of spinal cord lesions was associated with worse disability up to seven years following a CIS (Swanton *et al.*, 2009). In another study, which examined patients with disease duration of up to 10 years, patients with less

disability, compared to SPMS, had significantly less lesions in the cervical cord (Bonek *et al.*, 2007). A recent paper has demonstrated that presence of contrast enhancing brain lesions and spinal cord lesions correlate with the development of SPMS by 15 years following a CIS (Brownlee *et al.*, 2019). The association between presence of early gadolinium enhancing lesions and subsequent disability has also been demonstrated in other studies (Di Filippo *et al*., 2010).

Following the acute inflammatory phase at lesion onset, some MS lesions are described as chronic active, slowly expanding, and 'smouldering' (Frischer *et al*., 2015). These have been demonstrated, in both pathological and in vivo studies, to be associated with a more disabling disease course, with patients reaching higher EDSS and worse cognitive measures, or reaching earlier secondary progression (Luchetti *et al*., 2018; Absinta *et al*., 2019).

As mentioned previously, GM pathology, reflected by atrophy, is now recognised to be relevant to longer-term disability in MS. Following a CIS, early brain atrophy (in particular GM) appears to be more rapid in people at high risk of developing MS (Dalton *et al.*, 2004; Calabrese *et al.*, 2011, 2013; Pérez-Miralles *et al.*, 2013) and 20 years following onset, GM atrophy distinguishes people with a benign and non-benign outcome (Fisniku *et al.*, 2008b). In progressive MS, brain and spinal cord atrophy have been shown to correlate more closely than the number or volume of brain lesions with evolving disability (Stevenson *et al.*, 2000; Bermel *et al.*, 2003; Kearney *et*

al., 2014). In the spinal cord, by measuring the upper cervical cord cross-sectional area, patients with SPMS have been demonstrated to have significantly more atrophy than ambulatory patients (Filippi *et al.*, 1996). Measures of magnetization transfer ratio (MTR) in MRI, currently limited to research, is thought to be an estimate of tissue damage, being influenced by demyelination, axonal loss, and gliosis (Schmierer *et al.*, 2004). MTR values in a group of BMS patients (EDSS <3, at least 15 years disease duration), compared to a group of EDSS-matched RRMS patients, early in their disease course, were significantly less in both white matter and grey matter (De Stefano, 2006). This suggests that brain tissue damage is less extensive in BMS patients, and/ or that they have more effective compensatory mechanisms for such damage.

It has been suggested that the lack of disability in some patients may be due to the relative sparing of clinically eloquent regions in the brain (Rovaris *et al.*, 2009). In the same paper, it was also postulated that tissue damage within lesions may be less pronounced in such patients, due to the observation that previously gadolinium-enhancing lesions less frequently become black holes in people with less disability, compared to those with SPMS.

1.3.3 'Benign' Multiple Sclerosis

It is apparent that some patients with MS who, despite the passage of time, do not develop significant neurological impairment. This is a controversial

entity that is relatively understudied, and is termed ‘benign’ MS.

It is probable that previous natural history studies may have been biased towards those who were accessing healthcare more frequently, and therefore not capturing a subset of MS patients with relatively mild disease. The diagnostic threshold for MS has also changed over the years. With the evolution of the MS diagnostic criteria (Section 1.2.4), a diagnosis of MS is now more likely to be made, in someone with less prominent symptoms and signs. The study of those with favourable long-term outcomes may offer insights into MS pathogenesis.

Perhaps part of the controversy of BMS is that its definition and existence are often debated (Amato and Portaccio, 2012; Correale *et al.*, 2012b; Hawkins, 2012).The idea of BMS was first described in the literature by McAlpine, who defined a group of patients who were ‘without restriction of activity for normal employment and domestic purposes, but not necessarily symptom-free after a follow-up of more than 10 years’ (McAlpine, 1961). In 1996, an international survey conducted by the National MS Society defined BMS as ‘a disease in which the patient remains fully functional in all neurological systems, 15 years after disease onset’ (Lublin, F, Reingold, 1996). Unfortunately, the term ‘fully functional’ was not distinctly defined.

1.3.4 Current definitions and caveats

The controversy surrounding the existence of BMS may in part be due to the

lack of an agreed definition. Some definitions require 10 or more years since disease onset, while others require 15 years (Ramsaransing and De Keyser, 2006; Correale *et al.*, 2012b). Longitudinal studies have also shown that some people initially thought to have BMS go on to develop disability later, suggesting that it is a temporary status only, and that longer term follow-up is required to better estimate the true prevalence of BMS (Sayao *et al.*, 2007; Costelloe *et al.*, 2008; Hirst *et al.*, 2008). This would be logical given that the progressive phase of the disease typically occurring in the second decade after onset of symptoms (Confavreux and Vukusic, 2006; Scalfari A, Knappertz V, Cutter G, Goodin DS, Ashton R, 2013). In addition, nearly all existing definitions focus on a low level of physical disability as determined by the Expanded Disability Status Scale (EDSS) score, usually ≤2, ≤3, or ≤4 (Kurtzke, 1983). However, while the EDSS is the most widely used measure of disability in MS, it does have its shortfalls: it is heavily weighted towards motor disability, particularly ambulation, with non-motor features such as cognition, bladder function and fatigue, common issues in MS, only minimally considered. The scale ranges from zero to 10, with increments representing worse disability. Below a score of four, the assessment is based on neurological examination and can be subjective. Substantial impairments, for example in vision or arm function, make only limited differences to the overall scores. Scores of 5 and above are defined by impairment in mobility (Table 1.2)(Kurtzke, 1983).

In the past two decades, multiple studies have been carried out in an attempt to estimate the prevalence of BMS, with substantially differing results. The

exact proportion of people considered to have BMS depends on the definitions used, ranging between 3% and 64%, with an average of approximately 25%, 10 or more years after disease onset (Table 1.3)(Hawkins and Mcdonnell, 1999; Ramsaransing and De Keyser, 2006). This variation is most likely reflective of differences between cohorts, different definitions used, study design, follow-up duration, and changes in MS diagnostic criteria.

Table 1.2 Kurtzke Expanded Disability Status Scale (EDSS)(Kurtzke, 1983)

Score	Description
0	Normal neurological examination
1.0	No disability, minimal signs in one FS
1.5	No disability, minimal signs in more than one FS
2.0	Minimal disability in one FS
2.5	Minimal disability in two FS
3.0	Moderate disability in one FS, or mild disability in three or four FS; fully ambulatory
3.5	Moderate disability in one FS and more than mild disability in several others; fully ambulatory
4.0	Significant disability but self- sufficient; ambulatory without aid or rest ≥ 500 metres
4.5	Significant disability but up and about much of the day, able to work a full day; ambulatory without aid or rest ≥ 300 metres
5.0	Ambulatory without aid or rest ≥ 200 metres
5.5	Ambulatory without aid or rest ≥ 100 metres
6.0	Unilateral assistance to walk 100 metres with or without resting
6.5	Bilateral assistance to walk 20 metres without resting
7.0	Unable to walk beyond 5 metres with aid; essentially restricted to wheelchair; wheels self and transfers alone; up and about in wheelchair ~12 hours a day
7.5	Unable to take more than a few steps; restricted to wheelchair; may need aid in transferring and in wheeling self
8.0	Essentially restricted to bed or chair or pushed in a wheelchair; out of bed most of the day; retains many self-care functions; generally has effective use of arms
8.5	Essentially restricted to bed much of the day; has some effective use of arms, retains some self-care functions
9.0	Confined to bed; can communicate and eat
9.5	Confined to bed and fully dependent; unable to communicate effectively or eat/ swallow
10.0	Death due to MS

FS= functional system; these include: visual, brainstem, pyramidal, cerebellar, sensory, bowel and bladder, and cerebral.

Table 1.3 Definitions and frequencies of ‘benign’ multiple sclerosis

Year of publication	Country	Study type	MS population (n)	MS diagnostic criteria	BMS definition	BMS frequency (%)	Reference
1961	England	Natural history	241	Uncertain	Without restriction of activity for normal employment and domestic purposes, but not necessarily symptom-free, after at least 10 years	32	(McAlpine, 1961)
1977	USA	Natural history	234 (all male)	Schumacher	EDSS ≤2 after >10 years	20	(Kurtzke *et al.*, 1977)
1979	Scotland	Natural history	557	Uncertain	Without restriction of activity for normal employment and domestic purposes, but not necessarily symptom-free, after at least 10 years	7.2	(Shepherd, 1979)
1986	Ireland	Hospital based	400	Poser	EDSS ≤3 after >10 years	42	(Thompson AJ, Hutchinson M, Brazil J, Feighery C, 1986)
1987	Germany	Retrospective	363	Poser	EDSS ≤2 after >10 years	19	(Lauer *et al.*, 1987)
1999	Northern Ireland	Prevalence	259	Poser	EDSS ≤3 after >10 years	20	(McDonnell *et al.*,1999)

1999	Italy	Natural history	224	Poser	EDSS <3 after >10 years	29	(Amato *et al.*, 1999)
2001	Martinique	Population based	62	Poser	EDSS ≤3 after ≥10 years	19.4	(Cabre *et al.*, 2001)
2001	Italy	Hospital based	500	Uncertain	EDSS ≤3 after ≥15 years, with normal neuropsychological examination	6	(Perini *et al.*, 2001)
2002	Iceland	Population based	372	Poser	EDSS ≤3.5 after ≥15 years	64	(Benedikz *et al.*, 2002)
2003	Iran	Hospital based	200	Poser	EDSS ≤3 after ≥10 years, or EDSS ≤2 after ≥5 years	14	(Kalanie *et al.*, 2003)
2010	Crete	Natural history	587	McDonald 2005	EDSS ≤3 after ≥20 years	30	(Mastorodemos *et al.*, 2010)
2016	USA	Clinics based registry	6258	Poser	EDSS ≤2 after ≥10 years EDSS ≤3 after ≥15 years EDSS ≤3 after ≥10 years	19.8* 20* 33.3*	(Zivadinov *et al.*, 2016)
2018	Wales	Population based registry	1049	Poser	EDSS ≤3 after ≥15 years, with no significant fatigue, mood disturbance, cognitive impairment or disrupted employment, and no DMT treatment	3	(Tallantyre *et al.*, 2018)

*38-40% of patients in these groups were treated with a DMT

1.4 Rationale for this study

In the last two decades, there has been rapid growth in the area of MS disease modifying therapies (DMTs). From the approval of the first DMTs in the 1990's, there is now a total of 13 licensed DMTs for RRMS in the UK, with likely more in the future. The management of MS has evolved significantly with these welcomed developments. The treatments primarily aim to reduce inflammatory relapses and the development of new MRI lesions in the RR phase of MS. They vary in their efficacy, route and frequency of administration, mechanism of action, likelihood and severity of adverse effects, and monitoring requirements. First-line, lower efficacy, DMTs, such as the interferons and glatiramer acetate, have good long-term safety profiles, and potential adverse effects are mostly reversible. However, as injections they can be associated with intolerable side-effects, and in one large North American study, between 20% and 40% of people stopped after one year of treatment due to concerns regarding efficacy, tolerability or safety (Fox *et al.*, 2013). At the other end of the DMT spectrum, the monoclonal antibodies, while offer higher efficacy, can be associated with significant adverse effects including secondary autoimmunity and opportunistic infections. In addition, their long-term safety profiles are not yet known. Some of the newer DMTs, such as alemtuzumab and cladribine, now aim to induce sustained disease remission (Fox and Rhoades, 2012; Edan and Le Page, 2013). However, treatment with such induction agents may carry more risks, which are potentially irreversible. For example, treatment with alemtuzumab is associated with a 20-30% risk of autoimmunity and a

~3% risk of immune thrombocytopenia (Coles, 2013). MS is also a life-long condition and treatment is likely to be long-term. Continuous treatment therefore potentially exposes patients to varying degree of risks. These risks need to be set in the context of likely benefit, particularly as a substantial proportion of people who have a CIS will never develop MS (Fisniku *et al.*, 2008a) and those who do may accrue little or no disability in the long term (Hawkins and Mcdonnell, 1999; Ramsaransing and De Keyser, 2006).

There is increasing interest in very early treatment of people following a CIS, or immediately after a diagnosis of MS was made, in the hope that this will prevent or reduce long-term disability (Comi *et al.*, 2001, 2009; Kappos *et al.*, 2006; Freedman *et al.*, 2014). While this is undoubtedly necessary in those with aggressive disease, early identification of those likely to remain with a benign disease course in the long-term, and therefore less to gain from DMTs, would be useful. The approach to treatment should, as far as possible, involve a well-informed risk-benefit analysis. Ideally, this analysis is personalized and as a minimum should seek to identify those who are likely to do well without treatment, and those who have the most to gain from it. Currently BMS is only defined retrospectively, when the window of opportunity to treat early, or not, has passed. Data on prediction of long-term outcomes in MS remains lacking. Using a radiological-clinical CIS cohort with the longest known follow-up duration, this study will aim to answer the questions proposed in section 1.1.

Chapter 2: The London CIS cohort and study design

2.1 Introduction

To address the aims of this thesis, detailed demographic, clinical and radiological information, from early on after a CIS, is required. Starting from after a CIS rather than a diagnosis of clinically definite MS is important for several reasons: firstly, clinical conversion to MS, based on a second relapse, may occur a decade or more after a CIS (Fisniku *et al.*, 2008a). Secondly, some people may develop SPMS without having a secondary clinical relapse. Thirdly, since the introduction of the original McDonald diagnostic criteria in 2001, there have been three revisions, in 2005, 2010, and most recently, 2017 (McDonald *et al.*, 2001; Polman *et al.*, 2005, 2011; Thompson *et al.*, 2018a). These have been described in more details in section 1.2.4. With these revisions, a diagnosis of MS can be reached earlier and more frequently following a CIS, with increased sensitivity and specificity (Dalton *et al.*, 2002b; Swanton *et al.*, 2007; Montalban *et al.*, 2010; Brownlee *et al.*, 2015). For example, a diagnosis of MS can be made in those who have had a CIS and who fulfil the MRI and CSF criteria for RRMS. Thus, studies that only follow-up individuals with an established diagnosis of MS, made at different earlier time points, is likely to significantly underestimate the true risk of conversion to MS following a CIS, and particularly may exclude a group of people with relatively inactive disease. Lastly in clinical practice, treatment decisions are being made increasingly early, sometimes following a CIS, and as such prognostic factors identified relative to a CIS

may be more useful than those identified only following a clinical diagnosis of MS, i.e. following a further relapse or the onset of progression.

2.2 Initial recruitment

The London CIS cohort was first recruited prospectively between May 1984 and July 1987 (Miller *et al.*, 1988, 1989). One hundred and forty people with CIS were recruited at the National Hospital for Neurology and Neurosurgery, Queen Square, London, and Moorfields Eye Hospital. All individuals were assessed and diagnosed by a neurologist, with the most likely cause of their symptoms being a CIS. Eight individuals were subsequently excluded as they were found to have alternative diagnoses: pontine arterio-venous malformation, pontine haematoma, Leigh's disease, myasthenia gravis, HIV-related CNS complications, systemic lupus erythematosus and two individuals had had strokes. CIS were classified as being an optic neuritis, transverse myelitis, or brainstem syndrome, based on clinical features. Patient above 50 years of age at baseline were excluded, due to the more frequent occurrence of WM lesions of vascular aetiology (Miller *et al.*, 1989). Table 2.1 shows the baseline cohort demographic and clinical features.

The demographic profile was representative of people presenting with CIS: the mean age was 32 years and approximately 2/3 female. There was a slight over-representation of people presenting with optic neuritis: 52% compared to 35-40% in other large prospective CIS cohort series (Eriksson

et al., 2003; Tintore *et al.*, 2015)

Table 2.1 Baseline demographic and clinical features of the London CIS cohort, at recruitment

Participants assessed (n)		132
Mean age at presentation (years)		32*
Female (n, %)		80 (61)
Male (n, %)		52 (39)
CIS type	Optic neuritis (n, %)	69 (52)
	Spinal cord syndrome (n, %)	36 (27)
	Brainstem syndrome (n, %)	27 (20)

*Based on 127 participants; CIS onset date was not available in five individuals, all of whom were subsequently lost to follow-up

2.3 Earlier follow-up time-points

Since their recruitment at initial presentation, the cohort has been followed up on five occasions – at approximately 1 year, 5 years, 10 years, 14 years, and at 20 years following first symptom onset. This study is the updated 30-year follow-up of the cohort. Clinical information and MRI scans were gathered at each time-point, except at one year, when only MRI scans were performed. MRI scans and clinical assessments were all performed in a single centre, at the UCL Institute of Neurology. Studies on this cohort have so far provided major insights into the risk factors for conversion to MS following a CIS (Miller *et al.*, 1988, 1989, 2002), and have directly contributed to the McDonald MS diagnostic criteria (McDonald *et al.*, 2001; Polman *et al.*, 2005, 2011). Table 2.2 shows the number of participants assessed or with a known outcome at each time point, and Table 2.3 shows more detailed demographic and outcome data on participants with known outcomes.

Table 2.2 Number of participants assessed or with a known outcome at each time-point

		Follow-up time-point (years)					
		0	**1**	**5**	**10**	**14**	**20**
Available data	Clinical only	0	0	3	14	13	27
	MRI only	0	108	0	0	0	0
	Clinical & MRI	132	0	91	66	55	77
	Alive but otherwise unknown	0	0	0	0	1	1
	Deceased	0	0	3	5	7	10
	Unknown	0	24	35	47	56	18

Table 2.3 Demographic and outcome data on participants with known outcomes, at each earlier time-point

		Follow-up time-point (years)					
		0	**1**	**5**	**10**	**14**	**20**
Participants assessed (n)		132	108	94	80	68	104
Of those assessed	Mean age at presentation (years)	32*	32	31	32	32	32
	Female (n,%)	80 (61)	67 (62)	53 (56)	53 (66)	47 (69)	69 (66)
	Optic neuritis (n, %)	69 (52)	52 (48)	46 (49)	43 (54)	35 (51)	53 (51)
	Transverse myelitis (n, %)	36 (27)	31 (29)	27 (29)	25 (31)	23 (34)	29 (28)
	Brainstem syndrome (n, %)	27 (20)	25 (23)	21 (22)	12 (15)	10 (15)	22 (21)
Deaths (n)				3	5	7	10
Of those deceased	Death related to MS (n)				1	3	3
	CIS, death unrelated to MS			3	4	4	7
Of those known to be alive	CIS (n)	132		44	27	17	30
	RRMS (MRI, n)	0		3	2	2	6
	RRMS (clinical, n)	0		44	40	37	43
	SPMS (n)	0		3	11	12	25
Total EDSS assessment (n)		118#	NA	94	80	68	104
	Telephone EDSS assessment (n)	0	0	0	0	11	27

*based on 127 patients; CIS onset date was not available in five individuals, all of whom were subsequently lost to follow-up

#determined retrospectively

2.4 Study design

2.4.1 Background

Based on data from natural history studies, the median time to death is around 30 years from disease onset (Compston and Coles, 2008). A recent review of survival data from people with MS also reports that the mean disease duration at death was 35 years (Scalfari A, Knappertz V, Cutter G, Goodin DS, Ashton R, 2013). In a population-based study, which matched 5,797 individuals with MS with the general population, survival in those with MS had a median of 7 years lower, matched for age, sex, and socioeconomic status: median survival in the MS population was 75.9 years, versus 83.4 years in the matched population (Marrie *et al*., 2015).

Natural history studies have demonstrated that in the majority of people with RRMS, it took over a decade for their mobility to become limited, and over two decades before they were immobile without aids (Confavreux, 2003). Given this, assessment of the relationships between early prognostic features and later outcomes ideally requires clinical follow-up of two decades or more. These indicated that extending the follow-up duration to 30 years was essential if trying to identify those who have a genuine benign outcome. As MRI first became available in the 1980's, this cohort provides a uniquely long-term view on the relationship between early radiological and clinical features, their evolution over time, and their ultimate prognostic value,

following a CIS.

2.4.2 Recruitment

Members of the original cohort were identified, contacted, and invited to take part. Contact details, where available from previous time-point(s), were used. Where contact details were unavailable or out of date, the National Health Service (NHS) tracing service was used. For those who have deceased, where possible death certificates were obtained via the UK General Register Office.

The study has been approved by the UCL institutional ethics committee and the National Research Ethics committee (15/LO/0650). At 30 years, all participants gave informed consent, written if they attended in person, or verbal if they provided clinical information via telephone only.

2.4.3 Clinical assessment

As the inception of the cohort pre-dated the development of the EDSS, baseline EDSS scores were not recorded prospectively and so were determined retrospectively by review of available notes (103 participants) or participant recall (15 participants). Baseline EDSS scores could not be determined in 14 individuals due to absence of notes and unclear recall. Post-CIS ‘nadir’ EDSS scores, either at the nadir where clinical improvement had plateaued, or at one year, whichever was earlier, was also determined in

114 subjects. EDSS data from adjacent time-points, and from clinical records, where available, were used. At five and ten years, EDSS scores were determined by clinical examination. At 14, 20 and 30 years, participants who attended our unit were interviewed face-to-face and clinically assessed in person. In those who were unable to attend in person, a telephone interview was performed. EDSS scores were obtained via clinical assessment, and in those who were unable to attend, via telephone interview using a validated questionnaire (Lechner-Scott *et al.*, 2003). At 30 years, detailed neurological history, including original CIS presentation, any new or progressive symptoms since their last review, any neurological investigations or treatments, details of other medical conditions, particularly vascular risk factors, and regular medications, were collated from all participants.

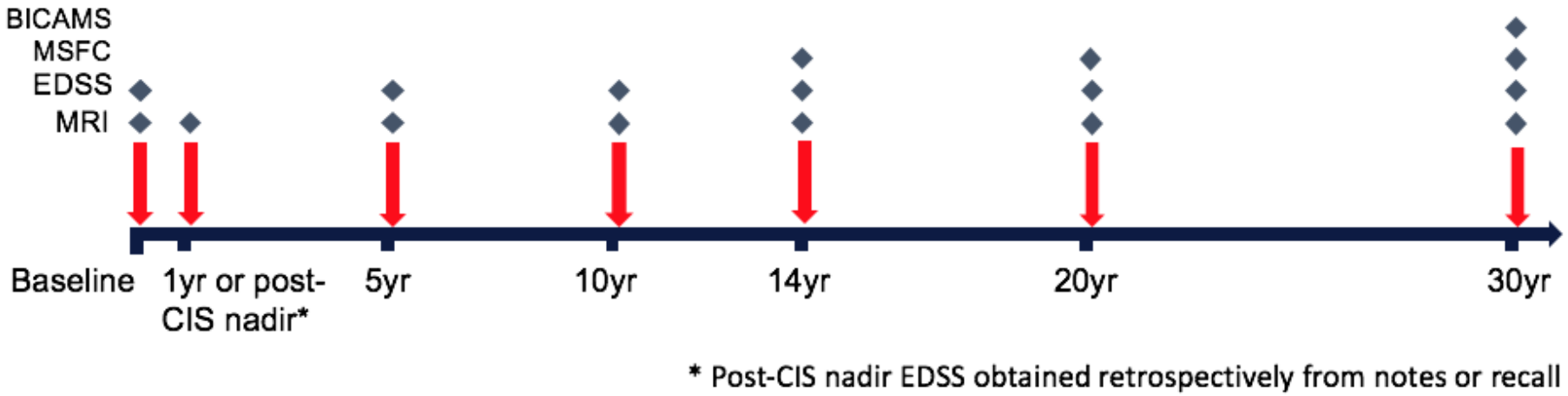
From 14 years onwards, Multiple Sclerosis Functional Composite (MSFC) scores, which includes timed 25-foot walk, 9-hole peg test and paced auditory serial addition test (PASAT) scores (Fischer *et al.*, 1999), were assessed in those who attended the UCL Institute of Neurology for review. As highlighted in chapter 1.3, cognitive impairment is poorly captured by EDSS. However, it is increasingly recognized as a symptom in MS, and is estimated to effect 40-65% of patients (Amato *et al.*, 2006; Correale *et al.*, 2012a). Thus at 30-year follow-up, the brief international cognitive assessment for MS (BICAMS) was additionally performed. This consists of three components: the symbol digit modalities test (SDMT) for information processing speed, California verbal learning test (CVLT) for verbal memory, and the brief visuospatial memory test (BVMTR) for visual memory (Langdon

et al., 2012). Employment status data were collected as a marker for assessing the social impact MS has on an individual (Glad *et al.*, 2011). The National Adult Reading Test (NART) was used to estimate pre-morbid intelligence (Nelson, 1982). In addition, the Hospital Anxiety and Depression Scale was used to measure mood (Zigmond and Snaith, 1983), and fatigue was measured using the fatigue severity scale (Krupp *et al.*, 1989) and the Cognitive leisure activities questionnaire (Sumowski *et al.*, 2010).

Figure 2.1 gives an overview of the data obtained over the 30-year follow-up time span.

Assessment at all time-points were conducted at a single university site.

Figure 2.1 Overview of clinical and radiological data obtained at each time-point



2.4.4 MRI acquisition

At baseline, 1 year, and 5 years, MRI scans were obtained using a Picker 0.5- Tesla system. At 10, 14 and 20 years, a 1.5-Tesla General Electric Signa scanner, and at 30 years, a 3-Tesla Philips Achieva. Slice thickness at baseline was either 10mm for the earlier scans, or 5mm for the later scans.

All scans obtained were brain scans; spinal cord images were not consistently undertaken in this cohort. At baseline, 1 and 5 years, PD-weighted scans were obtained. T2-weight scans were obtained from 10 years onwards. T1-weighted and FLAIR sequences were additionally obtained at 14 years; at 20 years, FLAIR sequences were not obtained. At 30 years, in addition to more standard sequences including PD/T2-weighted, T1-weighted and FLAIR, two novel methods were also applied: phase sensitive inversion recovery (PSIR) and high-resolution brain MTR. The former method aims to measure GM lesion loads, and the latter, to assess GM and WM demyelination (Calabrese *et al.*, 2007; Samson *et al.*, 2014).

Table 2.4 shows details of MRI acquisition and sequences at each follow-up time-point. Figure 2.2 shows representative axial MRI scans from each time-point, demonstrating the changing quality of images over time.

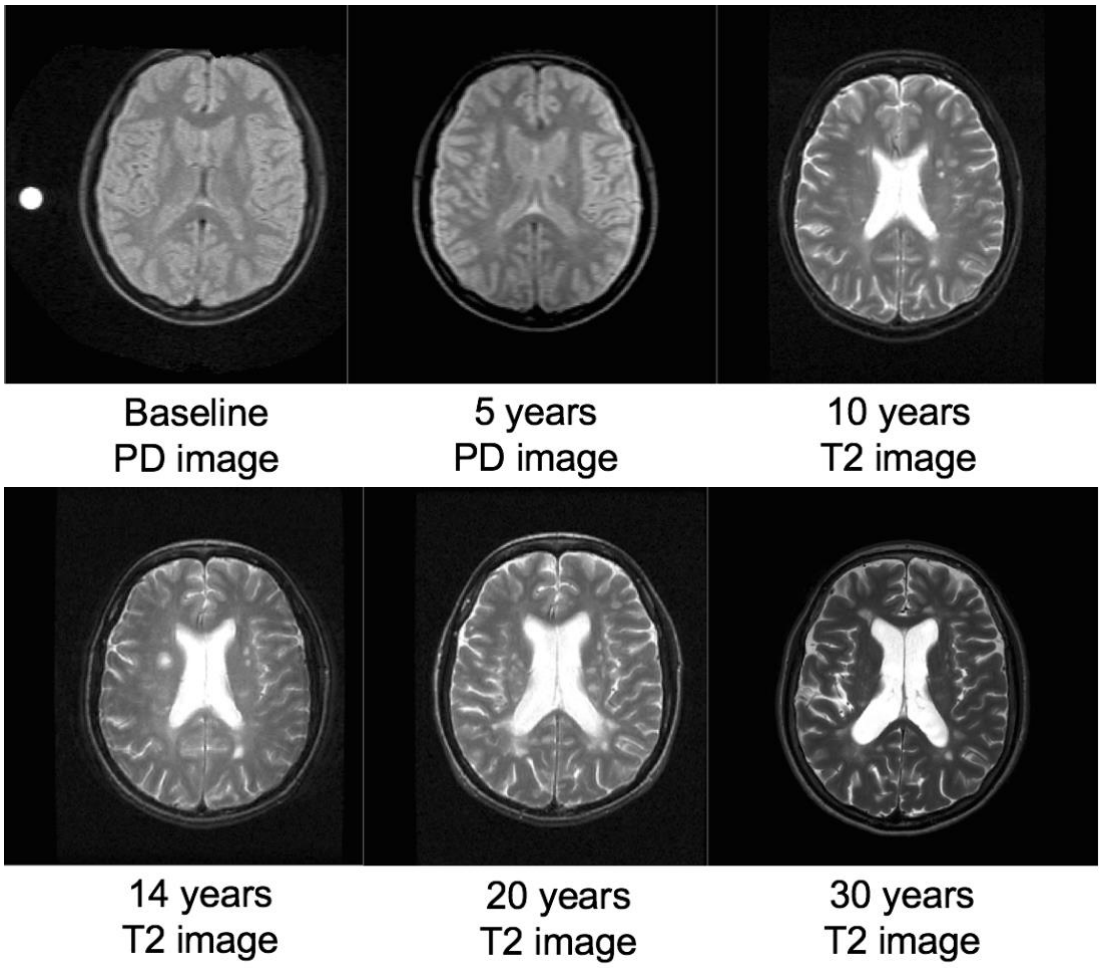
All participants were asked to complete our research unit's MRI safety checklist. Those who were unable or unwilling to have MRI were still able to participate in the study, by providing clinical information and being

assessment, without an MRI scan.

Table 2.4: MRI acquisition details and sequences obtained at each time point

		Follow-up time-point (years)						
		0	1	5	10	14	20	30
Scanner	Manufacturer and field strength	Picker 0.5T	Picker 0.5T	Picker 0.5T	GE Sigma 1.5T	GE Sigma 1.5T	GE Sigma 1.5T	Philips Achieva 3T
PD/T2-weighted	Slice thickness (mm)	5 or 10	5	5	5	5	5	3
	In plane resolution (mm x mm)	1.2 x 1.2	1.2 X 1.2	1.2 x 1.2	1.0 x 1.0	1.0 x 1.0	1.0 x 1.0	0.5 x 0.5
	TE (ms)	60	60	60	30/90	14/98	17/102	85
	TR (ms)	2000	2000	2000	2000	2000	2000	4375
Acquired sequences		PD	PD	PD	PD/T2	PD/T2/T1 FLAIR	PD/T2/T1	T2/T1 FLAIR PSIR MTR

Figure 2.2: Representative axial brain images from one participant, acquired at baseline, 5, 10, 14, 20 and at 30 years



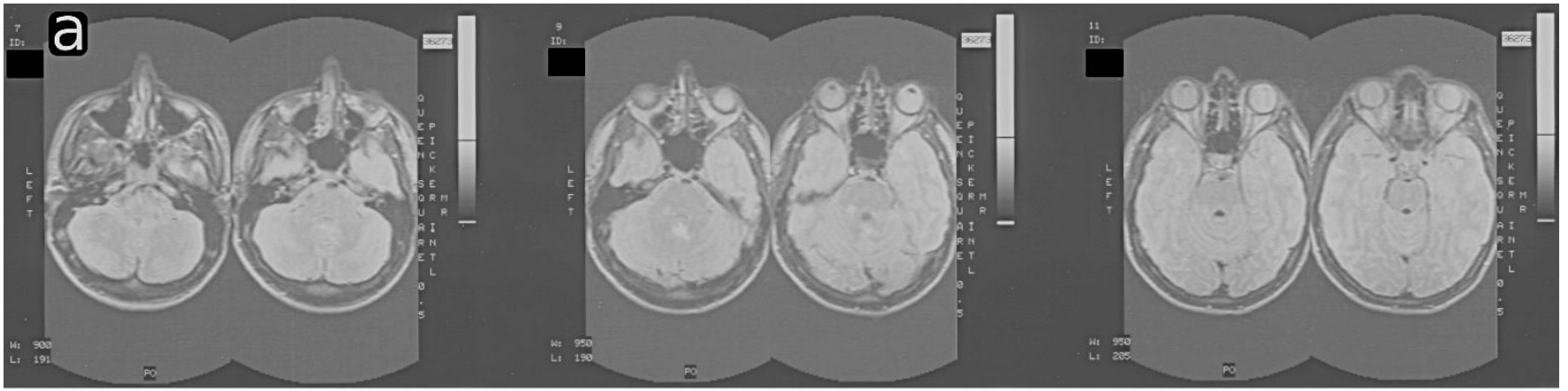
2.4.5 Clinical outcomes

All available notes, research records and MRI scans from previous time-points were reviewed. The 2010 McDonald criteria was applied prospectively and retrospectively, on scans and clinical information from earlier time points, to determine if an individual had MS or not (Polman *et al.*, 2011). Participants were classified as either remaining as CIS, or having MS. In those with MS, they were further sub-classified as having this either on clinical grounds, where an individual had a further relapse or clinical progression, or on radiological grounds, where new demyelinating lesions were seen on MRI, in the absence of further clinical symptoms. They were also sub-classified as RRMS or SPMS, with the latter defined by those who had had gradual disability worsening and progression that was irreversible (Lublin *et al*., 2014). In those who have deceased, the cause of death was determined as related or unrelated to MS, by consensus review of death certificates and notes, where available, by myself and Dr. Declan Chard. Death due to MS, which also equates to an EDSS score of 10 (see Table 1.2), was determined when MS was either given as the cause of death, or a clear contributing factor. For example, aspiration pneumonia in someone with advanced MS, or a pulmonary embolus from a deep vein thrombosis, secondary to chronic immobility.

2.5 MRI analysis

Only a subset of the MRI scans acquired at baseline, 1, 5 and 10 years were digitally archived, and were otherwise available as hard copies on films, where axial images were printed sequentially on a single, or multiple film(s) (Figure 2.3, Table 2.5). Film prints from earlier time-points were re-digitised using Vidar Diagnostic Pro Advantage film digitiser and processed with the Hipax Diagnostic Workstation medical image viewer software, and subsequently exported as DICOM files. In order to reconcile and organize all previous MRI data in a format that allowed us to undertake longitudinal analyses, these digitised scans were then processed to reconstruct a digital image stack, comparable to native stacks, available at later time-points (Figure 2.4) (Ebner *et al.*, 2018). The reconstructions would allow more advanced image analyses to be applied to the older images, where this was previously not possible. Table 2.5 shows the number of scans and formats available for the analyses undertaken.

Figure 2.3: Example of earlier scans, where axial images were printed sequentially on a single, or multiple X-ray film(s)



The example given here shows axial PD images with 5mm slice thickness, displayed sequentially side-by-side

Figure 2.4: Overview of volumetric reconstruction of older scans, printed on films (Ebner *et al.*, 2018)

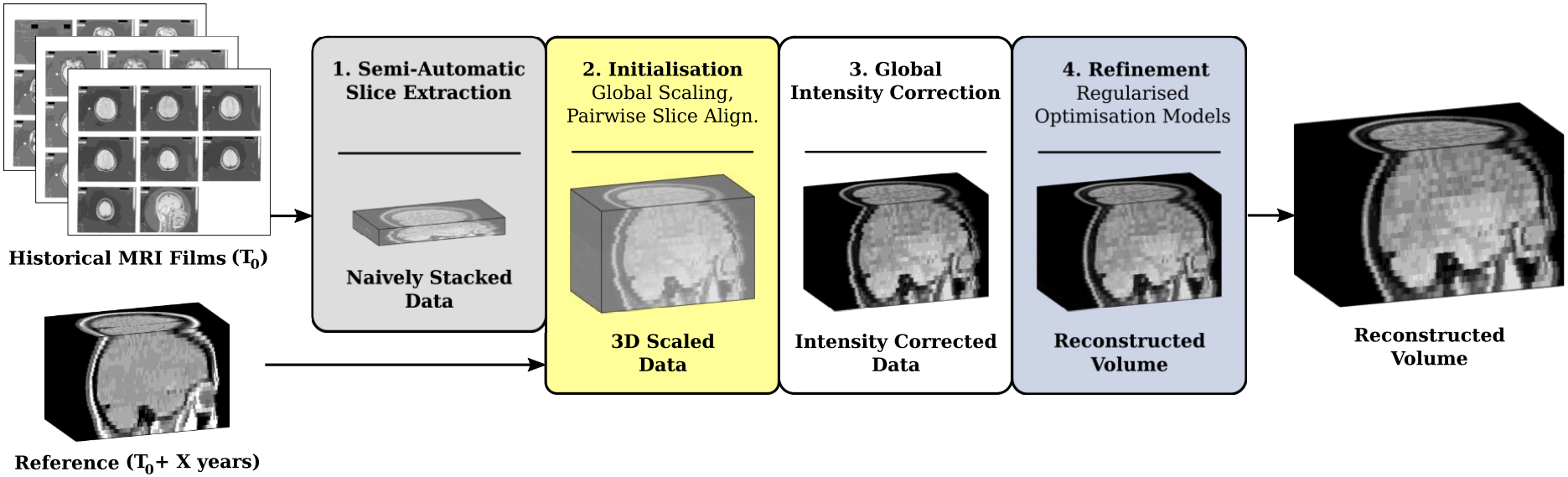


Table 2.5: Number of scans and formats available at each follow-up time-point

		Follow-up time-point (years)						
		0	**1**	**5**	**10**	**14**	**20**	**30**
Total MRI brain scans performed (n)		132	108	91	66	55	77	63
MRI scan availability and format	Original digital (n)	42	0	48	63	55	75	63
	Film prints (n)	61	95	38	3	0	0	0
	Reconstructed (n)	39	71	31	2	0	0	0
	Missing (n)	29*	13+	5^	0	0	0	0
	Total available for WM lesion count (n)	103	95	86	66	55	75	63
	Total available for TVW atrophy measures (n)	88	86	71	62	48	71	63

WM= white matter; TVW= third ventricular width

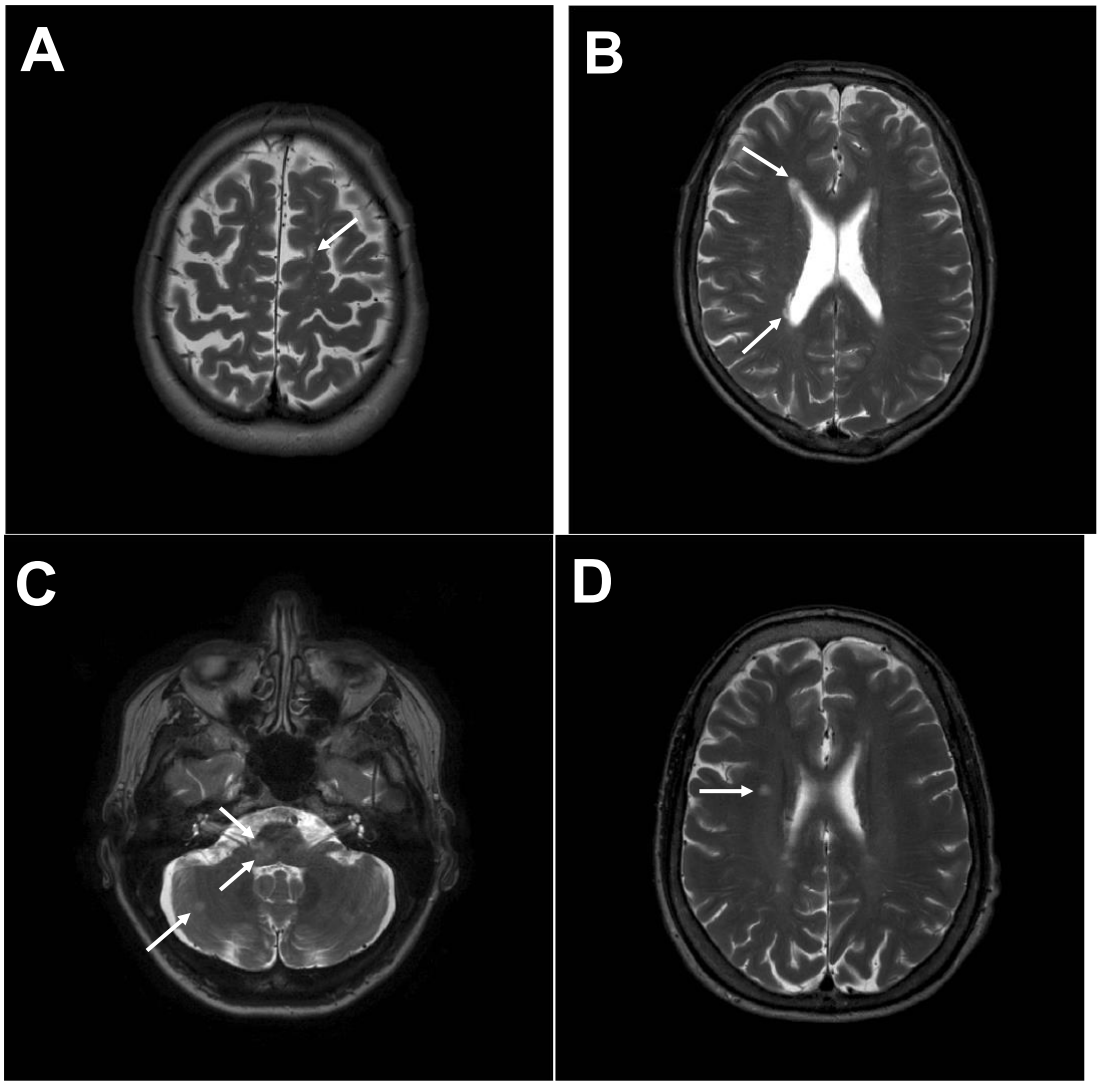
*historical lesion count data available in 16 participants; lesion locations not known; +historical lesion count data available in 9 participants; lesion locations not known; ^historical lesion count available in all 5 participants; lesion locations not known

2.5.1 White matter lesion number and location

For each participant, all available scans were reviewed side by side, using 3D Slicer version 4.4 (3D Slicer. https://www.slicer.org). All scans were reviewed by myself with Professor Frederik Barkhof, or Dr Declan Chard, or both. Both Professor Barkhof and Dr Chard were blinded to the subjects' clinical details. PD- or T2-hyperintense WM lesions were marked by consensus, with reference to preceding or subsequent scans, where available. Marked lesions were then manually counted by me. Other than total whole brain lesion count, lesions were also quantified based on locations defined in the MS diagnostic criteria: juxta-cortical (JC), periventricular (PV), infra-tentorial (IT), and deep white matter (DWM). DWM lesions were defined as supratentorial lesions that were neither juxta-cortical nor periventricular. Figure 2.5 demonstrates examples of WM lesions in the four anatomical locations.

Results relating to white matter lesion count and location are presented in Chapters 3 and 4.

Figure 2.5 Examples of white matter lesions in the four anatomical locations



White matter lesions demonstrated (indicated by arrows) on T2-weighted images, at (A) juxta-cortical, (B) periventricular, (C) infra-tentorial, and (D) deep white matter locations.

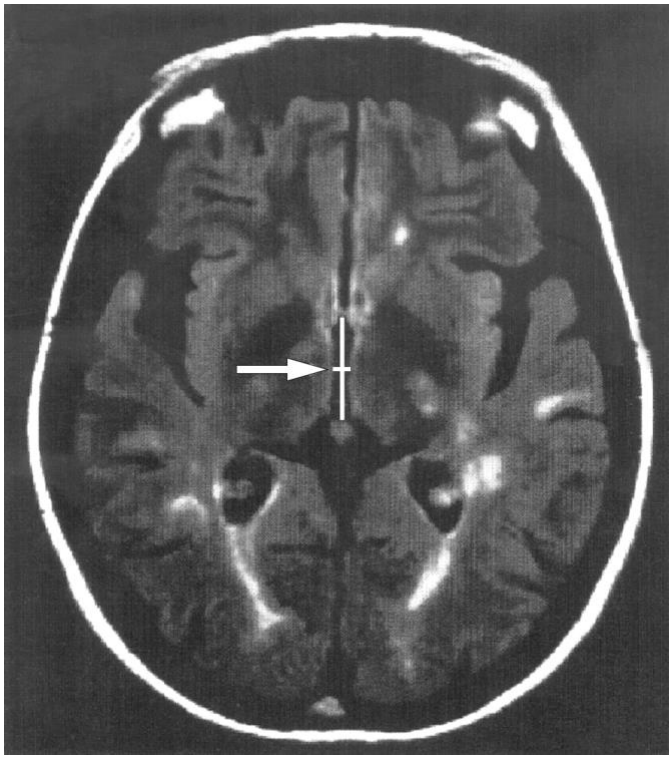
2.5.2 Brain atrophy

As the earlier images were two dimensional only, third ventricular width (TVW), a linear method for atrophy measurement, was used (Turner *et al.*, 2001; Sánchez *et al.*, 2008). All TVW measurements were performed on axially acquired images. At baseline, 1 year, and 5 years, PD images were used; at 10, 14 and 20 years, measurements were made on both PD and T2-weighted images; at 30 years, on T2-weighted images (Tables 2.4 and 2.5).

The TVW was measured as the width of the third ventricle at the midpoint of a line, running parallel to the long axis of the ventricle (Figure 2.6) (Benedict *et al.*, 2004). All available images were assessed for TVW by myself (KC) independently on two occasions, by Dr Lukas Haider and by Miss Giselle Birch. Dr Declan Chard (DC) also measured TVW in a randomly selected 10% of subjects.

Dr Declan Chard and I used Jim (version 6.0) (Xinapse Systems, Northants, United Kingdom. http://www.xinapse.com) to view the images, and Dr Haider and Miss Birch used FIJI (ImageJ) (Schindelin *et al.*, 2013). Pearson correlation coefficients were used to calculate intra- (KC/ KC) ($r= 0.930$, $p <0.001$) and intra-rater (KC/ DC) variabilities ($r= 0.826$, $p< 0.001$).

Figure 2.6 Method of third ventricular width measurement



The third ventricular width was measured at the midpoint of a line, running parallel to the long axis of the ventricle (Benedict *et al.*, 2004)

2.6 Statistical analysis

In Chapter 3, most of the results presented were descriptive statistics, for the purposes of more simple comparisons. Mean and age-adjusted comparisons of cognitive outcomes between groups at 30-years, presented in section 3.4.2, were performed using multiple regression of the cognitive measure on group indicators with age as covariate. For the comparison of CIS versus benign MS, presented in section 3.4.4, the two sub-groups were analysed using the unpaired t-test.

To test the association of TVW changes over the course of 30 years, a mixed effect model was used in R studio (RStudio, 2016). Logistic regression models were used to analyse the relationships between early atrophy measures, and the binary outcomes of 30-year EDSS ≤ 3.5 vs >3.5 (including deaths due to MS, i.e. EDSS= 10, by 30 years); this EDSS cut-off was chosen a priori as more clinically meaningful and objective than the >3.0 vs ≤3.0 threshold. Analysis of variance (ANOVA) model was used to analyse TVW measurements at given time-points, against the 30-year classifications of CIS, MS with EDSS ≤3.5, MS with EDSS >3.5, and EDSS 10. Changes in TVW measurements between time-points were calculated by subtracting the more recent value from an earlier time-point.

For the predictive models presented in Chapter 4, univariable and multivariable logistic regression analyses were used to identify early (baseline, one year and five year) predictors of the following three

binary 30-year outcomes: i) 30-year EDSS ≤ 3.5 versus 30-year EDSS > 3.5 (including deaths due to MS); as already mentioned, this EDSS cut-off was chosen a priori as it is more clinically meaningful and objective than the >3.0 vs ≤3.0 threshold; ii) SPMS diagnosis by 30 years, including SPMS deaths, versus CIS and RRMS at 30 years; iii) death due to MS by 30 years versus all participants who were still alive at 30yrs. Early prediction models were fitted from the perspective of earlier time-points, when future or final outcomes and diagnostic classifications were unknown, and therefore, unless otherwise stated, include all available subjects, including those who remained classified as CIS.

Independent variables analysed include: CIS type, age at onset, disease duration, gender, early EDSS scores and interval EDSS changes between follow-up time-points, number of relapses within the first five years, and early total and location-specific lesion counts.

For MS-associated death, a Cox proportional hazards model was used to identify best predictors. All deceased participants, regardless of MS status and cause of death, contributed to the Cox survival analysis, being censored at the times of death. Individuals whose deaths were unrelated to MS were not included in the models for 30-year outcomes.

Lesion count predictors were categorised or dichotomised either to generate approximately equal frequencies, or because of a priori importance (e.g. 1+ vs 0 lesions); resulting categories and thresholds are reported where

relevant. When the ordinal lesion variables did not predict materially better than binary, models with binary lesion predictors were reported. For multivariable logistic and Cox models, manual backwards stepwise elimination of variables with $p > 0.05$ was used to identify the best subset of independent predictors, and discarded variables were re-introduced singly in the final model to confirm their lack of contribution.

Accuracy, positive predictive value (PPV), negative predictive value (PPV), sensitivity, and specificity, were calculated, for individual binary predictors, directly from the relevant 2x2 classification tables. For multivariable logistic models these were calculated from the model predicted probability, using a 0.5 probability classification cut-off. Model predicted probabilities are also reported for the various combinations of predictor values.

Unless otherwise stated, analyses were performed in SPSS version 24.0 (IBM Corp) and Stata 15.1 (StataCorp, 2017), and statistical significance is reported at $p < 0.05$.

Chapter 3: Clinical outcomes thirty years following a clinically isolated syndrome

3.1.1 Cohort overview

At 30 years, nine of the 132 initial participants were not traceable. Three people were alive but declined to participate. Outcome data, including deaths, were obtained in 120 out of the original 132 participants (91%). Twenty-nine individuals were deceased. 66 individuals attended our department for clinical assessment, and 63 underwent MRI scanning; MR imaging was contraindicated in the other three individuals. BICAMS scores were obtained in 61 participants, 41 with MS and 20 with CIS. EDSS assessment via telephone interviews were carried out in 25 individuals. The mean follow-up duration was 30.9 years.

In four individuals with 30-year outcomes, data on their baseline brain scans, including historical lesion counts, were unavailable. Table 3.1 summarises the data obtained at 30 years.

Table 3.1 Summary of MRI and EDSS data obtained at 30 years

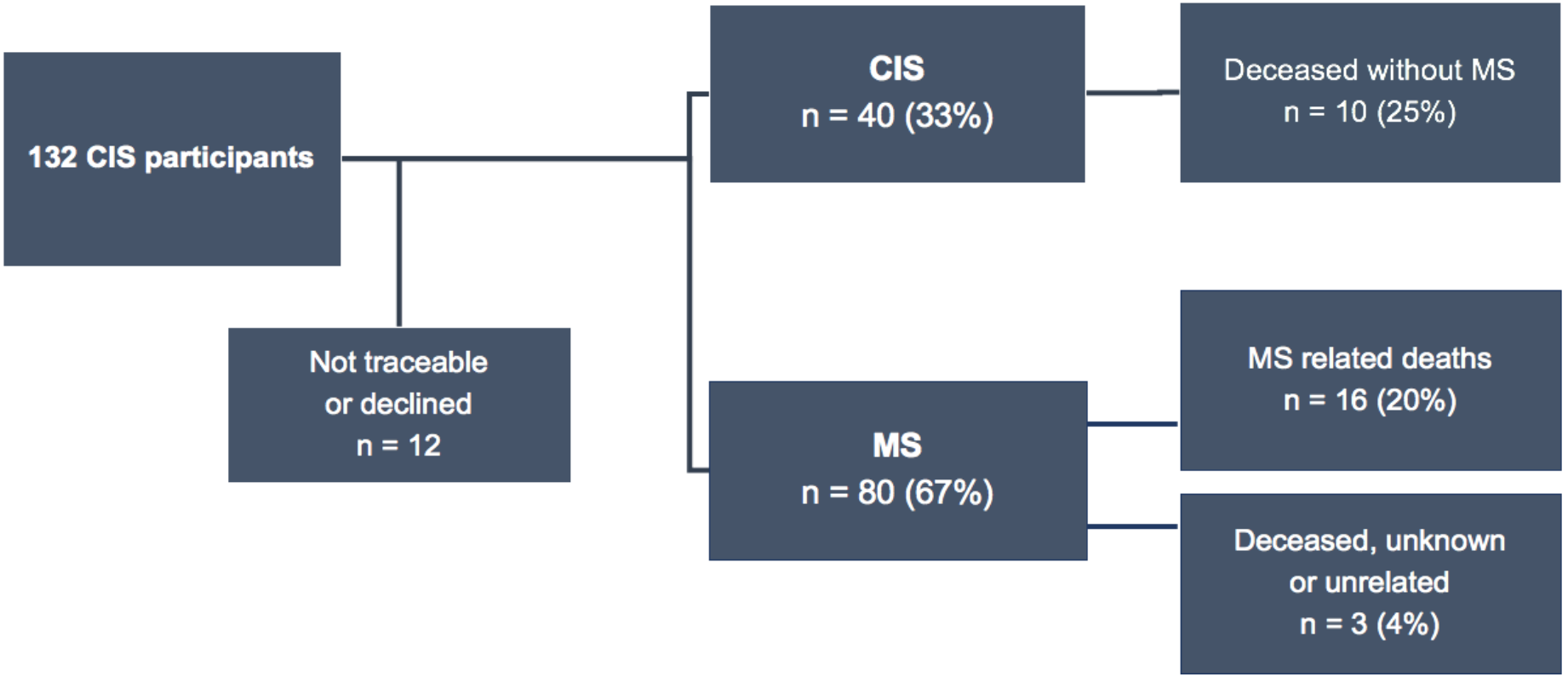
		Number of participants (n)
MRI brain scans		63
Total EDSS assessment excluding deaths related to MS		91
Clinical EDSS assessment		66
Telephone EDSS assessment		25
MS-related deaths; EDSS= 10		16
Deceased, not MS-related	Deceased with CIS	10
	Deceased with MS (not due to MS or unknown)	3
Unknown outcome		12

Of the 91 individuals who were alive, 30 remained CIS with no further relapse or progression since their initial presenting CIS. Of these 30 individuals, 21 had a brain MRI scan and nine had telephone or clinical assessments, without scans.

3.1.2 Deceased participants

In total, of the 120 participants with known outcomes, 29 individuals were deceased. Nineteen were known to have MS, and the other ten passed away with last known classification as CIS. Of these ten participants, three were last assessed at 20-years, one at ten years, two at five years, two at one year, and two at baseline. 27 out of 29 death certificates were obtained. In the 19 known to have MS, MS led to and/ or contributed to death in 16, all of whom had developed SPMS at the time of death. Two individuals died of unrelated conditions, ovarian carcinoma and stroke, and the cause of death was undetermined in one individual (Figure 3.1). All three were assessed and documented to have RRMS at 20 years, with EDSS scores of 2.5, 3.0 and 6.0. Thirty-year outcomes were not allocated to these three individuals. In total 30-year EDSS and clinical outcomes were available in 107 of the original cohort.

Figure 3.1 Clinical outcomes of 120 participants, including deaths



3.1.3 Comparison to baseline cohort

Baseline clinical and demographic characteristics of these 107 individuals were similar to the original baseline cohort (Table 3.2). In those alive at 30 years, the mean (SD) age was 61.6 (7.4) years, with 59 (65%) female and 32 (35%) male. Clinical and demographic characteristics of those who did not contribute to 30-year analyses were also included.

Table 3.2 Comparison of demographic and clinical data between original baseline cohort, the 107 participants with known 30-year outcomes, and those who did not contribute to 30-year analyses

		Baseline	30-year follow-up	Deceased not relating to MS	Unknown outcome
Participants assessed (n)		132	107	13	12
Mean age at presentation (years)		32*	31	36	31
Female (n, %)		80 (61)	69 (64)	6 (46)	5 (42)
Male (n, %)		52 (39)	38 (36)	7 (54)	7 (58)
CIS type	Optic neuritis (n, %)	69 (52)	56 (52)	5 (38)	8 (67)
	Spinal cord syndrome (n, %)	36 (27)	30 (28)	3 (23)	3 (25)
	Brainstem syndrome (n, %)	27 (20)	21 (20)	5 (38)	1 (8)

*Based on 127 participants; CIS onset date was not available in five individuals, all of whom were subsequently lost to follow-up

3.1.4 MS cohort

80 participants in total were known to have developed MS by 30 years: 61 were alive at 30 years and 19 were deceased. As described in section 3.1.1, the cause of death was unknown or unrelated in three people. Of the 61 who were alive, 26 had SPMS and 35 had RRMS (Figures 3.2 & 3.3).

Eleven (14%) individuals had a DMT at some stage, all of which were first-line injectable drugs, i.e. glatiramer acetate or interferons, with the earliest commencing ten years after their MS diagnosis when DMTs first became available in the UK. Of these 11 individuals, seven had SPMS at 30 years, and four had RRMS.

3.1.5 Clinical classifications

Figure 3.3 summarizes the numbers of participants with each clinical classification: CIS, RRMS on MRI grounds, RRMS on clinical grounds, SPMS, and death related to MS (EDSS 10), at each follow-up time-points.

Figure 3.2 Clinical outcomes and classification of the MS cohort at 30 years

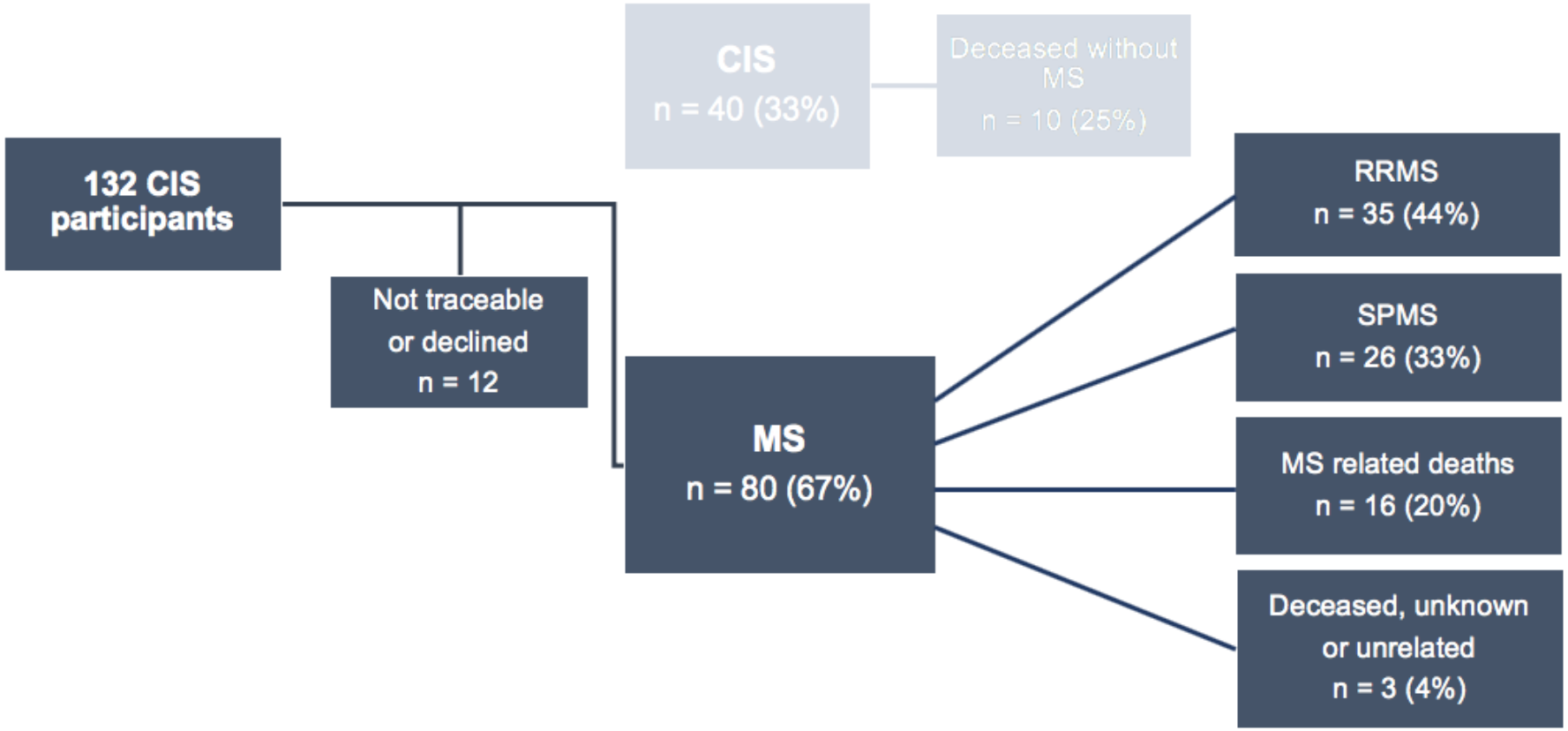
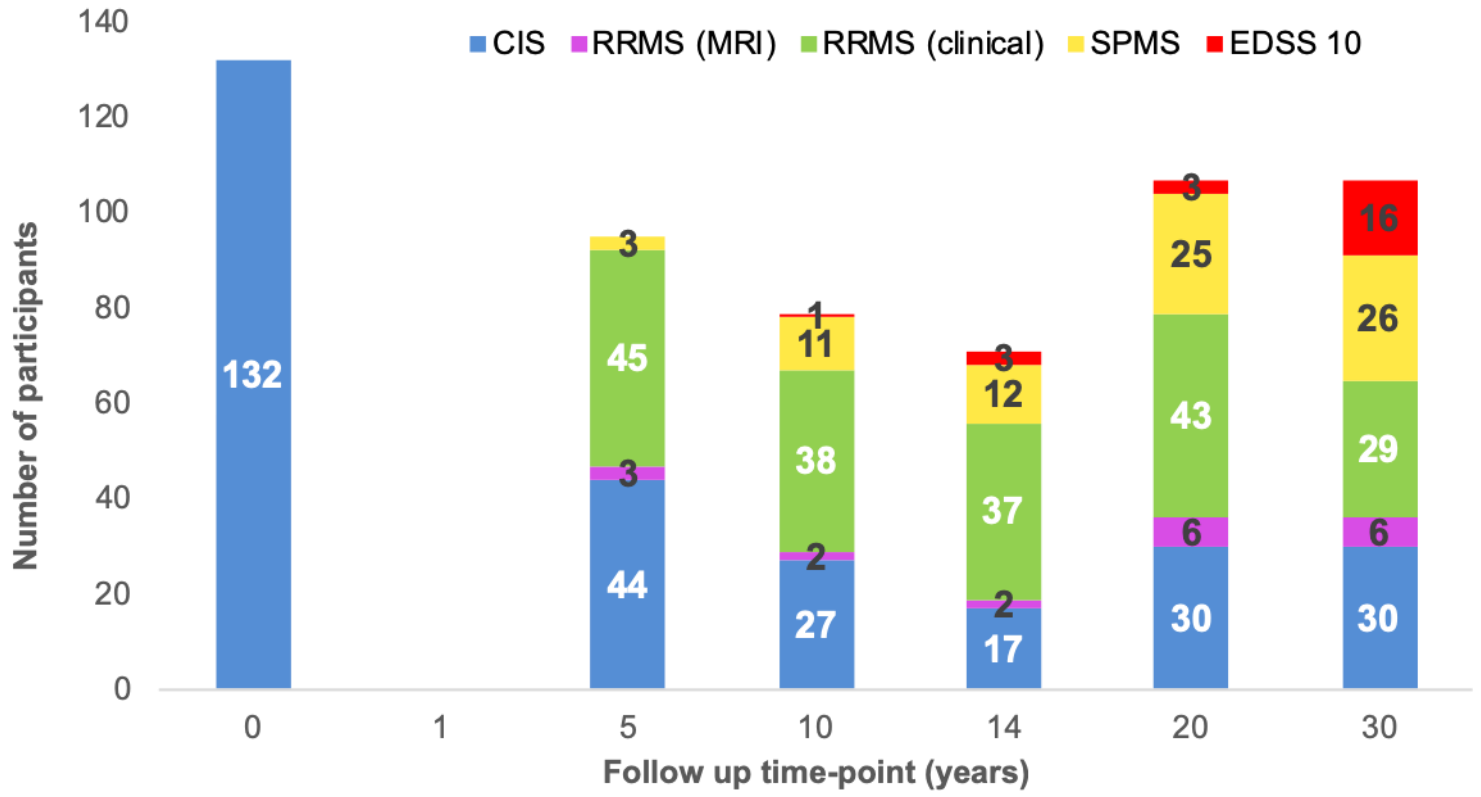


Figure 3.3 Number of participants per clinical classification, at each follow-up time-point



3.2 EDSS outcomes

3.2.1 EDSS outcomes at 30 years

Results from the last section show that, in this cohort, 30 years following a CIS, clinical outcomes were diverse. This was also reflected by EDSS scores and this section will focus on 30-year EDSS outcomes.

At 30 years, EDSS peaks for the whole cohort were observed at 0 (70% CIS, 30% RRMS), 2.0 (44% CIS, 56% RRMS), 6.0 (100% SPMS) and 10 (all of whom had SPMS), with the lowest points at 4.0 and 9.5 (Figure 3.4). Of the 30 classified as CIS at 30 years, 14 (47%) had an EDSS score of 0, with the highest EDSS score from residual disability following a CIS, 4.0.

Among the MS group (n=80), all of the 26 with SPMS (33%) had an EDSS score of >3.5. Of the 35 (44%) with RRMS, 32 (40%) had an EDSS score of ≤3.5. The six people who fulfilled the 2010 McDonald criteria on radiological rather than clinical grounds, all had EDSS ≤3.5.

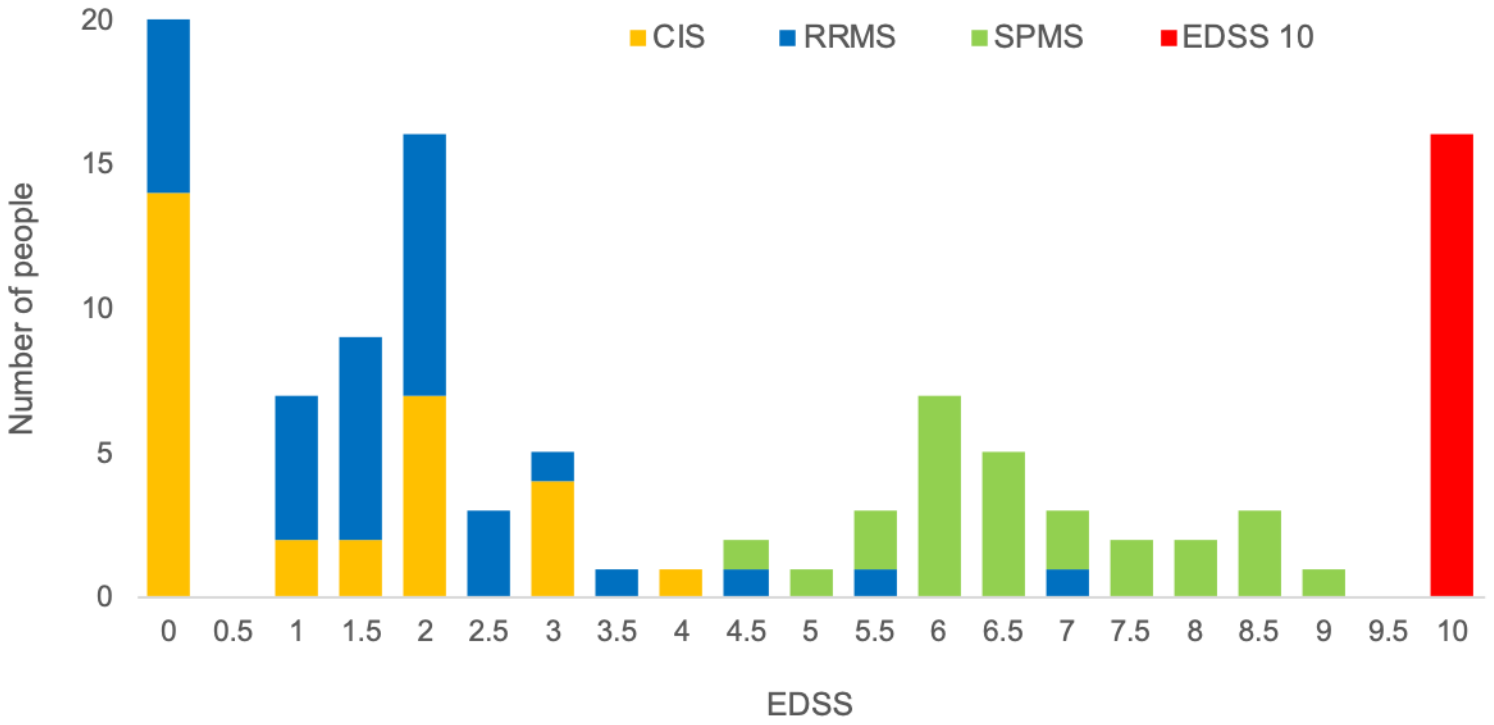
3.2.2 Working EDSS threshold for benign MS

Based on the EDSS outcomes, a working ‘benign’ EDSS threshold of ≤3.5 was confirmed for use in subsequent analyses, as this threshold appeared to distinctly separate two groups of people with MS in this cohort. This EDSS threshold is also clinically meaningful and objective, as it distinguishes

people who remain fully ambulatory, with or without abnormal neurological findings on examination, from those with impaired mobility.

At 30 years disease course was classified as CIS, MS with EDSS ≤3.5, MS with EDSS >3.5 (of which 90% had SPMS), and death relating to MS (EDSS 10).

Figure 3.4 EDSS outcomes at 30 years, whole cohort



EDSS scores obtained from 107 individuals at 30 years. An EDSS score of 10 was assigned to those where MS was known to have contributed to death.

3.2.3 EDSS trajectories over 30 years

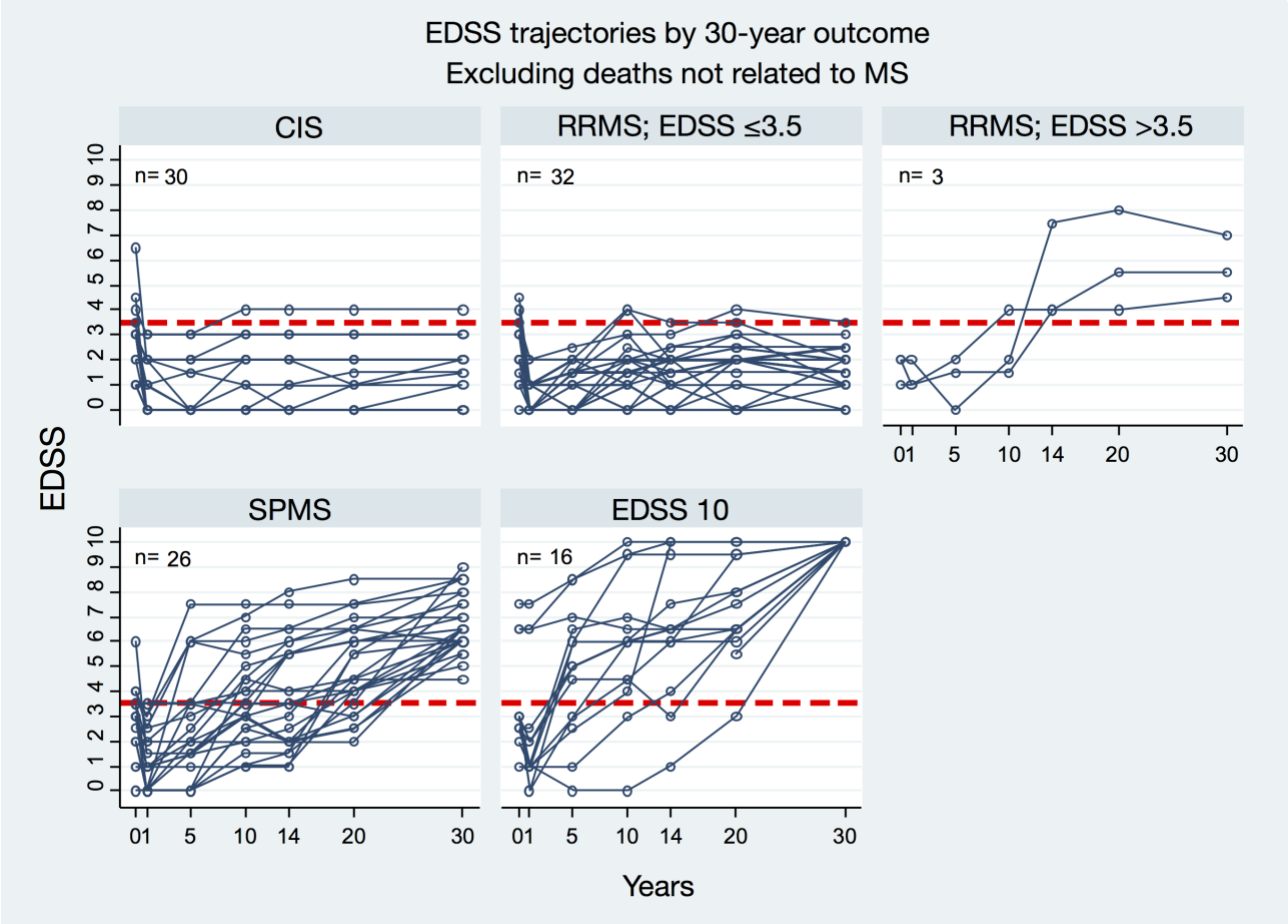
Other than EDSS outcomes at 30 years, how the cohort's EDSS scores changed over the course of 30 years, was also studied. Figure 3.5 shows the EDSS trajectories over 30 years, by 30-year classification. At baseline, and at nadir, EDSS values were indistinguishable between the 30-year outcome groups. A clear separation of EDSS values can be observed from five years, with a continuing rising pattern between EDSS values at 10-, 14- and 20-year, and eventual 30-year outcomes (Table 3.3). For those with less favourable long-term outcomes, namely the SPMS and EDSS 10 subgroups, disability accrual started early on, by or around 5 years from CIS onset. In contrast, there was no significant difference in EDSS trajectories between the CIS group, and the non-disabled (EDSS ≤3.5) RRMS group.

Table 3.3 EDSS scores at each assessment time point, categorised by 30-year classification

30-year outcome	**EDSS mean (SD) at each follow-up time-point (years)**					
	0	**1**	**5**	**10**	**14**	**20**
CIS	2.73 (1.1)	1.22 (1.2)	1.01 (1.2)	1.12 (1.3)	1.16 (1.3)	1.14 (1.3)
RRMS	2.47 (1.2)	0.81 (1.1)	1.13 (1.3)	1.72 (1.4)	1.81 (1.7)	2.32 (1.8)
SPMS	2.97 (1.5)	1.46 (1.7)	3.09 (2.5)	4.36 (2.4)	4.85 (2.7)	5.95 (2.2)

It is worth noting that, at 5 years, 103 people had EDSS ≤3.5, of which 61 (59%) remained EDSS ≤3.5 at 30 years. At 20 years, of the 72 people who had EDSS scores ≤3.5, seven (10%) deteriorated to EDSS >3.5 by 30 years, and 65 (90%) remained below the EDSS ≤3.5 threshold at 30 years.

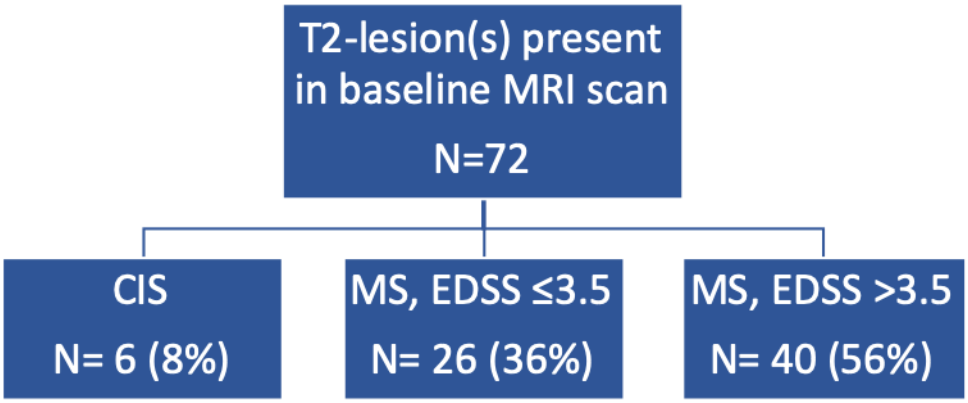
Figure 3.5 EDSS trajectories over 30 years, by 30-year status; lines represent individual participants and overlap in those with the same scores



3.3 MRI brain lesions

Baseline brain MRI scans were normal in 31 people with known 30-year outcomes. 21 of the 31 (68%) with a normal baseline MRI brain scan remained classified as CIS by 30 years; the other ten (32%) subsequently developed MS, of which five had EDSS ≤3.5 by 30 years. T2-hyperintense lesion(s) were present in baseline brain MRI scans in 72 people with known 30-year outcomes: six (8%) had CIS, with 66 (92%) developed MS, of which 26 had EDSS ≤3.5 and 40 had EDSS >3.5 (Figure 3.6).

Figure 3.6 Thirty-year clinical outcomes in those with T2-hyperintense lesion(s) in baseline MRI scans, n= 72



Tables 3.4 and 3.5 shows the total number of MRI brain lesions, and the number of lesions per anatomical region, according to 30-year clinical classification. At all of the time points, there was a general trend for total number of MRI brain lesions and EDSS to correlate, but the pattern was not striking. When lesion counts were analysed per the four anatomical regions, however, the number of infratentorial lesions gave the clearest rising pattern.

Figure 3.7 demonstrates the whole brain lesion count trajectories over the 30-year span, presented by 30-year clinical outcome.

It can be difficult to differentiate, on MR imaging, demyelinating lesions from those arising from small vessel disease. Figure 3.8 shows the total number of intracranial lesions at 30-years, by number of known vascular risk factors in participants. Vascular risk factors included: hypertension, hypercholesterolaemia, diabetes, ischaemic heart disease, and current smoker status. Former smoker status was not included, and participants were not routinely asked if they had a history of migraines. The potential that some of the lesions, particularly at 30 years, were of vascular aetiology, is discussed in section 5.5.3.

Table 3.4: Lesion count based on classification at 30 years (mean ± SD, range)

30-year classification	**Follow-up** (years, number with known lesion count)						
	Baseline (119)	**One** (102)	**Five** (91)	**Ten** (66)	**Fourteen** (55)	**Twenty** (77)	**Thirty** (63)
All	11.2 ± 19.4 (0-100)	14.1 ± 21.4 (0-84)	29.5 ± 41.4 (0-196)	41.4 ± 54.4 (0-331)	51.4 ± 53.7 (0-228)	52.0 ± 54.6 (0-244)	83.8 ± 88.0 (0-346)
CIS	0.60 ± 1.33 (0-5)	0.53 ±1.61 (0-6)	1.1 ± 1.87 1.2 (0-6)	3.1 ± 5.33 (0-18)	5.3 ± 7.7 (0-21)	6.8 ± 9.7 (0-32)	19.1 ± 36.2 (0-152)
RRMS, EDSS ≤3.5	11.2 ± 18.0 (0-80)	10.8 ±16.0 (0-82)	21.1 ± 18.8 (0-88)	31.8 ± 26.9 (2-104)	36.4 ± 25.0 (3-114)	49.1 ± 32.3 (5-143)	88.5 ± 61.8 (5-244)
RRMS, EDSS >3.5	12.3 ± 11.5 (1-24)	10.0 ± 8.2 (3-19)	30.7 ± 21.4 (6-43)	50.0 ± 56.6 (10-90)	77.0 ± 77.8 (22-132)	87.0 ± 74.6 (28-171)	142.7 ± 128.0 (43-287)

SPMS	15.9 ± 23.0 (0-77)	22.4 ± 26.0 (0-84)	46.0 ± 43.1 (2-151)	60.3 ± 47.7 (3-140)	78.8 ± 59.9 (5-195)	80.9 ± 61.8 (6-210)	155.3 ± 106.3 (28-346)
MS contributed to death	24.3 ± 26.4 (0-100)	29.8 ± 26.4 (0-80)	73.8 ± 68.7 (13-196)	86.5 ± 98.5 (23-331)	95.1 ± 68.0 (36-228)	103.1 ± 67.1 (39-244)	

Historical lesion counts were used in 16 participants at baseline, nine at one year, and in five people at five years.

Table 3.5 Lesion count by anatomical location, based on classification at 30 years (mean ± SD, range)

Anatomical location	30-year outcome	**Follow-up** (years)						
		Baseline	**One**	**Five**	**Ten**	**Fourteen**	**Twenty**	**Thirty**
Infratentorial	CIS	0	0.05 ±0.22 (0-1)	0	0	0	0.06 ± 0.24 (0-1)	0.11 ± 0.46 (0-2)
	RRMS, EDSS ≤3.5	0.25 ± 1.0 (0-5)	0.16 ±0.5 (0-2)	0.67 ± 1.0 (0-3)	0.93 ± 1.5 (0-5)	1.0 ± 1.8 (0-6)	1.3 ± 2.0 (0-9)	2.1 ± 3.8 (0-17)
	RRMS, EDSS >3.5	0	0	0.3 ± 0.6 (0-1)	1.0 ± 1.4 (0-2)	2.5 ± 2.1 (1-4)	3.7 ± 4.6 (1-9)	5.3 ± 5.9 (1-12)
	SPMS	0.7 ± 1.0 (0-3)	1.0 ± 1.3 (0-5)	1.4 ± 1.8 (0-7)	1.7 ± 2.0 (0-7)	2.6 ± 2.5 (0-8)	2.9 ± 3.5 (0-14)	5.3 ± 6.6 (0-24)
	EDSS 10	1.9 ± 2.9 (0-11)	3.1 ± 3.6 (0-12)	6.0 ± 6.3 (0-20)	6.7 ± 8.4 (0-27)	6.9 ± 6.4 (0-15)	6.3 ± 6.3 (0-17)	
Peri-ventricular	CIS	0.20 ± 0.82 (0-4)	0.26 ± 0.93 (0-4)	0.29 ± 0.9 (0-4)	0.6 ± 1.0 (0-3)	0.9 ± 0.9 (0-2)	0.7 ± 0.9 (0-2)	1.4 ± 2.9 (0-13)

	RRMS, EDSS ≤3.5	3.4 ± 4.3 (0-19)	3.4 ± 3.6 (0-15)	5.6 ± 4.1 (0-16)	7.6 ± 4.9 (0-16)	8.4 ± 5.5 (0-21)	11.3 ± 7.3 (0-32)	14.3 ± 7.9 (1-31)
	RRMS, EDSS >3.5	3.0 ± 2.6 (1-6)	2.3 ± 3.2 (0-6)	2.0 ± 2.6 (1-6)	6.0 ± 4.2 (3-9)	10.5 ± 9.2 (4-17)	9.3 ± 9.7 (1-20)	11.7 ± 12.5 (3-26)
	SPMS	3.9 ± 5.8 (0-22)	5.6 ± 7.1 (0-22)	10.5 ± 10.3 (0-36)	12.7 ± 9.2 (2-28)	17.9 ± 14.1 (2-49)	15.2 ± 10.5 (0-33)	20.1 ± 12.5 (6-50)
	EDSS 10	5.6 ±4.6 (0-17)	8.4 ± 6.1 (0-17)	11.9 ± 8.4 (0-27)	13.0 ± 11.5 (1-33)	17.7 ± 12.7 (5-36)	22.7 ± 13.5 (8-46)	
Juxtacortical	CIS	0.04 ± 0.2 (0-1)	0	0	0	0	0.05 ± 0.22 (0-1)	0.05 ± 0.2 (0-1)
	RRMS, EDSS ≤3.5	0.5 ±1.5 (0-6)	0.76 ±2.1 (0-10)	2.3 ± 4.0 (0-19)	5.9 ± 10.8 (0-46)	5.9 ± 10.8 (0-45)	7.1 ±9.2 (0-45)	10.0 ± 12.0 (1-52)
	RRMS, EDSS >3.5	0.33 ± 0.6 (0-1)	1.0 ± 1.0 (0-2)	2.7 ± 2.5 (0-5)	5.5 ± 7.8 (0-11)	4.5 ± 6.4 (0-9)	6.0 ± 5.3 (0-10)	14.0 ± 13.7 (2-29)

	SPMS	1.2 ± 2.2 (0-7)	2.0 ± 3.2 (0-11)	4.3 ± 6.4 (0-25)	8.8 ± 11.7 (0-38)	12.1 ± 13.6 (0-47)	11.9 ± 14.3 (0-52)	14.7 ± 19.3 (0-71)
	EDSS 10	1.5 ±2.4 (0-7)	1.9 ± 2.8 (0-7)	8.3 ± 17.6 (0-63)	6.0 ± 6.6 (0-19)	8.6 ± 7.0 (1-22)	8.9 ± 6.7 (2-20)	
Deep white matter	CIS	0.28 ±0.7 (0-3)	0.21 ± 0.92 (0-4)	0.81 ± 1.5 (0-5)	2.5 ± 4.5 (0-15)	4.4 ± 7.0 (0-19)	6.0 ± 9.2 (0-31)	17.1 ± 33.4 (0-139)
	RRMS, EDSS ≤3.5	7.0 ± 12.9 (0-51)	6.1 ± 11.4 (0-57)	12.6 ± 14.0 (0-64)	17.6 ± 15.2 (2-64)	21.1 ± 12.2 (2-42)	29.5 ± 20.0 (2-74)	62.3 ± 47.4 (2-175)
	RRMS, EDSS >3.5	9.0 ±11.5 (0-22)	6.7 ± 4.5 (2-11)	23.7 ± 16.3 (5-35)	37.5 ± 43.1 (7-68)	59.5 ± 60.1 (17-102)	68.0 ± 57.7 (20-132)	111.7 ± 97.2 (32-220)
	SPMS	11.3 ± 16.4 (0-55)	14.9 ±17.4 (0-57)	29.6 ± 29.2 (1-104)	37.8 ± 35.8 (1-115)	48.0 ± 39.2 (2-133)	52.8 ± 46.4 (3-162)	115.4 ± 86.7 (10-253)
	EDSS 10	15.3 ±20.3 (0-82)	17.4 ± 18.0 (0-59)	47.7 ± 46.2 (5-124)	60.6 ± 75.6 (16-257)	62.0 ± 47.3 (19-160)	65.3 ± 47.7 (16-161)	

Where scans were missing, historical lesion counts were used and regional lesion data were not available. Historical lesion counts were used in 16 people at baseline, nine at one year, and in five at five years.

Figure 3.7 Whole brain lesion count trajectories over 30 years, by 30-year status

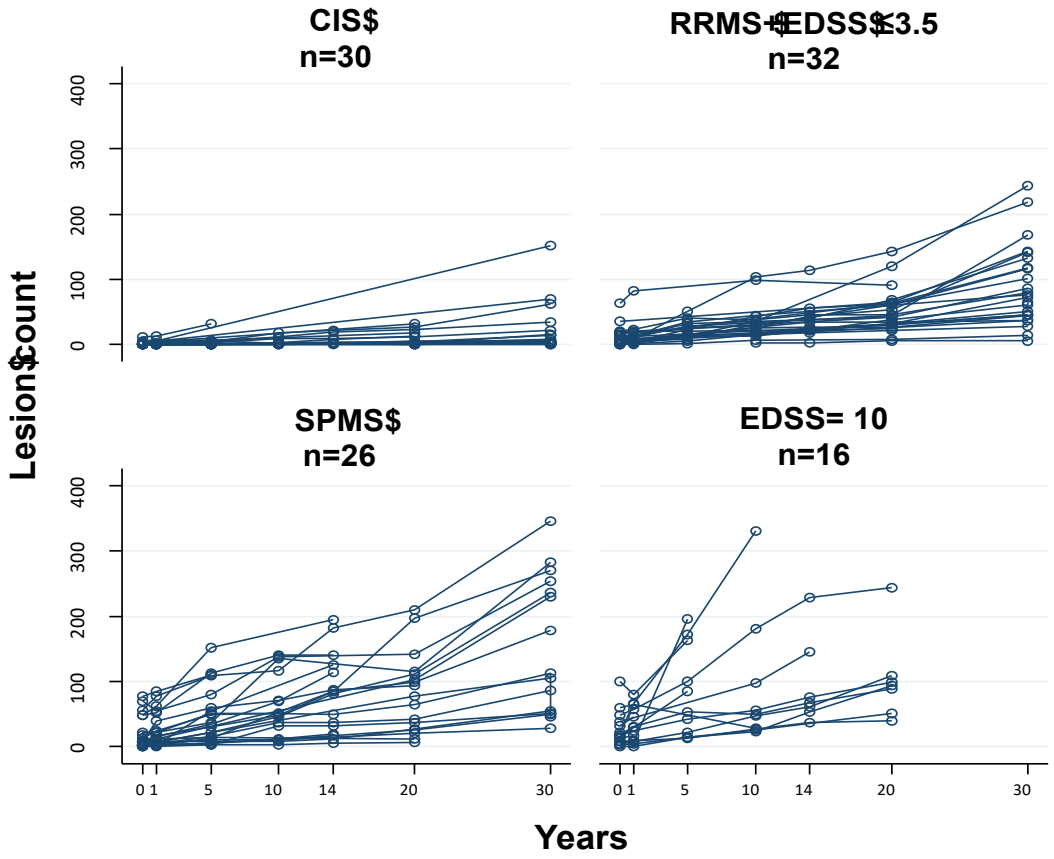
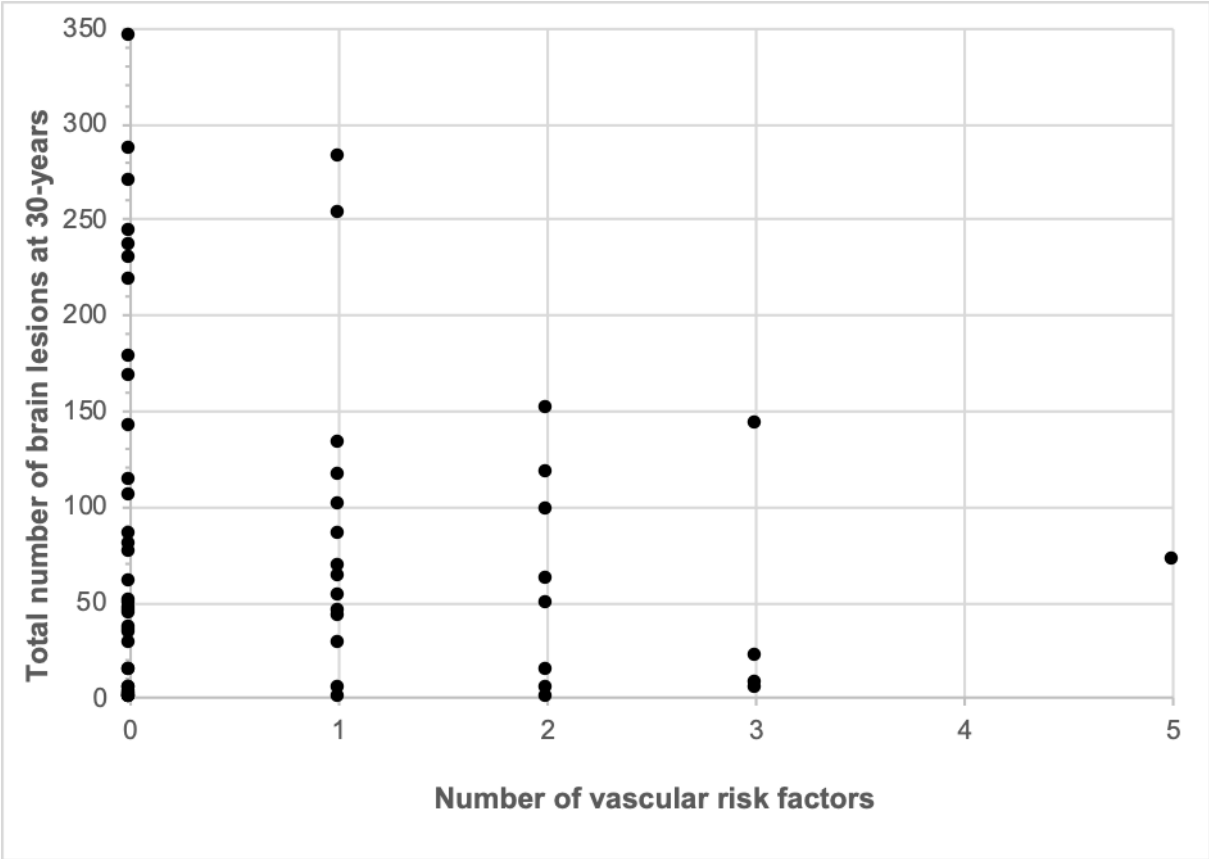


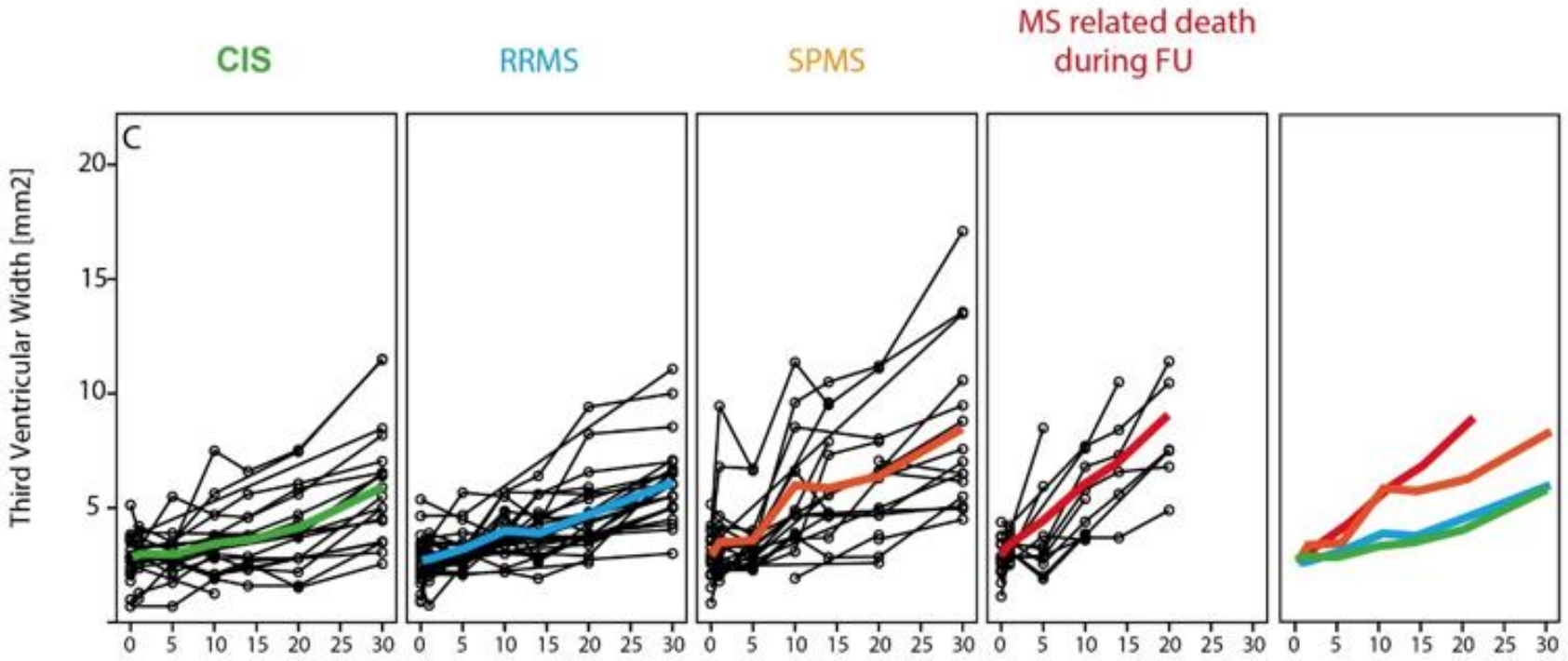
Figure 3.8 Number of total intracranial lesions at 30-years, by number of known vascular risk factors in participants, n= 63



3.4 MRI atrophy measures

Figure 3.9 shows the available TVW measurements over 30 years, by 30-year classification. On inspection, those who were classified as remaining CIS or as RRMS at 30 years, followed a similar course. In contrast, those who had developed SPMS by 30 years, or in whom MS contributed to their death, had noticeably more tissue loss within five to ten years, and the rate of atrophy continued to be more pronounced. Atrophy as measured by TVW, was associated with higher EDSS scores, at cross-sectional time-points over the 30-year follow-up duration, ß= 0.35, $p < 0.001$.

Figure 3.9 Third ventricular width measurements over 30 years, by 30-year status



The mean for each group is shown in green (CIS), blue (RRMS), orange (SPMS) and red (death related to MS), with a plot of overlaid mean value curves provided at the furthest right panel.

3.5 Benign MS

This section focuses on the sub-group of people with MS who were fully ambulatory at 30 years after the onset of their first MS symptoms. A working definition of 'benign' or non-disabling MS at 30 years was an EDSS score of ≤3.5. This EDSS threshold reflects a person who is fully ambulatory, with or without symptoms, and with or without abnormal neurological findings on examination (see Table 1.2).

3.5.1 Clinical information

All those with an EDSS of ≤3.5 remained in the relapsing-remitting phase of the condition, and they accounted for 91% (32 out of 35) of those who were classified as RRMS at 30 years.

Three people in the EDSS ≤3.5 group had been treated with a DMT, compared to one in the RRMS with EDSS >3.5 group, and seven in the SPMS or EDSS 10 groups.

Due to the relatively long gaps between follow-up time-points, relapse occurrences and frequencies were not consistently documented, and patient recall was felt to be not reliable enough. Relapse rate was therefore not included in analyses. With this limitation in mind, however, it is perhaps still worthwhile reporting that of the 32 with MS and EDSS ≤3.5 at 30 years, six individuals (19%) reported a relapse between 20- and 30-years; the other 26

individuals did not report or recall a clinical relapse in the past decade.

Of the 32 individuals, eight (25%) are reviewed by a MS health care professional at intervals, at least annually: six by neurologists and two by MS specialist nurses; the other 24 individuals (75%) do not routinely consult a neurology health care professional.

In terms of presenting CIS, Table 3.6 compares the number of participants in each 30-year outcome groups, based on their CIS presentation. The proportion of people presenting with optic neuritis and who developed RRMS, and had a 30-year EDSS of ≤3.5, was 47%; this was not significantly different from the 52% who presented with ON in the original cohort.

Table 3.6 Number of participants with each CIS presentation, at baseline and at 30-years

		CIS presentation n (% of each outcome classification)		
		Optic neuritis	Brainstem	Transverse myelitis
Baseline		69 (52)	27 (20)	36 (27)
30-year outcome	CIS	19 (63)	3 (10)	8 (27)
	RRMS EDSS ≤ 3.5	15 (47)	7 (22)	10 (31)
	RRMS EDSS > 3.5	2 (66)	0	1 (33)
	SPMS	13 (50)	5 (18)	8 (31)
	Deceased with CIS	4 (44)	3 (33)	2 (22)
	Deceased due to MS	7 (44)	6 (38)	3 (19)
	Deceased with MS (not due to MS)	0	2 (100)	0
	Deceased (unknown cause)	1 (50)	0	1 (50)
	Declined	2 (50)	1 (25)	1 (25)
	Unknown	6 (75)	0	2 (25)

3.5.2 Cognition

As the EDSS takes little account of cognition, we determined the proportion in each group who were found to have cognitive impairment on BICAMS.

Mean NART, an estimate of pre-morbid intelligence, and years of education, were similar in all the outcome groups.

BICAMS scores were obtained in 41 people with MS who attended in person (26 with EDSS≤3.5, 15 with EDSS>3.5), and BICAMS z-scores (adjusted for age, sex and years of education) were available in 31 subjects who were ≤65 years of age (Table 3.7). Subjects who did not complete the cognitive tests tended to be more physically disabled, both early (mean nadir EDSS 1.6 vs 0.8 for subjects who completed assessment) and later in the disease course (mean 30-year EDSS 6.0 vs 4.1).

In the 32 participants with MS and EDSS ≤3.5 at 30 years, BICAMS was performed in 26, and validated z-scores available in 21. Impairment is defined as a z-score of <-1.5 in any of the three BICAMS modalities (Langdon, 2016). Two out of 21 people had a BICAM modality z-score of <-1.5, one with EDSS 2.0 and one with EDSS 3.5.

Table 3.7: Number of BICAMS assessments with validated z-scores (where available), performed on people with MS, categorized by EDSS at 30 years

EDSS	**0**	**0.5**	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**	**4.5**
BICAMS tested/ number of people	4/6	0	5/5	7/7	6/9	3/3	0/1	1/1	0	2/2
BICAMS z-scores (impaired*/ available)	0/2		0/5	0/7	1/4	0/2		1/1		0/1
EDSS	**5**	**5.5**	**6**	**6.5**	**7**	**7.5**	**8**	**8.5**	**9**	**9.5**
BICAMS tested/ number of people	0/1	3/3	5/7	2/5	3/3	0/2	0/2	0/3	0/1	0
BICAMS z-scores (impaired*/ available)		2/2	1/3	2/2	0/2					

Z-scores adjusted for sex, age and years of education available in those ≤65 years of age. *Impaired as defined by a z-score of <-1.5 in any of the three modalities (Langdon, 2016)

Comparing between the 30-year outcome groups, age-adjusted cognitive measures in the group with MS and EDSS ≤3.5 were not significantly different from the CIS group: for PASAT, the MS with EDSS ≤3.5 group was 9% worse than the CIS group, p= 0.26; for BVMTR, the MS with EDSS ≤3.5 group was 0.2% better, p= 0.97; for CVLT, the MS with EDSS ≤3.5 group was 1% worse, p= 0.83; for SDMT, the MS with EDSS ≤3.5 group was 7% worse, p= 0.096. In the group with MS and EDSS >3.5, cognitive measures were more substantially and significantly worse than the CIS group (except for CVLT): for PASAT, the EDSS >3.5 group was 23% worse, p= 0.008; for BVMTR, 21% worse, p= 0.006; for CVLT, 9% worse, p= 0.23; and for SDMT, 24% worse, p<0.001 (Table 3.8).

Table 3.8 Comparison of age-adjusted cognitive measurements between MS EDSS ≤3.5 and EDSS >3.5 groups, with the CIS group

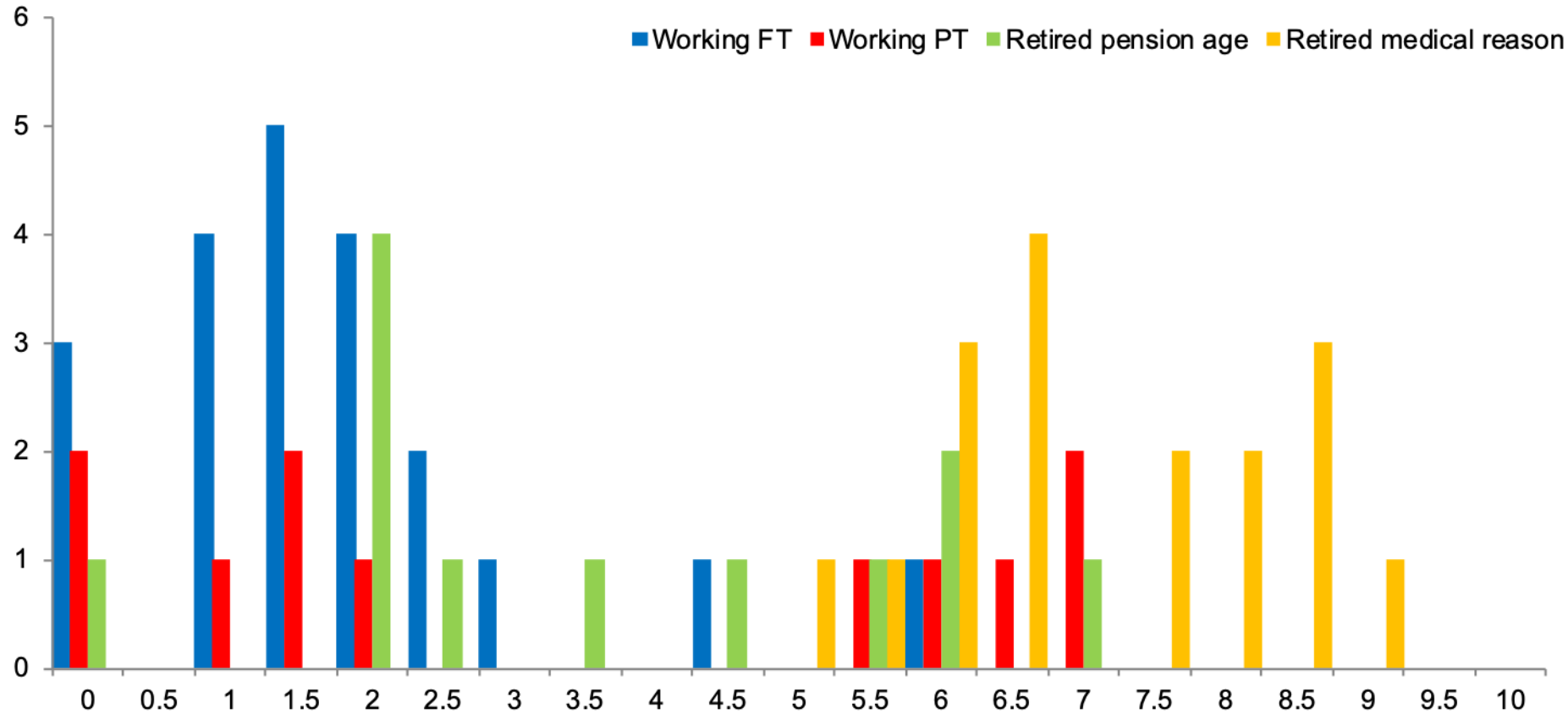
	CIS at 30-years	**MS with EDSS ≤3.5 at 30-years**			**MS with EDSS >3.5 at 30-years**		
	Adjusted mean	*Adjusted mean*	*Compared to CIS group (%)*	*p-value*	*Adjusted mean*	*Compared to CIS group (%)*	*p-value*
PASAT	46.31	42.32	-9	0.26	35.63	-23	0.008
BVMTR	25.84	25.89	+0.2	0.97	20.40	-21	0.006
CVLT	52.41	51.70	-1	0.83	47.92	-9	0.23
SDMT	54.39	50.40	-7	0.096	41.49	-24	<0.001

3.5.3 Occupational status

Employment status can be a reflection of the social impact MS has on an individual (Glad *et al.*, 2011), and was included in McAlpine's original definition of benign MS (Mcalpine, 1961). The proportion who remained in employment or who had retired at the national state pension age of 60 years, was therefore also determined at 30 years.

Of the 32 with EDSS ≤3.5, none had retired early for medical reasons, and all remained in employment (full time or part time), or retired at the national state pension age (Figure 3.10). Reasons for working part time were varied, with MS being a contributing factor in three of the six individuals (50%). It was not explored whether anyone altered their choice of career or employment, as a result of their MS. In contrast, early retirement due to medical reasons was observed in people with EDSS scores of 5 and upwards. An EDSS score of 6 marks the threshold when an individual with MS requires the use of a mobility aid: of the 23 people with MS and an EDSS ≥6.0, only 5 (22%) remained in employment and one (5%) in full time employment.

Figure 3.10 Employment and retirement status in people with MS, by EDSS at 30 years. N= 61



3.5.4 Comparison between the CIS and benign MS groups

Some additional comparisons between the CIS and benign MS subgroups were made. Overall, there was no significant difference between the two groups in any of the measures.

For measures of physical ability, there was no significant difference between the two groups. The mean (SD) EDSS for the CIS group was 1.17 (1.3), and for the BMS group, 1.48 (0.9); t(60)= -1.14, p= 0.26. In the timed 25-foot walk, where two readings were averaged, the mean (SD) speed for the CIS group was 5.87 (1.9) seconds, and BMS group 5.65 (0.9) seconds; t(45)= 0.54, p= 0.59. In the 9-hole peg test for upper limb function, where two readings were again averaged, in the dominant hand, the CIS group achieved a mean (SD) of 21.1 (2.3) seconds, and the BMS group 22.1 (4.2) seconds; t(45)= -0.98, p= 0.33. In the non-dominant hand, the measurements were 21.4 (3.4) seconds for the CIS group and 23.3 (4.3) seconds for the BMS group; t(45)= -1.75, p= 0.087.

In terms of non-physical measures, cognition has already been compared in section 3.4.2. There was negligible difference between PASAT scores, and BICAMS raw and z-scores, between the CIS and benign MS groups (Table 3.8)

For mood and fatigue, the mean (SD) score for the Hospital Anxiety Scale was 5.0 (3.3) for the CIS group, and 5.0 (2.7) for the BMS group; t(42)=

0.045, p= 0.96. For the Hospital Depression Scale, the results were 2.3 (2.0) for the CIS group, and 2.1 (1.7) for the BMS group; t(42)= 0.29, p= 0.77. The fatigue severity scale scores, where a higher score denotes worse fatigue, were 21.5 (10.0) for the CIS group, and 25.7 (14.4) for the BMS group; t(42)= -1.09, p= 0.28. The mean (SD) cognitive leisure activities scores, where a higher score reflects more frequent cognitive leisure activities, were 20.1 (6.0) for the CIS group, and 20.0 (4.6) for the BMS group; t(42)= 0.19, p= 0.85.

In terms of occupation status, this was similar between the two groups, with the BMS groups no less ‘active’ than the CIS group in terms of work.

Chapter 4: Relationship between early features and long-term clinical outcome

This chapter will focus on *early* (within five years of onset) prognostic features of 30-year outcomes following a CIS.

4.1 Demographic and clinical features

Table 4.1 shows the early univariable predictors of the three binary 30-year outcomes, looking at demographic and clinical features.

Relationships between CIS type, age at onset, gender, disease duration, and 30-year outcomes, were assessed. There was no association of gender and disease duration with any of the outcome groups at 30 years.

In terms of age at onset, older people at presentation were at greater risk of MS-related mortality, with HR 1.07 per year (p= 0.04).

4.1.1 Clinically isolated syndrome presentation

People presenting with a brainstem CIS had higher EDSS scores early on, compared to those presented with optic neuritis (ON) or transverse myelitis (TM). The mean (SD) nadir EDSS based on CIS subgroups were: ON 0.86 (0.9), TM 1.44 (1.4), and brainstem 1.46 (2.0), and for 5-year EDSS, ON 1.39 (1.6), TM 1.75 (2.0), and brainstem 2.72 (2.7). Those with a brainstem CIS were at greater risk than those presenting with either ON or TM, of MS-

related death, with a hazard ratio (HR) of 2.87, p= 0.04. This was consistent with the higher proportion of brainstem CIS subjects with a baseline IT lesion present (41%), compared to those with TM as CIS (19%), and ON as CIS (11%; chi-squared test, p= 0.009). The brainstem CIS group had the lowest proportion of people who remained CIS or had benign MS by 30 years (Table 3.6). The change in EDSS from nadir to five-years was also largest in the brainstem CIS group (mean= 1).

Table 4.1 Early univariable predictors of 30-year outcomes: demographic and clinical features

Predictor		30-year EDSS ≤3.5 vs 30-year EDSS >3.5			SPMS by 30 years			Death related to MS by 30 years		
		n	*Odds ratio (95% CI)*	*P-value*	*n*	*Odds ratio (95% CI)*	*P-value*	*n*	*Odds ratio (95% CI)*	*P-value*
CIS type	TM vs brainstem	107	0.61 (0.20, 1.87)	0.383	119	0.62 (0.21, 1.82)	0.385	119	0.27 (0.06, 1.21)	0.087
	ON vs brainstem	107	0.63 (0.23, 1.74)	0.375	119	0.67 (0.26, 1.71)	0.397	119	0.33 (0.10, 1.11)	0.074
Age at onset, OR per year		102	1.05 (1.00, 1.11)	0.060	112	1.02 (0.97, 1.07)	0.411	117	1.08 (1.00, 1.16)	0.036
Gender, male vs female		107	1.12 (0.50, 2.48)	0.787	119	0.97 (0.45, 2.14)	0.949	119	0.98 (0.33, 2.92)	0.978
Disease duration, OR per year		102	1.54 (0.59, 4.01)	0.378	111	1.40 (0.69, 2.88)	0.353	110	0.69 (0.10, 4.63)	0.702

TM= transverse myelitis; ON= optic neuritis

4.2 EDSS scores as predictor

Figure 3.5 in the last chapter shows the EDSS trajectories of all subjects over the 30-year time span, based on 30-year outcome. Separation between the groups became apparent at five years, and was more pronounced at subsequent time-points. Baseline EDSS scores were not significantly associated with 30-year outcome. EDSS scores at nadir (where clinical improvement had plateaued, or at one year, whichever was earlier) and at five years were stronger predictors.

4.2.1 Prediction of 30-year EDSS ≤3.5 versus EDSS >3.5

For prediction of a 30-year EDSS ≤3.5 versus a 30-year EDSS >3.5 outcomes, EDSS scores at nadir and five years were significant, more so than EDSS changes between these time-points. The predictive value of EDSS became greater with time, which would be expected. People with a nadir EDSS of ≥2.5 had OR 6.44 of having 30-year EDSS >3.5, compared to those with a nadir EDSS of <2.5, p= 0.007, PPV 79%, NPV 64%, sensitivity 25%, specificity 95%, and accuracy 66% (95% CI 56%, 75%). At five years, an EDSS of ≥2.5 gave a corresponding OR of 15.6, p<0.001, PPV 85%, NPV 73%, sensitivity 52%, specificity 93% and accuracy 76% (95% CI 67%, 84%) (Tables 4.2 and 4.4).

Table 4.2 Contingency tables of early EDSS scores as predictors of 30-year EDSS ≤3.5 versus EDSS >3.5

	Nadir EDSS ≥2.5	*Nadir EDSS <2.5*	Total
30-year EDSS >3.5	11	33	44
30-year EDSS ≤3.5	3	58	61
Total	14	91	105

PPV 79%, NPV 64%, sensitivity 25%, specificity 95%, and accuracy 66% (95% CI 56%, 75%)

	5-year EDSS ≥2.5	*5-year EDSS <2.5*	Total
30-year EDSS >3.5	57	4	61
30-year EDSS ≤3.5	21	23	44
Total	78	27	105

PPV 85%, NPV 73%, sensitivity 52%, specificity 93% and accuracy 76% (95% CI 67%, 84%)

4.2.2 Prediction of SPMS

For prediction of a SPMS outcome following a CIS, there was no significant association with either EDSS at baseline or at nadir, but a 5-year EDSS of ≥2.5 was a significant predictor, with an odds ratio (OR) 9.43, p< 0.001, PPV 72%, NPV 79%, sensitivity 50%, specificity 90% and accuracy 77% (95% confidence interval (CI) 68%, 85%). EDSS change from nadir to 5-years was a slightly stronger predictor than either the absolute 5-year EDSS value, or the change from baseline to 5-years: people who progressed by ≥ 2 EDSS points between nadir and 5-years had an OR of 40.7 greater than those with <2 EDSS change, p< 0.001, PPV 93%, NPV 76%, sensitivity 36%, specificity 99%, accuracy 78% (95% CI 69%, 85%) (Table 4.3).

4.2.3 Prediction of MS-related death

For MS-related death, five-year EDSS was the strongest early predictor in a survival analysis (there was no MS-related death by five years). At five-years, EDSS ≥2.5 versus <2.5 gave a HR of 23.6 (p <0.001). Having a five-year EDSS of ≥2.5 predicted death by 30 years with PPV 43%, NPV 98%, sensitivity 86%, specificity 83% and accuracy 84% (95% CI 75%, 90%) (Table 4.3).

Table 4.3 Contingency tables of early EDSS scores as predictors of SPMS, and of MS-related death, by 30 years

	5-year EDSS ≥2.5	*5-year EDSS <2.5*	Total
CIS or RRMS by 30 years	7	66	73
SPMS by 30 years	18	18	36
Total	25	84	109

PPV 72%, NPV 79%, sensitivity 50%, specificity 90% and accuracy 77% (95% confidence interval (CI) 68%, 85%)

	EDSS change ≥2 from nadir to 5-years	*EDSS change <2 from nadir to 5-years*	Total
CIS or RRMS by 30 years	1	72	73
SPMS by 30 years	13	23	36
Total	14	95	109

PPV 93%, NPV 76%, sensitivity 36%, specificity 99%, accuracy 78% (95% CI 69%, 85%)

	5-year EDSS >2.5	*5-year EDSS ≤2.5*	Total
MS-related death by 30 years	12	2	14
Alive at 30 years	17	86	103
Total	29	88	117

PPV 43%, NPV 98%, sensitivity 86%, specificity 83% and accuracy 84% (95% CI 75%, 90%)

Table 4.4 Early EDSS values as univariable predictors of 30-year outcomes

Predictor	30-year EDSS ≤3.5 vs 30-year EDSS >3.5			SPMS by 30 years			Death related to MS by 30 years		
	n	*Odds ratio (95% CI)*	*P-value*	*n*	*Odds ratio (95% CI)*	*P-value*	*n*	*Odds ratio (95% CI)*	*P-value*
Baseline EDSS ≥2.5 vs <2.5	100	1.89 (0.81, 4.42)	0.140	110	2.57 (1.06, 6.20)	0.036	108	4.00 (0.84, 19.05)	0.082
Nadir EDSS ≥2.5 vs <2.5	105	6.44 (1.68, 24.77)	0.007	115	2.55 (0.90, 7.25)	0.079	111	3.12 (0.84, 11.69)	0.090
5-year EDSS ≥2.5 vs <2.5	105	15.61 (4.83, 50.48)	<0.001	109	9.43 (3.41, 26.06)*	<0.001	110	30.00 (6.12, 147.15)	<0.001
EDSS change baseline to 5-year ≥0 vs <0	100	8.84 (3.44, 22.76)	<0.001	103	4.98 (2.03, 12.19)*	<0.001	105	8.02 (2.05, 31.45)	0.003
EDSS change nadir to 5-year ≥2 vs <2	105	31.03 (3.91, 246.49)	0.001	109	40.70 (5.05, 328.17)*	<0.001	110	19.80 (5.34, 73.46)	<0.001

*excluding three subjects who had already developed secondary progression by 5 years

4.3 White matter lesions as predictors

Overall, WM lesions detected on MRI proved to be more effective predictors of 30-year outcomes than EDSS, and lesion location was more important than lesion number. In multivariable models including both lesion and EDSS variables, the latter no longer contributed significantly, with their coefficients substantially reduced. Table 4.5 shows early WM lesion variables as univariable predictors of 30-year outcomes.

4.3.1 Prediction of 30-year EDSS ≤3.5 versus EDSS >3.5

Total baseline lesion count was a significant predictor ($p < 0.001$) of 30-year EDSS >3.5, with an OR 1.84 ($p < 0.001$), per adjacent higher lesion number category (0, 1-3, 4-10, 11-20, 21+); or the reciprocal OR, 0.54 for 30-year EDSS ≤3.5.

The strongest early MRI predictors of 30-year EDSS >3.5 were the presence of: i) baseline IT lesions, OR 12.4, $p < 0.001$, accuracy 71% (95% CI 61%, 80%), ii) one-year DWM lesions, OR 10.7, $p < 0.001$, accuracy 68% (95% CI 57%, 78%), and iii) one-year IT lesions, OR 11.1, $p < 0.001$, accuracy 73% (95% CI 62%, 82%) (Table 4.6).

For predicting 30-EDSS ≤3.5, these odds ratios apply to the absence of these lesions.

Table 4.5 Early WM lesion variables as univariable predictors of 30-year outcomes

Predictor	30-year EDSS ≤3.5 vs 30-year EDSS >3.5			SPMS by 30 years			Death related to MS by 30 years		
	n	*Odds ratio (95% CI)*	*P-value*	*n*	*Odds ratio (95% CI)*	*P-value*	*n*	*Odds ratio (95% CI)*	*P-value*
Lesion counts									
Total baseline lesion count, 0/1-3/4/10/11-20/21+	103	1.84 (1.35, 2.52)	<0.001	114	1.86 (1.37, 2.52)	<0.001	108	2.03 (1.32, 3.12)	0.001
Total baseline lesion count, 1+ vs 0	103	5.21 (1.91, 14.2)	0.001	114	7.63 (2.47, 23.62)	<0.001	108	7.62 (0.96, 60.40)	0.054
Total 1-year lesion count, 0/1-4/5-15/16+	86	2.51 (1.60, 3.93)	<0.001	96	2.67 (1.70, 4.17)	<0.001	92	2.48 (1.31, 4.68)	0.005
Total 1-year lesion count, 1+ vs 0	86	9.26 (2.48, 34.51)	0.001	96	12.87 (2.83, 58.58)	0.001	92	5.60 (0.69, 45.19)	0.106
Total 5-year lesion count, 0/1-10/11-20/21-40/41+	81	2.57 (1.69, 3.90)	<0.001	88	2.30 (1.57, 3.39)	<0.001	85	2.26 (1.27, 4.02)	0.005
Total 5-year lesion count, 0-10 vs 11+	81	6.84 (2.38, 19.70)	<0.001	88	6.70 (2.25, 19.92)	0.001	52	See ^ below	<0.001
Total lesion change baseline to 1-year, 1+ vs 0	86	3.41 (1.38, 8.45)	0.008	96	3.95 (1.59, 9.81)	0.003	91	3.41 (0.89, 13.08)	0.073
Total lesion change baseline to 5-year, 0/1-9/10+	81	5.90 (2.57, 13.53)	<0.001	88	4.79 (2.16, 10.63)	<0.001	85	4.70 (1.18, 18.72)	0.028

Predictor	30-year EDSS ≤3.5 vs 30-year EDSS >3.5			SPMS by 30 years			Death related to MS by 30 years		
	n	*Odds ratio (95% CI)*	*P-value*	*n*	*Odds ratio (95% CI)*	*P-value*	*n*	*Odds ratio (95% CI)*	*P-value*
Lesion locations									
Baseline PV lesion count, ≥1 vs 0	97	3.91 (1.63, 9.36)	0.002	107	4.48 (1.85, 10.88)	0.001	102	6.38 (1.37, 29.77)	0.018
Baseline DWM lesion count, ≥1 vs 0	97	4.67 (1.83, 11.92)	0.001	107	7.65 (2.67, 21.92)	<0.001	102	4.80 (1.03, 22.47)	0.046
Baseline JC lesion count, ≥1 vs 0	97	4.05 (1.41, 11.62)	0.009	107	3.67 (1.41, 9.58)	0.008	102	4.00 (1.28, 12.53)	0.017
Baseline IT lesion count, ≥1 vs 0	96	12.41 (3.35, 45.98)	<0.001	106	20.27 (5.43, 75.65)	<0.001	101	4.67 (1.51, 14.41)	0.007
1-year PV lesion count, 1+ vs 0	82	3.69 (1.34, 10.16)	0.11	92	5.41 (1.84, 15.94)	0.002	87	3.52 (0.73, 16.94)	0.116
1-year DWM lesion count, 1+ vs 0	82	10.65 (2.85, 39.84)	<0.001	92	14.93 (3.27, 68.18)	<0.001	87	7.19 (0.89, 58.12)	0.064
1-year JC lesion count, 1+ vs 0	82	4.76 (1.70, 13.31)	0.003	92	3.80 (1.48, 9.73)	0.005	87	2.84 (0.88, 9.16)	0.080
1-year IT lesion count, 1+ vs 0	82	11.1 (3.32, 37.23)	<0.001	92	19.29 (5.67, 65.58)	<0.001	87	5.93 (1.75, 20.09)	0.004
5-year PV lesion count, 1+ vs 0	81	4.43 (1.73, 11.31)	0.002	88	4.28 (1.69, 10.86)	0.002	85	4.06 (1.02, 16.26)	0.047
5-year DWM lesion count, 1+ vs 0	81	7.75 (2.83, 21.23)	<0.001	88	6.43 (2.36, 17.49)	<0.001	85	5.43 (1.11, 26.52)	0.037
5-year JC lesion count, 1+ vs 0	81	6.00 (2.26, 15.93)	<0.001	88	5.40 (2.05, 14.23)	0.001	85	12.62 (1.55, 102.85)	0.018
5-year IT lesion count, 1+ vs 0	78	8.85 (3.07, 25.50)	<0.001	85	8.21 (3.01, 22.38)	<0.001	82	6.13 (1.51, 24.83)	0.011

For all ordinal variables, ORs apply to per adjacent category.

Table 4.6 Contingency tables of best early MRI predictors of 30-year EDSS ≤3.5 versus EDSS >3.5

	Presence of baseline IT lesion	*No baseline IT lesion*	Total
30-year EDSS >3.5	19	25	44
30-year EDSS ≤3.5	3	49	52
Total	22	74	96

PPV 86%, NPV 66%, sensitivity 43%, specificity 94%, accuracy 71% (95% CI 61%, 80%)

	Presence of DWM lesion at 1 year	*No DMW lesion at 1 year*	Total
30-year EDSS >3.5	35	3	38
30-year EDSS ≤3.5	23	21	44
Total	58	24	82

PPV 60%, NPV 88%, sensitivity 92%, specificity 48%, accuracy 68% (95% CI 57%, 78%)

	Presence of IT lesion at 1 year	*No IT lesion at 1 year*	Total
30-year EDSS >3.5	20	18	38
30-year EDSS ≤3.5	4	40	44
Total	24	58	82

PPV 83%, NPV 69%, sensitivity 53%, specificity 91%, accuracy 73% (95% CI 62%, 82%)

4.3.2 Prediction of SPMS

Given 91% of those with 30-year EDSS>3.5 had SPMS, the strongest MRI predictors for SPMS were the same; the presence of: baseline IT lesion, OR 20.3, p<0.001, accuracy 78% (95% CI 69%, 86%), DWM lesion at one-year, OR 14.9, p<0.001, accuracy 65% (95% CI 55%, 75%) and IT lesion at one-year, OR 19.3, p<0.001, accuracy 65% (95% CI 55%, 75%)(Tables 4.5 and 4.7).

Table 4.7 Contingency tables of best early MRI predictors of SPMS by 30 years

	Presence of baseline IT lesion	*No baseline IT lesion*	Total
CIS or RRMS by 30 years	3	64	67
SPMS by 30 years	19	20	39
Total	22	84	106

PPV 86%, NPV 76%, sensitivity 49%, specificity 96%, accuracy 78% (95% CI 69%, 86%)

	Presence of DWM lesion at 1 year	*No DMW lesion at 1 year*	Total
CIS or RRMS by 30 years	30	28	58
SPMS by 30 years	32	2	34
Total	62	30	92

PPV 52%, NPV 93%, sensitivity 94%, specificity 48%, accuracy 65% (95% CI 55%, 75%)

	Presence of IT lesion at 1 year	*No IT lesion at 1 year*	Total
CIS or RRMS by 30 years	4	54	58
SPMS by 30 years	20	14	34
Total	24	68	92

PPV 83%, NPV 79%, sensitivity 59%, specificity 93%, accuracy 80% (95% CI 71%, 88%)

4.3.3 Prediction of MS-related death

For MS-related mortality, in a survival analysis, the presence of baseline IT lesions had a HR of 3.9 (p= 0.007). The proportion of deaths in those with baseline IT lesions, was 8/23 (35%), whereas in those without these lesions, it was 8/78 (10%). The PPV was 35%, NPV 90%, sensitivity 50%, specificity 82%, accuracy 77% (95% CI 68%, 85%). The presence of 1-year IT lesions had a HR of 5.25 (p=0.003), with corresponding proportions of deaths in 9/26 (35%) of those with such lesions, vs 5/61 (8%) in those without; PPV 35%, NPV 92%, sensitivity 64%, specificity 77%, accuracy 75% (95% CI 64%, 83%).

Tables 4.8 and 4.9 summarize the strongest univariable predictors of EDSS ≤3.5 and of SPMS at 30 years.

Table 4.8 Univariable predictors of EDSS ≤3.5 at 30 years, a) at one year following presentation, and b) at five years following presentation

a) at 1 year	**OR**	**95% CI**	***p* value**
Post-CIS nadir EDSS	1.64	1.14-2.36	0.007
Absence of baseline IT lesion	12.4	3.35-46.0	<0.001
Absence of 1-year DWM lesion	10.65	2.84-39.84	<0.001
Absence of 1-year IT lesion	11.1	3.31-37.22	<0.001

b) at 5 years	**OR**	**95% CI**	***p* value**
5-year EDSS	2.12	1.53-3.00	<0.001
EDSS change from post-CIS to 5-years	2.73	1.61-4.63	<0.001
5-year DWM lesion, >5 vs ≤5	3.59	2.07-6.25	<0.001

Table 4.9 Univariable predictors of SPMS at 30 years, a) at one year following presentation, and b) at five years following presentation

a) at 1 year	**OR**	**95% CI**	***p* value**
Presence of baseline IT lesion	20.3	5.43-75.6	<0.001
Presence of 1-year DWM lesion	14.9	3.27-68.2	<0.001
Presence of 1-year IT lesion	19.3	5.67-65.6	<0.001

a) at 5 years	OR	95% CI	*p* value
5-year EDSS	1.80	1.34-2.43	<0.001
EDSS change from post-CIS to 5-years	2.85	1.42-5.72	<0.003
5-year DWM lesion, >5 vs ≤5	2.99	1.80-4.97	<0.001

4.4 Brain atrophy measures as predictors

Changes in supratentorial atrophy, reflected by changes in TVW, within the first five years, were not found to be significantly different between the 30-year outcomes groups. In prediction of 30-year EDSS ≤3.5 vs >3.5, changes in TVW measurements between baseline and five years did not reach statistical significance, ß= 0.53, p= 0.128.

For subjects classified in the four 30-year groups of CIS, MS with EDSS ≤3.5, MS with EDSS >3.5m, and EDSS 10, there was no significant difference in their early TVW measurements: at baseline, F(4, 65)= 2.70, mean square error (MSE)= 2.78, p= 0.08; at 1-year, F(4, 57)= 0.80, MSE 1.94, p= 0.53; and at 5-years, F(4, 66)= 2.35, MSE= 2.50, p= 0.06.

Potential limitations to these results are discussed in Chapter 5.

4.5 Combined predictive models

Table 4.10 shows the combined predictive models of: EDSS >3.5 at 30 years, EDSS ≤3.0 at 30 years, and SPMS status by 30 years. For each of the outcomes, two models are shown: best early predictors up to one year, and best early predictors up to five years. Overall, IT and DWM lesions were the best predictors, with the addition of post-CIS-nadir-to-five-years EDSS change in the SPMS prediction model. The up to one-year models show that subjects with neither baseline IT nor one-year DWM lesions had a 13%

probability of 30-year EDSS being >3.5 (so 87% probability of 30-year EDSS ≤3.5), while subjects with at least one lesion in both locations had 94% probability (95% CI 83%, 100%) of 30-year EDSS being >3.5, and 94% probability of SPMS by 30 years. In the up to five-years models, it is shown that subjects with five or less DWM lesions at five years and EDSS change of <2 points from nadir to five years, had a 11% probability of developing SPMS by 30 years; conversely, subjects with more than five DWM lesions and an EDSS change of ≥2 points, had a 96% probability (95% CI 86%, 100%) of developing SPMS.

For MS-related death, the best independent early predictors, up to one year, were having one or more IT lesions at one year, HR 3.6 (95% CI 1.1, 11.2; p= 0.03), and the post-CIS nadir EDSS, HR 1.5 per additional EDSS score unit (95% CI 1.2, 2.0; p= 0.003)

Table 4.10: Combined predictive models of 30-year EDSS score and 30-year SPMS status. All models include EDSS 10 at or before 30 years.

30-year EDSS >3.5, best independent predictors up to 1 year						
Predictor	**Odds ratio (95% CI)**	**P-value**	**Predictor combinations**			
Baseline infra-tentorial lesion count, ≥1 vs 0	16.8 (2.0, 139.7)	0.009	0	0	≥1	≥1
1-year deep white matter lesion count, ≥1 vs 0	6.7 (1.7, 26.0)	0.006	0	≥1	0	≥1
Model predicted probabilities for 30-year EDSS >3.5 (95% CI)			13% (0, 26)	49% (33, 64)	–[a]	94% (83, 100)
Model predicted probabilities for 30-year EDSS ≤3.5[b]			87%	51%	–	6%
Model n=80. Overall model accuracy using 0.5 probability cut-off (95% CI): 71% (60, 81).						
[a]There were no subjects with this lesion combination. [b]These probabilities and their CIs are 100% minus the >3.5 probabilities.						
30-year EDSS >3.5, best independent predictors up to 5 years						
Predictor	**Odds ratio (95% CI)**	**P-value**	**Predictor combinations**			
Baseline infra-tentorial lesion count, ≥1 vs 0	8.0 (1.5, 41.4)	0.013	0	0	≥1	≥1
5-yr deep white matter lesion count, >5 vs ≤5	5.1 (1.7, 15.6)	0.004	≤5	>5	≤5	>5
Model predicted probabilities for 30-year EDSS >3.5 (95% CI)			18% (5, 30)	52% (36, 71)	63% (22, 100)	90% (76, 100)

Model predicted probabilities for 30-year EDSS ≤3.5			82%	48%	37%	10%
Model n=79. Overall model accuracy using 0.5 probability cut-off (95% CI): 75% (64, 84).						
30-year SPMS, best independent predictors up to 1 year						
Predictor	**Odds ratio (95% CI)**	**P-value**	**Predictor combinations**			
Baseline infra-tentorial lesion count, ≥1 vs 0	26.0 (3.1, 215.0)	0.003	0	0	≥1	≥1
1-yr deep white matter lesion count, ≥1 vs 0	8.6 (1.8, 41.0)	0.007	0	≥1	0	≥1
Model predicted probabilities for 30-year SPMS (95% CI)			7% (0, 16)	38% (23, 53)	–	94% (83, 100)
Model n=89. Overall model accuracy using 0.5 probability cut-off (95% CI): 79% (69, 87)						
30-year SPMS, best independent predictors up to 5 years						
Predictor	**Odds ratio (95% CI)**	**P-value**	**Predictor combinations**			
5-yr deep white matter lesion count, >5 vs ≤5	5.3 (1.7, 16.6)	0.005	≤5	≤5	>5	>5
EDSS score change from nadir to 5-year, ≥2 vs <2	31.1 (3.5, 279.9)	0.002	<2	≥2	<2	≥2
Model predicted probabilities for 30yr SPMS (95% CI)			11% (2, 21)	80% (46, 100)	41% (24, 57)	96% (86, 100)
Model n=85. Overall model accuracy using 0.5 probability cut-off (95% CI): 78% (67, 86)						

Chapter 5: Summary and Conclusions

5.1 Overview

This cohort provides a unique perspective on the very long-term clinical and MRI evolution of relapse-onset MS. As MR imaging first became available in the 1980's, and DMTs in the late 1990's, it is highly unlikely that such long-term, natural history data, from an essentially untreated cohort of people with MS, can be obtained again.

The study has several additional unique advantages: firstly, the prospective design from the start of the study would minimise poor reliability of data captured early on, although some limitations are noted (discussed in Section 5.5). Secondly, the long 30-year time span increases the likelihood of reaching the end-point (Compston and Coles, 2008; Scalfari A, Knappertz V, Cutter G, Goodin DS, Ashton R, 2013). Lastly, the high ascertainment of the cohort and data, particularly at baseline and at 30 years. The participants with outcome data at 30 years had similar demographic features to the original cohort in terms of age at onset, gender, and CIS presentation (Table 3.2).

This study aimed to examine the very long-term clinical outcome following a CIS, with particular focus on the concept of so called 'benign' MS and its definition. It also aimed to identify early predictors of non-disabling and

disabling outcomes.

To meet the study's objectives, I will now discuss the questions originally posed in Chapter 1:

1. Does 'benign' MS exist? If so, what proportion of people with MS run a 'benign' clinical course over their lifetime?
2. How should we define 'benign' MS?
3. Can we predict a very long-term (30-years) outcome using early and intermediate clinical and radiological features?

5.2 Does benign MS exist?

Results from this study suggests that, 30 years following a CIS, the clinical spectrum was diverse. Approximately one third remained CIS, at least clinically, while in the MS group, there were three distinct outcomes: 1) an RRMS group with mild accrued disability (EDSS ≤3.5) and ran a very stable course, 2) an SPMS group who all had impaired mobility (EDSS >3.5); and 3) a group who have had their lives shortened by MS, all of whom had SPMS. In this study, within the available data, the corresponding percentages within the MS sub-group were: 1) 40% remained a relapsing-remitting course and were fully ambulatory, 2) 33% had SPMS associated with more severe disability with impaired ambulation, and 3) MS contributed to their death in 20%.

The results also suggest that, at 30 years, cognitive assessment scores in the EDSS ≤3.5 group, were not significantly different from the CIS group. Everyone in the EDSS ≤3.5 either remained in employment (full time or part time), or retired at the expected age. Measurements of other non-motor features, such as mood, fatigue, and occupational status, were also not significantly different between the MS EDSS ≤3.5 group and the CIS group.

Although the cohort was largely untreated, 11 people were treated with first-line injectable DMTs. While this is too small a sample for meaningful statistical interpretation, it is interesting to note that a larger proportion of those treated had SPMS by 30 years (7 out of 11), compared to remaining RRMS (4 out of 11). This may be due to reverse causality, suggesting that when the first generation DMTs became available in the late 1990's, those who were more severely affected, were more likely to be treated.

From this cohort, there is good evidence that a relatively long-term 'benign' clinical outcome in MS does exist, and in this cohort, using an EDSS threshold of ≤3.5, 40% of individuals with MS would meet this requirement. This is a higher proportion than in other longitudinal studies, and potential factors that have contributed towards this favourable outcome are discussed in Section 5.5.

5.3 How should we define benign MS?

This study used well-established clinical outcome measures. This is less

controversial for physically disabling outcomes such as SPMS or MS-related deaths, but what is considered ‘benign’ may differ substantially depending on whose perspective it is from, and patient-reported outcomes have not been assessed (Sayao *et al.*, 2007; Hawkins, 2012; Tallantyre *et al.*, 2018). While a low EDSS score indicates good ambulation and lack of severe disability, it can range from zero where an individual has no neurological symptoms and no signs on objective examination, to the upper limit of 3.5 where one may have moderate to severe symptoms and/or limitations in at least one or more functional systems (Table 1.2). A low EDSS score also does not mean that MS symptoms are not distressing or functionally significant. For example, we have not included detailed assessments of fatigue and mood, which may significantly affect functional outcomes in people with MS. We also did not explore whether any of the individuals opted for careers that would be considered less challenging or stressful, as a result of their MS. A recent population-based study demonstrated that agreement between patients and clinicians on the idea of ‘benign’ MS status was lacking (Tallantyre *et al.*, 2018).

Interestingly, in a study exploring health-related quality of life in patients with ‘benign’ MS, little difference was noted between MS patients with EDSS ≤3.0 and those with EDSS >3.0, in terms of cognitive function, sexual satisfaction, energy, emotional well-being, and overall quality of life. Depression and fatigue were the two metrics most associated with a lower quality of life (Bueno *et al.*, 2015). This would suggest that other than physical disabilities, mood and fatigue are also key factors in determining quality of life in people

with MS.

From a practical point of view, however, as DMTs do not specifically treat symptoms such as depression and fatigue, these findings are unlikely to alter decisions regarding DMT management of such patients. Instead, management of non-motor symptoms will depend on their severity, and management will be independent of the individual's DMT treatment.

The other main caveat of trying to define 'benign' MS, is that these observations are currently made retrospectively, and it is the ability to predict these outcomes early on in the disease course, that is more needed in clinical practice. The latency to the onset of SPMS is also highly variable, and remains possible beyond 30 years. Although this current study described a very long duration, it remains difficult to confidently state when an individual's outcome is truly stable or not: while the likelihood of secondary progression beyond 30 years is presumed to be less, it remains a possibility. Indeed in the Gothenburg population-based, natural history cohort, some people were observed to transition to the secondary progressive phase of the condition, more than 40 years after first symptom onset (Skoog *et al.*, 2012).

It is perhaps worth considering the clinical outcomes of CNS demyelination, as along a continuous clinical spectrum: from incidental post-mortem findings of MS plaques, to sub-clinical radiologically isolated syndromes (RIS), to people with a monophasic CIS, through to those who develop highly active

and disabling MS early on. In hospital-based autopsy series, it is estimated that 0.06% to 0.7% of patients without previously known neurological symptoms during life and without a diagnosis of MS, were found to have pathologically confirmed demyelination (Gilbert and Sadler, 1983; Forslin *et al.*, 2016). With brain imaging increasingly being performed for an array of symptoms, an increasing number of people are being identified with RIS, where there is evidence of demyelinating lesions on MRI suggestive of MS, in the absence of clinical symptoms. Less than half such cases develop symptoms within a 5-year follow-up period (Okuda, 2017). Similarly, approximately one third of people in the present cohort appeared to only have had a monophasic neuro-inflammatory episode. Indeed, the lack of significant differences between the 30-year CIS and RRMS groups, would further lend to this idea that they are part of the same continuum, instead of distinct separate entities.

A concise, reliable and clinically practical definition of ‘benign’ MS has yet to be established. Several definitions have been suggested but none are without their flaws. While those based on disability (defined by EDSS) after a given disease duration may better reflect the functional consequences of MS, they do not necessarily guarantee a similar favourable outcome in the future. In contrast, definitions based on clinical stability do not necessarily reflect the residual impact of a CIS. Patient reported outcomes have also not been considered. In the disease course and phenotype classifications recently proposed by Lublin (Lublin, 2014), more favourable outcomes may be better defined as ‘non-progressive (i.e. with a low EDSS), inactive, with or without

functionally significant deficits arising from previous relapses'.

With these in mind, the term 'benign' should be used with caution, as it has a tendency to undermine non-motor symptoms, and can potentially cause complacency in both the patient and the physician. Instead, the non-disabling MS sub-group identified in this study may more pragmatically be considered to be people with MS who have consistently low levels of disability with no progression, and therefore less to gain from DMTs, rather than those who have no ill effects, or symptoms, from MS.

5.4 Can we predict a very long-term (30-years) outcome using early and intermediate clinical and radiological features?

After allowing for other factors, 30-year outcomes were not independently associated with age at onset, gender, baseline EDSS and CIS type. Some of the features, such as early age at symptom onset and female sex, which were previously considered good prognostic markers (Confavreux and Vukusic, 2006), were found to be lesser predictors of long-term outcomes, than EDSS and radiological biomarkers. While brain atrophy has been shown to occur very early in the course of MS (Brex *et al.*, 2000; Dalton *et al.*, 2002a), it remains unknown at which point it becomes clinically relevant. In this study, early atrophy was not significantly associated with 30-year outcomes, although the association of atrophy and disability becomes much more apparent, later on in the disease course. While the findings may be related to study limitations (discussed in Section 5.5), it may also be possible

that processes underlying brain atrophy are initiated soon after the clinical onset of MS, but has less clinical impact until a certain threshold is reached.

In this study, 30-year outcomes could, in part, be predicted by early EDSS scores and more robustly by MRI-derived regional white matter lesion counts: in multivariable models including both lesion and EDSS variables, the latter no longer contributed significantly, with their coefficients substantially reduced. As changes in lesion counts require data from two adjacent time-points, there was more missing data for changes in lesion counts than for absolute lesion counts, which only require data from one time-point (Table 4.5); this may be a factor in why new early lesions were not as predictive. The same would apply to changes in TVW measurements.

Interestingly, while PV and JC lesions are highly relevant in the diagnosis of MS (Polman *et al.*, 2011), it was early IT and DWM lesions that had the greatest long-term prognostic value. For example, in people with at least one baseline IT lesion and at least one DWM lesion by one year, the chances of having SPMS were 94%, while conversely in those with neither baseline IT nor 1-year DWM lesion, the probability of having EDSS ≤3.5 by 30 years, was 87%. Those with one or more IT lesions by one year, were five times more likely to have died due to, or related to, MS than the rest of the cohort. Infratentorial lesions have previously been linked with less favourable outcomes in people with MS, after a mean follow-up of 7.7 years (Tintore *et al.*, 2010a).

Considering the potential application of these results, treatment decisions are often made prospectively and increasingly early in the disease course, prognostic factors identified within a year of symptoms onset may prove more useful than those identified within five years. However, favourable prognostic features at one-year may also not impact significantly on choices; instead the emergence of markers suggestive of more disabling outcomes later in the disease course, may carry more weight.

5.5 Shortfalls

There are several limitations in this study, and they are discussed in this section.

5.5.1 Changes in the McDonald diagnostic criteria

Since this cohort was first established, the diagnostic criteria for MS have been revised on several occasions (see section 1.2.4 and Table 1.1), and most significantly an MS diagnosis can now be made in people after only a single clinical episode of symptoms, but who fulfil MRI and/ or CSF criteria for dissemination of lesions in space and time (Polman *et al.*, 2011; Thompson *et al.*, 2018a). In one study of CIS patients, some 11-15% remained CIS clinically with no second episode of symptoms, but developed new brain or spinal cord lesions on MR imaging, thus fulfilling radiological criteria for MS (Chard *et al.*, 2011). In the present study, six individuals had MS diagnosed on radiological grounds only, all of whom had EDSS ≤3.5 at

30 years, compared with 37% if diagnosed on clinical grounds. Therefore, they appear to represent a clinically silent end of the MS spectrum, who would previously have been overlooked. In the 30 people classified in the 30-year CIS outcome group, nine were not imaged, and there was little data on their CSF oligoclonal bands status. It is therefore possible that they may in fact fulfil diagnostic criteria and be classified as having MS, and may have been overlooked. If this is the case, it is possible that the more favourable end of the scale has been relatively under-estimated, although this is unlikely to change the main findings of this study.

The revisions of the McDonald diagnostic criteria have led to the earlier diagnosis of MS following a CIS, with increased sensitivity and specificity (Montalban *et al.*, 2010; Brownlee *et al.*, 2015). These changes will likely shift and increase the proportion of people with MS that are at the clinically more silent end of the spectrum, and are more likely to fare well over time. Additionally, with the now routine use of DMTs, the long-term evolution of MS may be changing: in a large cohort of RRMS patients, where 62% were treated with DMTs, only 11.3% transitioned to SPMS after a 17-year period (Cree *et al.*, 2016).

5.5.2 EDSS data

With regard to EDSS data, particularly early in the follow-up time-points, these were not captured consistently. In order to minimize inaccuracies, data from adjacent time-points, and from clinical records, where available, were

also used. However, it is worth noting that EDSS scores of ≤3.5 are derived from symptoms and examination findings, while scores from ≥4 upwards represent thresholds of ambulatory impairment. Although a >3.0 vs ≤3.0 threshold has been proposed in the literature on 'benign' MS, our use of an EDSS threshold of ≤3.5 vs >3.5 at 30-years is more objectively interpretable, should minimize the impact of any inter- or intra-rater variabilities, and be more reliable in the predictive models. Further, only one participant in our sample would have been reclassified if a >3.0 vs ≤3.0 dichotomy was used, with little impact on the main results.

In a small number of individuals, symptoms or signs due to non-neurological causes that would otherwise impact on an EDSS score, were omitted. For example, in one participate who had a recent hip injury, the need for a walking aid was temporary and therefore did not contribute to her EDSS score. As the mean age at 30-years was 62 years of age, some symptoms, for example sensory or bladder symptoms, may not be uncommon and can often be difficult to differentiate whether they were directly MS-related or not. It is therefore possible that symptoms not due to MS have contributed towards individuals' EDSS scores, particularly when EDSS was ≤3.5, due to reasons outlined above.

5.5.3 Radiological factors

At the inception of this cohort, MR imaging was a new technique, and was only at the very early stages of routine clinical use. Section 2.4.4

demonstrated the image acquisition details and sample images at different time-points. In particular, some of the very early scans had slice thickness of 10mm. Significant progress and improvement in scanner technology have been made since, and the earlier image quality was therefore not as good as is achievable now. Given this, analyses of the earlier images will be less reliable than later ones. However, with the digital reconstruction of old films (described in section 2.5), use of digital software and availability of 3T images obtained at 30 years, we were able to retrospectively visualise lesions on scans obtained from earlier time-points, much more easily.

The current routine methods used to measure whole brain atrophy could not be applied to the early scans, and therefore we were only able to measure atrophy using linear measurements. While linear measures can approximate volumetric measures, they have nevertheless proven to be less sensitive to cross-sectional differences and longitudinal changes (Turner *et al.*, 2001; Butzkueven *et al.*, 2008). Additionally, as differences in longitudinal measurements were being calculated, TVW measurements were undertaken on PD/T2-weighted images, as these were the only images available at all time-points. The lack of significance between early TVW measurements and long-term outcome, could be related to these limitations.

Post-gadolinium sequences were not obtained at any time-point, and only limited T1-weighted images were obtained at 14- and 20-years, and as such we have not been able to assess active lesion inflammation at the time of scanning, or assess the early relevance of T1 ‘black holes’ for longer-term

outcomes.

MS commonly affects the spinal cord, with spinal cord lesions found in up to ~90% of patients with known MS (Kidd *et al.*, 1993), and we now know the increasing relevance of lesions and atrophy in the spinal cord in MS (Casserly *et al.*, 2018; Brownlee *et al.*, 2019). In this cohort, however, spinal cord imaging was not routinely obtained.

At 30 years, the mean age of the cohort was 61 years, it was therefore likely that at least some of the white matter lesions visualised, had a vascular aetiology. However, there was no clear correlation between the total number of intracranial lesions at 30-years, and known vascular risk factors (Figure 3.8). Despite consultation with experts in the Neuroradiology fields, it remained impossible to determine whether some lesions were vascular or demyelinating in aetiology. However, it was felt that as every participant in the cohort had gained the same number of years in age, the net effect of inaccurate vascular lesions contributing to the results, would be minimal. Furthermore, the value of early MR findings remained more important in long-term prognostication.

5.5.4 Statistics

It should be highlighted that, in the predictive models, instead of only looking at the subgroup which eventually had MS by 30 years, people classified as CIS were included throughout. This was felt to be essential, as this reflects

clinical practice: at any given time-point, those remaining in CIS may be indistinguishable from those who eventually become diagnosed as MS. This may also be increasingly relevant clinically, as there is now increasing tendency to start a DMT after an episode of CIS only, before a diagnosis of MS is reached.

It should be noted that classification properties of our multivariable models might be improved with probability cut-offs different from 0.5; however, we believed this optimisation may not be reliably generalizable, and preferred not to screen for the best classification.

5.5.5 Presenting CIS

It is worth noting that optic neuritis as a presenting CIS has been associated with a more favourable outcome when compared with other CIS presentations (Ramsaransing and De Keyser, 2006). In this cohort, there was a slight over-representation of people presenting with optic neuritis: 52% of the participants presented with ON, compared to 35-40% in other large prospective CIS series (Eriksson *et al.*, 2003; Scalfari *et al.*, 2014; Tintore *et al.*, 2015). Although there was no evidence of significant association between ON and 30-year outcomes in this study, it is worth noting that whilst 52% of the participants presented with ON at baseline, a higher proportion (63%) remained CIS by 30-years, and there may still be some bias towards more favourable outcomes (Table 3.6).

5.5.6 Missing data and miscellaneous

Other limitations of the study included the small cohort size, and missing data.

We were not able to obtain outcome data for 12 of the 132 original participants. If an assumption is made that all 12 had SPMS with EDSS >3.5, then the corresponding figures would adjust to (figures in brackets refer to current study figures, presented in Section 5.2): 1) 35% (40%) remained a relapsing-remitting course and were fully ambulatory, 2) 41% (33%) had SPMS associated with more severe disability with impaired ambulation, and 3) MS contributed to their death in 20%.

We were also not able to obtain MRI scans in nine of the 30 classified as CIS at 30-years: it is possible that some of these individuals would fulfil MS diagnostic criteria, although this is unlikely to change the main findings of this study.

Twelve participants did not contribute to early MRI information, of whom four were classified at CIS, one had RRMS, and seven were lost to 30-year follow-up; it is likely that this missing radiological data has reduced our power to detect early predictors. It should also be noted that the cohort originated from one neurosciences centre, and therefore there may be limitations in generalizability.

Several comorbidities have been found to be associated with an increased risk of death in MS: depression, diabetes, ischaemic heart disease, depression, anxiety and chronic respiratory disease (Marrie *et al*., 2015). In this study, those with known MS and who were deceased, cause(s) of death were obtained from death certificates in 27 out of 29 individuals. Data on comorbidities were not available, unless they were considered a primary or secondary cause of death. It is therefore feasible that comorbidities contributed towards an increased hazard in mortality but were missed or under-recognised.

5.6 Conclusions

In conclusion, the results of this study suggest a divergence of natural outcomes in people with relapse-onset MS, 30 years after symptom onset: those with SPMS, who have developed greater disability and have a significant risk of their life being shortened by MS; and those classified as having RRMS, who remained fully ambulatory, with no significant cognitive impairment, and who remained employed or retired at the expected age. The results also indicate that for less favourable outcomes, the die may be cast early. This suggests that there are people with MS who have more to gain from earlier use of higher efficacy DMTs, although also counsels caution when considering the blanket use of DMTs in early MS or following a CIS. It is not understood why some patients fare well over time, and the underlying factors that contribute or lead to the development of early disability accrual, and the factors that may have a neuroprotective role, resulting in a more

favourable disease course, remain areas of ongoing research. A key goal of future research is to determine what pathologically differentiates progressive from persistently non-progressive MS, with a view to targeting treatments that would substantially increase the chances a person with MS follow a less-disabling clinical course.

The limited ability to predict long-term disability, or relative absence of it, is one of the greatest challenges clinicians face. The predictive models developed by this study include features that can be obtained in routine clinical practice, and so hopefully may inform risk-benefit analyses when considering treatment options.

References

Absinta M, Sati P, Masuzzo F, Nair G, Sethi V, Kolb H, et al. Association of chronic active multiple sclerosis lesions with disability in vivo. *JAMA Neurol*. 2019; 76(12):1474–1483.

Amato MP, Ponziani G, Bartolozzi ML, Siracusa G. A prospective study on the natural history of multiple sclerosis: Clues to the conduct and interpretation of clinical trials. *J. Neurol. Sci*. 1999; 168: 96–106.

Amato MP, Portaccio E. Truly benign multiple sclerosis is rare: let's stop fooling ourselves-Yes. *Mult. Scler. J*. 2012; 18: 13–14.

Amato MP, Zipoli V, Goretti B, Portaccio E, De Caro MF, Ricchiuti L, et al. Benign multiple sclerosis: Cognitive, psychological and social aspects in a clinical cohort. *J. Neurol.* 2006; 253: 1054–1059.

Ascherio A, Munger K, others. Epidemiology of multiple sclerosis: from risk factors to prevention. *Semin. Neurol.* 2008; 28: 17–28.

Bagnato F. Evolution of T1 black holes in patients with multiple sclerosis imaged monthly for 4 years. *Brain* 2003; 126: 1782–1789.

Barkhof F. MRI in multiple sclerosis: correlation with expanded disability status scale (EDSS). *Mult. Scler*. 1999; 5: 283–286.

Benedict RHB, Weinstock-Guttman B, Fishman I, Sharma J, Tjoa CW, Bakshi R. Prediction of neuropsychological impairment in multiple sclerosis: comparison of conventional magnetic resonance imaging measures of

atrophy and lesion burden. *Arch. Neurol.* 2004; 61: 226–30.

Benedikz J, Stefánsson M, Guðmundsson J, Jónasdóttir A, Fossdal R, Gulcher J, et al. The natural history of untreated multiple sclerosis in Iceland. A total population-based 50 year prospective study. *Clin. Neurol. Neurosurg.* 2002; 104: 208–210.

Bergsland N, Horakova D, Dwyer MG, Dolezal O, Seidl ZK, Vaneckova M, et al. Subcortical and cortical gray matter atrophy in a large sample of patients with clinically isolated syndrome and early relapsing-remitting multiple sclerosis. *AJNR Am. J. Neuroradiol*. 2012; 33: 1573–1578.

Bermel RA, Bakshi R. The measurement and clinical relevance of brain atrophy in multiple sclerosis. *Lancet Neurol.* 2006; 5: 158–170.

Bermel RA, Sharma J, Tjoa CW, Puli SR, Bakshi R. A semiautomated measure of whole-brain atrophy in multiple sclerosis. *J. Neurol. Sci.* 2003; 208: 57–65.

Bitsch A, Kuhlmann T, Stadelmann C, Lassmann H, Lucchinetti C, Brück W. A longitudinal MRI study of histopathologically defined hypointense multiple sclerosis lesions. *Ann. Neurol.* 2001; 49: 793–796.

Bonek R, Orlicka K, Maciejek Z. Demyelinating lesions in the cervical cord in multiple sclerosis 10 years after onset of the disease. Correlation between MRI parameters and clinical course. *Neurol. Neurochir. Pol.* 2007; 41: 229–233.

Brex PA, Jenkins R, Fox NC, Crum WR, O'Riordan JI, Plant GT, et al.

Detection of ventricular enlargement in patients at the earliest clinical stage of MS. *Neurology* 2000; 54: 1689–1691.

Brownlee WJ, Altmann DR, Prados F, Miszkiel KA, Eshaghi A, Gandini Wheeler-Kingshott CAM, et al. Early imaging predictors of long-term outcomes in relapse-onset multiple sclerosis. *Brain* 2019; 142: 2276–2287.

Brownlee WJ, Swanton JK, Altmann DR, Ciccarelli O, Miller DH. Earlier and more frequent diagnosis of multiple sclerosis using the McDonald criteria. *J. Neurol. Neurosurg. Psychiatry* 2015; 86: 584–585.

Bueno A-M, Sayao A-L, Yousefi M, Devonshire V, Traboulsee A, Tremlett H. Health-related quality of life in patients with longstanding 'benign multiple sclerosis'. *Mult. Scler. Relat. Disord.* 2015; 4: 31–38.

Butzkueven H, Kolbe SC, Jolley DJ, Brown JY, Cook MJ, van der Mei IAF, et al. Validation of linear cerebral atrophy markers in multiple sclerosis. *J. Clin. Neurosci.* 2008; 15: 130–137.

Cabre P, Heinzlef O, Merle H, Buisson GG, Bera O, Bellance R, et al. MS and neuromyelitis optica in Martinique (French West Indies). *Neurology* 2001; 56: 507–514.

Calabrese M, Rinaldi F, Mattisi I, Bernardi V, Favaretto A, Perini P, et al. The predictive value of gray matter atrophy in clinically isolated syndromes. *Neurology* 2011; 77: 257–263.

Calabrese M, Romualdi C, Poretto V, Favaretto A, Morra A, Rinaldi F, et al. The changing clinical course of multiple sclerosis: A matter of gray matter.

Ann. Neurol. 2013; 74: 76–83.

Calabrese M, De Stefano N, Atzori M, Bernardi V, Mattisi I, Barachino L, et al. Detection of Cortical Inflammatory Lesions by Double Inversion Recovery Magnetic Resonance Imaging in Patients With Multiple Sclerosis. *Arch. Neurol.* 2007; 64: 1416.

Casserly C, Seyman EE, Alcaide-Leon P, Guenette M, Lyons C, Sankar S, et al. Spinal Cord Atrophy in Multiple Sclerosis: A Systematic Review and Meta-Analysis. *J. Neuroimaging* 2018; 28: 556–586.

Chard DT, Dalton CM, Swanton JK, Fisniku LK, Miszkiel K, Thompson AJ, et al. MRI only conversion to multiple sclerosis following a clinically isolated syndrome. *J. Neurol. Neurosurg. Psychiatry* 2011; 82: 176–179.

Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann. Neurol.* 1983; 13: 227–231.

Coles A. Alemtuzumab Treatment of Multiple Sclerosis. *Semin. Neurol.* 2013; 33: 066–073.

Comi G, Filippi M, Barkhof F, Durelli L, Edan G, Fernández O, et al. Effect of early interferon treatment on conversion to definite multiple sclerosis: a randomised study. *Lancet* 2001; 357: 1576–1582.

Comi G, Martinelli V, Rodegher M, Moiola L, Bajenaru O, Carra A, et al. Effect of glatiramer acetate on conversion to clinically definite multiple sclerosis in patients with clinically isolated syndrome (PreCISe study): a

randomised, double-blind, placebo-controlled trial. *Lancet* 2009; 374: 1503–1511.

Compston A, Coles A. Multiple sclerosis. *Lancet* 2008; 372: 1502–17.

Confavreux C. Early clinical predictors and progression of irreversible disability in multiple sclerosis: an amnesic process. *Brain* 2003; 126: 770–782.

Confavreux C, Vukusic S. Natural history of multiple sclerosis: a unifying concept. *Brain* 2006; 129: 606–616.

Correale J, Peirano I, Romano L. Benign multiple sclerosis: a new definition of this entity is needed. *Mult. Scler. J*. 2012a; 18: 210–218.

Correale J, Ysrraelit MC, Fiol MP. Benign Multiple Sclerosis: Does it exist? *Curr. Neurol. Neurosci. Rep.* 2012b; 12: 601–609.

Costelloe L, Thompson A, Walsh C, Tubridy N, Hutchinson M. Long-term clinical relevance of criteria for designating multiple sclerosis as benign after 10 years of disease. *J. Neurol. Neurosurg. Psychiatry* 2008; 79: 1245–1248.

Cree BAC, Gourraud PA, Oksenberg JR, Bevan C, Crabtree-Hartman E, Gelfand JM, et al. Long-term evolution of multiple sclerosis disability in the treatment era. *Ann. Neurol.* 2016; 80: 499–510.

Dalton CM, Brex PA, Jenkins R, Fox NC, Miszkiel KA, Crum WR, et al. Progressive ventricular enlargement in patients with clinically isolated syndromes is associated with the early development of multiple sclerosis. *J.*

Neurol. Neurosurg. Psychiatry 2002a; 73: 141–147.

Dalton CM, Brex PA, Miszkiel KA, Hickman SJ, MacManus DG, Plant GT, et al. Application of the new McDonald criteria to patients with clinically isolated syndromes suggestive of multiple sclerosis. *Ann. Neurol.* 2002b; 52: 47–53.

Dalton CM, Chard DT, Davies GR, Miszkiel KA, Altmann DR, Fernando K, et al. Early development of multiple sclerosis is associated with progressive grey matter atrophy in patients presenting with clinically isolated syndromes. *Brain* 2004; 127: 1101–1107.

Di Filippo M, Anderson VM, Altmann DR, Swanton JK, Plant GT, Thompson AJ, et al. 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* 2010; 81: 204–208.

Degenhardt A, Ramagopalan S V, Scalfari A, Ebers GC. Clinical prognostic factors in multiple sclerosis: a natural history review. *Nat. Rev. Neurol.* 2009; 5: 672–682.

Dobson R, Ramagopalan S, Giovannoni G. The effect of gender in clinically isolated syndrome (CIS): a meta-analysis. *Mult. Scler.* 2012; 18: 600–4.

Ebner M, Chung KK, Prados F, Cardoso MJ, Chard DT, Vercauteren T, et al. Volumetric reconstruction from printed films: Enabling 30 year longitudinal analysis in MR neuroimaging. *Neuroimage* 2018; 165: 238–250.

Edan G, Le Page E. Induction Therapy for Patients with Multiple Sclerosis:

Why? When? How? *CNS Drugs* 2013; 27: 403–409.

Eriksson M, Andersen O, Runmarker B. Long-term follow up of patients with clinically isolated syndromes, relapsing-remitting and secondary progressive multiple sclerosis. *Mult. Scler. J.* 2003; 9: 260–274.

Eshaghi A, Prados F, Brownlee WJ, Altmann DR, Tur C, Cardoso MJ, et al. Deep gray matter volume loss drives disability worsening in multiple sclerosis. *Ann. Neurol.* 2018; 83: 210–222.

Filippi M, Campi A, Mammi S, Martinelli V, Locatelli T, Scotti G, et al. Brain magnetic resonance imaging and multimodal evoked potentials in benign and secondary progressive multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry*. 1995; 58: 31–37.

Filippi M, Campi A, Martinelli V, Colombo B, Scotti G, Comi G. Brain and spinal cord MR in benign multiple sclerosis: A follow-up study. *J. Neurol. Sci.* 1996; 143: 143–149.

Filippi M, Iannucci G, Tortorella C, Minicucci L, Horsfield MA, Colombo B, et al. Comparison of MS clinical phenotypes using conventional and magnetization transfer MRI. *Neurology* 1999; 52: 588–594.

Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *Lancet Neurol.* 2016; 15: 292–303.

Fischer JS, Rudick RA, Cutter GR, Reingold SC. The Multiple Sclerosis Functional Composite measure (MSFC): an integrated approach to MS

clinical outcome assessment. *Mult. Scler. J*. 1999; 5: 244–250.

Fisniku LK, Brex PA, Altmann DR, Miszkiel KA, Benton CE, Lanyon R, et al. Disability and T2 MRI lesions: A 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain* 2008a; 131: 808–817.

Fisniku LK, Chard DT, Jackson JS, Anderson VM, Altmann DR, Miszkiel KA, et al. Gray matter atrophy is related to long-term disability in multiple sclerosis. *Ann. Neurol.* 2008b; 64: 247–254.

Forslin Y, Granberg T, Jumah AA, Shams S, Aspelin P, Kristoffersen-Wiberg M, et al. Incidence of Radiologically Isolated Syndrome: A Population-Based Study. *Am. J. Neuroradiol.* 2016; 37: 1017–1022.

Fox EJ, Rhoades RW. New treatments and treatment goals for patients with relapsing-remitting multiple sclerosis. *Curr. Opin. Neurol.* 2012; 25: S11–S19.

Fox RJ, Salter AR, Tyry T, Sun J, You X, Laforet G, et al. Treatment Discontinuation and Disease Progression with Injectable Disease-Modifying Therapies. *Int. J. MS Care* 2013; 15: 194–201.

Freedman MS, Comi G, De Stefano N, Barkhof F, Polman CH, Uitdehaag BMJ, et al. Moving toward earlier treatment of multiple sclerosis: Findings from a decade of clinical trials and implications for clinical practice. *Mult. Scler. Relat. Disord.* 2014; 3: 147–155.

Frischer JM, Weigand SD, Guo Y, Kale N, Parisi JE, Pirko I, et al. Clinical and pathological insights into the dynamic nature of the white matter multiple

sclerosis plaque. *Ann Neurol*. 2015; 78: 710–721.

Frohman EM, Racke MK, Raine CS. Multiple Sclerosis — The Plaque and Its Pathogenesis. *N. Engl. J. Med.* 2006; 354: 942–955.

Gilbert JJ, Sadler M. Unsuspected Multiple Sclerosis. *Arch. Neurol.* 1983; 40: 533–536.

Glad SB, Nyland H, Aarseth JH, Riise T, Myhr K-M. How long can you keep working with benign multiple sclerosis? *J. Neurol. Neurosurg. Psychiatry* 2011; 82: 78–82.

Hawkins S. Truly benign multiple sclerosis is rare: let's stop fooling ourselves – No. *Mult. Scler. J.* 2012; 18: 11–12.

Hawkins SA, Mcdonnell G V. Benign multiple sclerosis? Clinical course, long term follow up, and assessment of prognostic factors. *J Neurol Neurosurg Psychiatry* 1999; 67: 148–152.

Hirst C, Ingram G, Swingler R, Compston DAS, Pickersgill T, Robertson NP. Change in disability in patients with multiple sclerosis: a 20-year prospective population-based analysis. *J. Neurol. Neurosurg. Psychiatry* 2008; 79: 1137–1143.

IBM Corp. IBM SPSS Statistics for Windows.

Jacobsen C, Hagemeier J, Myhr KM, Nyland H, Lode K, Bergsland N, Ramasamy DP, Dalaker TO, Larsen JP, Farbu E, Zivadinov R. Brain atrophy and disability progression in multiple sclerosis patients: a 10-year follow-up

study. *J. Neurol. Neurosurg. Psychiatry* 2014; 85: 1109–1115

Kalanie H, Gharagozli K, Kalanie AR. Multiple sclerosis: report on 200 cases from Iran. *Mult. Scler. J.* 2003; 9: 36–38.

Kappos L, Polman CH, Freedman MS, Edan G, Hartung HP, Miller DH, et al. Treatment with interferon beta-1b delays conversion to clinically definite and McDonald MS in patients with clinically isolated syndromes. *Neurology* 2006; 67: 1242–1249.

Kearney H, Rocca M a, Valsasina P, Balk L, Sastre-Garriga J, Reinhardt J, et al. Magnetic resonance imaging correlates of physical disability in relapse onset multiple sclerosis of long disease duration. *Mult. Scler.* 2014; 20: 72–80.

Kidd D, Thorpe JW, Thompson AJ, Kendall BE, Moseley IF, MacManus DG, et al. Spinal cord MRI using multi-array coils and fast spin echo: II. Findings in multiple sclerosis. *Neurology* 1993; 43: 2632–2632.

Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch. Neurol.* 1989; 46: 1121–3.

Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983; 33: 1444–1452.

Kurtzke JF, Beebe GW, Nagler B, Kurland LT, Auth TL. Studies on the natural history of multiple sclerosis—8. *J. Chronic Dis*. 1977; 30: 819–830.

Langdon D. A useful annual review of cognition in relapsing MS is beyond most neurologists - NO. *Mult. Scler. J.* 2016; 22: 728–730.

Langdon DW, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson S, et al. Recommendations for a brief international cognitive assessment for multiple sclerosis (BICAMS). *Mult. Scler. J.* 2012; 18: 891–898.

Lauer K, Firnhaber W. Epidemiological investigations into multiple sclerosis in Southern Hesse. V. Course and prognosis. *Acta Neurol. Scand*. 1987; 76: 12–7.

Lechner-Scott L, Kappos L, Hofman M, Polman CH, Ronner H, Montalban X, et al. Can the Expanded Disability Status Scale be assessed by telephone? *Mult. Scler.* 2003; 9: 154–159.

Lublin, F, Reingold S. Defining the clinical course of multiple sclerosis: results of an international survey. *Neurology* 1996; 46: 907–911.

Lublin FD. New Multiple Sclerosis Phenotypic Classification. *Eur. Neurol.* 2014; 72: 1–5.

Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* 2014; 83: 278–286.

Luchetti S, Fransen NL, van Eden CG, Ramaglia V, Mason M, Huitinga I. Progressive multiple sclerosis patients show substantial lesion activity that correlates with clinical disease severity and sex: a retrospective autopsy

cohort analysis. *Acta Neuropathol*. 2018; 135: 511–528.

Marrie RA, Elliott L, Marriott J, Cossoy M, Blanchard J, Leung S, et al. Effect of comorbidity on mortality in multiple sclerosis. *Neurology* 2015; 85: 240–247.

Mastorodemos V, Nikolakaki H, Tzagournissakis M, Kotzamani D, Panou T, Spanaki C, et al. Benign multiple sclerosis in Crete. *Mult. Scler. J*. 2010; 16: 701–706.

Mcalpine D. The benign form of multiple sclerosis. A study based on 241 cases seen within three years of onset and followed up until the tenth year or more of the disease. *Brain* 1961; 84: 186–203.

McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. *Ann. Neurol*. 2001; 50: 121–127.

McDonnell GV, Mawhinney H, Graham CA, Hawkins SA, Middleton D. A study of the HLA-DR region in clinical subgroups of multiple sclerosis and its influence on prognosis. *J. Neurol. Sci.* 1999; 165: 77–83.

Miller DH, Brex PA, Sailer M, O'Riordan JI, Ciccarelli O, Thompson AJ. A Longitudinal Study of Abnormalities on MRI and Disability from Multiple Sclerosis. *N. Engl. J. Med.* 2002; 346: 158–164.

Miller DH, Ormerod IE, McDonald WI, MacManus DG, Kendall BE, Kingsley DP, et al. The early risk of multiple sclerosis after optic neuritis. *J Neurol*

Neurosurg Psychiatry 1988; 51: 1569–1571.

Miller DH, Ormerod IEC, Rudge P, Kendall BE, Moseley IF, McDonald WI. The early risk of multiple sclerosis following isolated acute syndromes of the brainstem and spinal cord. *Ann. Neurol.* 1989; 26: 635–639.

Montalban X, Tintore M, Swanton J, Barkhof F, Fazekas F, Filippi M, et al. MRI criteria for MS in patients with clinically isolated syndromes. *Neurology* 2010; 74: 427–434.

Nelson H. The National Adult Reading Test. Windsor: NFER-Nelson; 1982.

Noseworthy JH, Lucchinetti C, Rodriguez M, Weinshenker BG. Multiple Sclerosis. *N. Engl. J. Med.* 2000; 343: 938–952.

Okuda DT. Radiologically Isolated Syndrome. *Neuroimaging Clin. N. Am.* 2017; 27: 267–275.

Pérez-Miralles F, Sastre-Garriga J, Tintoré M, Arrambide G, Nos C, Perkal H, et al. Clinical impact of early brain atrophy in clinically isolated syndromes. *Mult. Scler.* 2013; 19: 1878–86.

Perini P, Tagliaferri C, Belloni M, Biasi G, Gallo P. The HLA-DR13 haplotype is associated with ‘benign’ multiple sclerosis in northeast Italy. *Neurology* 2001; 57: 158–9.

Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. *Ann. Neurol.* 2011; 69: 292–302.

Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 Revisions to the 'McDonald Criteria'. *Ann. Neurol.* 2005; 58: 840–846.

Ramsaransing GSM, De Keyser J. Benign course in multiple sclerosis: A review. *Acta Neurol. Scand.* 2006; 113: 359–369.

Rocca MA, Battaglini M, Benedict RHB, De Stefano N, Geurts JJG, Henry RG, et al. Brain MRI atrophy quantification in MS. *Neurology* 2017; 88: 403–413.

Roosendaal SD, Bendfeldt K, Vrenken H, Polman CH, Borgwardt S, Radue EW, et al. Grey matter volume in a large cohort of MS patients: Relation to MRI parameters and disability. *Mult. Scler. J.* 2011; 17: 1098–1106.

Rovaris M, Barkhof F, Calabrese M, De Stefano N, Fazekas F, Miller DH, et al. MRI features of benign multiple sclerosis: Toward a new definition of this disease phenotype. *Neurology* 2009; 72: 1693–1701.

RStudio. RStudio Team, Version: 3.5.2, http://www.rstudio.com/ 2016 Available from: http://www.rstudio.com/

Samson RS, Cardoso MJ, Muhlert N, Sethi V, Wheeler-Kingshott CA, Ron M, et al. Investigation of outer cortical magnetisation transfer ratio abnormalities in multiple sclerosis clinical subgroups. *Mult. Scler. J*. 2014; 20: 1322–1330.

Sánchez MP, Nieto A, Barroso J, Martín V, Hernández MA. Brain atrophy as a marker of cognitive impairment in mildly disabling relapsing-remitting

multiple sclerosis. *Eur. J. Neurol*. 2008; 15: 1091–1099.

Sayao A-L, Devonshire V, Tremlett H. Longitudinal follow-up of 'benign' multiple sclerosis at 20 years. *Neurology* 2007; 68: 496–500.

Scalfari A, Knappertz V, Cutter G, Goodin DS, Ashton R EG. Mortality in patients with multiple sclerosis. *Neurology* 2013; 81: 184–92.

Scalfari A, Neuhaus A, Daumer M, Muraro PA, Ebers GC. Onset of secondary progressive phase and long-term evolution of multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* 2014; 85: 67–75.

Schindelin J, Arganda-carreras I, Frise E, Kaynig V, Pietzsch T, Preibisch S, et al. Fiji - an Open Source platform for biological image analysis. *Nat. Methods* 2013; 9: 676–682.

Schmierer K, Scaravilli F, Altmann DR, Barker GJ, Miller DH. Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain. *Ann. Neurol*. 2004; 56: 407–415.

Schmacher GA, Beebe G, Kibler RF, Kurland LT, Kurtzke JF, McDowell F, et al. Problems of experimental trials of therapy in multiple sclerosis: report by the panel on the evaluation of experimental trials of therapy in multiple sclerosis. *Ann. N. Y. Acad. Sci.* 1965; 122: 552–68.

Shepherd D. Clinical features of multiple sclerosis in adults. *Acta Neurol. Scand.* 1979; 60: 218–230.

Skoog B, Runmarker B, Winblad S, Ekholm S, Andersen O. A representative

cohort of patients with non-progressive multiple sclerosis at the age of normal life expectancy. *Brain* 2012; 135: 900–911.

StataCorp. Stata Statistical Software: Release 15. 2017

De Stefano N. Brain damage as detected by magnetization transfer imaging is less pronounced in benign than in early relapsing multiple sclerosis. *Brain* 2006; 129: 2008–2016.

Stevenson V, Miller D, Leary S, Rovaris M, Barkhof F, Brochet B, et al. One year follow up study of primary and transitional progressive multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* 2000: 713–718.

Summers M, Swanton J, Fernando K, Dalton C, Miller DH, Cipolotti L, et al. Cognitive impairment in multiple sclerosis can be predicted by imaging early in the disease. *J. Neurol. Neurosurg. Psychiatry* 2008; 79: 955–958.

Sumowski JF, Wylie GR, Gonnella A, Chiaravalloti N, DeLuca J. Premorbid cognitive leisure independently contributes to cognitive reserve in multiple sclerosis. *Neurology* 2010; 75: 1428–1431.

Swanton JK, Fernando KT, Dalton CM, Miszkiel KA, Altmann DR, Plant GT, et al. Early MRI in optic neuritis: The risk for disability. *Neurology* 2009; 72: 542–550.

Swanton JK, Rovira A, Tintore M, Altmann DR, Barkhof F, Filippi M, et al. MRI criteria for multiple sclerosis in patients presenting with clinically isolated syndromes: a multicentre retrospective study. *Lancet Neurol*. 2007; 6: 677–

686.

Tallantyre EC, Major PC, Atherton MJ, Davies WA, Joseph F, Tomassini V, et al. How common is truly benign MS in a UK population? *J. Neurol. Neurosurg. Psychiatry* 2018: 1–7.

Thompson AJ, Hutchinson M, Brazil J, Feighery C, Martin EA. A clinical and laboratory study of benign multiple sclerosis. *Q J Med*. 1986; 58: 69–80.

Thompson AJ, Kermode AG, MacManus DG, Kendall BE, Kingsley DP, et al. Pattern of disease activity: clinical and magnetic resonance imaging study. *BMJ*. 1990; 300: 631–634.

Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018a; 17: 162–173.

Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. *Lancet* 2018b; 391: 1622–1636.

Tintore M, Rovira A, Arrambide G, Mitjana R, R??o J, Auger C, et al. Brainstem lesions in clinically isolated syndromes. *Neurology* 2010a; 75: 1933–1938.

Tintore M, Rovira A, Arrambide G, Mitjana R, Rio J, Auger C, et al. Brainstem lesions in clinically isolated syndromes. *Neurology* 2010b; 75: 1933–1938.

Tintore M, Rovira À, Río J, Otero-Romero S, Arrambide G, Tur C, et al.

Defining high, medium and low impact prognostic factors for developing multiple sclerosis. *Brain* 2015; 138: 1863–1874.

Turner B, Ramli N, Blumhardt LD, Jaspan T. Ventricular enlargement in multiple sclerosis: A comparison of three-dimensional and linear MRI estimates. *Neuroradiology* 2001; 43: 608–614.

Wallin MT, Culpepper WJ, Nichols E, Bhutta ZA, Gebrehiwot TT, Hay SI, et al. Global, regional, and national burden of multiple sclerosis 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 2019; 18: 269–285.

Weinshenker BG, Bass B, Rice GP, Noseworthy J, Carriere W, Baskerville J, et al. The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. *Brain* 1989; 112: 133–146.

Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr. Scand.* 1983; 67: 361–70.

Zivadinov R, Cookfair DL, Krupp L, Miller AE, Lava N, Coyle PK, et al. Factors associated with benign multiple sclerosis in the New York State MS Consortium (NYSMSC). *BMC Neurol*. 2016; 16: 102.

3D Slicer. https://www.slicer.org/.

Xinapse Systems, Northants, United Kingdom. http://www.xinapse.com.