NEXT GENERATION BIOMARKERS TO UNDERSTAND EARLY MULTIPLE SCLEROSIS

Sarah-Jane Martin, MBChB MRCP

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Institute of Infection, Immunity and Inflammation
College of Medical, Veterinary & Life Sciences
University of Glasgow

March 2021

Abstract

Multiple sclerosis (MS) is increasingly treatable. However, highly efficacious treatments carry serious potential risks. Treatment decisions must therefore weigh up the risk of treatment with the risk of irreversible disability. Current prognostic tools do not fully capture the scope of the pathology of MS, particularly axonal loss (an important substrate of disability). As a result, the identification of individuals at greatest risk of poor prognosis is suboptimal, and treatment decisions can be difficult and inconsistent. There is therefore an unmet need for a clinical tool or biomarker which can be employed *early in the disease* to identify those at greatest risk of future disability. New technologies and techniques have the potential to address this unmet need but require careful analysis in large cohorts. The aim of this work was to establish a large cohort of relapsing remitting MS patients at the point of diagnosis and whilst treatment naïve, and then to explore the potential role of next-generation biomarkers in early relapsing remitting multiple sclerosis.

We focused on fluid biomarkers reflective of axonal damage, and in particular neurofilament (NfL). We evaluated the role of CSF NfL in MS subtypes through systematic review and meta-analysis and concluded that NfL has utility as a biomarker of acute disease activity. We then employed a single molecule array (Simoa) to demonstrate NfL can be measured in blood, and that blood NfL levels correlate with CSF NfL levels.

The extent to which demyelination drives axonal loss in MS is unknown. We combined analysis of blood NfL levels with advanced magnetic resonance imaging (MRI) techniques of myelin integrity- the MRI g-ratio, to examine, *in vivo*, the relationship between myelin integrity and axonal damage. The MRI g-ratio was higher (suggesting loss of myelin integrity) in MS lesions compared with normal appearing white matter, but varied between individuals. We showed an association between lesion volume, lesion MRI g-ratio, and blood NfL levels. This demonstrates how blood-based biomarkers can be combined with advanced imaging biomarkers to gain insights into clinically relevant biology of disease.

Finally, we asked whether the additional measurement of other brain proteins, such as glial fibrillary acidic protein (GFAP), could provide further insights into disease biology and clinical outcomes. Longterm follow-up of the Future MS cohort will identify whether the biomarker trends found in this work continue to be relevant in the identification of patients at the greatest risk of poor prognosis from relapsing remitting multiple sclerosis.

Publications

Cerebrospinal fluid neurofilament light chain in multiple sclerosis and its subtypes: a meta-analysis of case-control studies. Martin SJ, McGlasson S, Hunt D, Overell J. Journal of Neurology, Neurosurgery & Psychiatry, September 2019:90:1059-1067

MRI-derived g-ratio and lesion severity in newly diagnosed multiple sclerosis. Martin SJ, York EN…. Waldman AD, Hunt DPH. Under review.

Table of contents

List of tables

List of figures

Acknowledgements

This research project was made possible by the funding provided by the Anne Rowling Regenerative Neurology Clinic in Edinburgh and the multiple sclerosis research endowment fund from NHS Greater Glasgow and Clyde.

I would like to express my deepest gratitude to the many people who offered support and advice throughout my time in research.

Foremost amongst these are my supervisors, Professor David Hunt and Dr James Overell. Dr James Overell is an excellent clinician. I learnt a great deal from him about the management of patients with MS and this has moulded my clinical practice. Professor David Hunt is a talented clinician and researcher, and I benefited greatly academically from his advice and help. He guided me through this journey step by step, from zero research experience to being able to complete this thesis. I will be forever grateful to them both for their belief in me over the past four years.

Professors Siddharthan Chandran and Hugh Willison have both also played significant roles in securing my time in research and the coming together of this thesis and I am extremely grateful.

There are many people without whom the Future MS study would not have been possible - the team at the Anne Rowling clinic, particularly Professor Chandran, Dr Pete Conick, Shuna Collville, Chris Batchelor, Mary Monaghan, Dr Patrick Kearns and numerous others who have been involved in the set-up, recruitment and review of the Future MS cohort over the past four years. The clinical MS team in Glasgow - Dr. Stewart Webb, Dr Pushkar Shah, Dr. Colin O'Leary, Dr Niall McDougall, Dr. Paul Gallagher, Lynn Cherry, Kate McGuigan for their support, insights and advice, and for referring patients to me for the study. The entire Glasgow research neuroradiology department were extremely helpful in coordinating the research visits and a real pleasure to work with. In particular, Dr Natasha Fullerton was responsible for clinically reporting all the research scans in Glasgow.

The neuroimmunology laboratory in Glasgow - Professor Willison, Patricia Thomson, Dr John Goodfellow and all the staff who were so helpful in allowing me to use their lab space and equipment.

The Edinburgh research imaging team reported all the research scans for the Future MS study. In particular, I would like to acknowledge Professor Adam Waldman and Elizabeth York, with whom we collaborated with, and who both provided helpful insights into this work.

Dr Sarah McGlasson, who taught me most of the basic science and laboratory skills that I have learnt over the past four years and offered nothing but support and encouragement throughout - I am immensely grateful. Edwin Carter, who was extremely helpful in coordinating and analysis of Simoa data, and a pleasure to work with.

The acknowledgement cannot be complete without thanking my family, and in particular my husband, Fraser, for his support, patience and the many wonderful dinners he cooked!

Authors declaration

I declare that, except where explicit reference is made to the contribution of others, this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature _____

Printed name: Sarah-Jane Martin

List of abbreviations

AB	Antibody
AEB	Average Enzymes per bead
ANOVA	Analysis of variance
AVF	Axonal volume fraction
BBB	Blood brain barrier
BMI	Body mass index
BP	Blood pressure
CD	Cluster of differentiation
CDMS	Clinically definite multiple sclerosis
CHI	Community health index
CIS	Clinically isolated syndrome
CNS	Central nervous system
CRF	Clinical research facility
CRP	C-reactive protein
CSF	Cerebrospinal fluid
CV	Coefficient of variation
DMT	Disease modifying treatment
DTI	Diffusion Tensor Imaging
DW	Diffusion weighted
DWI	Diffusion weighted imaging
ECL	Electrochemiluminescence
EDSS	Expanded Disability Scoring system
ELISA	Enzyme-linked immunosorbent assay
FACS	Fluorescence-activated cell sorting
FLAIR	Fluid-attenuated inversion recovery
FMS	Future MS
GAD	Gadolinium
GFAP	Glial fibrillary acidic protein
HC	Healthy control
HIV	Human immunodeficiency virus
ICV	Intracranial volume
IF	Interferon
IFNB	Interferon-beta
IMP	Investigational medicinal product
IQR	Inter-quartile range
IRAS	Integrate research application system
LP	Lumbar puncture
MAG	Myelin association glycoprotein
MAGNIMS	Magnetic resonance imaging in multiple sclerosis
MBP	Myelin basic protein
MOG	Myelin oligodendrocyte glycoprotein
MPRAGE	Magnetization-prepared rapid gradient-echo
MR	Magnetization ratio

MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MSFC	Multiple sclerosis functional composite score
MSIS	Multiple sclerosis impact scale
MT	Magnetization transfer
MTR	Magnetization transfer ratio
MTsat	Magnetization transfer saturation
MVF	Myelin volume fraction
NAWM	Normal appearing white matter
NFL	Neurofilament light chain
NHS	National health service
NINDC	Non-inflammatory disease control
NODDI	Neurite Orientation Dispersion and Density Imaging
NOS	Newcastle-Ottawa Scale
OCB	Oligoclonal band
PASAT	Paced auditory serial additions test
PBMC	Peripheral blood mononuclear cells
PDSS	Patient derived severity score
PHQ	Patient health questionnaire
PLP	Proteolipid like protein
PML	Progressive multifocal leukoencephalopathy
PPMS	Primary progressive multiple sclerosis
PRISMA	Preferred reporting in Systematic reviews and Meta-analyses
PRMS	Progressive relapsing multiple sclerosis
PROSPERO	Prospective Register of Systematic Reviews
REC	Research ethics committee
RIS	Radiologically isolated syndrome
RRMS	Relapsing remitting multiple sclerosis
SD	Standard deviation
SDMT	Symbol digit modality test
SDS-PAGE	Sodium dodecyl Sulfate polyacrylamide gel electrophoresis
SMT	Spherical mean technique
SMS	Stratified medicine Scotland
SNR	Signal-to-noise ratio
SOP	Standard operating procedures
SPMS	Secondary progressive multiple sclerosis
UCH-L1	Ubiquitin carboxy-terminal hydrolase L1
VEP	Visual evoked potentials
WM	white matter
WML	white matter lesion
9-hpt	9-hole peg test

Chapter One

Introduction: Multiple sclerosis, a complex and variable disease

1. Introduction: Multiple sclerosis, a complex and variable disease

Multiple sclerosis (MS) is a serious and unpredictable neurological disease with complex pathophysiology (Compston and Coles, 2008; Jones and Coles, 2010). MS is an immunological disease of the central nervous system, characterised by prominent loss of myelin from nerve axons, often referred to as demyelination (Trapp and Nave, 2008; Lassmann, 2019).

Inflammation and demyelination can occur anywhere within the central nervous system in people with MS and can result in any one of a variety of clinical symptoms. Patients also differ remarkably in the severity of clinical relapses, the frequency with which they recur, and the extent to which they recover. As such, multiple sclerosis is a heterogeneous disease in both clinical presentation and prognosis.

MS is a major cause of disability in young adults, but is increasingly treatable with a number of high efficacy treatment options now available (Reich, Lucchinetti, *et al.* 2018). Disease modifying treatments with greater efficacy carry greater potential risk (Soleimani, Murray, et al. 2019); however, lower efficacy and potentially less risky treatments run the risk of inadequate disease control and accumulation of disability.

A critical issue in the management of patients with multiple sclerosis is therefore choosing the most appropriate treatment option for an individual. This in turn is dependent upon identifying patients with the greatest long-term risk of developing serious neurological deficits and disability. The prognostic markers currently used to make these decisions in routine clinical practice are inadequate, resulting in variability in therapeutic approaches amongst neurologists (Cerqueira *et al.*, 2018).

1.1 Epidemiology of MS

1.1.1 Worldwide prevalence

Worldwide, prevalence of multiple sclerosis varies significantly between low-risk regions (Southern and Central America, Africa, East Asia) and high-risk regions (Northern America and Northern Europe)(Wallin *et al.*, 2019). Within geographical regions, incidence and prevalence also vary by age and sex. Children of both sexes have equal (low) risk; whereas after adolescence MS is up to twice as prevalent in women as it is in men (Wallin *et al.*, 2019).

Variations in incidence rates of MS are also seen on a regional scale. Within Europe, the UK has a particularly high incidence of MS, and within the UK Scotland has a higher incidence of multiple sclerosis than England, Northern Ireland or Wales (Visser *et al.*, 2012; Mackenzie *et al.*, 2014). Variations in the prevalence of MS among, and within, nations of largely European descent appear to show an association with latitude - 'latitudinal gradient'. The reasons for such variations remain unknown, but are likely to be the result of a combination of genetic influence of predisposition (such as the HLA-DRB1 allele) and environmental factors (such as ultraviolet radiation exposure and vitamin D synthesis)(Simpson *et al.*, 2011).

1.1.2 Scottish prevalence

In 2019 the overall incidence rate of MS within Scotland was 9.79 per 100,000(Public Health Scotland, 2020), but regional differences are seen. The average annual incidence between 2010 and 2017 ranged from 6.24 per 100,000 (NHS Borders) to 17.36 per 100,000 (NHS Orkney)(Kearns *et al.*, 2019), with a trend to a higher incidence in more northerly regions. This intra-regional variation in incidence is consistent with reports of latitudinal gradient from other countries (Simpson et al. 2011), and highlights heterogeneity in the disease even within small populations.

Multiple sclerosis is a disease of young people, and in particular young women. In Scotland the average age of diagnosis is 41 years old and the female to make sex ratio is 2.3:1(Kearns *et al.*, 2019). This average age of diagnosis is similar to other northern European countries (Sweden 40.7 years, Denmark 38 years, Netherlands 39.9 years), but slightly older than certain southern European countries (France 35.1 years, Spain 32.1 years, Italy 34.2 years)(Kobelt *et al.*, 2017).

The age of onset of any disease has important and far-reaching socio-economic implications. Patients with MS, even early in the disease course, are more likely to reduce working hours or retire early on medical grounds, thus impacting gross income and pensions (Kobelt *et al.*, 2017). Later in the disease course, physical and/or cognitive disability may mean that patients require additional support, either directly provided by, or paid for, by family members (McCrone *et al.*, 2008). Over the lifetime of the patient these factors can significantly reduce household income. Furthermore, GP and specialist appointments, occupational and physiotherapy input, hospital inpatient stays and

medications are costly for the national health service, and therefore society in the whole (McCrone *et al.*, 2008; Kobelt *et al.*, 2017).

The recognition of multiple sclerosis as an important issue for Scottish health services resulted in the introduction of the NHS Scottish MS register in January 2010. The NHS Scottish MS register is implemented across all of Scotland's 14 heath-boards and mandates the collection of baseline demographic data for all new diagnoses of MS. A national report is released each year to audit the diagnosis and clinical care of MS patients, to identify treatment targets, and to allocate resources. The Scottish MS register is well implemented throughout Scotland and has a 99% completion rate for required data (Kearns *et al.*, 2019). As a result, Scotland is the only place in the UK where standardised, high-quality baseline demographic data are available on all MS diagnoses.

1.2 Clinical presentation and subtypes

Historically, the disease entity of multiple sclerosis was subdivided into four main groups based on the tempo of clinical relapses and the presence or absence of clinical disease progression between relapses, *see Figure 1-1*. Formal definitions of each were described in 1996 by the US National Multiple Sclerosis Society (NMSS) advisory Committee on clinical trials in MS.

This classification system was based on clinical disease activity alone and is now considered too restrictive and outdated. Instead, a more descriptive terminology should be employed which encompasses disease activity (both clinical and radiological) and disease progression (Lublin *et al.*, 2014).

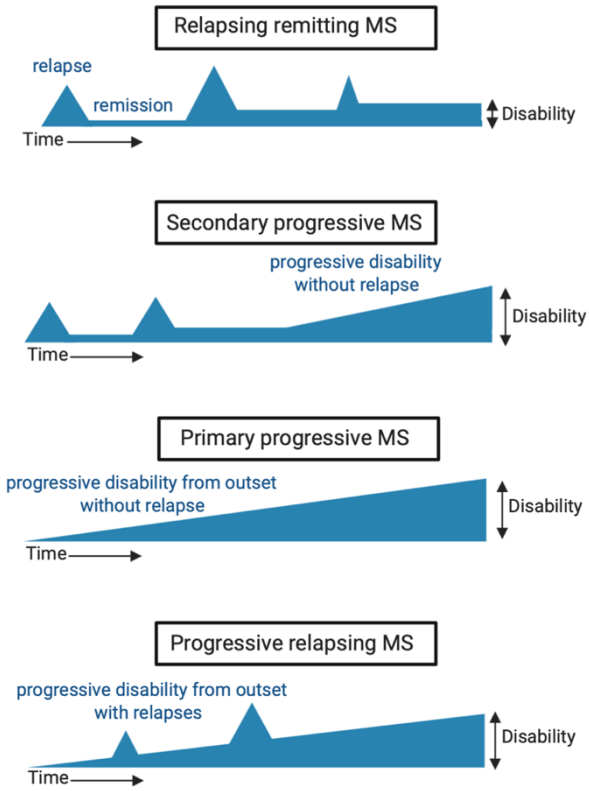


Figure 1-1 Historical subdivisions of MS based on clinical phenotype

The majority of individuals begin with a relapsing-remitting phase (RRMS) and later develop secondary progression (secondary progressive MS, SPMS). In a small subset, progression is seen from the outset without defined relapses (primary progressive MS, PPMS), or with superimposed clinical relapses (progressive relapsing MS, PRMS).

1.2.1 RRMS

Relapsing-remitting MS (RRMS), with or without disease activity, is the commonest phenotype and comprises approximately 85% of all patients at diagnosis. The hallmark of relapsing-remitting MS is clinical stability between relapses (irrespective of whether or not a full recovery was achieved from the previous relapse). Over time, the frequency of relapses diminishes, but despite this disability accrues. At this point the patient may then be described as developing progressive multiple sclerosis (PMS), with or without disease activity.

1.2.2 Progressive MS

Progressive MS was historically subdivided into primary or secondary. Secondary progressive MS (SPMS) described patients with a history of relapsing neurological symptoms, but who then experienced a gradual accrual of symptoms which limited function. Primary progressive MS (PPMS) was defined as ‘disease progression from the

outset, with occasional plateaus and minor improvements allowed'. A less commonly diagnosed subtype, progressive-relapsing MS (PRMS), allowed for acute relapses, with or without full recovery, on a background of disease progression from onset.

This nomenclature was based on clinical data alone. Magnetic Resonance Imaging (MRI) parameters were purposely excluded from the definitions as it was felt that at that time it could not differentiate between clinical courses. Much has changed since the original definitions were first proposed. The widespread routine use of MRI in disease monitoring enables identification of clinically silent disease activity, challenging the concept that relapses occur only infrequently in progressive disease. The ability to describe in the nomenclature whether disease activity is present or not, is important for identifying individuals who may benefit from disease modifying treatment. For these reasons, the original definitions have been superseded in light of emerging data about natural history of the disease (Lublin *et al.*, 2014).

Progressive MS is now categorised according to whether there is disability accumulation from outset ('primary progressive') or following an initial relapsing course ('secondary progressive'), but also whether there is current disease activity or ongoing disease progression – 'active, with progression', 'active without progression', 'not active but with progression' and 'not active and without progression – stable disease'(Lublin *et al.*, 2014).

1.2.3 CIS & RIS

"Clinically Isolated Syndrome" is an initial presentation with clinical symptoms and signs suggestive of CNS demyelination, which may show evidence of dissemination in space, but does not yet show evidence of dissemination in time, e.g. an initial single MS-like presentation which does not yet fulfil the MS diagnostic criteria.

Radiologically isolated syndrome (RIS) is a more simple concept in which MRI appearances may be consistent with those commonly seen in MS patients, but occur in the absence of clinical symptoms. RIS is therefore only described in patients who have undergone CNS imaging for symptoms *not* consistent with multiple sclerosis. A proportion of individuals with RIS will however go on to develop clinical symptoms consistent with an MS relapse within 10 years of diagnosis of RIS (Lebrun-Frenay *et al.*, 2019).

1.3 Pathophysiology

The pathological hallmarks of multiple sclerosis were described by the French neurologist and professor of anatomical pathology Jean-Martin Charcot in 1880. These include inflammation, demyelination, and critically, the persistence of axons as a character belonging to multiple sclerosis, suggesting a degree of axonal sparing (Kornek and Lassmann, 1999).

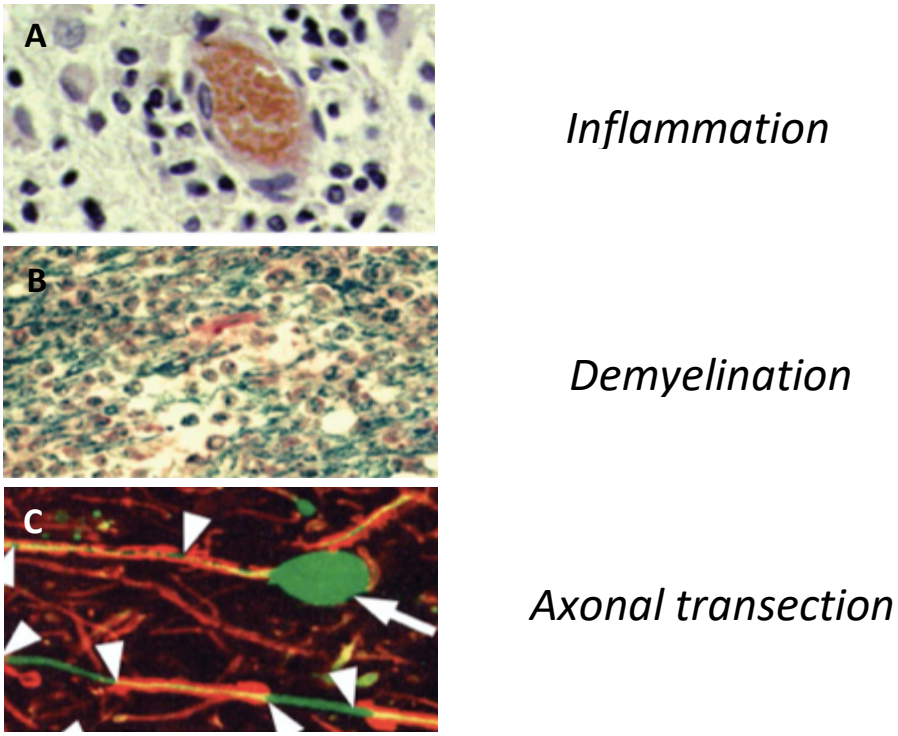


Figure 1-2 The pathological hallmarks of an MS lesion

Reproduced with permission from 'Multiple sclerosis', Compston, A. and Coles, A. Lancet 2008; 372: 1502–17, Copyright Massachusetts Medical Society.

(A) Perivascular lymphocytes, particularly T cells, make up a core component of the inflammatory infiltrate (haematoxylin and eosin stain). (B) Myelin loss in the early stages of lesion formation demonstrated by lack of staining with luxol fast blue stain. (C) Demyelination of axons is demonstrated by loss of myelin (in red, staining myelin basic protein) between arrow heads. Non-phosphorylated neurofilament is stained green with SMI32 antibodies, and axonal transection is demonstrated by the presence of a terminal axonal ovoid (arrow).

Charcot however also noted that, within some areas of demyelination, axons appeared to be entirely destroyed (Kornek and Lassmann, 1999). In 1906 an Austrian neurologist, Otto Marburg, found axonal loss to be more prominent than was the general consensus, and Marburg stressed the importance of 'relative' axonal sparing (Kornek and Lassmann, 1999)

Myelin loss (‘demyelination’) in multiple sclerosis occurs in discrete foci, producing irregular, sharply demarcated ‘MS plaques’ or ‘’lesions’. Although studies continue to show that the pathological hallmark of MS is focal demyelination, there is now a greater appreciation that disease is not limited to cerebral white matter lesions. Lesions can occur anywhere within the central nervous system, with a predilection for the optic nerves, periventricular white matter, cortical grey matter, brainstem, cerebellum and subpial spinal cord (Popescu and Lucchinetti, 2016).

Cortical lesions (subpial, intracortical and leukocortical) are an important element of whole brain axonal and neuronal loss (Kutzelnigg *et al.*, 2005), and have been shown to occur early in the disease process and to increase with disease duration (Lucchinetti *et al.*, 2011). Sex differences have also been noted, with males shown to have a higher incidence of cortical grey matter lesions than females (Luchetti *et al.*, 2018). However cortical lesions are not well visualised using conventional MRI techniques (Bjartmar and Trapp, 2001) and as a result are significantly under-reported when compared with pathological analysis (Lassmann, 2019).

Furthermore, diffuse disease is recognised throughout both the white and grey matter. Diffuse injury is prominent in non-lesional ‘normal appearing’ white matter (NAWM), particularly later in the disease course (Kutzelnigg *et al.*, 2005). Disease of the NAWM is thought to occur as a consequence of secondary Wallerian degeneration from remote focal white matter lesions or from neuronal loss in cortical lesions (Bjartmar and Trapp, 2001). However, the extent of diffuse NAWM injury does not correlate well with either focal white matter or cortical lesion load (Kutzelnigg *et al.*, 2005).

1.3.1 Inflammation

Cells of both the innate immune system (microglia and macrophages) and adaptive immune system (B and T lymphocytes) form a perivascular inflammatory infiltrate which disperses into surrounding parenchyma (Lassmann, 2019). Active lesions can be differentiated pathologically into ‘early’ or ‘late’ stage active lesions according to the myelin debris within macrophages - macrophage within early active lesions contain debris from minor myelin proteins (MOG/MAG) whereas macrophages within late active lesions contain debris from major myelin proteins (PLP/MBP) (Popescu and Lucchinetti, 2016).

For many years MS was considered to be a primarily T cell-mediated disease. This was due to the prominence of T cell infiltration seen within demyelinating lesions and GWAS studies highlighting many genes associated with adaptive cell-mediated immunity (Sawcer *et al.*, 2011). Recently, the identification of persistent B-cell follicle-like structures within the CNS, combined with the high efficacy of anti-CD20+ therapies (Hauser *et al.*, 2017) have suggested that B-cells play an important role.

CD4+ T cells may be involved in the initiation of the adaptive immune response. However cells from B cell lineage are a major component in the brain and spinal cord and are likely to be important drivers of inflammation, demyelination and neurodegeneration in MS (Serafini *et al.*, 2004; Jones and Coles, 2010; Lassmann, 2019). Particular ways in which B cells may contribute to inflammation in MS include in their role as professional antigen-presenting cells and by secretion of pro-inflammatory cytokines.

Both acute and chronic forms of inflammation can occur in multiple sclerosis (Frischer *et al.*, 2015; Kuhlmann *et al.*, 2017; Lassmann, 2019).

i. Acute inflammation is most significant in early and relapsing forms of MS, and is thought to occur as a result of the bulk invasion of B and T lymphocytes across a dysfunctional blood-brain-barrier. This type of inflammation results in the formation of the classic white matter demyelinated plaque and tends to decline with disease duration. This type of inflammation can have clinical correlate as a relapse.

ii. Chronic inflammation does not require a damaged blood-brain-barrier to gain entry to the central nervous system. Instead, B and T lymphocytes slowly accumulate in connective tissues spaces of the brain such as the meninges and perivascular spaces (Serafini *et al.*, 2004). Aggregation of lymphocytes in subpial spaces results in the development of cortical grey matter lesions and neuronal loss. This type of inflammation occurs with increasing disease duration (Frischer *et al.*, 2015).

It should be emphasised that although *in general* inflammation declines with disease duration, this primarily reflects a decline in significant bouts of acute inflammation. Inflammation in a chronic form, as seen in progressive disease, remains significant (Frischer *et al.*, 2015; Luchetti *et al.*, 2018).

1.3.2 Demyelination

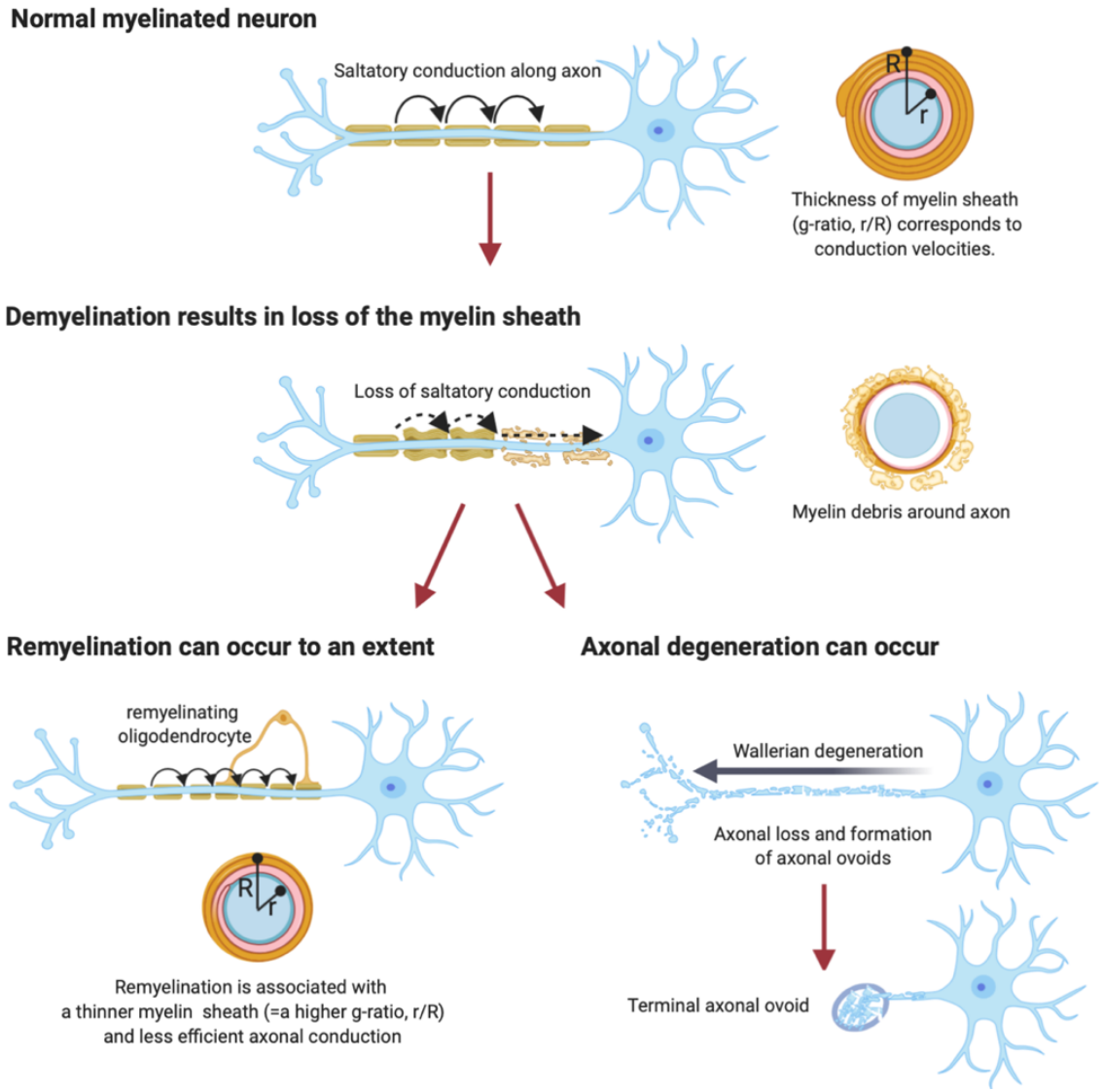


Figure 1-3 Demyelination of a neuron

(A) In a normal myelinated axon, an action potential is propagated by saltatory conduction between Nodes of Ranvier, and the thickness of the myelin sheath reflects the speed of conduction velocities. The thickness of the myelin sheath can be described by its g-ratio. (B) Myelin damage disrupts normal conduction (C) If an oligodendrocyte precursor cell is recruited remyelination may occur. However, remyelination results in a thinner myelin sheath and shorter internodes. (D) If remyelination does not occur the axon may undergo Wallerian degeneration. As axonal damage occurs, products of axonal transport accumulate, and terminal neuronal spheroids develop. These are a hallmark of axonal transection. Figure created in BioRender by SJM.

The link between inflammation and demyelination is hypothesised, but not proven, to occur as a result of an autoimmune response directed against myelin epitopes. One line of evidence for this is that MS-like disease can be mediated by adoptive transfer of myelin/MOG reactive T cells (Robinson *et al.*, 2014).

Despite the early insight into axonal loss by Charcot et al, historically the focus has remained on demyelination as the cause of functional impairment, with axonal destruction thought to occur late in the disease and secondary to chronic demyelination.

The myelinating cell of the central nervous system is the oligodendrocyte. Myelination begins *in utero* but continues for years after birth. In adulthood, oligodendrocytes perform myelin maintenance and provide structural, electrical and trophic support to multiple nerve axons (Podbielska *et al.*, 2013). This is a reciprocal relationship, as seen by the loss of oligodendrocytes in the context of demyelination (Podbielska *et al.*, 2013).

Demyelination describes degradation of the fatty myelin sheath surrounding a nerve axon and loss of the supporting oligodendrocyte, *see figure 1-3*. Demyelination slows the propagation of action potentials down the nerve axon as the saltatory conduction enabled by myelin Nodes of Ranvier is lost (Felts, Baker, et al. 1997). Neurons can recover to a certain extent, and some remyelination occurs, although this is highly variable between individuals, and rarely to pre-disease levels (Popescu and Lucchinetti, 2016).

Remyelination of an intact axon requires recruitment of oligodendrocyte progenitor cells, differentiation of into myelinating oligodendrocytes, and contact between the axon and the myelinating oligodendrocyte (Podbielska *et al.*, 2013). These are all potential points at which remyelination may fail.

If recovery does occur, the new myelin sheath surrounding the remyelinated axon remains thinner, and the internodes are shorter than in developmentally myelinated axons (Gledhill, Harrison, et al. 1973). The result in less efficient nerve conduction. Remyelinated plaques are identified histologically by a sharply demarcated ring of pale Luxol Blue staining - a ‘shadow plaque’. If remyelination does not occur, Wallerian degeneration may occur, *see figure 1-3*. Wallerian degeneration is the anterograde degeneration of the distal part of the axon which has been separated from the cell body (Dziedzic *et al.*, 2010).

Heterogeneity in the immunopathological appearances of active and early demyelinating lesions between, but not within, patients, was reported in a study by Lucchinetti et al in 2000. In that study, Lucchinetti et al proposed four main patterns of early demyelination - *see table 1-1*.

Table 1-1 Summary of four distinct immunopathological patterns of early demyelinating lesions proposed by Lucchinetti et al, 2000.

Pattern	***Peri-venous distribution?***	***Composition of inflammatory infiltrate***	***Additional features***	***Potential for remyelination?***
I	Yes	T-lymphocyte and macrophage dominated		Yes - High incidence of shadow plaques
II (most common)	Yes	T-lymphocyte and macrophages	Immunoglobulin (IgG) and complement deposition	Yes - High incidence of shadow plaques
III	No	T-lymphocyte and macrophages	Preferential loss of MAG protein seen in some.	No - Pronounced loss of oligodendrocytes with no shadow plaques
IV	Yes	T-lymphocyte and macrophages	Seen in PPMS. No preferential loss of particular myelin protein.	No - Oligodendrocyte death without shadow plaques

However, this description of four distinct pathological processes is highly disputed. Firstly, by nature of the cohort (autopsy and biopsy cases), these cases are unlikely to be truly representative. Furthermore, clinical follow-up data were only available for 43 of 51 biopsy cases, with at least one patient requiring repeat biopsies. Finally, not all the antibodies used in the histopathological analysis are commercially available, thereby limiting replication of the study (Lucchinetti *et al*, 2000). In contrast, other groups have reported immunopathological homogeneity between patients with established MS, suggesting that Lucchinetti's categorisation may simply represent different chronological stages of early lesion development (Breij *et al.*, 2008, Kuhlmann *et al.*, 2017).

In 2017 Kuhlmann et al proposed a simplified histological classification system for MS lesions. This differentiated between active, mixed active/inactive and inactive lesions (based on the presence/absence and distribution of macrophages), with or without ongoing demyelination. The aim was to develop a more descriptive classification that was simpler for neuropathologists to apply and thereby enabled easier comparison between studies (Kuhlmann *et al.*, 2017).

1.3.3 Neurodegeneration

If the neuronal axon remains demyelinated, transection and degeneration of the nerve axon may occur, as described above.

Histopathological studies since Charcot's early descriptions have demonstrated and reiterated that axon transection is abundant throughout MS lesions, and that axonal loss occurs to a greater extent within areas of active inflammation (Ferguson *et al.*, 1997; Bruce D. Trapp *et al.*, 1998). This suggests that axonal damage occurs, to a degree, as a consequence of demyelination.

Demyelinated axons become susceptible to inflammatory mediated damage from cytokines, proteolytic enzymes and free radicals. At the same time, they lose their structural, electrical and trophic support from oligodendrocytes. Loss of saltatory conduction redistributes sodium channels along the axon. Activity of Na+/K+ ATPase channels increases to counter an influx of sodium, augmenting the energy demands of the axon (Campbell, Worrall et al, 2014). In demyelinated axons, this increase in energy requirements causes mitochondria to gather in the axon (G. R. Campbell, Worrall, *et al.* 2014). However, these compensatory mechanisms often eventually fail, leading to axoplasmic influx of calcium and activation of proteolytic enzymes. As fast axonal transport is lost and the axon degenerates, amyloid precursor proteins accumulate. These are visualised as terminal axonal ovoids, which are considered a hallmark of axonal transection (Trapp and Nave, 2008).

At what stage in the disease process axonal damage begins has prognostic relevance. Axonal loss is thought to be the final common pathway in the development of disability, and once axonal injury occurs beyond the threshold for compensation it is the best pathological predictor of permanent neurological deficit (Bjartmar *et al.*, 2000). As such, the quantification of axonal loss, through imaging and body fluid biomarker approaches is increasingly relevant to clinical practice and will be discussed in sections 1.7.3 - 1.7.5.

1.4 Diagnosis of multiple sclerosis

A diagnosis of multiple sclerosis is based upon clinical history and examination alongside paraclinical tools such as magnetic resonance imaging (MRI), CSF analysis or visual evoked potentials (VEPs). Diagnostic criteria exist to ensure diagnosis is consistent between neurologists.

The diagnostic criteria are dependent upon showing that episodes of demyelination have occurred at least two time points, and in at least two areas of the central nervous system, i.e. that there is *dissemination of disease in time*, and *dissemination of disease in space*.

What is considered acceptable as proof of dissemination in time/space has varied with different renditions of the diagnostic criteria, *see figure 1-4*. However, all have focused on this key point.

1.4.1 Previous diagnostic criteria

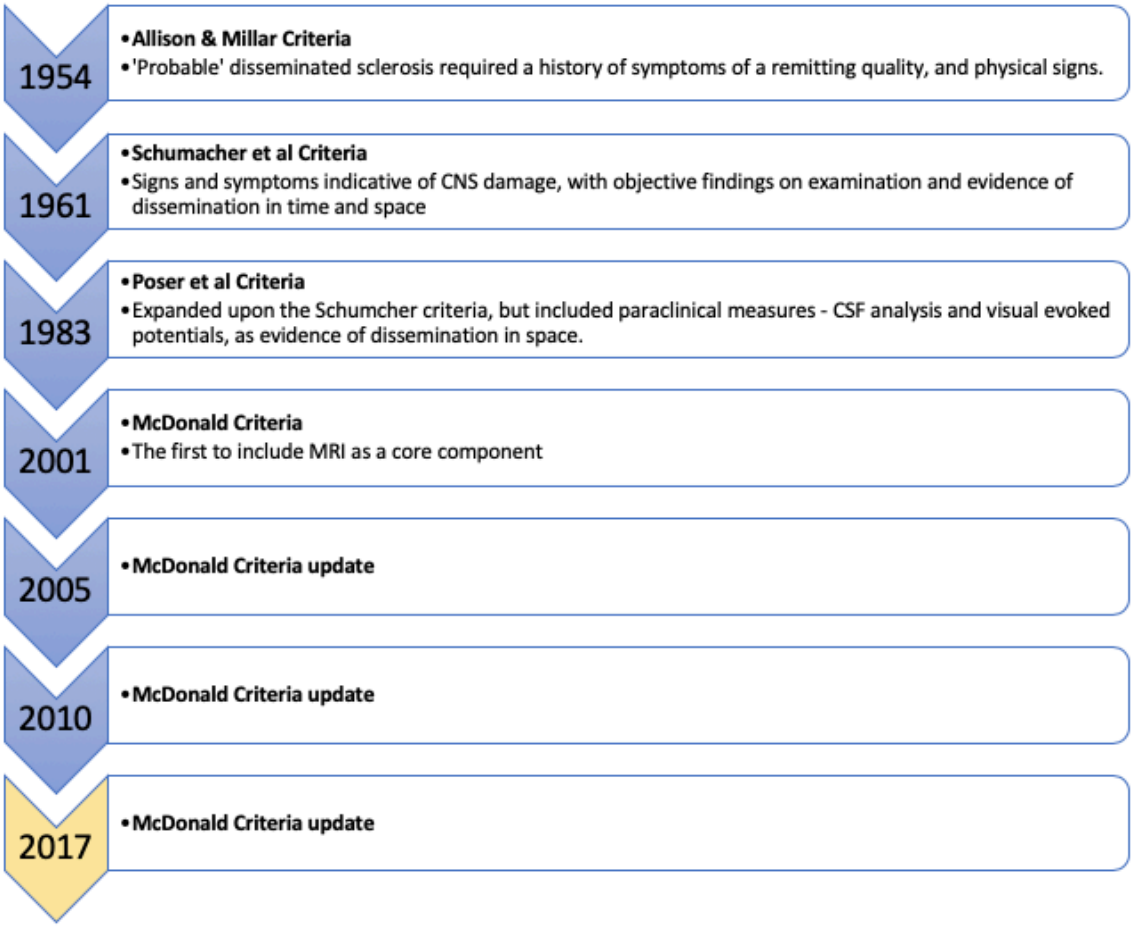


Figure 1-4 Multiple sclerosis diagnostic criteria to date

Timeline showing only the most commonly cited diagnostic criteria to date.

The first diagnostic criteria, published by the Schumacher Committee in 1965, were based solely on clinical findings (G.A. Schumacher, G. Beebe, *et al* 1965). All subsequent criteria have been based upon these. In 1983 the Poser criteria expanded upon the Schumacher criteria with the addition of paraclinical parameters and classifications of 'definite' or 'probable' multiple sclerosis (Poser *et al.*, 1983).

The 2001 McDonald Criteria were the first to include MR imaging as a core component (McDonald *et al.*, 2001). The McDonald criteria have now been revised three times since publication (Polman *et al.*, 2005, 2011; Thompson *et al.*, 2018). Between 2001 and 2010 the MRI criteria were simplified, and the European MAGNIMS criteria are now used (Swanton *et al.*, 2007).

1.4.2 Current diagnostic criteria

The major update in the 2017 McDonald criteria was the inclusion of unpaired oligoclonal bands in the CSF as evidence for dissemination in time.

Since publication of the original diagnostic guidelines, each subsequent review has aimed to reduce diagnostic ambiguity and allow diagnostic criteria to be fulfilled at an earlier stage in the disease. The effect of these changes is to allow earlier diagnosis, and therefore earlier treatment. The consequences are that a greater number of diagnoses are made and that MS populations may differ slightly according to the diagnostic criteria employed at the time. These are important points when comparing contemporary and historical MS cohorts, and are of relevance to interpreting the changing definition of MS used in clinical trials over the past four decades.

As RRMS is the most common form of MS we will focus here on the diagnostic criteria for RRMS, using the McDonald 2017 diagnostic criteria (Thompson *et al.*, 2018).

Diagnosis requires evidence of *dissemination in time* and *dissemination in space*

Dissemination in time can be demonstrated:

- CLINICALLY - A clinical history of two or more distinct episodes of neurological symptoms characteristic for MS, for example optic neuritis.
- RADIOLOGICALLY - MRI findings of the simultaneous presence of both Gadolinium contrast enhancing and non-enhancing T2 lesions; or the finding of new T2 hyperintense lesions on a follow-up MRI.
- CSF ANALYSIS - The finding of unpaired oligoclonal bands (OCB) in cerebrospinal fluid (CSF)

Dissemination in space can be demonstrated:

- CLINICALLY - A clinical history clearly reporting symptoms affecting distinct regions of the CNS, such as optic neuritis and symptoms of a spinal relapse (bilateral leg weakness and bladder symptoms).
- RADIOLOGICALLY - MRI findings of one or more T2 hyperintense lesion characteristic of MS, in at least two of four areas of the CNS: periventricular, cortical/juxta-cortical, infratentorial brain regions, and spinal cord.
- PARACLINICAL MEASURES - The finding of abnormal visual evoked potentials in a patient without prior clinical history of optic neuritis suggests previous asymptomatic optic nerve involvement. This can be considered alongside a secondary symptom or sign from a distinct neurological region.

1.4.3 Role of MRI in the diagnosis of MS

MRI is a non-invasive imaging modality that is particularly well-suited to imaging non-bony tissues such as the brain, and as such has revolutionised the field of neurology in general, and particularly for multiple sclerosis (McDonald, 1986). MRI is routinely used in the diagnosis of MS (McDonald *et al.*, 2001; Wattjes *et al.*, 2015), the monitoring of disease activity (Miller, 1994; Brex *et al.*, 2002), the assessment of treatment efficacy, and in identification of potential treatment side effects. Repeated imaging in this manner is possible because MRI uses non-ionising radiation.

An MRI machine produces a powerful magnetic field which causes all the protons in the body to align with that field. When a radiofrequency pulse is applied, the aligned protons

are disturbed. When the radiofrequency pulse is switched off, the protons swing back into alignment, releasing energy as they do so (National Institutes of Health Magnetic Resonance Imaging, 2020). The energy released by the protons and the time taken to re-align with the magnetic field varies between different molecules, allowing the MR sensors to differentiate between tissues, fluids, bone etc (National Institutes of Health Magnetic Resonance Imaging, 2020).

An 'MRI sequence' is a predetermined set of radiofrequency pulses used to produce images with a particular appearance (Murphy and Gaillard, 2020). Certain structures are better defined with one sequence than another, so in order to appreciate different tissues a combination of sequences are employed, *see figure 1-5.*

The commonly used sequences in an MRI demyelination protocol include:

- T1 weighted sequences - On T1 imaging fluid (such as CSF) is low intensity (appears black), fat is high intensity (appears white), and tissues such as the brain are intermediate intensity (shades of grey), *figure 1-5(A).*
- T2 weighted sequences - On T2 fluid and fat are both high intensity (appears white), and tissues are intermediate intensity (shades of grey), *figure 1-5(C).*
- Fluid attenuated (FLAIR) sequences - FLAIR is used to reduce the intensity signal from fluid*, figure 1-5(B) and (D).* An example of its use is in the identification of parenchymal oedema, where high T2 signal from nearby CSF may distorts images. By suppressing the high signal from CSF, FLAIR sequences allow for easier identification of parenchymal oedema.
- T1 with contrast enhancement - A Gadolinium-based contrast can be given intravenously to the patient prior to imaging. If there is breakdown of the blood-brain-barrier (BBB), the Gadolinium will leak through and appears bright within parenchymal tissue, *figure 1-5(F)*. As damage to the BBB heals within a matter of weeks, contrast enhancement only highlights acute lesions (approximately 2-6 weeks old). Gadolinium contrast is therefore used in the identification of new lesions or lesions of different ages. Recently, safety concerns surrounding the repeated use of Gadolinium contrast and its possible accumulation within the brain parenchyma have reduced its use in routine imaging and in research studies (Gulani *et al.*, 2017).

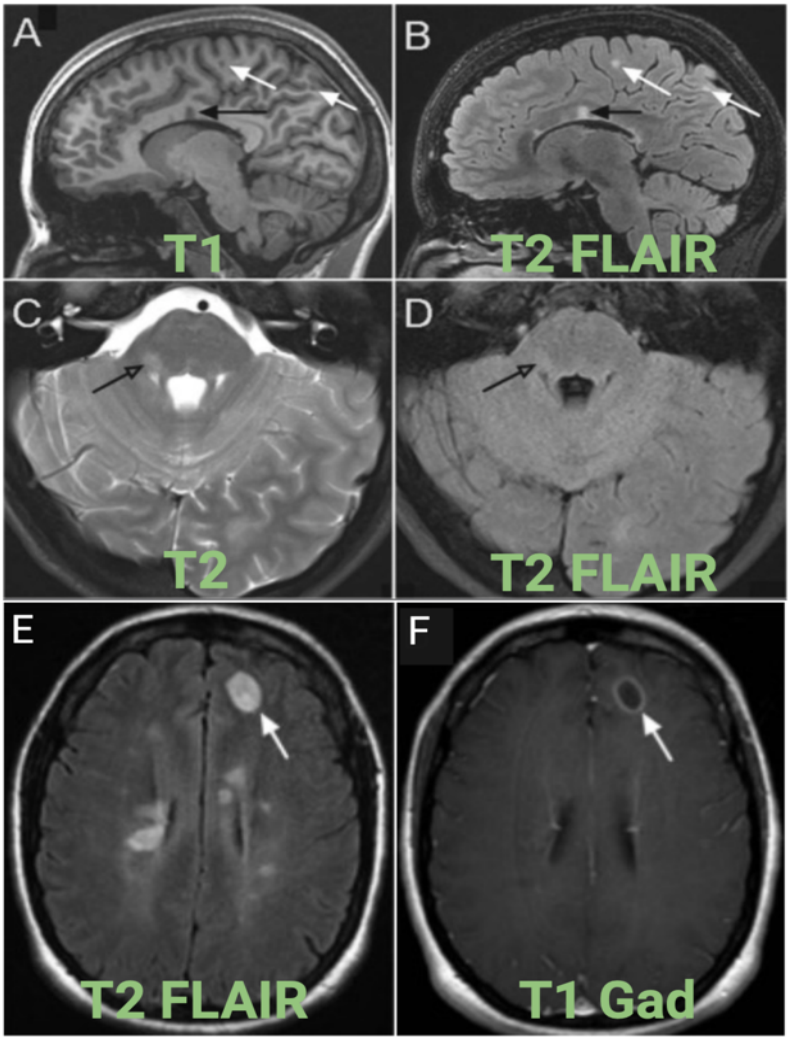


Figure 1-5 MRI sequences to demonstrate lesions in relapsing remitting MS

Reproduced with permission from Tillema, Jan-Mendelt, and Istvan Pirko. 'Neuroradiological evaluation of demyelinating disease' Therapeutic advances in neurological disorders vol. 6,4 (2013): 249-68.

(A) & (B) - sagittal images showing supratentorial white matter lesions as visualised on (A) T1-weighted and (B) T2 FLAIR sequences. T2 FLAIR has superior sensitivity in detection of juxtacortical (white arrows) and periventricular lesions (black arrows). (C) & (D) - axial images showing white matter lesions as visualised on (C) T2-weighted and (D) T2 FLAIR sequences, demonstrating the greater sensitivity of posterior fossa lesion detection using standard T2-weighted images over T2 FLAIR. (E) & (F)- axial images showing supratentorial white matter lesions on (E) T2 FLAIR and (F) T1-weighted post Gadolinium contrast sequences. The large frontal lesion (white arrow) demonstrates ring enhancement, thus differing it from the older, non-enhancing lesions seen on T2 FLAIR sequences (E) but not T1-weighted Gad sequences (F).

1.4.4 Role of CSF in the diagnosis of MS

Cerebrospinal fluid analysis is only used routinely at the point of diagnostic work-up. Several aspects of CSF analysis can be helpful in confirming or disputing a diagnosis, including the white blood cell count (typically <50 $x10^6$), CSF protein concentration (typically normal, but may be slightly elevated), and the assessment of intrathecal immunoglobulin synthesis.

Although not specific to MS, evidence of intrathecal antibody synthesis supports the diagnosis of MS (Thompson *et al.*, 2018). Intrathecal antibody synthesis is demonstrated by the finding of 2 or more oligoclonal bands (OCBs) in CSF but absent in paired serum.

Oligoclonal bands are bands of immunoglobulins that can be seen in CSF or blood by using a combination of isoelectric focusing and immunoblotting techniques. Standardised methods for the detection of OCB are commercially available, but interpretation is user-dependent. With experience however, the sensitivities for detection of OCB (using isoelectric focusing followed by immunoblotting) is greater than 95% (Thompson and Freedman, 2006).

Unpaired OCB are found in the CSF of approximately 95% of patients with multiple sclerosis (Thompson and Freedman, 2006). Therefore, although the absence of unpaired OCBs does not exclude a diagnosis of MS, caution is warranted, particularly if there are additional atypical clinical or imaging findings.

1.5 Treatment of multiple sclerosis

The number of treatments available for RRMS patients has increased exponentially over the past decade. The first disease modifying treatment (interferon beta-1b) was licensed in 1993 after being shown to reduce relapse rates by approximately 30% compared with placebo (Paty and Li, 1993)(PRISMS (Prevention of Relapses and Disability by Interferon-1a Subcutaneously in Multiple Sclerosis) Study Group 1998). In general, since then treatments have become ever more efficacious.

Disease modifying treatments (DMTs) can be divided (somewhat subjectively) into those which are considered lower in their efficacy of controlling disease, but with lower risk of potentially serious adverse effects; higher efficacy, which offer better disease control but

greater risk of potentially serious adverse events; and intermediate efficacy, which lie somewhere between the two.

In some cases, individuals at high risk of a specific potential adverse effect can be identified. For example, an older patients with a high JC virus load would be considered a high risk patient for receiving Natalizumab (Bloomgren *et al.*, 2012). However, in order for patients to make an informed decision about the risks and benefits of treatments, they need to be counselled on how severely their disease may affect them in the future. This requires clinical tools, applied early in the disease (and ideally at diagnosis), which would identify individuals at the greatest risk of developing fixed disability. Identification of this cohort could justify the use of higher efficacy treatments and at an earlier stage in the disease. Conversely, individuals deemed at low risk of a poor prognosis could avoid riskier treatments and the potential adverse effects associated with them.

At present such tools are crude, and are based upon old natural history studies of untreated MS which helped identify risk factors for poor outcomes (Weinshenker *et al.*, 1989). Consequently, treatment practices vary between clinicians, and both 'induction' and 'escalation' approaches are practiced (Ontaneda *et al.*, 2019). Inadvertently some patients may be unnecessarily exposed to potentially serious adverse effects of treatments, whilst others risk inadequate suppression of MS inflammatory activity.

1.6 Prognosis of MS

Whilst genome-wide association studies have identified more than two hundred gene variants which may increase the risk of developing multiple sclerosis, none have been shown to affect clinical course (Reich, Lucchinetti and Calabresi, 2018). Modifiable risk factors, such as cigarette smoking and obesity in young adulthood, are known to increase disease susceptibility, and may also have an impact on disease prognosis. For example, continued cigarette smoking after diagnosis of RRMS has been associated with quicker conversion to secondary progressive multiple sclerosis than those who stopped smoking at the point of diagnosis (Ramanujam *et al.*, 2015).

Current treatment decisions are primarily based upon the clinical history along-with the MR imaging findings. Relapse frequency and presence of MRI activity often features in the eligibility criteria for disease modifying therapies. Frequent relapses in early disease (Scalfari *et al.*, 2010) and a high burden of infratentorial and spinal lesions on MRI

(Langer-Gould *et al.*, 2006; Tintore *et al.*, 2010) are believed to be risk factors for poor outcome.

A greater frequency of relapses in years one and two has been shown to reduce the time taken to reach disability milestones (Scalfari *et al.*, 2010). However, this was reported in an *untreated* cohort, and is dependent on the identification or patient self-reporting of relapses and ability to define a 'start date' of the disease. In addition, using the number of relapses as a prognostic indicator inadvertently requires clinical disease activity to occur. An ideal prognostic tool would accurately determine risks of future neurological events without them actually happening (similar to stroke prevention).

1.6.1 Current MRI methods used in prognosis

Correlation between conventional MRI findings (such as T2 lesion volume) and clinical disability is weak. This is often referred to as the "clinic-radiological paradox" of multiple sclerosis (Barkhof, 2002; Brownlee *et al.*, 2019).

There are numerous potential reasons for this. These include a lack of sensitivity of measures of disability such as the EDSS, the fact that widely use standard MRI techniques do not visualise well cortical grey matter lesions (Geurts *et al.*, 2005; Mistry *et al.*, 2011), or because the spinal axis is often not imaged in full and with the addition of Gadolinium enhancement in order to optimise sensitivity (Wattjes *et al.*, 2015; Brownlee *et al.*, 2019).

An important additional explanation for the lack of correlation between MR imaging and the clinical prognosis in MS is that conventional MRI remains insensitive to microscopic pathology, as demonstrated by pathological evidence of axonal transection within 'normal appearing' white matter (NAWM)(B D Trapp *et al.*, 1998). There are therefore substrates of disability which are not so well visualised on conventional 1.5/3T MR imaging (Mistry *et al.*, 2011).

Brain or spinal cord atrophy are thought to reflect the accumulative, irreversible loss of neurons and glia cells as a consequence of multiple sclerosis. Atrophy measures are therefore an estimate of neurodegeneration (tissue destruction), and do correlate with longer term measures of clinical disability (Fisher *et al.*, 2000). However, atrophy measures are problematic to employ in standard practice in individual patients (De Stefano, Battaglini and Smith, 2007), and are insensitive to small changes in early disease. For example, "pseudoatrophy", describes a phenomenon whereby rates of brain atrophy

initially appear to accelerate after the initiation of disease modifying treatment. This is thought to represent an early reduction in white matter lesion oedema and inflammatory infiltrates as a result of a treatment-response; however its presence complicates the interpretation of treatment effects on brain volume (De Stefano and Arnold, 2015).

Conventional MRI can therefore be considered to be a sensitive biomarker of focal demyelination, but its role as a biomarker of neurodegeneration is currently limited. However, there are opportunities to develop aspects of advanced MR imaging which could confer greater information about MS lesion biology and provide better imaging biomarkers of neurodegeneration. These are described in the imaging biomarker section in 1.7.3.

1.6.2 Current CSF methods used in prognostication

Other than the role of OCB at diagnosis (Arrambide *et al.*, 2018), no CSF biomarkers are used in clinical practice for disease monitoring or prognostication. The main reason for this is likely to be the invasive nature of CSF sampling.

Previous studies have demonstrated there may be prognostic relevance in the identification of the subtype of OCB. IgG OCB are the common form that are tested by most laboratories and are found in approximately 95% of individuals with MS. IgM OCB against myelin lipids are less common. Their presence has been associated with a greater frequency of relapses and higher rates of disability that individuals with IgG OCB or IgM OCB that do not recognise myelin lipids (Villar *et al.*, 2008).

There may also be a prognostic relevance in the absence of CSF oligoclonal bands in individuals with MS. A greater proportion of patients with a diagnosis of benign MS (EDSS <3.5 at 10 years) were OCB negative compared with those diagnosed with severe MS (EDSS >7.5 at 10 years)(Avasarala, Cross, *et al,* 2001). However, as previously discussed, the absence of unpaired oligoclonal bands should also raises suspicion of alternative diagnoses.

In summary, the conventional MRI sequences and the routine CSF markers used in clinical practice have limited ability to predict an individual's future risk of a poor outcome from multiple sclerosis.

1.7 Biomarkers

1.7.1 Definition of a biomarker

A biomarker is an objective measure of a physiological or pathological parameter that enables characterisation of the health or disease state of an individual (Raphael *et al.*, 2014). Broadly speaking, biomarkers can be divided into clinical measures, radiological measures or laboratory analysis of a fluid or tissue (O'Connor *et al.*, 2006).

Biomarkers may be employed in different ways and at different stages in a disease - diagnosis; prognosis; identification of appropriate treatment groups; treatment monitoring for efficacy; identification of side effects; identification of a treatment end-point.

1.7.2 Clinical biomarkers for quantifying outcomes

When MR imaging biomarkers are measured against clinical phenotype there is greater disparity than when they are measured against pathological analysis. The accuracy of any biomarker is therefore ultimately dependent upon what it is measured against.

There are a myriad of clinical examination tools which have been developed specifically for use in the field of MS research. All share a common aim - to standardise the disability or clinical symptoms of a clinically heterogenous MS population. The most commonly used measure of disability is the Kurtze Expanded Disability Status Scale (EDSS). A commonly used measure of function is the Multiple Sclerosis Functional Composite measure (MSFC).

<u>Kurtze Expanded Disability Status Scale (EDSS)</u>

The routine clinical neurological examination was translated into a clinical examination disability scale specifically for use in MS - The Kurtze Expanded Disability Status Scale (EDSS) (Kurtzke, 1983). The EDSS divides the neurological examination into eight functional systems, assigning a score to each based on a combination of clinical signs and patients reported symptoms. The scores from each system are then combined to produce an overall EDSS score ranging from 0 (asymptomatic) to 10 (death as a result of MS).

The EDSS has well recognised limitations, primarily that the scoring is heavily dependent on mobility performance. In addition, it is considered to be insensitive to small changes in function (particularly changes in cognitive function) and is non-linear. That said, the EDSS

remains the most widely used standardised scoring system for describing clinical disability in MS and is included in all major studies and trials.

Multiple Sclerosis Functional Composite measure (MSFC)

The MSFC was developed to better reflect the overall impact of multiple sclerosis on function, particularly cognitive function. The MSFC is composed of three separate measures of performance - upper limb function, lower limb function and cognition. Upper limb and hand function are examined with the timed 9-hole peg test. Lower limb function and ambulation are assessed with a timed 25-foot walk. Finally, the Paced Auditory Serial Addition Test (PASAT) assesses auditory information processing speed and flexibility as well as calculation ability (Fischer *et al.*, 1999). The MSFC overall score is calculated from z-scores, whereby each component of an individual's score is measured in terms of standard deviations compared with the mean of the group. A lower zMSFC represented a poorer performance across MSFC tests.

The primary advantage of the MSFC over the EDSS is that it is less dependent on ambulation and therefore captures additional dimensions of the impact of multiple sclerosis on function. The disadvantages are that interpretation of the scoring is not intuitive and that it requires specific equipment, such as a 9-hole peg test (Fischer *et al.*, 1999).

1.7.3 Evolving concepts in MRI biomarkers

MRI is a cornerstone in the diagnostic criteria for multiple sclerosis. 7T MRI scanners and growing libraries of MRI sequences could enable a diagnosis of MS to be made with ever greater sensitivity and specificity (Sinnecker *et al.*, 2019). However, the role of MRI in the management of MS patients goes beyond diagnosis.

With an increasing number of treatment options available, MRI is now used in routine clinical practice as a biomarker of MS disease activity. For example, eligibility criteria for immunotherapies such as ocrelizumab depend on identification of active disease, on clinical or MRI criteria (Scottish Medicines Consortium, 2020). Radiological evidence of active MS may therefore result in the initiation or a change in therapy, despite an absence of clinical relapses. Routine MR imaging in individuals established on certain DMTs (Natalizumab, Fingolimod, Dimethyl-fumarate) also offers a secondary benefit as a safety biomarker in the identification of potentially fatal progressive multifocal leukoencephalopathy (PML).

For the reasons discussed above the role of *conventional MRI* as a prognostic biomarker in multiple sclerosis remains limited. However, technological advances in MR imaging which enable the quantification of both focal and diffuse structural changes, in white as well as grey matter, offer new opportunities. For example, advanced MRI techniques now enable the indirect measure of myelin integrity.

Quantification of myelin integrity provides an opportunity to identify myelin disruption at an earlier stage than previously possible, and prior to the development of large focal lesions. An *in-vivo* biomarker of myelin integrity could enable patients to be classified according to their ability to remyelinate damaged axons, and may also be of use in the identification of remyelination therapies (Compston and Coles, 2008; Podbielska *et al.*, 2013). One such method of quantification of myelin integrity is measurement of the MRI aggregate g-ratio (Stikov *et al.*, 2015; Campbell *et al.*, 2018).

The g-ratio is defined as the ratio of the inner axonal radius to the myelinated fiber radius. It is a measure of the degree of myelination of an axon (*figure 5-1*). The g-ratio is quantified by electron microscopic visualisation and measurement, however an aggregate g-ratio can be calculated using the advanced MR imaging techniques of Magnetization Transfer Saturation (MTsat) and Diffusion Weighted Imaging (DWI).

Magnetization transfer

(Schmierer *et al.*, 2004; Stikov *et al.*, 2015; St-Amant and O'Gorman, 2020)

Magnetization transfer refers to the exchange of energy between protons from three pools:

1) bound up in macromolecules in myelin, membranes and proteins

2) in 'free' water; and

3) in water layers between membranes.

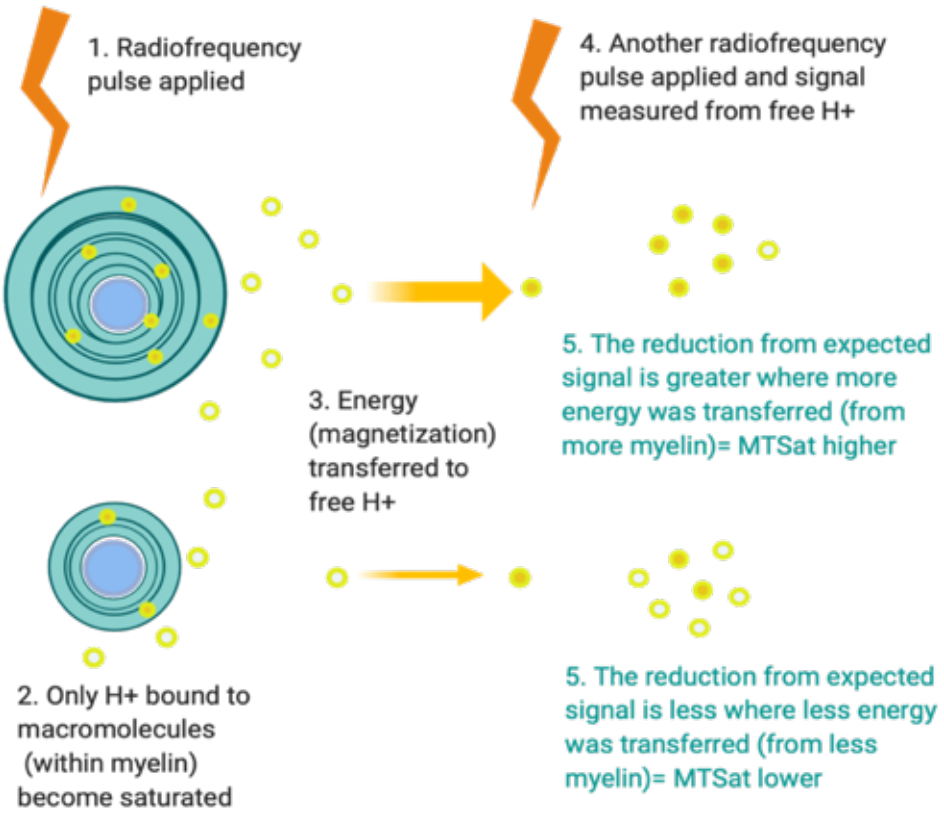


Figure 1-6 Magnetization Transfer

Myelinated axons have more macromolecules to absorb energy than demyelinated axons. After a radiofrequency pulse is applied, macromolecules transfer their energy (magnetization transfer) to free hydrogen ions. Axons with greater number of macromolecules will transfer more energy. Therefore, when a second radiofrequency pulse is applied, areas with more myelin will have already transferred greater energy, i.e. will have a greater magnetization saturation.(St-Amant and O'Gorman, 2020)

The difference between the free water signal detected with, and without, the preceding magnetization pulse produces the magnetization transfer ratio (MTR). A higher MTR occurs when more protons are bound within macromolecules such as myelin, whereas a lower MTR results from less magnetization transfer, i.e. less protons bound in macromolecules such as myelin. MTR has been shown to correlate well with histological reports of myelin content (Schmierer *et al.*, 2004). MTsat is a variation on MTR that corrects for T1 relaxation and B1 inhomogeneities.

Diffusion weighted imaging

Diffusion weighted imaging (DWI) is based on the principle that free water molecules move differently to those contained within cellular structures.

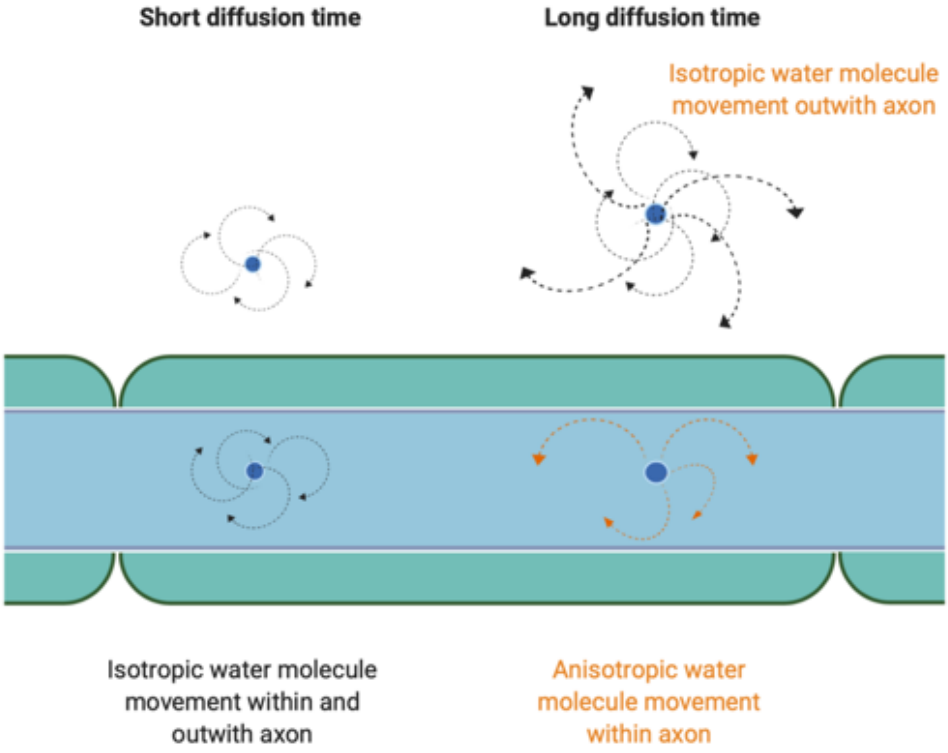


Figure 1-7 Diffusion weighted imaging

Within an axon, water molecules will show isotropic distribution at short diffusion times, but diffusion anisotropy at long diffusion times. Figure created in BioRender by SJM.

Diffusion models can be used estimate the size and integrity of a cellular compartment based on the movement patterns of free water molecules (Rovira, Auger and Alonso, 2013). For example, anisotropic water molecule movement within a MR voxel suggests that the water molecules are contained within cellular structures, such as an intact axon. Isotropic water molecule movement within an MR voxel may however suggest disruption

of tissue architecture, such as axonal damage. Diffusion weighted imaging has demonstrated isotropic patterns of water molecule movement within normal appearing areas of brain tissue in MS patients, suggesting subtle, diffuse injury to tissue structures (Rovira, Auger and Alonso, 2013).

1.7.4 Evolving concepts in CSF biomarkers

Even with the advances in MR imaging techniques in MS, there is still an unmet need for a biomarker of axonal loss. Due to its proximity to the pathology, CSF offers greater resolution with which to identify and study proteins of brain damage. By-products of neuronal damage, such as the release of intrinsic structural proteins, offer a source of potential markers with which to measure neuronal injury and axonal loss.

Intermediate filament proteins (IFs) provide structural support to eukaryotic cells. Intermediate filaments are so-called because of the diameter of the assembled filament (8-10 nm) when compared with the two other cytoskeleton structures - actin microfilaments and microtubules.

Different cell types produce distinct IFs to meet tissue-specific requirements, and six main subgroups have been categorised according to amino acid sequence (Petzold, 2005):

- IF type 1 - Acidic keratins
- IF type 2 - Basic and Neutral keratins
- IF type 3 - Vimentin, desmin, peripherin and glial fibrillary acidic protein (GFAP)
- IF type 4 - Neurofilaments and alpha internexin
- IF type 5 - Lamin A, B and C (the only IF to be nuclear as opposed cytoplasmic)
- IF type 6 - Nestin

Within the nervous system, intermediate filaments are found in astrocytes (IF type 3 – GFAP) and neurons (IF type 4 – Neurofilaments). Here we will focus on neurofilaments.

Neurofilaments are composed of three polypeptide subunits which differ in molecular weight and phosphorylation, known as neurofilament light, medium and heavy chains. Assembled subunits are orientated along the axon, and form an integral part of the extreme cytoarchitecture of neurons - the neuronal cell body measures between 0.001 to 0.005mm diameter, but may support an axon up to 1m in length (Petzold, 2005).

Both neurofilament light chain (NfL) and neurofilament medium chain (NfM) are coded on chromosome 8p21. Neurofilament light chain is composed of 543 amino acids and has a molecular mass of 61kDa. However, as a result of phosphorylation and glycosylation, NfL is often described as a molecular mass of 68kDa due to its slow migration in sodium dodecyl sulfate (SDS) polyacrylamide gels (PAGE). Neurofilament medium chain is composed of 916 amino acids, has a true molecular mass of 102.5kDa, and a molecular mass of 150kDa calculated from SDS-PAGE. Neurofilament heavy chain is coded on chromosome 22q12.2, is comprised of 1020 amino acids and has a true molecular mass of 111kDa. Neurofilament heavy chains have a long tail region with variable degrees of phosphorylation, which give NfH a SDS-PAGE calculated molecular mass of between 190-210kDa (Petzold, 2005).

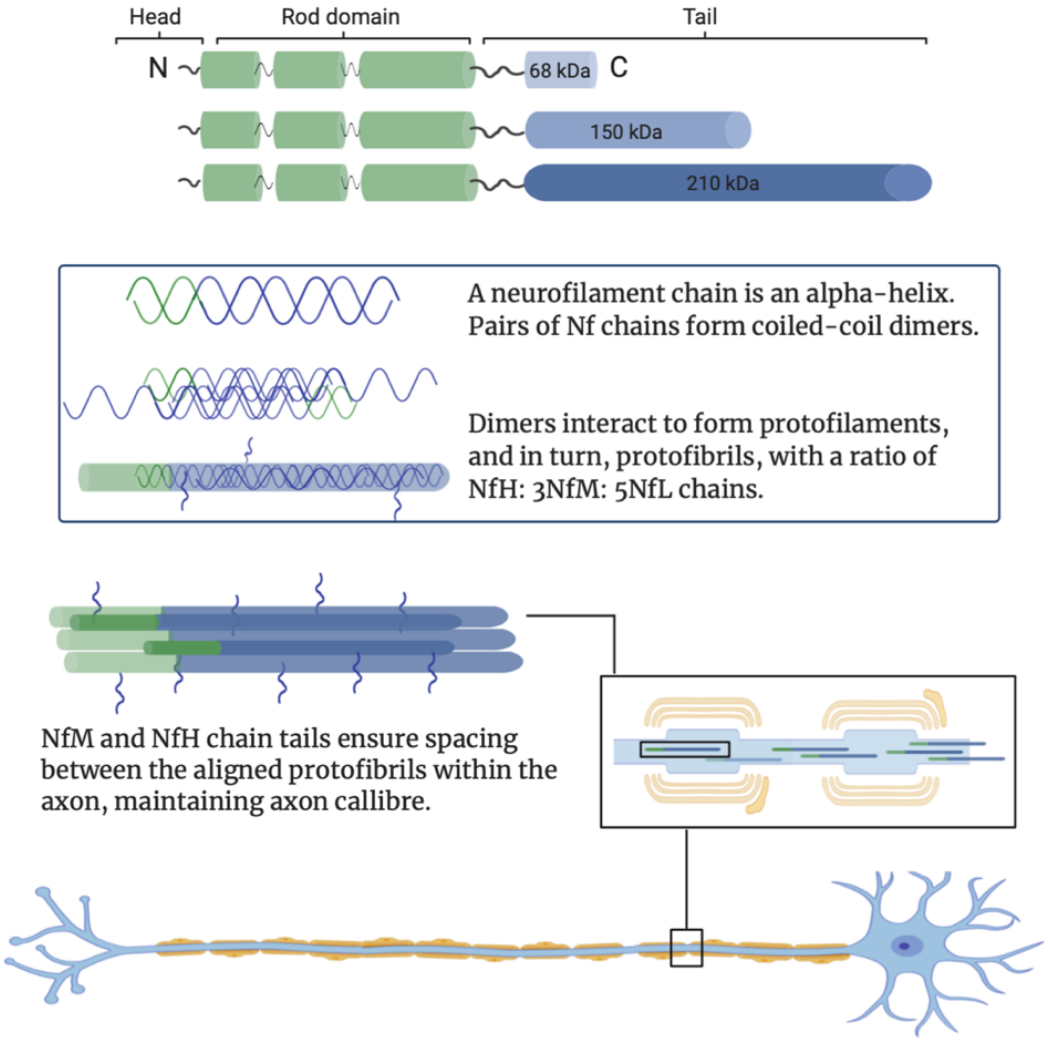


Figure 1-8 The structure of neurofilament

Neurofilaments are composed of light, intermediate and heavy chains in a ratio of 5:3:1. NfL, NfM and NfH differ in molecular weight and extent of phosphorylation. The subunits interact to form coiled-coil dimers, then protofilaments, the protofibrils. The final neurofilament is important for growth in axonal diameter during development and in maintaining axon calibre. Figure created in BioRender by SJM.

Each neurofilament subunit contains a central hydrophobic rod domain, a subtype specific globular head domain (N-terminus) and a variable tail domain (C terminus), *see figure 1-8.* Each domain plays a specific role. The head domain regulates microtubules within the axon. The rod domains allow interactions between the subunits. The primary function of the tail domain is to form extensions which increases the distance between neighbouring neurofilaments and microtubules, thus increasing axon calibre (Yuan *et al.*, 2012).

Neurofilament construction starts with the rod domains of two alpha-helix subunits forming a coiled-coil dimer, with head-to-head and tail-to-tail ends. Two dimers then

interact in an anti-parallel fashion to form a tetramer, with two head and two tail domains at either end. Adjacent tetramers form protofibrils. Protofibrils interact via the head domain to produce a neurofilament 10nm in diameter.

After a neurofilament is formed in the neuronal cell body it is transported into the axon. Neurofilaments are considered very stable proteins. Although they have bi-directional movement within the axon, their overall path is to move slowly along the axon from cell body to axonal terminal (Al-Chalabi and Miller, 2003). Under healthy conditions neurofilaments are degraded only when they reach the axon terminals (Lee and Cleveland, 1996). The stability of neurofilament is thought to occur to arise from their phosphorylation (Goldstein, Sternberger *et al*, 1987).

Depending on the length of the axon, neurofilaments may have a lifespan of up to two years (Lee and Cleveland, 1996). During this time neurofilaments can undergo inflammatory or oxidative damage which results in their accumulation. Accumulation of neurofilament proteins has been noted in neurodegenerative diseases other than MS, including amyotrophic lateral sclerosis (Rouleau *et al.*, 1996) and dementias (Petzold, 2005). Accumulation of aberrant neurofilaments can result in loss of axonal structural integrity and ultimately in axonal transection. Axonal transection, from any cause, releases neurofilament sub-units into the extracellular space, enabling their detection in CSF (Petzold, 2005), *see figure 1-9*. The slow turnover of neurofilaments under healthy conditions means that their elevated concentrations in the CSF mainly reflect aberrant release as a result of axonal transection. Neurofilament levels are therefore a biomarker of axonal damage.

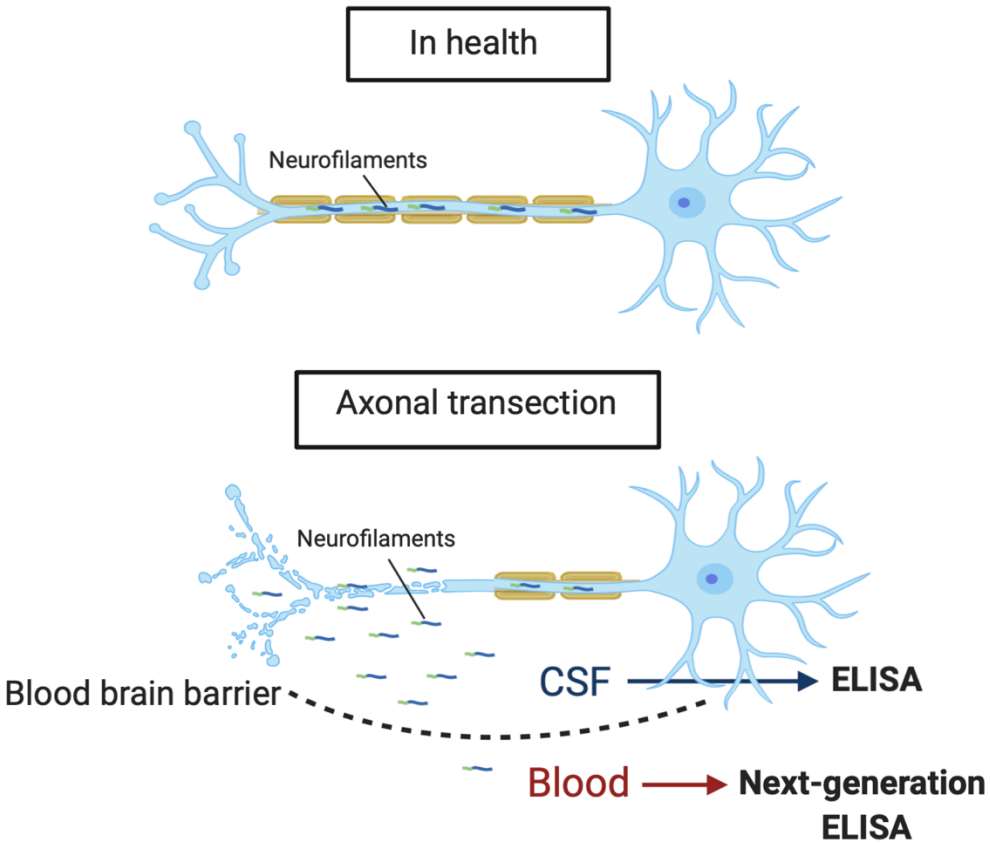


Figure 1-9 Neurofilament is released as a result of axonal transection

Axonal transection, from any cause, releases neurofilament subunits into the extracellular fluid thereby allowing their detection in CSF (using ELISA), or blood (using next-generation ELISA). Figure created in BioRender by SJM.

Neurofilament subunits are formed at a ratio of NfH: NfM: NfL of 1:3:5. Neurofilament light chain is the smallest, the most soluble, the least phosphorylated and the most plentiful, and as a result has been the main focus of biomarker research, particularly within the field of multiple sclerosis.

In 1996, Rosengren et al were the first to measure NfL in cerebrospinal fluid by adapting an enzyme-linked immunosorbent assay (ELISA) protocol for the measurement of GFAP (Rosengren *et al.*, 1996). The result was an ELISA able to quantify NfL levels ranging between 125-16,000 ng/L. Two years later, Lycke et al used the same assay to compare CSF NfL levels between 60 RRMS patients and 11 HCs. NfL was detectable in 78% of MS patients, but below the limit of detection in all HCs (Lycke *et al.*, 1998).

Detecting neurofilament by Enzyme-linked immunosorbent assay

Enzyme-linked immunosorbent assay (ELISA) is an analytical biochemistry assay which utilises an antibody to detect a specific antigen.

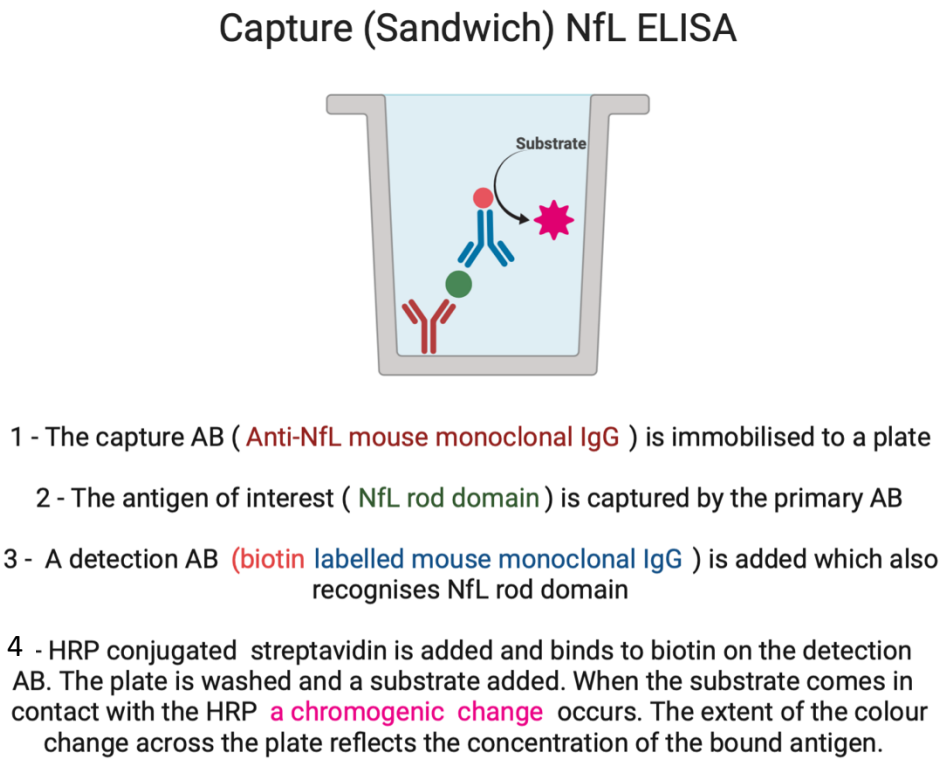


Figure 1-10 Capture (sandwich) neurofilament ELISA

Sandwich ELISAs are a commonly performed type of ELISA. The neurofilament ELISA described above employs capture and detection antibodies made commercially available through Uman Diagnostics.

There are several different methods of performing an ELISA. These differ, for example, in whether the antigen or the antibody is bound to the plate or in a solution (direct ELISA versus capture ELISA). All ELISAs however are fundamentally dependent on the antigen-antibody pairing, which is the key element determining the sensitivity and specificity of the assay.

In 2003, Norgren et al used hybridomas to identify antibody pairs with superior affinity, and by doing so were able to reduce the lower limit of detection of the NfL ELISA from 125ng/L to 60ng/L. Over the following 10 years Norgren's assay was also outdated by the development of a commercial ELISA with a sensitivity of 32pg/mL (UmanDiagnostics). Improving the sensitivity of the assay enabled CSF NfL to be accurately quantified in both MS patients and healthy controls.

A large body of work has examined the potential role of CSF NfL both in multiple sclerosis and in many other neurological diseases. This has been summarised in a large meta-analysis comprising approximately 10,000 individuals and 35 neurological diseases (Bridel *et al.*, 2019). CSF neurofilament levels differed significantly between different diseases. The highest NfL levels were found in cognitively impaired HIV positive patients (where levels were approximately 20 times higher than in healthy controls) and amyotrophic lateral sclerosis (approximately ten times higher than in healthy controls). In contrast, CSF NfL levels in multiple sclerosis were less than five times those seen in healthy controls (Bridel *et al.*, 2019). CSF NfL levels in frontotemporal dementia were significantly higher than other dementias, and may therefore have use in this context as a diagnostic biomarker (Bridel *et al.*, 2019). These results show that neurofilament light chain is a non-specific marker of axonal damage, and that *disease-specific* reference values are likely to be required to interpret clinically meaningful results.

Within the field of multiple sclerosis, levels of CSF NfL have been shown to be higher in MS patients compared with controls, to rise in the context of relapse, and to fall with certain disease modifying treatments (Lycke *et al.*, 1998; Novakova *et al.*, 2016; Novakova, Axelsson, *et al.*, 2017). Although further systematic review and meta-analysis of the data are required, results reported suggest that CSF NfL has the potential to be a clinically meaningful marker of disease activity. However, CSF is invasive to obtain, making it unsuitable for repeat or longitudinal sampling and less likely to be available from both patients and controls. This is probably the main reason for the lack of CSF biomarkers used clinically, not only in multiple sclerosis, but across neurodegenerative diseases as a whole.

1.7.5 Evolving concepts in blood biomarkers

Single molecule array (Simoa) is a type of next-generation, bead-based sandwich ELISA. Simoa technology which was originally described in 2010 (Rissin *et al.*, 2010).

In a Simoa, instead of the capture antibody being immobilised to a plate it is coupled to paramagnetic beads, *see figure 1-11.* When the beads are added to an experimental sample the protein of interest is therefore captured on the beads. Detection and secondary antibodies are then added, punctuated by wash cycles in a similar series of steps as a traditional ELISA. The entire antigen-antibody complex occurs on the bead.

In a traditional ELISA the antigen-antibody-detector complex is formed within a 360µL-volume well and the colour change of the reaction (the optical density) across the well is determined by spectrophotometry. In a Simoa assay the beads are separated into individual 50fL-volume wells spread over a disc. The wells are sealed with a layer of oil, thereby containing each fluorescent reaction within individual wells.

The disc is then imaged in several steps to determine which wells contain a bead, and which beads are 'on' – i.e. produce a fluorescent signal, or 'off'. A white light image is first used to identify the wells containing beads. Fluorescent images are obtained over a series of frames to determine whether beads are 'on' or 'off'. For a bead to be labelled as 'on', it must have a detectable fluorescent signal in all five frames taken, and the fluorescence must increase by greater than 20% over four frames. The proportion of wells containing 'on' beads compared to those containing 'off' beads is calculated to produce the 'average enzyme per bead' (AEB) value. Calibrators with predetermined concentrations enable the protein concentration of the experimental sample to be interpolated from the AEB in the same manner as a traditional ELISA.

The primary difference between a traditional ELISA and a Simoa is therefore the physical volume in which the detectable signal is contained and measured. By reducing the volume, and thus increasing the local concentration of fluorescent signal, a single molecule is sufficient to produce a detectable fluorescence. This technology has reduced the lower limit of detection of proteins a thousand-fold to femtomolar concentrations (10^{-15}) (Rissin *et al.*, 2010).

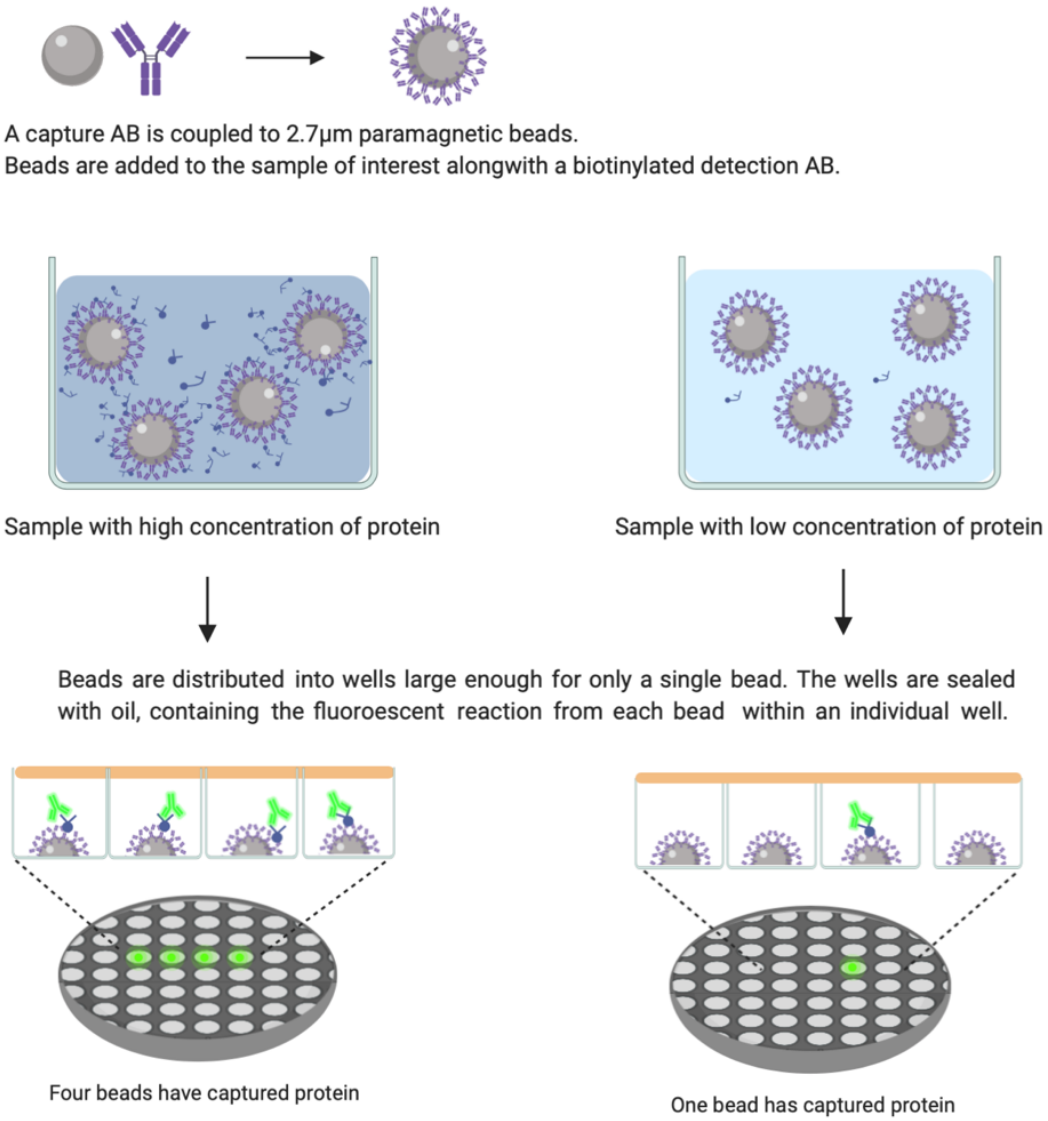


Figure 1-11 Principle of Ultrasensitive Single molecule array (Simoa)

(A) A capture AB (e.g. IgG1 monoclonal anti-NfL) is coupled to paramagnetic beads. (B) The beads are added to the sample of interest in a 96-well plate. A series of washes removes uncoupled protein from the beads, and then a detection AB (IgG1 monoclonal biotin-labelled anti-NfL) is added, followed by a fluorescent substrate and reporter enzyme. The samples are incubated before a further wash cycle. (C) Within the Simoa analyser the beads are distributed throughout a disc containing femtolitre seized wells. (D) The wells are imaged using white light and fluorescent images to determine how many wells contain a bead, and what proportion of beads have captured the protein of interest. Figure created in BioRender by SJM.

Next-generation ELISA technology has the potential to address significant gaps in the field of clinically relevant fluid biomarkers, not only in multiple sclerosis, but across neurodegenerative diseases as a whole.

Enabling biomarker research in blood as opposed CSF aids access to both disease and control samples, increases the volume of samples available for testing and allows for longitudinal sampling. As a result, since the introduction of Simoa blood levels of neuronal and glial proteins have been measured in various neurological diseases, including multiple sclerosis. To date these studies have often been performed retrospectively, using biobanked samples, and usually in heterogenous cohorts of relapsing and progressive MS patients.

1.8 Large cohorts are required for biomarker research

A major unmet need for multiple sclerosis patients currently is a biomarker which can be used *early in the disease course* to identify patients at risk of a poor outcome. The advances in MRI and fluid biomarkers described above have the potential to change this.

To fully address this issue requires a standardised and methodical approach. Firstly, a large, homogenous cohort of early MS patients is required. Ideally the cohort should also be treatment naïve. Certain disease modifying treatments are known to affect biomarkers, for example normalise MRI appearances and reduce neurofilament levels (Novakova *et al.*, 2016). Identification of a large cohort of treatment naïve, early MS patients is difficult in the treatment era, and requires recruitment of individuals *at the point of diagnosis*.

In addition, detailed clinical and radiological phenotyping of the cohort is needed to correlate biomarker outcomes with clinical measures. Specific technologies are required to perform next generation single molecule arrays or advanced MRI imaging. Finally, the cohort needs to be longitudinally followed up for many years in order to determine the true use of any potential novel biomarkers.

Summary of Chapter One

- Multiple sclerosis (MS) is a serious and unpredictable disease with complex pathophysiology, but it is also increasingly treatable. However, highly efficacious treatments carry serious potential risks.

- There is a paucity of clinical tools to identify patients at risk of a poor prognosis, and treatment decisions are therefore difficult and inconsistent.

- There is therefore an unmet need for a clinical tool or biomarker which can be used early in the disease to identify patients at risk of poor prognosis.

- Advances in MR imaging and next-generation single molecule arrays have the potential to address the unmet need but require careful analysis and validation in large cohorts.

1.9 Aims of thesis

Aims of Thesis

Aim One: To establish a large cohort of newly diagnosed, treatment naïve relapsing remitting multiple sclerosis patients - The Future MS study.

Aim Two: Use the Future MS study to evaluate the potential role of next-generation biomarkers in early RRMS.

Chapter Two

Materials and Methods

2 Materials and methods

This chapter provides detailed descriptions of the clinical and laboratory research methodology presented in this thesis.

2.1 Future MS Ethical approvals (including substudies)

Overarching study design and ethical approval for the Future MS study were sought and obtained by the Future MS study team at the Anne Rowling Regenerative Neurology Centre, led by Professor Siddharthan Chandran and Dr Peter Connick. Ethical approval reference REC 15/SS/0233.

Study design and ethical approval for the BEFORE-MS research substudy were carried out by the author of this thesis, with assistance from the Glasgow Neuroimmunology laboratory at the Queen Elizabeth University Hospital, Glasgow.
Ethics approvals REC 17/LO/1611 and REC 17/WM/0379.

2.2 Future MS protocols (including substudies)

2.2.1 Future MS Eligibility criteria and funding

Table 2-1 Inclusion and exclusion criteria for the Future MS study

Inclusion Criteria
1. A diagnosis of relapsing-remitting multiple sclerosis (RRMS), fulfilling the 2010, or updated 2017, McDonald Criteria
2. Diagnosis of multiple sclerosis within six months of baseline assessment
3. Aged 18 years or over
4. Capacity to provide written informed consent
Exclusion Criteria
1. Initiation of disease modifying treatment prior to baseline assessment
2. Participation in a clinical trial for an investigational medicinal product (IMP)
3. Contraindication to magnetic resonance imaging (MRI)

The study was funded by Stratified Medicine Scotland Innovation Centre (SMS-IC) with additional support from Biogen.

2.2.2 Recruitment sites

Participants were recruited from the fourteen NHS health boards and study visits were conducted across five main sites – Aberdeen Royal Infirmary; Ninewells Hospital, Dundee; The Anne Rowling Clinic, Edinburgh Royal Infirmary; Queen Elizabeth University Hospital, Glasgow; and Raigmore Hospital, Inverness.

2.2.3 Primary and secondary aims of the Future MS study

Table 2-2 Aims of the Future MS study

Primary Aim
• To develop predictive tools for focal neuroinflammatory disease activity over 12 months, based on baseline clinical, imaging, laboratory and genomic assessment.
Secondary Aims
• To develop predictive tools for neurodegenerative disease activity over 12 months, based on baseline clinical, laboratory and genomic assessment. • To develop predictive tools for clinical measures of disease activity. • To develop predictive tools for clinical measures of quality of life.

2.2.4 Study data recording

Each participant was identified by a unique Future MS study ID in the form 01-xxxx, where 01 was the site number (01 = Edinburgh; 02 = Glasgow; 03 = Dundee; 04 = Aberdeen, 05 = Inverness). Future MS study IDs were allocated in chronological order of recruitment, and were used to label all paperwork, biological samples and imaging results.

Clinical data were recorded in paper files which were stored in a locked clinical research facility. Additional demographic and clinical data from patient questionnaires were recorded on a web-based electronic data system. The electronic data system was designed and maintained by Aridhia Informatics, Edinburgh. Data extraction was performed by Dr Patrick Kearns.

2.2.4.1 Future MS study protocol

Potential participants were identified by clinical care teams and invited to participate. If they were interested in participating a member of the Future MS research team then contacted them directly to discuss participation and ensure they met inclusion criteria. A study information leaflet (version 1.3, May 2016) was provided by post or email. Potential

participants were followed up to ensure they had received the information leaflet and, if they wished to participate, to book a research visit.

Research visits were conducted by myself, Rowling clinical fellows and Future MS research nurses. Written informed consent was obtained, and a copy of the signed consent form given to all participants. Participants were made aware of their right to withdraw from the study at any time and for any reason. All clinical, laboratory and imaging assessments were performed in a standardised manner at each research visit. Data collected in the main study are outlined in *table 2-3*.

Four optional sub-studies were offered alongside the main Future MS research study. Separate patient information leaflets were provided for each substudy, and participation required written consent in the form of a dated participant signature. All substudies fell under the same ethical approval as Future MS, and additional ethical approval was obtained to allow access to linked serum/CSF samples as described above.

Follow-up 12 months review assessments were arranged by the clinical fellows and research nurses covering each site.

2.2.4.2 *Future MS study data collection*

Table 2-3 Future MS data collection at baseline and review visits

	Baseline visit	*12 month follow-up visit*
Demographics	Date of birth Sex Address GP Ethnicity Occupation	Updated where necessary
Medical history	Past medical history Medications 'Over-the-counter' medications and supplements Family history	Updated where necessary
MS history	Date of first symptoms Descriptions and dates of relapses Hospitalisations Use of oral/IV steroids Date of diagnosis	Disease modifying treatment Descriptions and dates of relapses Hospitalisations Use of oral/ IV steroids
Clinical measures	Blood pressure (BP) Body mass index (BMI)	Blood pressure (BP) Body mass index (BMI)
Clinical assessments	Expanded Disability Status Scale (EDSS) Timed 25-foot walk (T25W) Paced auditory serial addition test (PASAT) 9-hole peg test (9HPT) Symbol Digit Modalities test (SDMT)	Expanded Disability Status Scale (EDSS) Timed 25-foot walk (T25W) Paced auditory serial addition test (PASAT) 9-hole peg test (9HPT) Symbol Digit Modalities test (SDMT)
Patient completed questionnaires	MS Impact Scale (MSIS) CDC Health-related quality of life (CDC-HRQOL-4)	MS Impact Scale (MSIS) CDC Health-related quality of life (CDC-HRQOL-4)

	Patient determined disease steps (PDSS) Generalised Anxiety Disorder-7 (GAD-7) Patient Health Questionnaire 9-item depression scale (PHQ-9) Baecke Habitual Physical activity Cognitive leisure questionnaire Social and lifestyle questionnaire	Patient determined disease steps (PDSS) Generalised Anxiety Disorder-7 (GAD-7) Patient Health Questionnaire 9-item depression scale (PHQ-9) Baecke Habitual Physical activity Cognitive leisure questionnaire Social and lifestyle questionnaire
Clinical blood samples*	Full blood count Renal function Liver function Cholesterol HBA1c Vitamin D CRP	
Blood for storage*	PBMCs were stored for genetic analysis	
MRI	3T non-contrast MRI brain scan. Sequences included image localiser, volumetric (1mm³ isotropic) T1 / MP-RAGE, T2, and fluid attenuated inversion recovery (FLAIR)	3T non-contrast MRI brain scan. Sequences included image localiser, volumetric (1mm³ isotropic) T1 / MP-RAGE, T2, and fluid attenuated inversion recovery (FLAIR)

* Performed at baseline visit only.

2.2.4.3 Study blood sampling and storage

Clinical blood tests required x2 ochre 5.7ml serum separating tubes and x1 purple 4.0ml EDTA plasma tube and were performed by NHS laboratories following local protocols. Blood results were reviewed by the research teams at each site and any unexpected finding were relayed to the participant's GP.

Peripheral blood mononuclear cell (PBMC) extraction and freezing was performed at local CRFs using existing SOPs and required x8 PBMC extraction tubes (approximately 30ml).

Samples were batch transported for storage preparation, FACS sorting, and DNA and RNA extraction to the Wellcome Trust CRF Genetics Core Laboratory (Edinburgh). Samples were then cryopreserved (-196 °C) at a managed tissue bank in the Wellcome Trust CRF Genetics Core Laboratory (Edinburgh).

2.2.4.4 Study MRI acquisition

MR imaging was performed in dedicated research imaging facilities associated with the local clinical research facilities at each site. Images from each site were transferred to the University of Edinburgh imaging for processing. MRIs were also clinically reported at each site by the local neuroradiologist, and the clinical care team made aware of the MRI results.

2.2.5 Future MS substudies

Four optional sub-studies were offered alongside the main Future MS research study. Separate patient information leaflets were provided for each substudy, and participation required written consent in the form of a dated participant signature.

- Substudy One - Available to participants across all sites. Participation enabled the research team to store diagnosis and contact details on a secure NHS approved server within the University of Edinburgh, with the purpose of advising participants of additional research opportunities in the future.
- Substudy Two – *see section 2.2.5.1*
- Substudy Three – Available to participants at Edinburgh and Glasgow site. Participation involved detailed visual assessment and retinal imaging using optical coherence tomography to measure retinal nerve fibre layer (RNFL) thickness, ganglion cell and inner plexiform layer (GCIP) thickness, and inner nuclear layer (INL) thickness.
- Substudy Four – *see section 2.2.5.2*

2.2.5.1 Sub-study for biobanking of plasma samples

'Storage of donated tissue for future research purposes' (sub-study two) was open to individuals at all sites. Participation involved an addition 34ml of blood, drawn alongside the standard Future MS blood samples so that no additional venepuncture was required. Participants were made aware that samples were for research purposes only, and that no results would be made available to them.

Samples were processed in parallel with those for Future MS using SOPs for the separation and storage of plasma. In brief, vacutainers were centrifuged at 2000g for 15 minutes to separate the plasma, which was then pipetted out using a liquipipette into two 2.0ml cryovials. The cryovials were stored at -80°C. At a later stage the cryovials were aliquoted into 10 x 200µL aliquots. All samples had therefore undergone one freeze-thaw prior to analysis.

2.2.5.2 Sub-study for advanced MR imaging

'Extended MRI in Future MS' (substudy four) was available for patients recruited at the Edinburgh site between September 2017 and April 2020. Participants had addition magnetization transfer imaging (MTsat) and diffusion tensor imaging (DTI) sequences performed alongside their routine Future MS MRI brain scan. No additional visits were therefore required, but participation resulted in an additional 30 minutes of scanning time. Images were transferred to the University of Edinburgh imaging for processing.

2.2.5.3 BEFORE-MS substudy

'BEFORE-MS', Biomarker Evaluation For Outcome in RElapsing MS'- was devolved from the main Future MS study and had separate ethical approvals. BEFORE-MS was available only in Glasgow and enabled the acquisition of 'paired' CSF and serum specimens from two source - historical samples from the Glasgow neuroimmunology biobank for patients participating in Future MS, and prospective collection from individuals at the point of diagnostic work-up. Incorporation of BEFORE-MS alongside Future MS substudies enabled linkage of CSF and serum samples with clinical data.

Historical samples from Glasgow neuroimmunology biobank

Surplus samples (CSF and serum left over after OCB testing) are stored at -80°C by Glasgow Neuroimmunology Biobank at the Queen Elizabeth University Hospital. Specific ethics approvals were sought and acquired to enable the identification of biobanked specimens from patients who had gone on to participate in Future MS (IRAS ID 232152; REC 17/WM/0379; R&D GN17NE455). For inclusion we required that the surplus specimens of both CSF and serum be of sufficient volume (greater than approximately 100µL). Identified samples were stored at -80°C prior to transfer on dry ice to the Institute of Genetic and Molecular Medicine, University of Edinburgh for analysis.

Prospective samples collected at diagnostic work-up - BEFORE-MS

CSF, serum and urine samples were collected from patients undergoing routine diagnostic lumbar puncture at the Queen Elizabeth University Hospital, Glasgow (IRAS 224527; REC 17/LO/1611; R&D GN17NE373). All lumbar punctures were performed by myself. Clinical samples were collected first, followed by the research sample. Between 5-8ml of CSF was collected (depending on the CSF flow rate), into a polypropylene tube, and placed on ice (Teunissen *et al.*, 2009). Blood and urine specimens were collected immediately afterwards. All samples were prepared for storage according to a predetermined protocol. CSF was centrifuged at 500g for 10 mins at room temperature and the supernatant aliquoted into two cryovial tubes. Blood samples were centrifuged at 1000g for 10mins at room temperature. The serum was removed and divided between two cryovials for storage. Urine was not centrifuged, and 1ml of urine was pipetted into 2 cryovials for storage. All cryovials were stored at -80°C prior to transfer on dry ice to the Institute of Genetic and Molecular Medicine, University of Edinburgh for analysis.

2.3 Healthy controls

Sixty-five healthy controls for blood bio-banking were recruited from the University of Edinburgh. Controls were age-band matched and sex matched to the Future MS cohort, and specimens were handled according to the same SOPs as the Future MS blood samples.

2.4 Simoa protocols

Quanterix Simoa Neuro 4-plex A kit was used to measure neurofilament light chain (NfL), glial fibrillary acidic protein (GFAP), Tau and UCH-L1 enzyme levels in CSF and blood samples using a Quanterix SR-X benchtop instrument. Lot numbers 501888 and 502148.

2.4.1 Simoa laboratory protocol for CSF and blood specimens

Samples, reagents, calibrators and controls were equilibrated to room temperature. CSF and serum/plasma samples were centrifuged at 1100 rpm for 10 minutes at room temperature. 152µL of calibrators A-H were added in triplicate, and 38 uL of low/high controls in quadruplicate, as per the manufacturer's instructions. CSF samples (3.8µL) and serum/plasma samples (38µL) were pipetted in in duplicate. Control and serum/plasma samples were diluted with 114µL of supplied sample diluent (a dilution factor of x4). CSF was diluted with 148.2µL of supplied sample diluent, (a dilution factor of x40). 20µL of the supplied beads were added to every well, followed by 20µL of detector reagent. The plate was covered with a black lid and placed on an automated orbital plate shaker for a 30

minute incubation at 35°C. During the incubation time the SR-X plate set-up was performed (as per the ELISA plate layout depicted in *figure 4-6*), and the calibrator concentrations for each biomarker added using the lot-specific certificate of analysis. After the 30 minute incubation, the 96-well plate was transferred to an automated plate washer where 100µL of N4PA reagent was added, followed by a two-step was cycle. The plate was allowed to dry for 10 minutes and then transferred to the Quanterix SR-X benchtop instrument.

2.4.2 Data extraction and review

Data were provided by Quanterix SR-X benchtop instrument in the form of a run report. The run report included the following data: standard curve fit algorithm, fit equation and R^2 value, as well as replicate AEB values, AEB mean and standard deviations (SD) and AEB coefficient of variation (CV); fitted concentrations, fitted concentration mean and SD, and fitted concentration CV; for all four biomarkers, in every well.

Data were reviewed to determine the reliability of results. All biomarker results were above the lower limit of quantification. The standard curves were checked visually for goodness of fit and the coefficient of variation was calculated for each duplicate. This was a check of both laboratory technique and the interpolation of biomarker concentrations from the standard curves. High and low controls, supplied by the manufacturer and ran in quadruplicate, were within the expected ranges reported in the lot-specific certificate of analysis*, see figure 4-6*. CSF and serum duplicates with a coefficient of variation (CV%) of greater than 30% were excluded. Ideally, we would have repeated samples with a CV% greater than 20%, however a significant number of our samples had insufficient volume of either CSF or serum to allow this. We therefore compromised by excluding samples with a clearly unacceptable CV%, (>30%) but chose to include samples with a CV% up to 30% to include as many samples as possible. Sample volume constraints were less stringent with baseline blood samples, and therefore any plasma duplicates with a CV% greater than 20% in NfL, GFAP or Tau was repeated.

For two runs of plasma samples (dated 11/12/2019 and 19/12/19) the SR-X was unable to produce standard curves due to a technical error. However, as the raw AEB values had been calculated for the calibrators and samples we were able to derive standard curves and interpolate results. Calculations were performed by two researchers (SJM and EC) independently, and results compared.

2.5 Imaging protocols

All MRI data were processed, reviewed and analysed by the University of Edinburgh research imaging team (Professor AdamWaldman (AW), Beth York (BY), Rozanna Meijboom (RM), Dr. Daisy Mollison (DM), Nicole White (NW) and Agniete Kampaite (AK)). AW and BY specifically were involved in the processing, analysis and interpretation of MRI data from the extended imaging substudy.

2.5.1 Future MS

Future MS participants were assessed by MR brain imaging within four weeks of clinical assessment visits, and in the vast majority within the same day. MR imaging was performed in the dedicated research imaging facilities associated with local CRFs.

MRI sequences were obtained using Siemens magnetom prisma 3.0T systems and included: image localiser, volumetric (1mm^3 isotropic) T1 / MP-RAGE, T2, and Fluid attenuated inversion recovery (FLAIR). MR imaging took approximately 30 minutes per subject.

MRI outcomes at baseline included:

1. White matter lesion volumes, corrected for intracranial volume (processed by RM and AK)

MRI outcomes at 12 months included:

2. White matter lesion volumes, corrected for intracranial volume (processed by RM and AK)
3. Visual reads for ‘new/ enlarging T2 lesions - yes/no’ (performed by DM and AW).

White matter lesion volumes were corrected for intracranial volume in order to minimise potential confounders such as overall brain volume.

2.5.2 Extended imaging substudy

Participants in the extended imaging substudy had routine Future MS MR brain imaging performed (T1, T2 and FLAIR sequences) alongside advanced imaging sequences (MTsat and DWI). *The methods and analysis below were provided by AW and BY.*

Calculation of the myelin volume fraction

The MVF was calculated from MTsat, assuming a linear relationship (Campbell *et al.*, 2018), as:

$$MVF = \delta_{app} * k$$

where k is a constant calibrated from healthy control data (see supplementary methods).

Calculation of the axonal volume fraction

The AVF was calculated from NODDI data and scaled according to the MVF as:

$$AVF = (1 - MVF)(1 - v_{iso})v_{ic}$$

where v_{iso} and v_{ic} are the isotropic and restricted signal fractions derived from the NODDI model.

Calculation of the aggregate g-ratio

The aggregate g-ratio was derived using the following equation (Mohammadi *et al.*, 2015):

$$g = \sqrt{\left(1 + \frac{MVF}{AVF}\right)^{-1}}$$

2.6 Statistics and Graphics

With the exception of meta-analysis data, all data were analysed using GraphPad prism version 8.4.3. The normality of all data were assessed using a Shapiro-Wilk test. Additional statistical tests relevant to data are described in the figure legends. Statistical non-significance was defined as $p>0.05$. Meta-analysis data were analysed using R software with assistance from Dr Matthew Jamieson, University of Glasgow. All figures, unless otherwise stated, were designed by the author of this thesis using a combination of BioRender (www.biorender.com) and GraphPad prism version 8.4.3.

Chapter Three

Future MS Study
- Baseline characteristics of a newly diagnosed MS cohort

3 Future MS Study: Baseline characteristics of a newly diagnosed MS cohort

3.1 Introduction

There are currently fourteen treatments disease modifying treatments licensed for the treatment of MS in the UK. These treatments differ both in their efficacy in controlling the disease process and their safety profile. For this reason, it is increasing important to predict which individuals are at greatest risk of developing disability, in order to counsel patients and guide immunomodulatory treatment. However, multiple sclerosis is an unpredictable disease in clinical presentation and disease course. Determining the prognosis for an individual *early* in the disease course of relapsing remitting MS is difficult.

Current prognostic indicators are mainly coarse measures of disease activity that has already occurred. More sensitive clinical, imaging and biomarker tools, applicable at an earlier stage in the disease, are needed to prognosticate *on an individual level* to guide treatment decisions and ultimately change long-term disability rates in MS.

Current clinical practice is to consider MRI disease burden at baseline, demographic data and the number of clinical relapses as predictors of risk of accruing disability (Scolding *et al.*, 2015). A higher burden of infratentorial and spinal lesions on baseline MR imaging is associated with poorer long-term prognosis (Brownlee *et al.*, 2019). The evidence on sex as a prognostic indicator is conflicting (Degenhardt *et al.*, 2009), with studies both supporting (Confavreux, Vukusic, et al. 2003) and opposing (Kalincik *et al.*, 2013) the view that disability accrues more quickly in males than females, and may be confounded by a higher likelihood of progressive disease in men. An older age at onset of RRMS is associated with a shorter time spent in the relapsing-remitting phase, with quicker advancement to disability milestones (Confavreux and Vukusic, 2006; Degenhardt *et al.*, 2009). In this manner, older age could be construed as a poor prognostic factor. Although the time taken to reach disability milestones is longer for younger patients, the absolute age at which they are reached is similar irrespective of the age of disease onset (Confavreux and Vukusic, 2006). Age at diagnosis is therefore probably not useful in identifying an individual's long-term risk of poor prognosis.

A greater frequency of clinical relapses in years one and two is associated with a shorter time to reach certain EDSS milestones (Scalfari *et al.*, 2010). This has prognostic relevance but is dependent on the identification of relapses (by both the patient and

clinician) and ability to define a 'start date' of the disease. Consequently, it is usually the occurrence of relapses *after* year two which drive treatment decisions (but which have no clear association with prognosis (Scalfari *et al.*, 2010)). Clinical presentation of relapses may also carry prognostic significance. Although studies are conflicting as to whether a polysymptomatic onset is a poor prognostic factor (Degenhardt *et al.*, 2009) or not (Scalfari *et al.*, 2010), complete recovery from the initial relapse is associated with a longer interval to fixed disability endpoints (Confavreux, Vukusic, et al. 2003).

The identification of better prognostic tools requires large, homogenous cohorts of patients, recruited from disease onset and followed-up through life. However, this is an ideal - it is often not possible to identify patients at *disease onset* (as opposed to diagnosis), and it is difficult to categorise a *homogenous* cohort when the disease remains unpredictable.

Scotland has a high incidence of MS within a stable population of 5.4 million people. The Scottish national healthcare system (NHS) is linked across the fourteen health-boards. Every resident has a unique medical reference number (CHI) which is used at all interactions with health care services. This allows every attendance at primary or secondary care services to be tracked and linked with test results, imaging results, and prescriptions from across Scotland (ISD Data dictionary, CHI number).

Scotland is the only UK country to have a national register of all MS diagnoses, established by the Scottish Government in January 2010 as a public health tool. Because the NHS Scottish MS register is well implemented across all health-boards, standardised, high-quality baseline demographic data are available on all MS diagnoses throughout Scotland. These data are freely available as a public health resource against which research cohorts can be audited to ascertain whether they are representative of the Scottish MS population.

For these reasons Scotland offers many practical advantages for the long-term study of patients with MS. Detailed phenotypic studies of a large cohort, recruited at the point of diagnosis and then longitudinally followed-up, enables retrospective identification of groups of clinically, pathologically, or prognostically homogenous individuals. A longitudinal cohort study therefore offers a flexible way in which to study multiple different aspects of the disease, without the need to recruit different participants to separate studies.

Future MS is a Scotland-wide, prospective, observational cohort study with initial data collection points at baseline and one year. Future MS differs from other large MS cohorts in its specific recruitment of newly diagnosed and treatment naïve RRMS participants. This study provides the opportunity to identify potential determinants of prognosis at the earliest possible stage in the disease course, with minimal confounders, and then longitudinally follow the cohort in a 'real world' clinical setting.

Aim of Chapter Three

The aim of this chapter is to establish a cohort of RRMS patients at the point of diagnosis (the FUTURE MS study) and to explore the baseline clinical associations of the cohort.

3.2 Methods and substudy methods

The methodology of the Future MS study is detailed in chapter 2 (section 2.2) and the relevant details are summarised here.

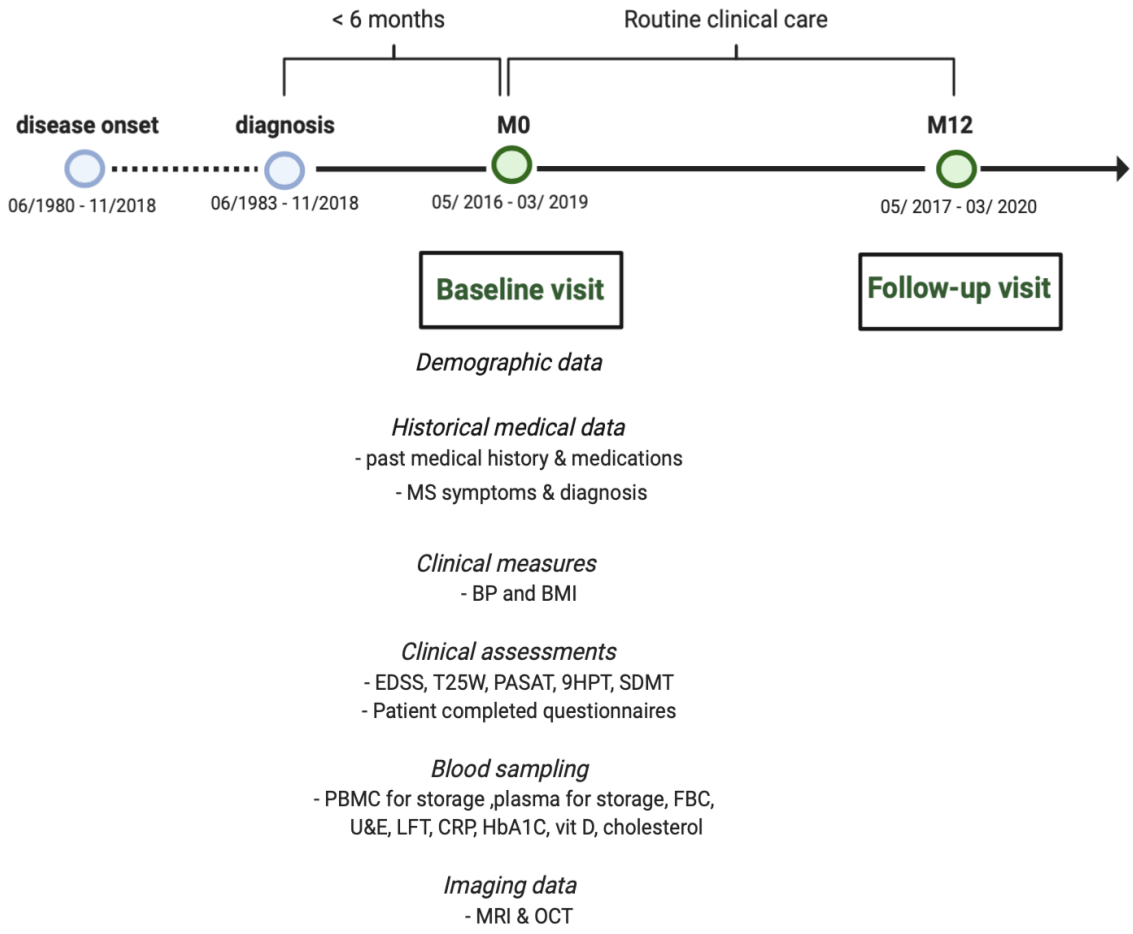


Figure 3-1 Overview of Future MS visits and the datapoints collected at each visit

The study was conducted throughout Scotland between February 2016 and April 2020. Participants were prospectively recruited from fourteen health boards and study visits were conducted across five main sites: Aberdeen Royal Infirmary; Ninewells Hospital, Dundee; The Anne Rowling Clinic, Edinburgh Royal Infirmary; Queen Elizabeth University Hospital, Glasgow; and Raigmore Hospital, Inverness. Participants were not restricted to only attend their local research site if, for example, an appointment date at different site was preferable.

Eligibility criteria included adults that were over 18 years old with a new diagnosis of relapsing remitting multiple sclerosis in accordance with the 2010 or 2017 McDonald criteria. Participants had to be reviewed for baseline assessment within six months of

diagnosis and before disease modifying treatment (DMT) was started. Follow-up review occurred 12 months later.

The recruitment target was 440 participants. Formal power calculations based on the methods of Dobbin & Simon (Biostatistics 2007;8:101-117) were performed by Dr Pete Connick to identify the sample size required for investigation of genetic variants (single nucleotide polymorphisms) associated with disease course within the study population.

Eligible patients were identified by their clinical care team and invited to participate. Research visits were appointed and conducted by Rowling clinical research fellows and Future MS research nurses. Data points collected are detailed below - see table. All participants provided written consent and the Future MS research visit (including MRI) lasted approximately 2 hours.

Participants in the main study were eligible to participate in optional sub-studies (overview in chapter 2; detailed specifically in chapters 4 and 6 (storage of CSF/blood samples); and chapter 5 (extended imaging substudy)), *see figure 3-2*. Participation in each substudy required separate, specific written consent.

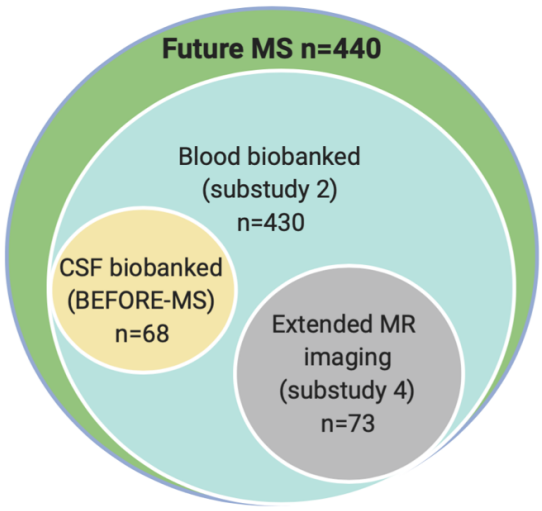


Figure 3-2 Substudies in Future MS

430 of 440 participants had plasma samples biobanked (substudy 2) at baseline visit. Paired CSF and serum biobanking from diagnostic work-up (BEFORE-MS) was only available at the Glasgow site. The extended imaging substudy (SUBSTUDY 4) was only available at the Edinburgh site.

3.3 Results

3.3.1 Overview of Future MS recruitment and demographics

Four hundred and forty participants (100% of the recruitment target) were recruited between May 2016 and March 2019. Four hundred and seven were followed-up at 12-month review, and 33 (7.5%) either had follow-up postponed as a result of the COVID-19 outbreak or were lost to follow-up.

Table 3-1 Recruitment of Future MS

Site	***Recruitment dates***	***Recruited and reviewed by***	***Number***
Glasgow	**Nov 2016 – Feb 2019**	**SJM**	**166**
Edinburgh	May 2016 - March 2019	ARC research team	184
Dundee	Dec 2016 – Dec 2018	**SJM,** clinical care teams and ARC research team	46
Aberdeen	March 2017- Jan 2019	**SJM,** clinical care teams and ARC research team	36
Inverness	Dec 2017- Oct 2018	Clinical care teams	8

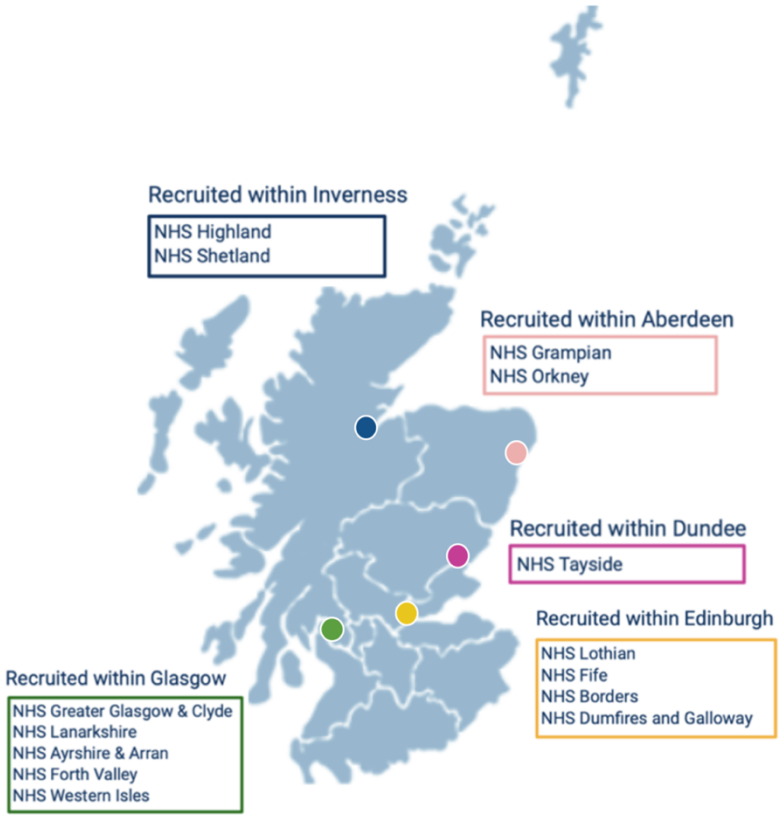


Figure 3-3 Map showing the recruitment strategy for covering the 14 Scottish NHS health boards

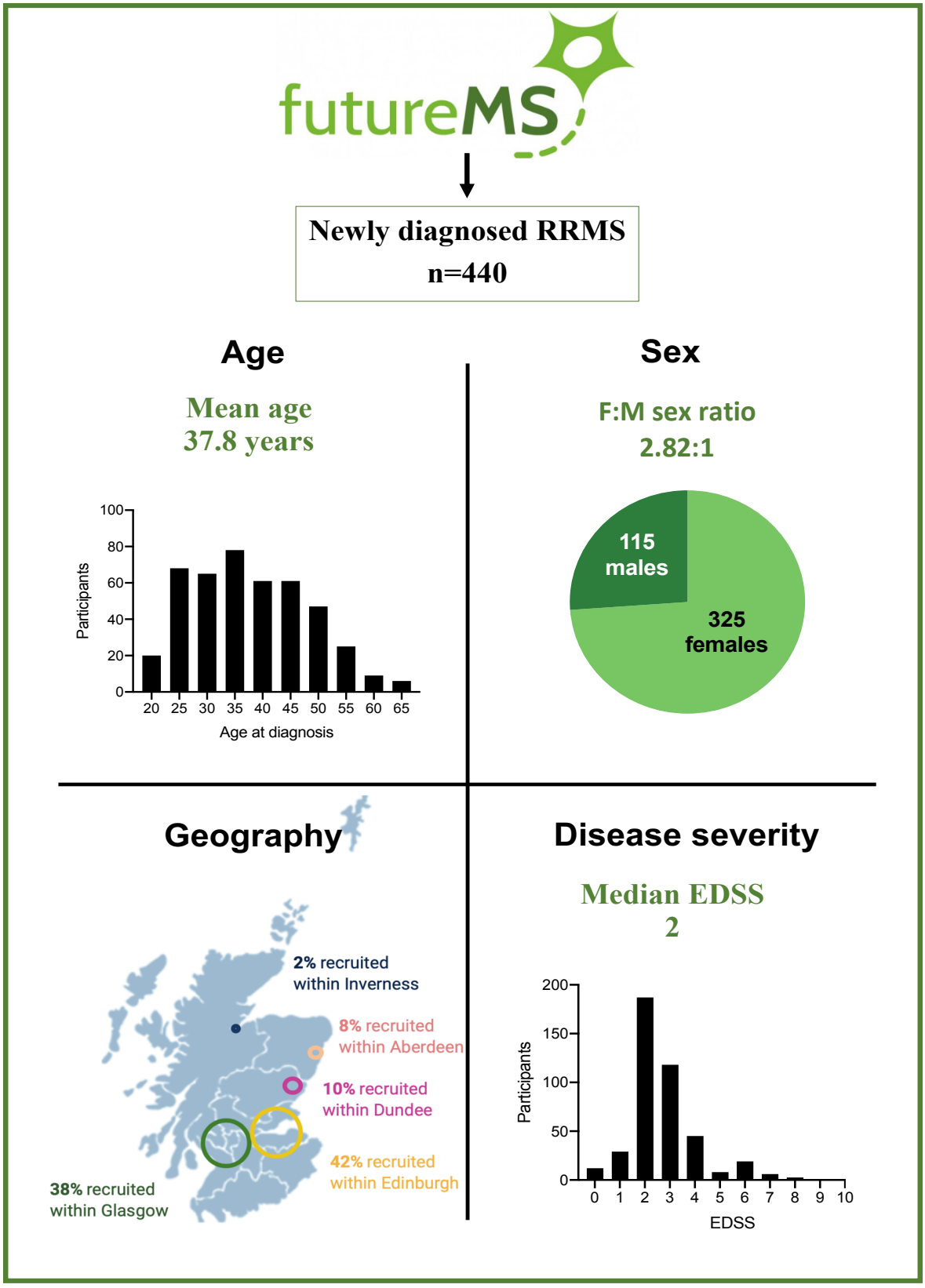


Figure 3-4 Summary figure of the baseline demographics and disease severity of the Future MS cohort

3.3.2 Age and sex distribution of the cohort

The Future MS cohort was primarily Caucasian, reflecting the Scottish population (92% Caucasian according to consensus data)(National Records of Scotland 2018). Basic demographic data for the cohort were compared with NHS incidence data. The Future MS cohort had a mean age at diagnosis of 37.8 years, with a range from 18.7 to 67 years old, and was made up of 325 females and 115 males. The proportion of females to males was therefore slightly higher than Scottish MS register data (Scottish MS register, 2019).

Table 3-2 Baseline demographics of the FMS cohort (newly diagnosed RRMS), with reference to "real-life" Scottish MS register data (all forms of MS)

	Future MS cohort	***Scottish MS resgiter data***
MS subtype	RRMS	All MS
Sex ratio, F:M	1: 2.82	1: 2.3
Age at diagnosis, mean (SD)	37.8 (10.51)	40

Age, baseline EDSS and time from the onset of first symptom attributed to MS to diagnosis were similar for males and females.

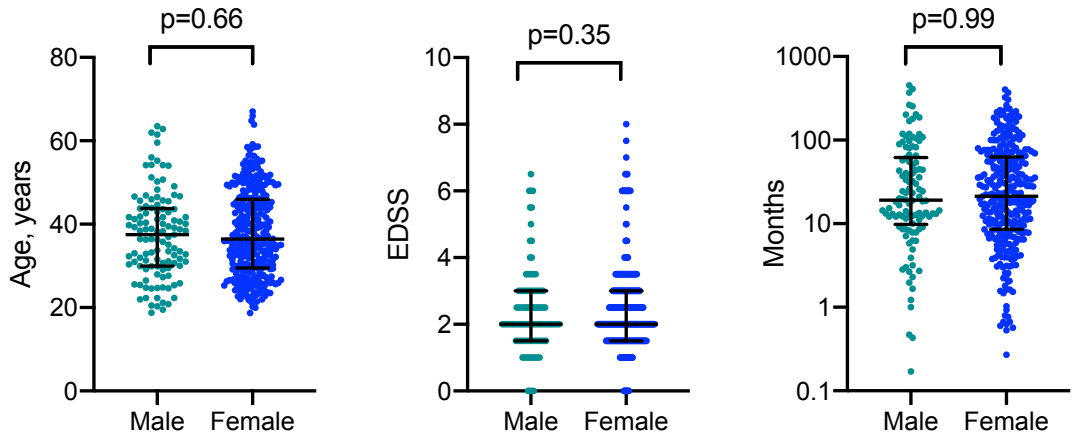


Figure 3-5 Comparison of baseline demographics between male and female participants in Future MS

No differences were found between male and female participants age, baseline EDSS score or duration from onset of first symptom attributable to MS to diagnosis.

3.3.3 Geographical distribution of cohort recruitment

Incidence rates of MS vary throughout Scotland (Kearns *et al.*, 2019). To assess if the Future MS cohort captured the geographical spread of patients, we compared the proportion of participants recruited at each site with Scottish MS register data for each site (Scottish MS register, 2019).

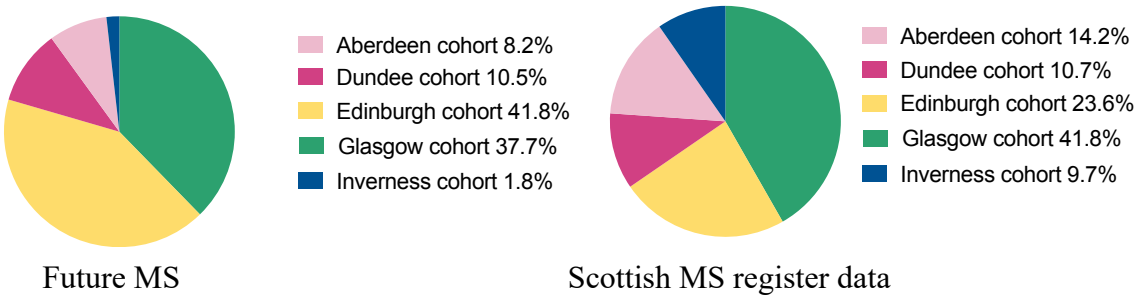


Figure 3-6 Comparison of geographical distribution of Future MS with Scottish MS register data

Pie chart showing a breakdown of the Future MS cohort according to site of recruitment, and comparison with SMSR data from 2019 according to side of diagnosis.

The majority of participants were recruited from sites with the largest population densities. Edinburgh was over-represented (41.8% of the study cohort but 23.6% of Scottish MS register data) and the northern centres underrepresented.

3.3.4 Baseline disability of the cohort

The Future MS cohort had a median baseline EDSS score of 2 (IQR 1.5-3). This is comparable to Scottish MS register data (Scottish MS register, 2019) and also to other large cohorts of MS patients early in the disease course.

3.3.5 Demographic and clinical data are similar between recruitment sites

To ascertain if the populations were similar across recruitment sites we compared basic demographic and clinical data, *see table 3-3*. No significant differences were found for age, sex, EDSS, employment status or smoking status between sites using univariate analyses.

Table 3-3 Demographic variations between Future MS sites

	Aberdeen	*Dundee*	*Glasgow*	*Edinburgh*	*Inverness*	
N	36	46	166	184	8	
Sex ratio, F:M	2.6	3.18	2.32	3.18	all female	**p=0.3**
Age at diagnosis,	36.2 (30.2 - 47.4)	39.3 (32 – 47.2)	35.5 (29.3 – 44.5)	36.7 (29.7 – 45)	31.1 (26.7 – 46.6)	**p=0.47**
EDSS at baseline	2 (1.5 - 3)	2.5 (1.6 - 3)	2 (2 - 3)	2 (1.5 - 3)	3 (2.5 - 3.3)	**p=0.37**
Percentage of cohort in employment	82.9	84.8	84.6	80.9	87.5	**p=0.9**
Percentage of cohort ever smoked	42.9	56.5	49.4	49.7	50	**p=0.68**

For sex ratio, employment and smoking history Chi-squared tests were used to assess the significance of differences between the groups. For age and EDSS, non-parametric ANOVA with Kruskal-Wallis test for multiple comparisons was used to assess the significance of differences between the groups.

3.3.6 Time from first symptom to diagnosis of MS averaged 21 months, but varied across Scotland

In order to participate in the study, patients had to be reviewed within six months of diagnosis. Future MS was therefore designed as an *early disease* cohort. However, disease onset may occur many years before formal diagnosis, and 'newly diagnosed' does not necessarily equate to early pathology (Wijnands *et al.*, 2017).

There is an increasing trend towards earlier diagnosis of multiple sclerosis, aided by the latest update to the McDonald diagnostic criteria, which allow cases previously classified as "CIS" to be called multiple sclerosis if OCBs are present in the CSF (Thompson *et al.*, 2018). Earlier diagnosis allows for earlier treatment, which may improve outcome (Coles *et al.*, 2008). Time to diagnosis is therefore of clinical importance.

To determine the likely duration of (clinical) disease we calculated the time interval from the onset of the first symptom attributable to MS (reported by the patient) to the date of formal diagnosis (documented by the physician). Data were available for a total of 424 participants.

The median time from onset of the first symptom attributable to MS to a formal diagnosis of RRMS was 21 months (IQR 8.6-62). Approximately half the participants (227, 53.5%) were diagnosed within two years of the onset of symptoms, however the range was large (from 5 days to 37.8 years). Forty-seven participants (11.1%) were diagnosed more than 10 years after symptom onset and 10 (2.4%) after an interval of greater than 20 years, *figure3-7(A)*.

We reviewed whether this time interval differed between the five research sites. We found that participants in Glasgow had a significantly shorter time from onset of first symptom to diagnosis than participants in Edinburgh, p=0.004, and Dundee/Aberdeen/Inverness combined, p=0.0008 (Non-parametric ANOVA with Kruskal-Wallis test for multiple comparisons, *figure 3-7(B) and (C).*

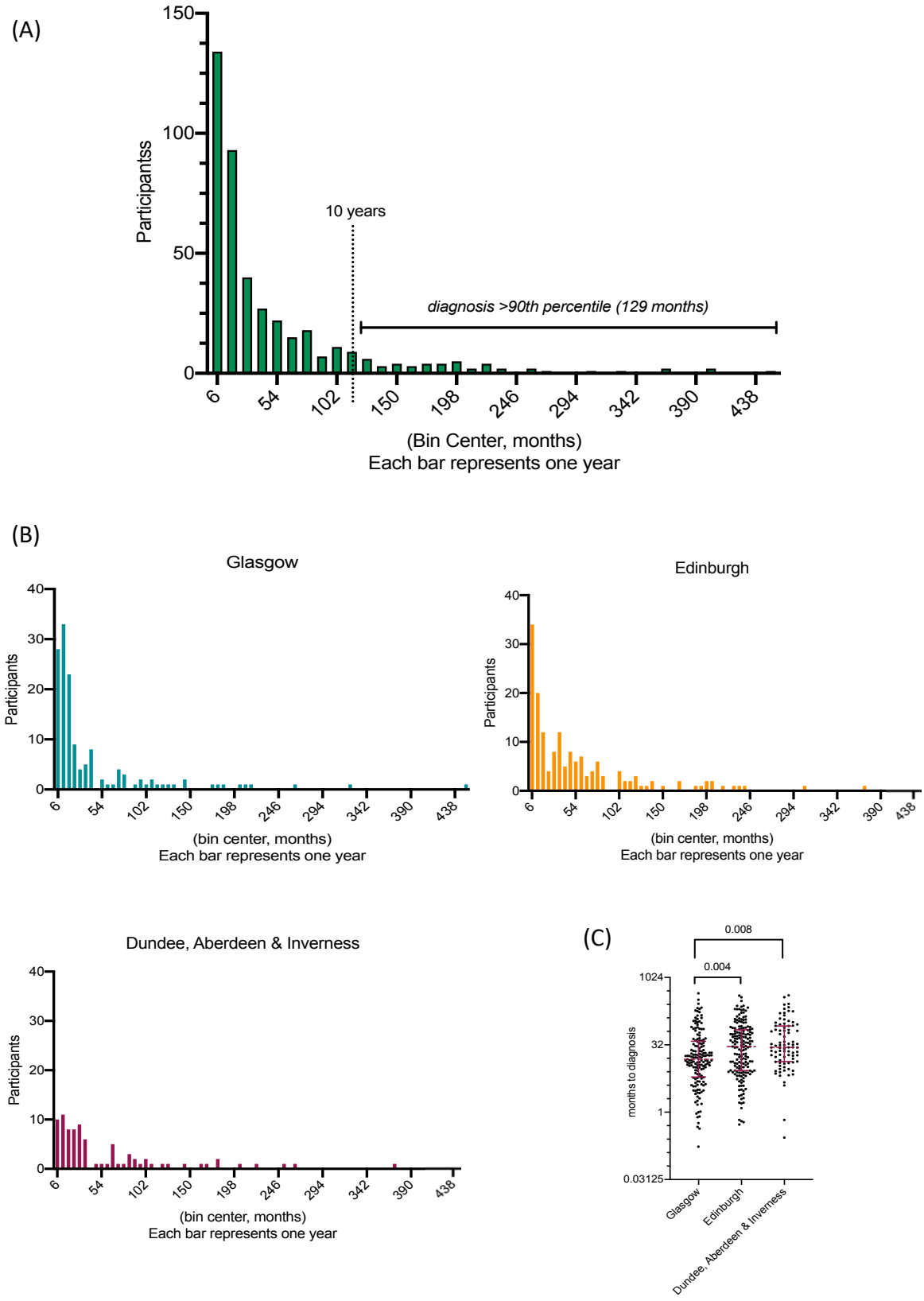


Figure 3-7 Histograms of time from symptom onset to diagnosis

(A) Histogram of time from first symptom onset to diagnosis for the Future MS cohort as a whole, (B) Histograms for Glasgow, Edinburgh and the northern sites individually, (C) statistical comparison using ANOVA with Kruskal-Wallis test for multiple comparisons showed that Glasgow had a significantly shorter time period between onset of symptoms and diagnosis than either Edinburgh (p=0.004) or the northern centres (p=0.008).

3.4 Discussion

Our aim was to develop a large cohort of newly diagnosed and treatment naïve RRMS patients from across Scotland. We met our recruitment target of 440 participants.

Baseline demographic and clinical data were compared with data from the Scottish MS registry (SMSR) and the cohort appeared generally representative of the newly diagnosed MS population in Scotland. The Future MS cohort was slightly younger (37.8 years versus 40 years) and had a higher proportion of females than SMSR data (male to female sex ratio for Future MS cohort 1: 2.82 versus 1: 2.3), (Scottish MS register, 2019). These observations are not unexpected as incidence data collected by the Scottish MS register encompasses *all diagnoses* of MS, including progressive forms. Individuals diagnosed with (primary or secondary) progressive MS are typically older than those diagnosed with RRMS (Scalfari *et al.*, 2011), and the sex ratio in primary progressive MS is more equal than for RRMS (Scalfari *et al.*, 2011). The inclusion of individuals with progressive MS in the Scottish MS registry is therefore likely to increase the age and reduce the female to male sex ratio of SMSR data.

There are several potential reasons for the differences seen in the geographical representation of the Future MS cohort versus real life incidence data. Edinburgh, as the primary research site, was open to recruitment for the longest period of time. The larger sites (Edinburgh and Glasgow) also had a full-time Future MS research presence and were therefore less dependent on clinical staff to identify and contact potential participants. Northern centres (NHS Orkney, NHS Shetland and NHS Highland) were unrepresented most likely as a result of the logistical difficulties involved for remote patients to attend mainland sites for research visits. These sites were also more restricted in terms of their availability for research appointments, and as a result a small number of participants chose to attend the Edinburgh site instead. As the cohort has been divided according to the site at which participants were reviewed for the study (as opposed their local NHS board) a very small number of participants will be incorrectly attributed to the wrong site.

The point of *disease (pathology) onset* in multiple sclerosis can be clinically silent, with a diagnosis not made until many years later. Recruiting a cohort of patients that are truly representative of early disease is therefore difficult, but the point of diagnosis represents an important and standardised landmark of early disease where dissemination in time and space has been demonstrated.

Whilst there are numerous cohort studies of multiple sclerosis, many include a mix of relapsing-remitting and progressive MS patients, recruited years after diagnosis and treated to differing extents. We recruited participants within an average of 2 months (IQR 1.2 to 3.1 months) of diagnosis **and before initiation of disease modifying treatment**.

Recruitment of individuals prior to starting treatment was important as certain treatments, such as Natalizumab, significantly affect neurofilament levels (Gunnarsson *et al.*, 2011). It was therefore preferable to obtain treatment naïve specimens in order to interrogate the potential use of neurofilament as a prognostic tool, without the confounding effects of treatments. Furthermore, by having a treatment naïve baseline neurofilament level against which to compare post-treatment levels we could potentially investigate the sensitivity of neurofilament in the monitoring of treatment efficacy.

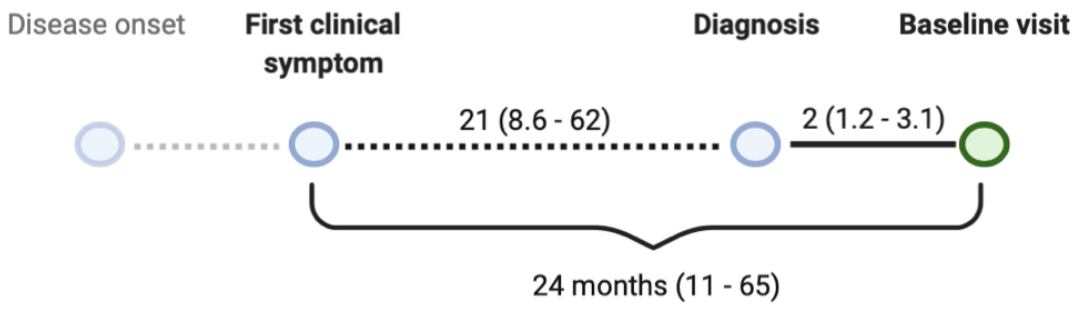


Figure 3-8 Median time from onset of clinical symptoms to diagnosis, and from diagnosis to baseline visit

The median time from first symptom to diagnosis was 21 months (IQR 8.6-62), and the median total time from the onset of first symptoms attributable to MS to baseline review in the study was 24 months (IQR 11 to 65 months).

To understand how well the Future MS study encapsulates an *early disease* cohort in the context of previously published work we compared baseline demographics with those from three other large observational cohort studies - CLIMB (Gauthier *et al.*, 2006), EPIC (Cree *et al.*, 2016) and SWIMS (Zajicek *et al.*, 2010), as well as the Danish MS registry (Brønnum-Hansen, Koch-Henriksen, *et al*. 2011), *table 3-4.* All patients in Denmark with a suspected diagnosis of MS have their hospital discharge letter mailed to the Danish MS registry office. Patients are therefore enrolled *at diagnosis.* We used the Danish MS registry as a benchmark for a newly diagnosed cohort.

Table 3-4 Comparisons between different cohorts/registries inclusion criteria, baseline demographic and clinical data

Cohort	*N (at year)*	*MS subtypes included*	*Treatment naïve at inclusion?*	*Age at inclusion Mean (SD)*	*Percentage female*	*EDSS at inclusion Med (IQR)*	*Time from first symptom onset to inclusion, years Med (IQR)/Mean (SD)*
Future MS, Scotland. Multi-centre, observational, longitudinal cohort study. 2016 onw	**440 (2020)**	**RRMS Enrolled within 6 months of diagnosis**	**Yes**	**37.8 (10.5)**	**73.9**	**2 (1-3)**	**2 (0.46-5.4)**
CLIMB, Boston, USA. Single centre, observational, longitudinal cohort study. 2000 onw	>2000 (2017)	All MS/CIS Enrolled within 3 years of diagnosis	No	39 (10.5)	76.1	1.5 (1-2)	7.6 (8.7) relapsing onset only N=1541
EPIC, San Francisco, USA. Single centre, observational, longitudinal cohort study. 2004 onw	>500	All MS/CIS (but ambulatory patients and those with 'recent onset' were preferentially recruited)	No	42.7 (9.9)	67.5	1.5 (1-3)	7 (2-13.5)
SWIMS, England. Patient-centred, prospective, observational longitudinal cohort study. 2004-2017	1600 (2017)	All MS/CIS No limit on time since diagnosis	No	51.6 (11.5)	75	Mean 4.3 (SD 2.3)	13.3 (6.8 - 24.5)
Danish MS registry, Copenhagen Epidemiological register with 91% completion rate. 1956 onw	>13,000 (2013)	All MS/CIS Individuals are enrolled in register at point of discharge from hospital - *i.e. at diagnosis*	No	37.6	68.9	2 (1-3)	2

Average age and EDSS at study inclusion were similar between the Danish MS registry, Future MS and two cohorts of established disease (EPIC and CLIMB), *figure 3-9(A)*.

Time from first symptom to study inclusion varied more notably. Participants in the EPIC and CLIMB cohorts had an average duration of symptoms of approximately 7-8 years prior to inclusion. Participants in the SWIMS study reported an average of thirteen years from symptom onset to study enrolment (Zajicek *et al.*, 2010). The median time from first symptom to study inclusion was the approximately 2 years for both the Future MS cohort and the Danish MS registry (Steenhof *et al.*, 2019), *figure 3-9(B)*.

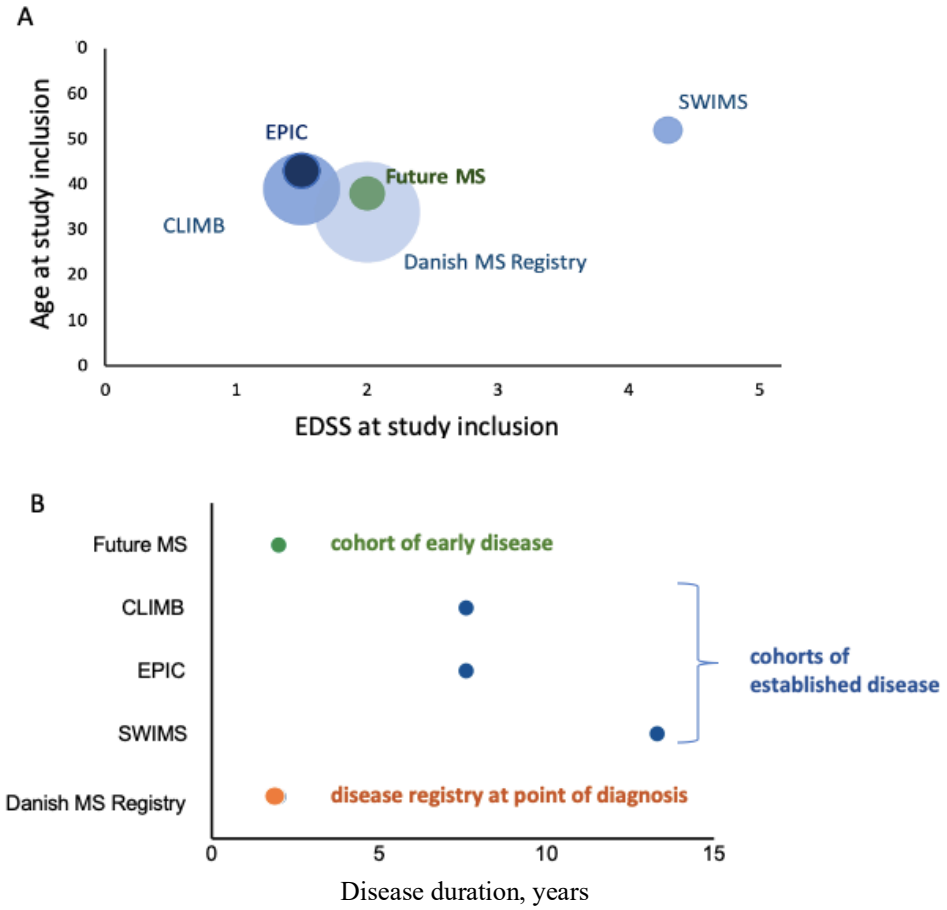


Figure 3-9 Comparison of baseline demographic and clinical data for several large MS cohorts (A) Bubble chart showing age and EDSS at study inclusion (where size of bubble represents size of cohort) are similar between several cohorts of established disease (EPIC & CLIMB), Future MS, and the Danish MS registry; (B) Disease duration varies between cohorts of established disease (EPIC, CLIMB and SWIMS), but it similar for Future MS and the Danish MS registry.

The established disease cohorts (EPIC and CLIMB) appeared to have similar average baseline demographics as participants at the point of diagnosis (the Danish MS registry and Future MS), yet with a markedly different 'disease duration' (time since first symptom). This is likely a reflection of the MS populations recruited.

With the exception of Future MS, all the studies included varying proportions of patients with clinically isolated syndrome and progressive MS. Approximately 16% of the EPIC and CLIMB cohorts were CIS patients (Bove *et al.*, 2018), compared with only 3.6% of the SWIMS cohort (Zajicek *et al.*, 2010). Conversely, approximately 40% of the SWIMS cohort reported a diagnosis of progressive MS (and 20% did not know)(Zajicek *et al.*, 2010) compared with 7% of CLIMB and 14% of EPIC (Bove *et al.*, 2018). Variations between the composition of cohorts will affect demographic and clinic results, with a larger proportion of progressive patients skewing the cohort towards an older age, longer disease duration and higher EDSS, and the inverse seen with CIS. This explains how cohorts of established disease appear to have similar baseline demographics to cohorts of early disease.

Our comparisons also show that participants in Future MS had a similar age, EDSS and duration of symptoms as those enrolled into the Danish MS registry, *figure 3-9*. This demonstrates that the baseline characteristics of the Future MS cohort are similar to established cohorts based around the study of early MS.

The proportion of patients receiving disease modifying treatment at inclusion varied from approximately 18% in the SWIMS study (Zajicek *et al.*, 2010) to approximately 80% in the CLIMB study (Raghavan *et al.*, 2015). Future MS is the only treatment naïve cohort. Whilst this reduces the confounding influence of treatments, by only including treatment naive individuals the Future MS study risked excluding patients with very active disease (who may be required to start treatment as an inpatient at time of diagnosis). We actively tried to minimise this by including participants who were current inpatients as a result of a severe first presentations of MS. Despite this, it is likely that some potential participants have been missed, for example as a result of being hospitalised in a district general hospital. However, the median EDSS of the Future MS cohort appears consistent with other studies and the range of EDSS scores within the Future MS cohort (0 - 7.5) demonstrates inclusion of individuals with significant disability.

In the context of previously published MS cohorts, Future MS therefore represents both *a homogenous* and *an early disease population* of RRMS patients.

Date of diagnosis does not equate to date of disease onset. However, a multiple sclerosis relapse can produce any one of a number of possible symptoms and identifying the earliest symptom truly attributable to MS is difficult. In order to gauge the potential duration of the disease it is common to ask patients if they have experienced similar symptoms. Whilst some may give a classical description of symptoms suggestive of optic neuritis, in other instances a recount of sensory disturbance affecting the hand for example, could be attributable to several potential causes. Whether a historical symptom is ultimately considered to be potentially relevant to a diagnosis of MS is therefore dependent on both the patient and the clinician. This is important to remember when interpreting results, such as when we reviewed how time from symptom onset to diagnosis differed between participants, and noticed variability between recruitment sites.

No differences in baseline demographic or clinical data between recruitment sites were apparent, therefore the significantly quicker diagnosis seen in the Glasgow cohort was unlikely to be a reflection of a younger MS population or the inclusion of individuals with particularly aggressive disease. We considered whether this could be the result of an ascertainment basis, as the Glasgow cohort were predominantly reviewed by one researcher (SJM). However, this seems unlikely given that the same researcher reviewed participants in other sites. A potential reason for the variation in time to diagnosis may be due to regional disparities in the use of cerebrospinal fluid in the diagnosis of RRMS, *figure 3-10.*

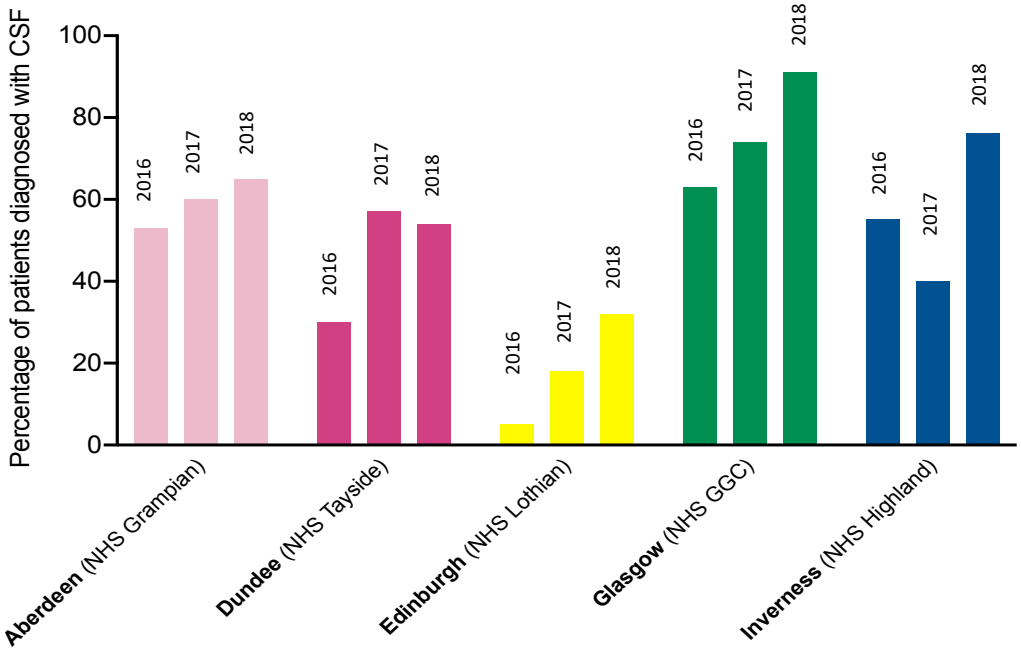


Figure 3-10 The percentage of patients across Scotland diagnosed with MS using CSF analysis Across all centres, there was a general trend towards increasing use of CSF in the diagnosis of multiple sclerosis, however variation was apparent, particularly between the two largest centres - Edinburgh and Glasgow.

The use of CSF in conjunction with the McDonald 2017 diagnostic criteria facilitates an earlier diagnosis of MS (Thompson *et al.*, 2018). NHS incidence data detailing the percentage of patients who had a lumbar puncture performed at diagnostic work-up in the years 2016, 2017 and 2018 was provided from the Scottish MS Register (Scottish MS Register, 2019).

In general, the use of CSF increased between 2016 and 2018, however, there was regional variability. Glasgow employed CSF analysis in the diagnostic work-up of MS significantly more frequently than other sites (Glasgow vs Edinburgh, p=0.035, ANOVA with Dunn's multiple comparisons). This may explain in part the shorter time between first attributable symptom onset and MS diagnosis to diagnosis seen in the Glasgow cohort. This is an area which requires further study, as the ability to reduce time to diagnosis has potential prognostic significance.

With increasing disease modifying treatment options available there is an increasing consensus towards earlier treatment (Cerqueira *et al.*, 2018). Delayed diagnosis risks delaying treatment options which could potentially alter the disease course and reduce the risk of disability.

Identification of the most appropriate treatment option for an individual requires weighing up the balance between potential risks of the medication against potential risk of disability if the patient remains untreated. Prognostic tools are therefore required in order to identify individuals most likely to develop permanent disability, in whom risker but higher efficacy treatments would be warranted. The Future MS cohort offers an opportunity to employ next-generation technologies to study potential prognostic markers of poor prognosis in an early disease cohort.

Summary of Chapter Three

- We successfully recruited a large, representative, cohort of treatment naïve individuals from across Scotland within an average of 2 months of diagnosis of RRMS.
- Analysis of baseline data identified regional similarities in demographics and clinical status, and regional variability in time to diagnosis.
- When compared with other large cohorts, the Future MS study has achieved recruitment of a particularly *homogenous*, *early disease* cohort of individuals.
- Longitudinal follow-up over many years will enable us to link baseline data with long-term outcomes.

Chapter Four

Neurofilament in CSF and blood in MS: Meta-analysis and validation

4 Neurofilament in CSF and blood in MS: Meta-analysis and validation

Part of this chapter has been published: Martin SJ *et al.* Cerebrospinal fluid neurofilament light chain in multiple sclerosis and its subtypes: a meta-analysis of case–control studies *Journal of Neurology, Neurosurgery & Psychiatry,* September 2019;90:1059-1067. This is an unofficial translation of a manuscript that has been accepted for publication by BMJ. Neither BMJ nor its licensors have endorsed this translation

4.1 Introduction

Axonal damage is thought to be an important substrate of disease progression in MS, and once axonal injury occurs beyond the threshold for compensation it is the best pathological predictor of permanent neurological deficit (Bjartmar *et al.*, 2000). Neurofilament appears to be promising biomarker of axonal damage.

MRI is currently the gold-standard biomarker for disease diagnosis, monitoring and prognostication in multiple sclerosis, and is non-invasive and readily available. However, MRI is insensitive to microscopic pathology, primarily capturing the macroscopic endpoint of white matter disease that has already occurred. A biomarker that reflects *current* fluctuations in disease pathology may enable more 'reactive' identification of disease activity.

Neurofilament is the major structural protein of neuronal axons in the central nervous systems (Petzold, 2005; Yuan *et al.*, 2012). Release of neurofilament subunits into extracellular fluid can occur consequent to any cause of axonal transection (Yuan *et al.*, 2012). Detection of elevated neurofilament levels is therefore non-specific and may require disease specific reference values (Bridel *et al.*, 2019). Neurofilament light chain is the most abundant subunit of neurofilament and is reliably quantified in CSF. Levels of NfL in the CSF have been shown to be higher in MS patients compared with controls, to rise in the context of relapse, and to fall with certain disease modifying treatments (Lycke *et al*. 1998; L. Novakova *et al.* 2016; Novakova, Axelsson, *et al*. 2017). This suggests that CSF NfL has the potential to be a clinically meaningful marker of axonal damage. However, CSF is invasive to obtain, making it unsuitable for repeat or longitudinal sampling and less likely to be available from both patients and controls.

Blood is a more acceptable biofluid to sample. The limiting factor in the detection of neuronal proteins in blood has always been the lower limit of quantification of the assays available. Conventional NfL ELISAs had a lower limit of detection of approximately 60ng/L (Norgren, Rosengren *et al*. 2003).

Next generation ELISA technology - single molecule array (Simoa), was developed in 2010 (Rissin *et al.*, 2010) and has reduced the lower limit of detection of proteins a thousand-fold, to sub-femtomolar concentrations ($<10^{-15}$)(Rissin *et al.*, 2010).

Simoa enables the quantification of NfL in blood in both MS patients and healthy controls (Rissin *et al.*, 2010; Kuhle *et al.*, 2016). The concentrations of neurofilament in blood seem to correlate with those seen in CSF (Kuhle *et al*. 2016; Novakova, Zetterberg, *et al*. 2017) and appear too to reflect MS disease activity. Blood NfL levels are greater in MS than controls, higher in relapse than remission and decrease with disease modifying treatment (Novakova, Zetterberg, *et al*. 2017). Higher blood NfL at baseline has been linked with more significant brain atrophy at ten years, suggesting that blood NfL levels may also have a prognostic relevance.

Current biomarker studies are often retrospective analyses of cohorts of MS patients with disease durations which range by decades. The proportion of the cohorts that are treated (and what disease modifying treatment they are receiving), vary with local practice. Finally, how researchers define 'relapse' (clinically versus radiologically), and thus 'remission', also differs between studies.

We sought to synthesise the published literature on CSF NFL. We performed a systematic review and meta-analysis in order to better understand the relevance of neurofilament measurement, and improve the power to detect the potential utility of NfL as a biomarker of axonal damage.

After reviewing the literature on CSF neurofilament in MS, we then demonstrated the use of next-generation ELISA technology. We used Simoa to measure neurofilament in paired CSF and blood samples from RRMS patients at the point of diagnosis in order to determine how well blood NfL reflects CSF NfL levels in *early RRMS*.

Aims of chapter Four

Aim 1 - To evaluate the role of CSF NfL in MS subtypes through systematic review and meta-analysis.

Aim 2 - To demonstrate that, within the Future-MS study, plasma NFL can be measured using single molecule ELISA.

4.2 Meta-analysis of the potential utility of CSF NfL in MS

No meta-analyses of neurofilament light chain in multiple sclerosis were published or registered with PROSPERO (*National Institute for Health Research PROSPERO*, 2017) in September 2017. We therefore designed a meta-analysis, which we registered with PROSPERO (*National Institute for Health Research PROSPERO*, 2017) (ID CRD42017078996) and conducted according to a predetermined protocol. The research question was defined as 'Can CSF NfL levels differentiate people with MS (pwMS) from controls (either healthy controls or controls with other diseases)?'. If so, 'Do CSF NfL levels differentiate between different MS disease stages or states?'.

Any original study quantifying NfL in CSF of people with MS was identified. No language or publication date restrictions were imposed. Patients of any age were included, with no restrictions on disease duration or subtype, time since relapse, disability, co-morbidities or treatment. Diagnosis had to be stated with reference to established diagnostic criteria. Where cohorts were not differentiated by MS subtype, they were named accordingly, for example 'clinically definite MS (CDMS). Clinical and radiologically isolated syndromes were excluded in order to reduce the heterogeneity of the overall cohort. Each MS cohort required a control comparator. Ideally, studies should reference guidelines on defining control groups, but this was not an inclusion criterion (Teunissen C *et al*., 2013). Studies could be retrospective, cross-sectional or prospective. CSF collection and bio-banking were required to meet criteria proposed by BioMS-EU (Teunissen C *et al*., 2009). If these criteria were not referenced, the paper was required to describe CSF sampling, pre-analytical handling and storage techniques applied to ensure the samples used were of sufficient quality. Studies also had to use a validated assay or describe the ELISA technique to satisfy inclusion. Assays with a coefficient of variation >25% were excluded, as were studies where NfL was detectable in less than 85% of either comparator group.

One author (SJM) searched electronic databases for published and unpublished 'grey' literature using the terms ["MS" OR "Multiple sclerosis"] AND ["NfL" OR "Neurofilament light"] AND ["CSF" OR "cerebrospinal fluid"] *(see appendix 1).* The search was performed on 8th September 2017. Detailed review of potentially eligible papers followed, as per Preferred Reporting Items for Systematic Reviews and Meta-Analyses, 2009, (PRISMA) guidelines, *figure 4-1*.

Sixty-eight duplicates were removed, and 100 records excluded on abstract alone *(see appendix 2*). Of the remaining 51 papers, 17 studies had no control group, three measured NfL in a categorical way, and one measured serum NfL levels. Seven studies were excluded on the basis that CSF NfL level was detectable in less than 85% of one comparator group. These were older studies that used a less sensitive assay (Malmeström C et al., 2003, Norgren N *et al*., 2004). Three papers were excluded as NfL levels for the MS cohort had previously been published. One paper used previously published data from a control cohort, but compared it with a new MS cohort, and was included (Novakova L *et al*., 2017). Twenty studies met our criteria for inclusion *(see appendix 2).*

Within the studies identified, data were reported either as mean and standard deviation (SD), or as median and range or interquartile range. To improve comparison between datasets we contacted authors and asked them to provide raw data.

We applied a weighted fixed effect model to estimate standardised mean differences in CSF NfL level between groups (with 95% CI, and corresponding p value). Heterogeneity between studies was documented as a Q test statistic and corresponding p value. Publication bias was assessed using funnel plots. Demographic differences between cohorts were tested for significance using two-way T tests and Z scores.

Required data were available in the original paper or provided by the authors in 14 of the 20 studies. This equated to a total of 805 MS patients (638 RRMS, 104 SPMS and 63 PPMS) and 435 controls (332 non-inflammatory neurological disease controls (NINDCS) and 103 healthy controls).

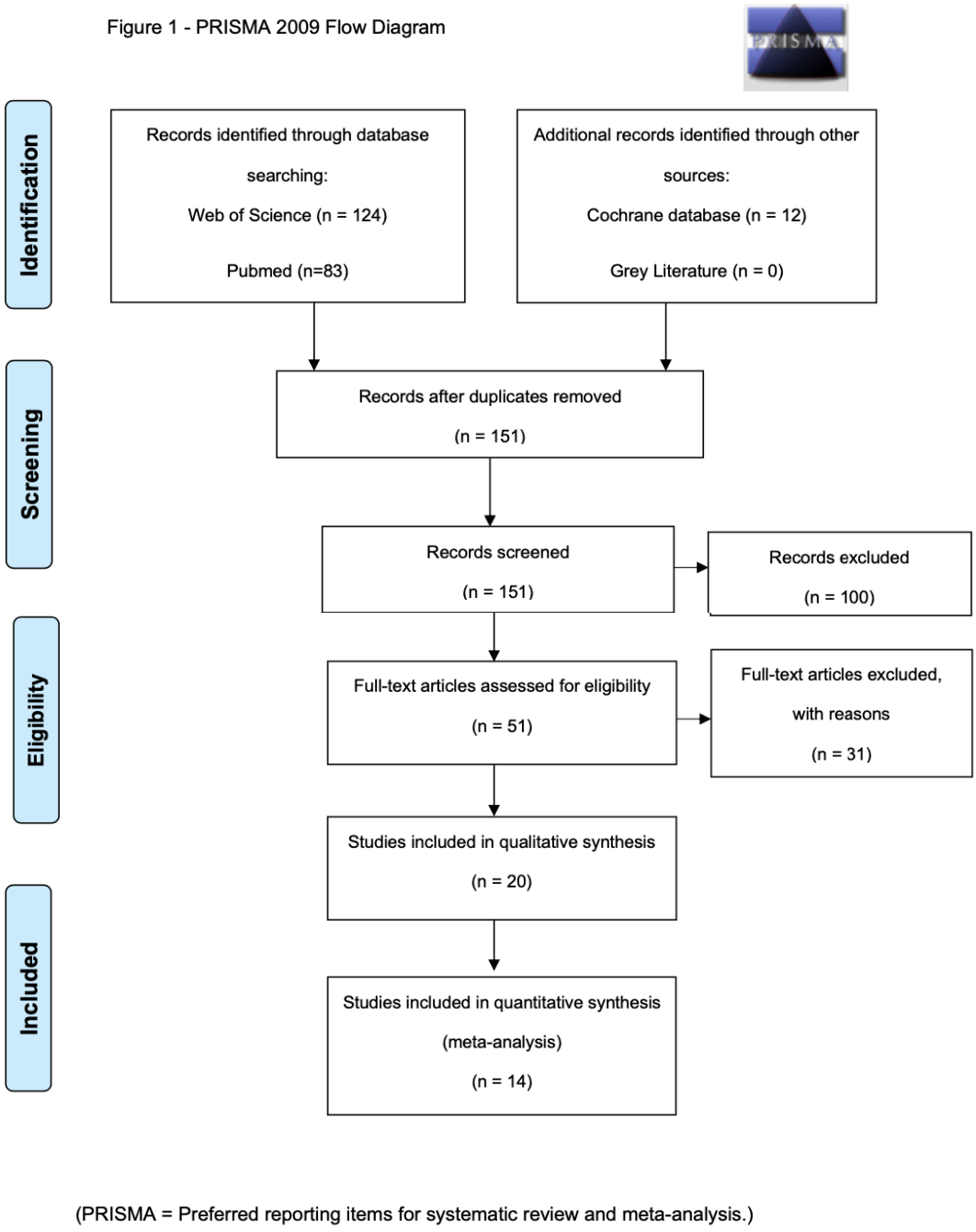


Figure 4-1 PRISMA 2009 (Preferred reporting in Systematic reviews and Meta-analyses)

Seven studies were retrospective, four prospective, and three were cross-sectional analyses. Seven studies referenced BioMS-EU guidelines on biobanking, and two referenced guidelines relating to definition of control populations (Teunissen *et al.*, 2013; Trentini *et al.*, 2014; Stilund *et al.*, 2015). All 14 studies used the commercially available Uman NfL ELISA to measure CSF NfL, which has a lower limit of detection of 31ng/L documented by the manufacturer. The intra-assay coefficient of variations reported by the authors (between 3.5% and '<15%') were all below our predetermined cut-off of '<25%'. Seven papers explicitly reported that the analysis was blinded (Gunnarsson *et al.*, 2011; Axelsson *et al.*, 2014; Burman *et al.*, 2014; Trentini *et al.*, 2014; Aeinehband *et al.*, 2015; Stilund *et al.*, 2015; Hakansson *et al.*, 2017).

The Newcastle-Ottawa Scale (NOS) is a scoring system designed to assist with quality assessment of non-randomised research. We used this as a framework when reviewing the design of the studies included in our meta-analysis. However, as the NOS has not been validated, no articles were excluded based on this score (Stang, 2010).

Table 4-1 Summary table of studies included in meta-analysis

1st Author; Year	***No MS***	***Mean age of MS patients***	***No controls***	***Mean age of controls***	***Quality (NOS)***		
					Selection (4*)	Comparability (3*)	Exposure (2*)
Piehl; 2017	39	39.6	27	35.2	**		***
Trentini; 2014	31	49.6	15	39	***		**
Novakova; 2017a	59	37	39 (dup)	33.6	****		***
Novakova; 2017b	43	39.7	39	33.6	****	*	***
Hakansson; 2017	22	unknown	22	32	****	**	****
Bergman; 2016	110	37.7	113	40.2	***		**
Lam; 2015	59	45.7	44	40.4	**		**
Stilund; 2015	59	41.2	39	40.7	***		**
Villar; 2015	127	33.6	37	34.6	**	*	***
Aeinehband; 2015	48	41.2	18	30.4	***		**
Burman; 2014	63	43.8	15	40.2	**		***
Axelsson; 2014	35	48	14	42	****	*	***
Fialová; 2013	18	38	24	33	**		**
Gunnarsson; 2011	92	37.3	28	43	***		**

Quality (NOS) - Each study is scored using a star system based on three domains

1) selection of study groups (cases and controls) – maximum 4;*

2) comparability of the groups – maximum 2;*

3) ascertainment of outcome – maximum 3.*

(NOS = Newcastle-Ottawa Scale)

4.3 Results of meta-analysis

4.3.1 CSF NfL levels are higher in MS compared with controls

CSF NfL levels were approximately three times higher in 746 'clinically definite' MS patients (CDMS) than in 435 controls (1965.8 ng/L, SD=3102.5 vs. 578.3 ng/L, SD=1212.3). Meta-analysis revealed a statistically significant moderate effect size, 0.61, p<0.00001. MS and control groups were comparable in age (41.3 v 37.3 years, respectively) and sex (63% and 62% female). A funnel plot showed spread around the observed outcome, making publication bias unlikely.

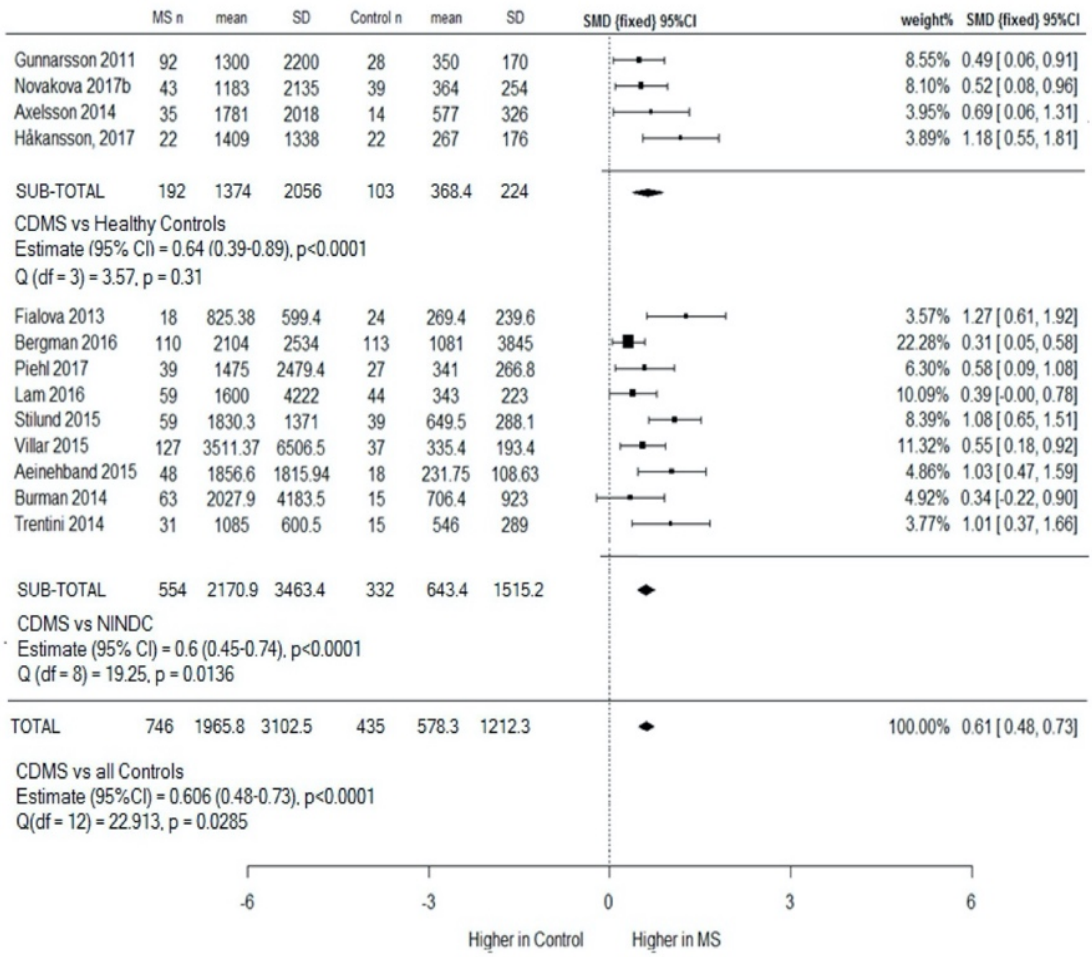


Figure 4-2 Clinically definite MS versus controls (HCs and NINDCs), subgroup and combined meta-analysis

Four studies compared CSF NfL levels in patients with MS with HCs, and nine studies used NINDCs. One group (Novakova et al) used the same control cohort in two papers (Novakova a and b). (Novakova a) was therefore excluded from the overall analysis to avoid duplication. This study is used in later sub-analyses, and a sensitivity analysis including it did not alter results. CSF NfL levels are higher in MS than healthy and disease controls. CDMS, clinically definite MS; CSF, cerebrospinal fluid; HCs, healthy controls; MS, multiple sclerosis; NINDCs, non-inflammatory neurological disease controls; SMD, standard mean deviation.

Heterogeneity between study outcomes was significant in the CDMS v NINDCs meta-analysis but not the CDMS v HC meta-analysis. The MS cohorts were comparable (mean age 41.7 years; 56% and 65% female), with approximately a third of subjects in relapse. This suggested that there were differences between the control populations, and we found that the mean CSF NfL level in the 332 NINDCs was double that of the 103 healthy controls, 643.4 ng/L, SD=1515.2 vs. 368.4 ng/L, SD=224.

4.3.2 CSF NfL levels are higher in relapse compared to remission, irrespective of MS subtype

CSF NfL was higher in RRMS than controls. This was significant during both relapse and remission, $p<0.00001$, but with a larger effect size during relapse (1.13), than remission (0.67).

RRMS patients in remission had CSF NfL levels five times greater than NINDCs, (1896.4 ng/L, SD=3371.4, v 365.1 ng/L, SD=281.3). Heterogeneity was not significant. When RRMS patients in relapse were compared with the same control population, CSF NfL levels were nine times higher in the MS patients compared to the controls, (3272.2 ng/L, SD=5164.8 vs 364.9 ng/L, SD=275.3), but heterogeneity was detected, p=0.0008.

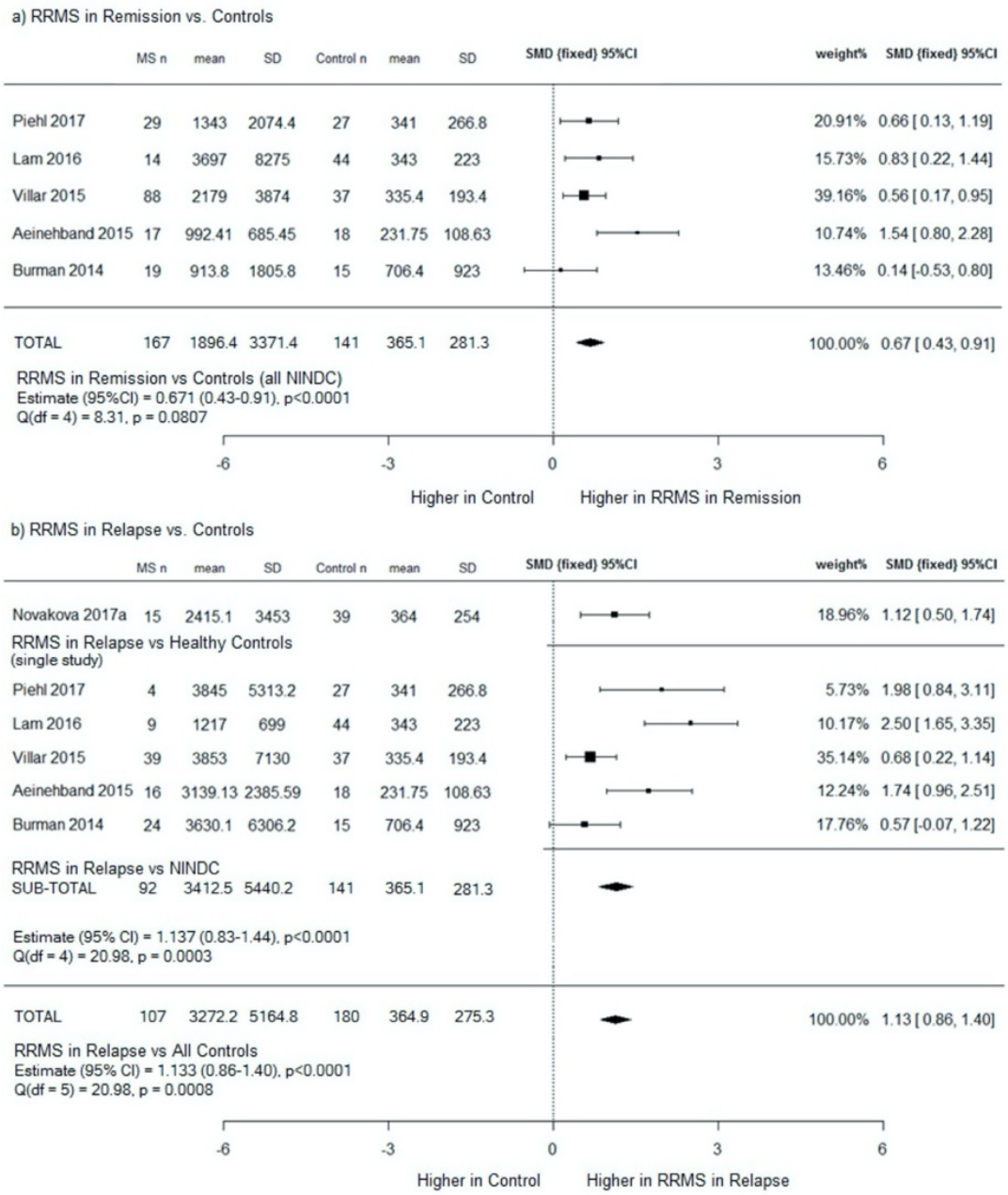


Figure 4-3 (A) RRMS in remission versus controls (NINDCs) and (B) RRMS in relapse versus controls (HCs and NINDCs), subgroup and combined meta- analysis

(A) CSF NfL levels were higher in patients with RRMS in remission than disease controls, with a moderate effect size of 0.67; (B) CSF NfL levels were higher in patients with RRMS in relapse than both healthy and disease controls, with a large effect size of 1.13; however, heterogeneity between studies was also significant. CSF, cerebrospinal fluid; HCs, healthy controls; NfL, neurofilament light chain; MS, multiple sclerosis; NINDCs, non-inflammatory neurological disease controls; RRMS, relapsing remitting MS; SMD, standard mean deviation.

Meta-analysis of 75 SPMS compared with 48 PPMS showed no difference in NfL levels (*see appendix 3*), and they were therefore combined as 'progressive MS' for further analyses. Mean CSF NfL levels were three times higher in progressive MS than controls, and meta-analysis showed a significant effect size of 0.96, p<0.00001. However, progressive MS patients were older (52.6 versus 38.4 years, p<0.001), and the sex distribution was unequal, (50.4% female versus 67.1% female, p=0.0047). When progressive patients were included alongside RRMS patients in an analysis of the effect of relapse on CSF NfL, NfL levels remained approximately twice as high in relapse populations compared to remission, (3080.6 ng/L, SD 4715.9 vs 1541.7 ng/L, SD 2406.5).

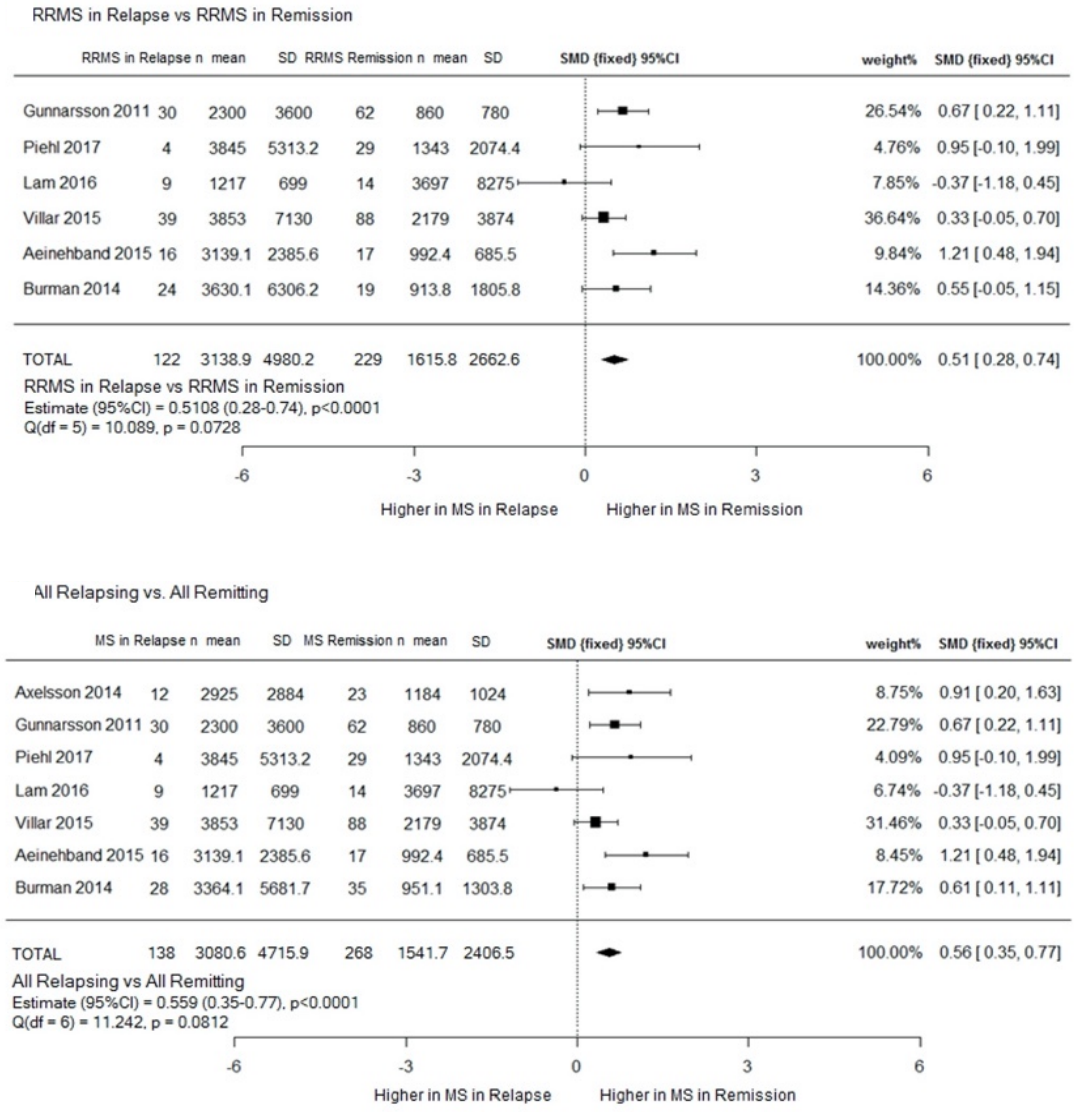


Figure 4-4 CSF NfL in relapse v remission for A) RRMS, and B) with the addition of progressive MS patients.

(A) CSF NfL levels are higher in RRMS in relapse than in remission; (B) when progressive and patients with RRMS are combined as 'all patients', CSF NfL levels remain higher in relapse patients than remission patients. SMD, standard mean deviation.

4.3.3 CSF NfL levels are higher in RRMS than progressive MS, but do not discriminate the two

Meta-analysis of five studies showed a higher CSF NfL in 176 RRMS patients compared with 92 progressive patients, (2124.8 ng/L versus 1121.4 ng/L). The effect size was small (0.34), but statistically significant, p=0.0108.

Demographic data were available for 87% of all subjects in this analysis. RRMS patients were younger (40.1 years versus 53.4 years), with a greater proportion female (74% versus 53%). Relapse data were available for 75% of the RRMS cohort (of which 40.2% were in relapse), and 83.7% of the progressive cohort (of which 5.2% were in relapse). There was also a marked difference in disease duration between the cohorts - 69.9 months for RRMS patients compared to 167.5 months for progressive patients. Limited subgroup data prevented us from being able to analyse the relationship between disease duration and NfL levels. Within the progressive cohort we did however note that SPMS patients had a longer mean disease duration than PPMS patients (204.3 versus 59.5 months), and that mean CSF NfL levels between the two did not differ.

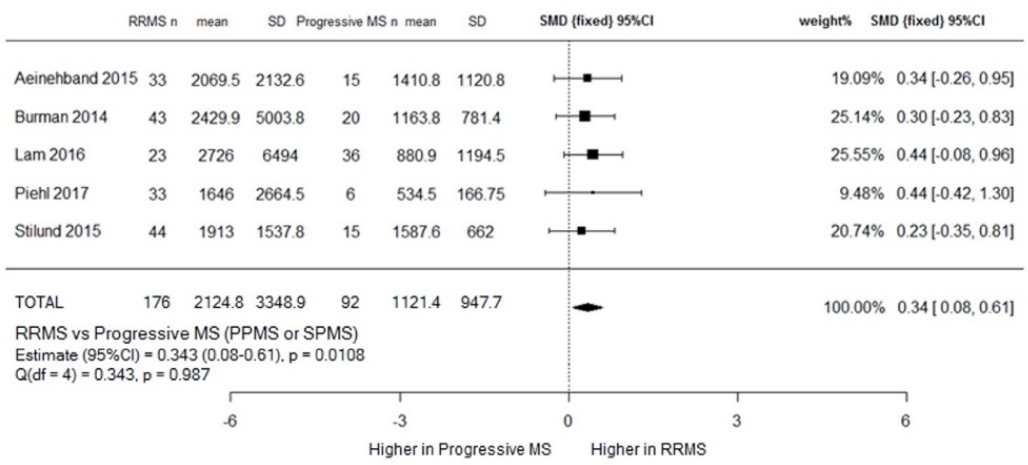


Figure 4-5 RRMS versus progressive MS (PPMS and SPMS combined) meta-analysis

We then asked whether CSF NfL might discriminate RRMS in remission from progressive MS. Although CSF NfL levels remained higher in RRMS in remission, the difference was not significant, suggesting that CSF NfL cannot be used to discriminate relapsing from progressive disease.

4.3.4 CSF NfL levels may be influenced by disease modifying treatments

Meta-analysis of 163 treated and 70 untreated patients with MS showed no effect of treatment on CSF NfL levels (*see appendix 4).* This was potentially a reflection of the fact that the majority (78%) of the treated cohort were on less efficacious DMT (interferon-beta, glatiramer acetate or teriflunomide). Less efficacious DMTs are recognised to reduce neurofilament levels to a lesser extent than highly efficacious DMTs, despite the fact they have a significant impact upon the development of new T2 lesions on MRI and reduce clinical relapse rates by approximately a third. Further analyses in larger cohorts are required to investigate the impact of different DMTs on neurofilament levels over both the short and long-term.

4.4 Use of next-generation single molecule array (Simoa) to measure NFL in blood

Our meta-analysis suggested that CSF NfL is a potentially useful marker of disease activity, and our findings are in keeping with other published meta-analyses (Cai et al, 2018; Bridel C et al., 2019).

However, CSF analysis in large powerful cohorts is limited by the small proportion of individuals who undergo CSF sampling as part of routine diagnosis. Further interrogation is required and measurement of NfL in blood is preferable to enable validation in large numbers of patients and controls.

Quantification of NfL in blood samples from an MS cohort was first performed in 2016 (Kuhle *et al.*, 2016). Since then, and the development of a commercially available assay, measurement of neurofilament in blood has been demonstrated by different groups, typically in cohorts of established MS, and using a Quanterix HD-1 Simoa analyser, which is fully automated and requires no sample or reagent manipulations (Cantó et al. 2019; Disanto et al. 2017; Novakova, Zetterberg, et al. 2017).

4.4.1 Measurement of NFL in paired serum-CSF samples in individuals with MS

We used a Simoa assay to ascertain whether we could detect neurofilament in the blood of individuals newly diagnosed with RRMS, and subsequently explored whether blood neurofilament levels reflected CSF levels.

Sixty-eight ‘paired’ CSF and serum specimens (i.e. samples taken sequentially) were acquired. All samples were taken during the diagnostic work-up for MS in patients who

had gone on to participate in the Future MS study. The population from whom the samples were obtained had baseline demographics typical of an RRMS disease cohort at the point of diagnosis, with a 2.05: 1 female to male predominance, and a median age of 34 years, IQR 30-43 years. The median EDSS score was 2.5 (IQR 2-3), reflecting the fact that the cohort had already developed mild disability by diagnosis.

Simoa Neuro 4-plex A kit was used with a Quanterix SR-X instrument. The SR-X instrument is a smaller, semi-automated benchtop version of the Quanterix HD-1 instrument. The SR-X requires manipulation of samples and reagents prior to analysis and is therefore more user-dependent than the HD-1 instrument.

CSF and serum samples from each participant were measured simultaneously, and in accordance to the manufacturer's instructions (see methods for detailed lab protocol), *figure 4-6(A).* All samples were measured in duplicate and researchers were blinded to sample identification during the experiment and initial data analysis. The mean fitted concentrations of each biomarker were calculated by Quanterix SR-X software, *figure 4-6(B).*

Data were reviewed to determine the reliability of results. All biomarker results were above the lower limit of quantification. High and low controls (provided by the manufacturer and run in quadruplicate) produced results within the expected ranges, *figure 4-6(C)*. Eight of the 68 paired samples had a coefficient of variation between duplicates of greater than 30% and were therefore excluded.

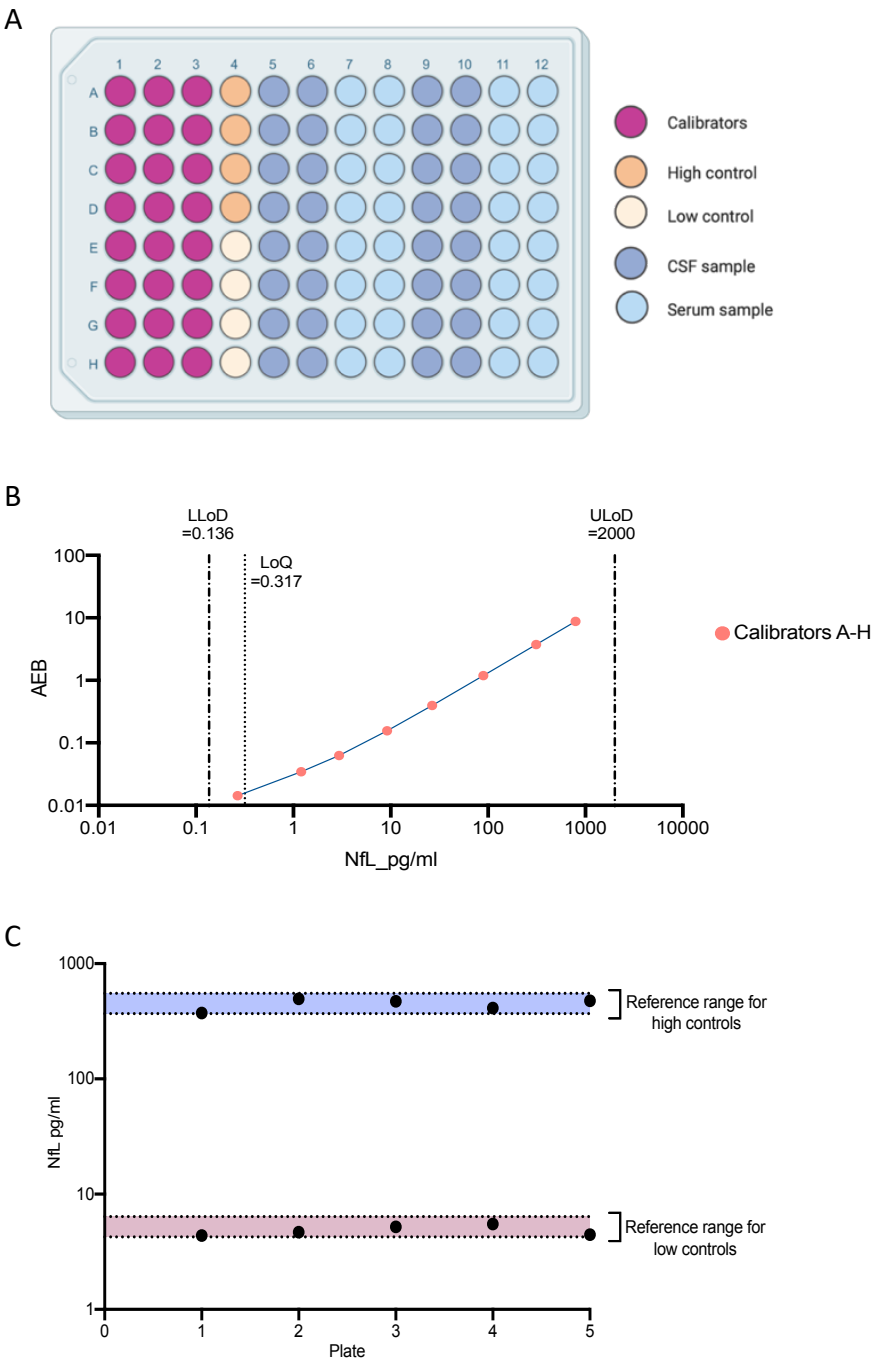


Figure 4-6 Simoa ELISA plate layout, standard curve and review of control results

(A) ELISA plate layout for simultaneous measurement of neurofilament in CSF and serum samples (B) Simoa NfL standard curve with annotations showing the dynamic range of the assay and lower limit of quantification (C) Interpolation of NfL concentration for high and low controls for each plate produced results within the expected ranges reported by the manufacturer.

4.4.2 Blood NfL levels correlate with CSF NfL levels

The median CSF NfL was 689 pg/ml (IQR, 418-1190), and the median serum NfL was approximately 60 times lower, 10 pg/ml (6.1-21), *figure 4-7(A)*.

Spearman rank correlation analysis revealed a strong positive correlation between CSF and serum NfL, r_s=0.69, which was statistically significant, p<0.001, *figure 4-7(B)*.

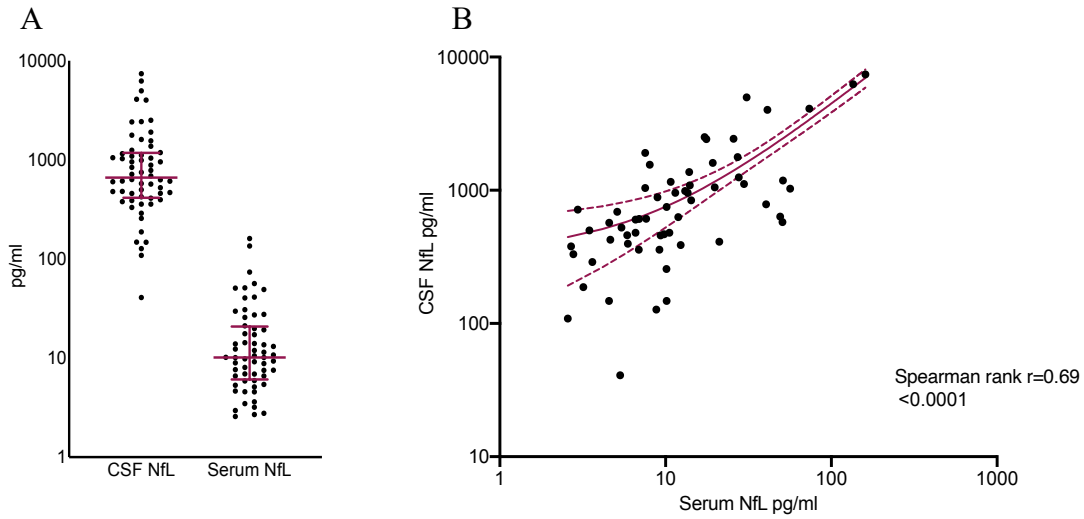


Figure 4-7 CSF and corresponding blood neurofilament concentrations

(A) CSF and serum neurofilament values shown with median and interquartile range. (B) Paired samples from 60 RRMS patients at the point of diagnostic work-up showed a strong correlation between CSF and serum results, r=0.69, p<0.0001.

4.4.3 Neurofilament levels did not differ between sexes or correlate with age

No significant correlation was found between age and neurofilament levels in CSF or serum in this cohort of newly diagnosed MS patients, *figure 4-8(A*). Neither CSF or serum NfL levels differed significantly between the sexes, p=0.30 and p=0.11, respectively, *figure 4-8(B)*.

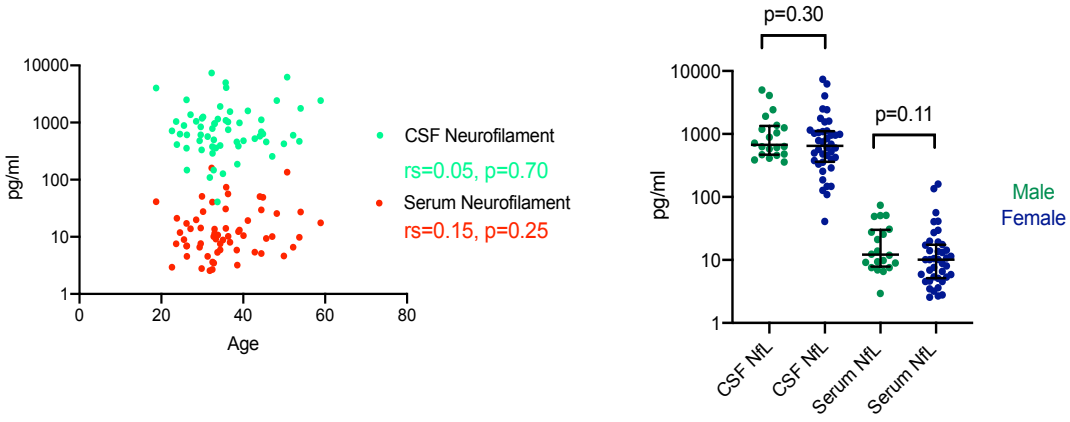


Figure 4-8 Association between CSF/serum neurofilament and age and sex

(A) Neurofilament levels did not show a significant correlation with age, in either CSF or serum. (B) Neurofilament levels did not differ between the sexes in either CSF or serum, Mann Whitney T test.

4.5 Discussion

Neurofilament is a non-specific marker of axonal injury. Consequently, neurofilament has been studied as a potential biomarker of neurodegeneration across many diseases, in different populations and using different assays (Bridel *et al.*, 2019). To better understand the potential role of neurofilament as a biomarker in multiple sclerosis we performed a meta-analysis of the literature on CSF neurofilament in multiple sclerosis.

In our meta-analysis we noted that levels of CSF NfL were elevated in all patients with MS, suggesting that axonal damage occurs *throughout the disease*. Our finding that relapse appears to be a stronger driver of CSF NfL levels than progressive disease may mean that NfL concentrations correlate more closely with acute than chronic axonal loss. This would suggest that transection of axons within an acute inflammatory lesion releases greater quantities of neurofilament subunits than insidious axonal damage as a consequence of chronic inflammation. This is demonstrated by the observation that on sequential analysis, CSF NfL levels within individuals are not *persistently significantly elevated* (Lycke *et al.*, 1998), although may remain higher than healthy controls.

At present, CSF NfL levels are thought to peak during acute relapse and decline within 3 months (Malmeström *et al.*, 2003). However, the dynamic changes of neurofilament levels in response to acute disease activity has not yet been established and will require serial samples in large cohorts over many years.

Limitations of our meta-analysis include the fact that a single author (SJM) reviewed the literature, and that studies were excluded due to insufficient data. We contacted the authors but had to exclude six papers. A further limitation was that we used raw data (uncorrected for age). The reason for this was that only some studies corrected data for age (when analysing results), and in others the demographic data provided was not sufficiently detailed to allow correction for age within subgroups. Despite these limitations, our results suggest that neurofilament could be a useful marker of acute axonal damage in MS. Our results are similar to other meta-analyses, both specifically in multiple sclerosis (Cai *et al*, 2018) and where multiple sclerosis has been included as part of a larger meta-analysis of neurofilament levels across different neurological diseases (Bridel C *et al*., 2019).

As CSF is invasive to obtain, the comparative ease of blood sampling is of significant advantage both in the validation of a fluid biomarker and its accessibility in routine clinical practice. Reliably quantifying a protein of CNS neuronal damage in blood was previously not possible due to the extremely low concentration of neuronal proteins in blood. By reducing the lower limit of quantification to sub-femtomolar levels, next-generation ELISA technology has changed that.

We successfully measured neurofilament levels in paired CSF and serum samples from patients at the point of diagnosis of RRMS. We have demonstrated that, in our hands, the Simoa neurofilament assay works with the Quanterix SR-X instrument and produces results in line with the published literature.

The median CSF NfL of 698pg/ml (418-1190pg/ml) is similar to other studies (Kuhle et al. 2016; Novakova, Zetterberg, *et al.* 2017) The reports surrounding blood NfL levels are less consistent, and our results of a median serum NfL of 10pg/ml (6.1-21pg/ml) are comparable with certain studies (Quanterix, 2018) but lower than others (Novakova, Zetterberg, *et al*. 2017). This may partly be a reflection of the disease activity of the participants sampled, but is also likely a result of variations in the techniques used (Quanterix semi-automated SR-X (our data); Quanterix fully-automated HD-1 analyser (Novakova, Zetterberg, *et al.* 2017); ECL assay using Uman AB (Kuhle *et al.*, 2016)).

Although neurofilament levels were approximately 60 times lower in serum than CSF, we found a moderate-strong correlation between the two, $r=0.69$, $p<0.0001$. We were also able to demonstrate this on an individual patient level, whereby the three individuals with the highest concentrations of CSF NfL also had the highest serum NfL levels (data not shown).

Using Simoa, we have demonstrated that neurofilament can be measured in patients *at the point of diagnosis of MS.* Our findings of a good correlation between CSF and blood neurofilament levels are consistent with published data and adds to the body of evidence of the use of blood in place of CSF in the measurement of neurofilament (Abdelhak *et al*. 2018; Kuhle *et al.* 2016; Novakova, Zetterberg, *et al*. 2017).

To address the extent to which neurofilament levels are elevated in early disease, and in what contexts, requires analysis in large, well-defined cohorts of early MS patients, and comparison with healthy controls. This is now feasible by using a Simoa assay to measure neurofilament levels in the Future MS cohort.

Measurement of neurofilament levels within the Future MS cohort also provides an opportunity to combine this blood biomarker of axonal damage with other biomarkers reflecting different aspects of the pathology of MS, such as an imaging biomarker of myelin integrity.

Summary of Chapter Four

- Metanalysis of systemically reviewed published studied shows that CSF Neurofilament has utility as a biomarker of acute disease activity in multiple sclerosis.

- Using Simoa, we have demonstrated that neurofilament can be detected in the blood at the point of diagnosis of MS, and that blood NfL levels correlate with CSF NfL levels.

Chapter Five

Evaluating myelin and axonal integrity in early MS lesions

5 Evaluating myelin and axonal integrity in early MS lesions

This chapter has been submitted for publication and is currently under review (joint first author with Elizabeth York). *This work was performed in close collaboration with Professor Adam Waldman (Professor of Neuroradiology, University of Edinburgh) and Elizabeth York. I was involved in the recruitment and review of the patients, the collection of samples, and the running and analysis of Simoa NFL assays. AW and EY were involved in the collection and analysis of raw imaging data. All parties contributed equally to the subsequent analysis of results.*

5.1 Introduction – Towards non-invasive evaluation of lesion biology

The pathological endpoint in multiple sclerosis is a combination of myelin damage, myelin repair and axonal loss. Multiple sclerosis is unpredictable but treatable. Individuals can develop irreversible disability over highly variable periods of time. Converging lines of experimental, clinical and trial evidence suggest that axonal loss is an important contributor to this disability (Ferguson *et al.*, 1997; Trapp and Nave, 2008; Gunnarsson *et al.*, 2011; Cantó *et al.*, 2019).

If prevention of permanent disability is to be achieved, **identification and quantification of factors which influence axonal damage *early* in the disease course is required to aid prognostication.**

Neuropathological studies

Neuropathological studies provide the greatest resolution in which to study the relationship between loss of myelin and axonal damage. Studies have demonstrated that axon transection occurs to a greater extent within areas of active inflammation (Ferguson *et al.*, 1997; Bruce D. Trapp *et al.*, 1998). This suggests that axonal loss occurs, to a degree, as a consequence of demyelination, and may therefore begin at disease onset. Histopathological studies have previously suggested that the *pattern* of demyelinating lesions may vary between, but not within, individuals; particularly early in the disease course (Lucchinetti *et al.*, 2000). This has been debated. Subsequent studies have not replicated these findings, suggesting that histopathological variation between demyelinating lesions may simply represent different time points of lesion formation (Breij *et al.*, 2008, Kuhlmann *et al.*, 2017), However, if the *severity* of demyelination varies between individuals, the extent of axonal damage may also differ, particularly early in the disease course when inflammation is most active.

By nature of the typical cohort (obtained post-mortem or on biopsy), histopathological studies lend themselves best to studying *established, chronic disease* or *fulminant severe disease*. If pathological heterogeneity between individuals is greatest in early disease (Breij *et al.*, 2008), then findings from post-mortem or lesion biopsy cases may not be representative of the majority of patients with early MS. *In vivo* biomarkers are therefore required to study the relationship between loss of myelin and loss of neuronal axons in *early or 'typical' disease*.

Magnetic resonance imaging is non-invasive and readily available (Wattjes *et al.*, 2015), and is the gold-standard clinical tool for the identification of focal areas of demyelination. A greater number of MS lesions at baseline has been associated with a higher chance of conversion from clinically isolated syndrome to clinically definite multiple sclerosis (Fisniku *et al.*, 2008) and a shorter time to reaching disability milestones (Tintore *et al.*, 2015). Ultimately however MRI is insensitive to microscopic pathology, primarily capturing the macroscopic end-point of white matter disease that has already occurred (Rovira, Auger, *et al.* 2013). Lesion appearances are diverse both within, and between patients; and clinical outcomes differ between patients with similar lesion burdens (Barkhof, 2002). Conventional MRI metrics therefore do not fully account for the clinical variability seen between individuals (Barkhof 2002; Rovira, Auger, *et al.* 2013), and lack sensitivity in estimating the risk of long-term disability for a particular patient.

The use of MRI to evaluate g-ratio

Advanced MR imaging techniques have been developed which offer opportunities for the noninvasive study of myelin integrity. One such method is the determination of an aggregate g-ratio (Stikov *et al.*, 2015; Campbell *et al.*, 2018). The g-ratio is defined as the ratio of the inner axonal radius to the fiber radius, quantified *in vitro* by electron microscopic visualisation and measurement. The g-ratio is therefore a measure of how well myelinated an axon is, and will vary during myelination, demyelination and remyelination, *figure 5-1*. Spatial resolution of MRI does not allow for individual axon g-ratios to be calculated. Instead, MRI can estimate an 'aggregate' g-ratio, within a voxel, using Magnetization Transfer Saturation (MTsat) and Diffusion Weighted Imaging (DWI) imaging modalities (see also sections 1.7.3 and 5.2.2.).

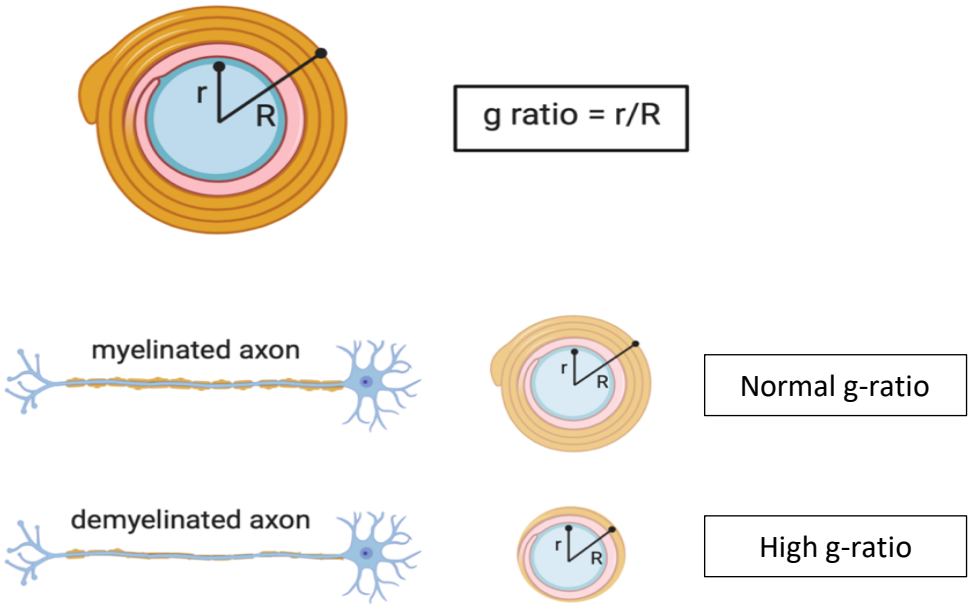


Figure 5-1 The g-ratio

The g-ratio is the ratio of the inner axonal radius to the myelinated fibre radius. The g-ratio is higher for demyelinated axons than for myelinated axons. Figure created in BioRender by SJM.

Combining MRI g-ratio with a blood biomarker of axonal damage

Neurofilament is released as a consequence of axonal transection (Lycke *et al.*, 1998; Petzold, 2005). We have demonstrated that, in our laboratory, Simoa can be used to accurately quantify NfL in blood and that levels correlate with those seen in CSF (Kuhle *et al*. 2016; Novakova, Zetterberg, *et al.* 2017).

Classical histopathological studies have demonstrated *ex vivo* that inflammation and axonal damage co-exist in *late disease,* and that axonal damage is significant in areas of acute inflammation. Based on the this, we hypothesized that in *early disease,* axonal damage is also likely to occur secondary to inflammation, and that individuals with more significant inflammatory disease activity (demyelination) may have more extensive axonal damage.

To investigate our hypothesis in newly diagnosed MS patients requires *in vivo* models. By combining advanced imaging metrics and blood neurofilament measurement we are able to simultaneously measure myelin integrity and axonal loss in a non-invasive way, *figure 5-2*.

Hypothesis

In newly diagnosed MS, individuals with more extensive inflammatory disease (demyelination) have greater axonal damage.

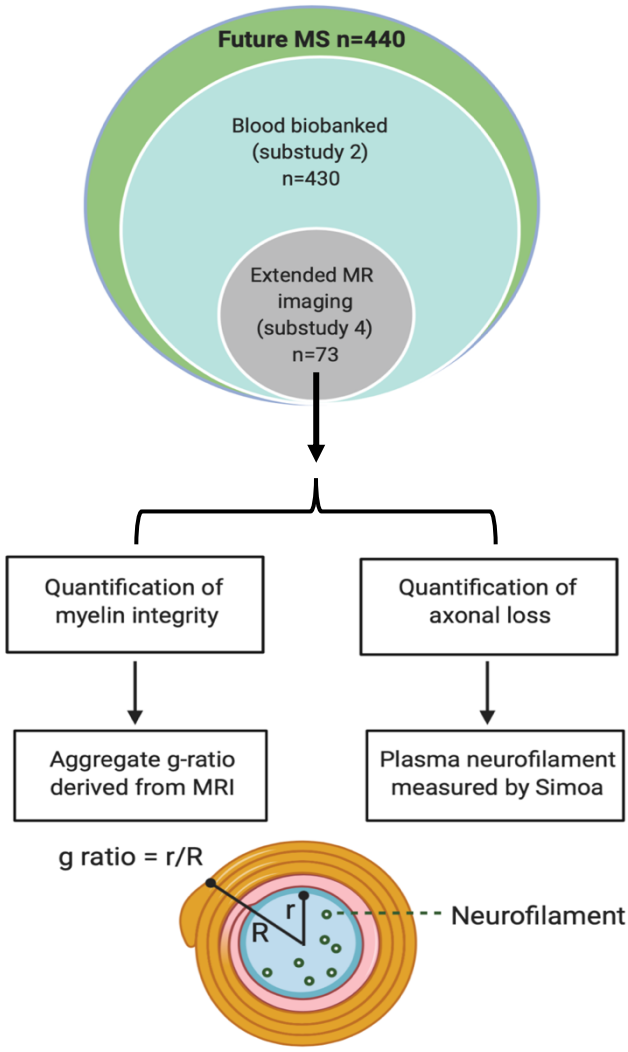


Figure 5-2 Study overview: Measuring myelin integrity and axonal damage in early MS

Within the Future MS cohort, 73 individuals participated in a substudy which enabled the simultaneous measurement of the MRI g-ratio (as a measure of myelin integrity) and blood neurofilament (as a measure of axonal damage).

Aim of chapter five

To examine the relationship between myelin integrity within MS lesions and axonal damage, at the point of diagnosis, in 73 individuals with early RRMS by combining advance MR imaging and blood-based biomarkers.

5.2 Methods

Within the larger Future MS cohort, 73 participants recruited at the Edinburgh site participated in an extended imaging sub-study which included the MRI sequences necessary to derive an aggregate g-ratio. All 73 had also participated in the sub-study biobanking plasma samples. Plasma samples were drawn at the same visit as MRI and were stored at -80ºC. As per the Future MS inclusion criteria, all participants had been diagnosed with RRMS according to the MacDonald 2010 or 2017 diagnostic criteria no more than six months prior to baseline review and were treatment naïve. Baseline demographic and clinical data were comparable to NHS incidence data, confirming that the study cohort was representative of the newly diagnosed MS population in Scotland.

5.2.1 Measurement of plasma NfL as a biomarker of axonal damage

We measured plasma neurofilament in 73 RRMS patients and 65 age-matched, healthy controls (HCs) recruited from the University of Edinburgh. The control samples were handled according to the same protocols as the Future MS samples.

All blood samples had undergone one freeze-thaw prior to thaw for measurement of plasma NfL. We used a Simoa Neuro 4-plex A kit along-with a Quanterix SR-X^{TM} benchtop instrument. The assay was performed by researchers (SJM and EC) who had previously validated its use in the measurement of CSF and serum NfL, and ran in accordance with the manufacturer's instructions (see method for lab protocol). Samples were measured in duplicate and researchers were blinded to sample identification during the experiment and in the initial data analysis.

The mean fitted concentrations were calculated by Quanterix SR-X software and provided in the form of a run report. Data were reviewed to determine the reliability of results.

All results were above the lower limit of quantification. High and low controls (provided by the manufacturer and ran in quadruplicate) produced results within the expected ranges. The average intra-assay coefficient of variation (CV) between duplicates was 9.4% for the 73 MS samples and 10.2% for the 65 HC samples. As a quality control step, only samples with a CV of less than 20% were included in analyses. Ten MS samples and 12 HC samples were therefore repeated. All repeat results were satisfactory for inclusion.

5.2.2 Measurement of the aggregate g-ratio as a biomarker of myelin integrity

The 73 participants in the extended imaging substudy had routine MR brain imaging performed (T1, T2 and FLAIR sequences) alongside advanced imaging sequences (MTsat and DWI). MRI data were acquired using Siemens magnetom prisma 3.0T systems and the images were processed by the University of Edinburgh research imaging team.

Anatomical images (specifically T2 FLAIR) were used for segmentation of cerebral white matter lesions (WML), *figure 5-3.* Lesion volumes (but not corresponding blood neurofilament levels) were corrected for intra-cranial volume (ICV). The lesion masks were visually inspected and corrected where necessary by the research imaging team. The remaining unmasked cerebral tissue was considered normal appearing white matter (NAWM).

The aggregate g-ratio was derived from MTsat and DWI data (neurite orientation dispersion and density imaging, NODDI) according to the following equation (Stikov *et al.*, 2015):

$$g = \sqrt{\left(1 + \frac{MVF}{AVF}\right)^{-1}}$$

where MVF is the myelin volume fraction and AVF is the axonal volume fraction. Myelin volume fraction was calculated from MTsat and axonal volume fraction from diffusion data (see also section 1.7.3).

The resulting g-ratio map was masked with the white matter lesion segmentations, and then mean aggregate g-ratios were determined for lesions and NAWM.

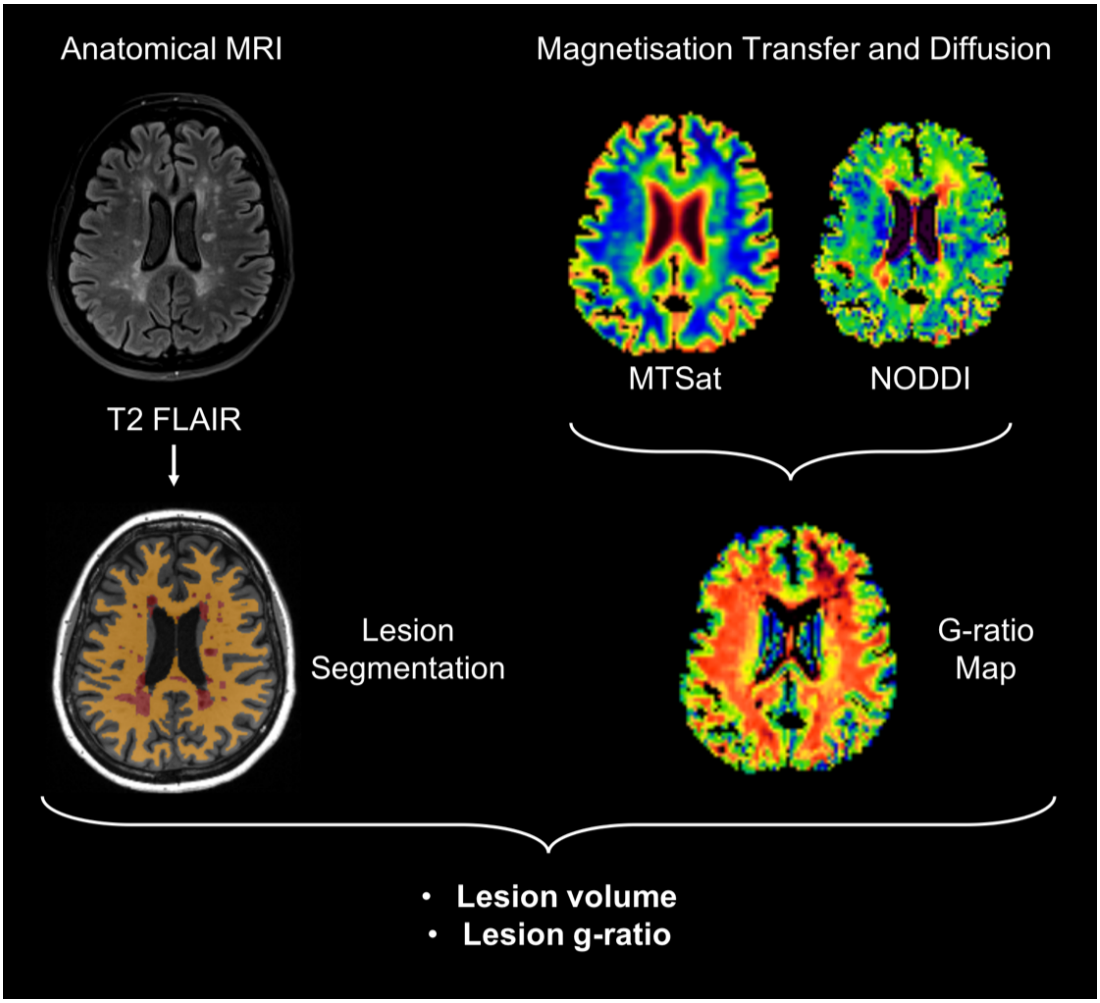


Figure 5-3 MRI analysis of the aggregate g-ratio within cerebral white matter lesions

Myelin integrity can be quantified as an "aggregate" MR g-ratio, derived from magnetization transfer imaging (MTsat) and diffusion (NODDI) MRI data, according to the equations shown in the text. In order to determine the aggregate g-ratio within lesions and within NAWM, anatomical MRI sequences (T2 FLAIR) were used to identify lesions, which were then segmented, and the segmentation map applied to the aggregate g-ratio map.

The normality of data were assessed using the Shapiro-Wilk test. Demographic, clinical, lesion volume and NfL data were not normally distributed and were therefore described by median and interquartile range. Correlations were analysed with Spearman rank correlation coefficient, and possible differences between groups were determined using Kruskal-Wallis one-way analysis of variance (ANOVA) with Dunn's multiple comparisons test or Mann-Whitney T-tests. All WML volumes used in analyses were corrected for intra-cranial volume.

5.3 Results

5.3.1 The study cohort was representative of the newly diagnosed RRMS population

We have previously demonstrated that the Future MS cohort appears representative of the Scottish newly diagnosed MS population (chapter 1). Demographic and clinic data of the 73 participants in the extended imaging substudy did not differ significantly from the Future MS cohort as a whole.

	This cohort	*Future MS cohort*
Participants	73	440
Sex ratio, F:M	3.6:1	2.82:1
Age at diagnosis median (IQR)	33 (28-45)	37 (30-45)
Time from symptom onset to inclusion in study (months)	34 (11-71) (N=67/73)	22 (10-62) (N=422/440)
Baseline EDSS	2 (1.5-3)	2 (1.5-3)

B

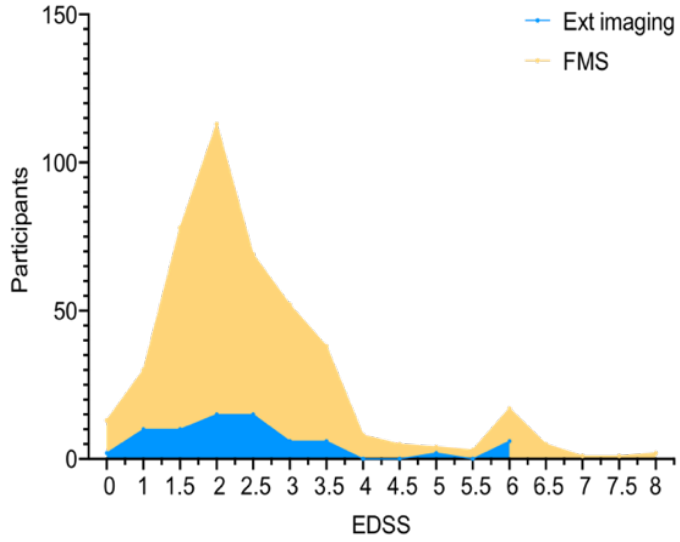


Figure 5-4 Comparison between the extended imaging cohort and the Future MS cohort as a whole

(A) Table comparing baseline demographic and clinic data between the extended imaging cohort and Future MS cohort as a whole, and (B) The distribution of baseline EDSS scores were similar for the extended imaging cohort (in blue) and the Future MS cohort as whole (in yellow).

5.3.2 The aggregate g-ratio was significantly higher in white matter lesions than NAWM

The MRI aggregate g-ratio was significantly higher in white matter lesions (median 0.607, range 0.538-0.682) than in NAWM (median 0.574, range 0.540-0.615), Mann-Whitney T test, p<0.0001, *figure 5-5*.

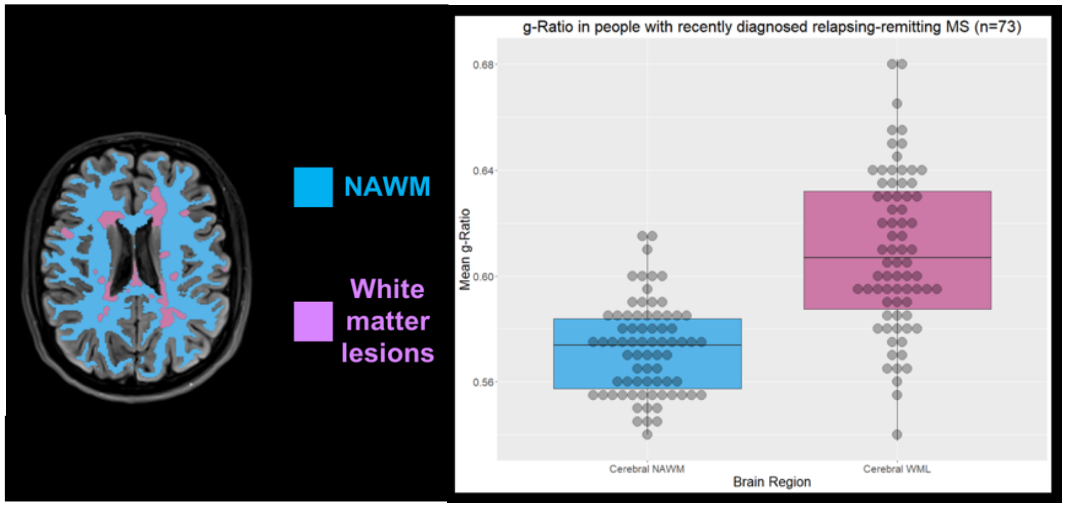


Figure 5-5 MRI aggregate g-ratio is higher in lesions than normal appearing white matter

Segmentation strategy for anatomical MRI scans shown, where lesions are purple, and NAWM is blue. The g-ratio was higher in lesions compared to NAWM.

From this we concluded that the MRI g-ratio can identify reduced signal within voxels/areas which are known to have myelin disruption (white matter lesions) compared with areas of normal appearing white matter. Although paired neuropathological analysis would be required to confirm that this reduced signal equates to myelin loss, the wide range of results suggests that the MRI aggregate g-ratio might be a marker of lesion severity.

We also noted that the g-ratio in white matter lesions not only varied between individuals, but appeared too to vary between individuals with similar structural MR imaging, *figure 5-6*.

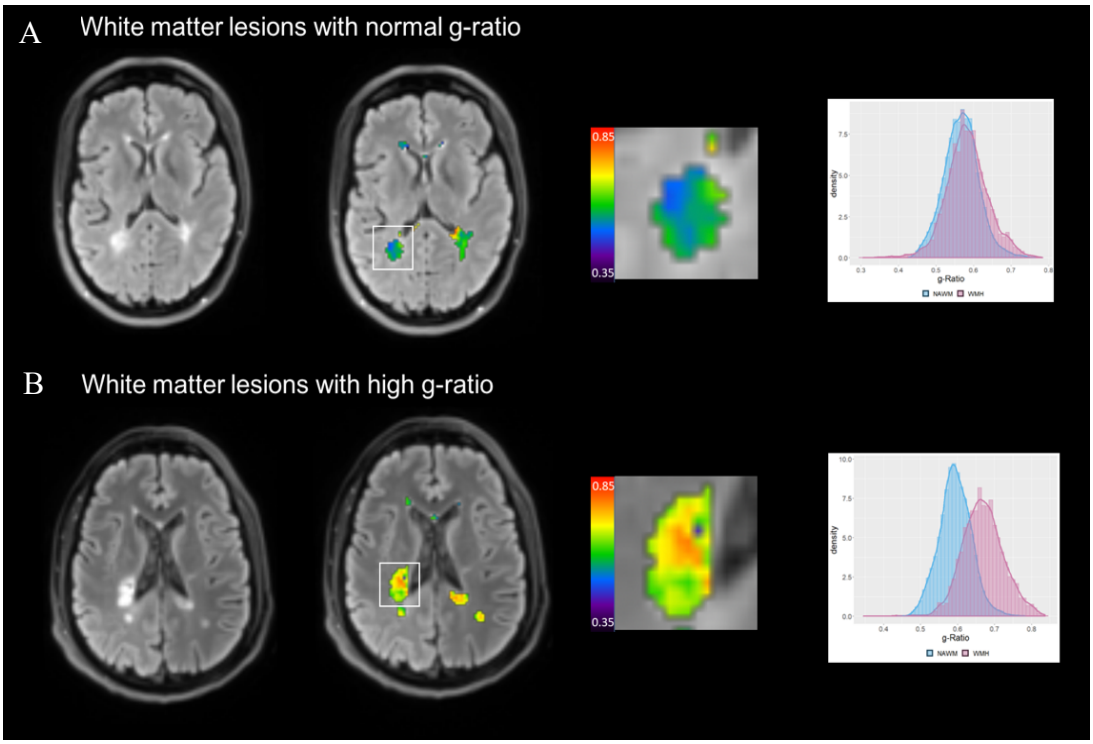


Figure 5-6 G-ratio in MS lesions at the point of diagnosis

(A) *Representative patient with significant lesion load and normal g-ratio within white matter lesions. On the left the T2 FLAIR image is shown, with hyperintense white matter lesions consistent with MS. The aggregate g-ratio from these lesion segmentations was evaluated (right hand MRI) and compared with the g-ratio from NAWM. This is shown as a histogram where the g-ratio from lesions (purple) is the same as the g-ratio from NAWM (blue).* ***(B)*** *Representative patient with significant lesion load and high g-ratio within white matter lesions. The lesion shown has a high g-ratio (right hand MRI) and the resulting individual histogram from the patient shows separation of g-ratios derived from white matter lesions (purple) and NAWM (blue).*

Based on these results we hypothesised that the aggregate g-ratio might have utility as a marker of lesion severity, (high G-ratio = more demyelination = more severe and vulnerable to axonal loss). To investigate this we examined the relationship between white matter lesion volume, the extent of myelin loss within lesion (aggregate g-ratio) and axonal loss (plasma NfL).

5.3.3 Plasma neurofilament levels are higher in MS than in age-matched healthy controls

In order to interpret blood NfL levels in MS patients we first measured plasma NfL in 65 healthy controls to identify a range of 'normal' values.

The healthy controls had a median age of 33 years old (IQR 28-48) and a female to male sex ratio of 3.3:1. The MS patients had a median age of 33 years old (IQR 28-45) and a female to male sex ratio of 3.6:1.

The median plasma NfL levels in HCs was 4.5pg/ml (IQR 3.5-6.5). The median plasma NfL in the 73 MS patients at the point of diagnosis was 7.5pg/ml, (IQR 4.9-12). We defined a 'high' NfL level as three standard deviations above the mean of HCs - 12.3pg/ml.

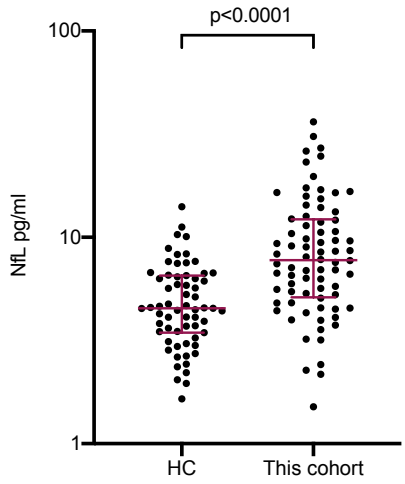


Figure 5-7 Comparison of plasma neurofilament levels between MS patients at point of diagnosis and age-matched controls

Data were non-parametric and a Mann-Whitney T-test was used to test significance. Plasma neurofilament levels were significantly higher in MS patients at the point of diagnosis than heathy controls, p<0.0001.

5.3.4 Plasma neurofilament is associated with both the size of lesions and the extent of myelin loss within lesions

The relationship between plasma NfL levels and white matter lesion volumes and g-ratios were analysed using Spearman's rho test for correlation. Plasma NfL levels showed a significant correlation with white matter lesion volume, r_s=0.39, p=0.007, and white matter lesion g-ratio, r_s=0.24, p=0.04, *figure 5-8.*

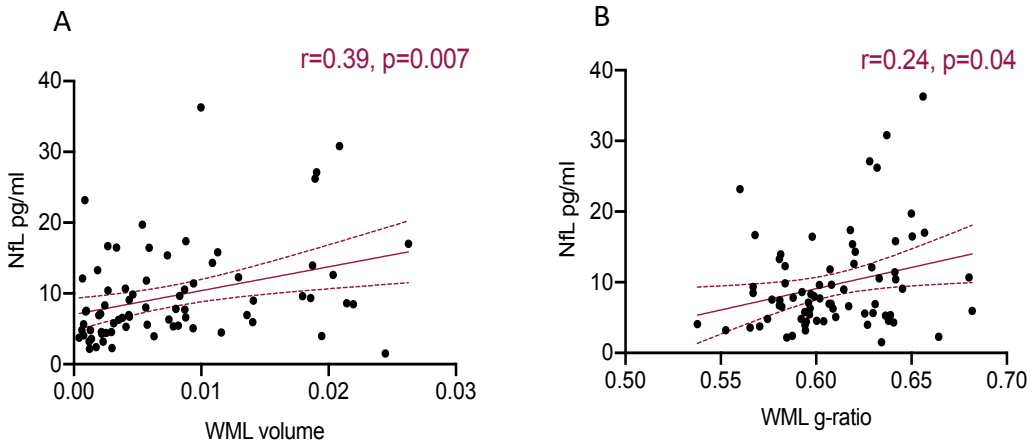


Figure 5-8 Correlation between plasma NfL and (A) lesion volume and (B) g-ratio

Plasma neurofilament showed a significant correlation with both white matter lesion volume and white matter lesion g-ratio when assessed using Spearman's rho test for correlation

This suggests that there is a relationship between plasma NfL and both lesion volume and lesion g-ratio. However, multivariate analysis revealed that only lesion volume was significant, p=0.01 (not lesion g-ratio, p=0.08) in determining plasma NfL.

Based on these results, we considered that axonal loss may be primarily driven by lesion size, but that the extent of myelin loss within lesions exerts an additional influence. To account for lesion volume and to examine the influence of g-ratio on axonal damage in patients with similar lesion loads, we divided the study cohort into two groups based on the median WML volume of the cohort – 'low lesion load' and 'high lesion load'. 'Low lesion load' was defined as a total white matter lesion volume below the median (WML volume </= 0.5% of intracranial volume), and 'High lesion load' as a WML volume above the median (>0.5% of intracranial volume). Within those groups we then investigated the relationship between MRI aggregate g-ratio and plasma NfL levels, *figure 5-9(A).*

‘Normal’ g-ratio’ and ‘High’ g-ratio were defined according to the median (0.607) in the same manner. Within the cohort with a low lesion load at diagnosis, 14% with a normal g-ratio and 8% with a high g-ratio had abnormally elevated plasma NfL levels (defined as >12.3pg/ml), 3/22 versus 1/13, no significant difference using Fisher’s Exact test. However, within the cohort with a high lesion load at diagnosis, 48% with a high g-ratio had abnormally raised plasma NfL levels, compared to 13% with a normal g-ratio, 11/23 versus 2/15, p=0.04, Fisher’s Exact test, *figure 5-9(B).*

The combination of a high lesion load *and* a high g-ratio therefore resulted in a significantly greater proportion of patients with abnormally raised blood NfL levels compared with those who had either a low lesion load, or a high lesion load but low g-ratio.

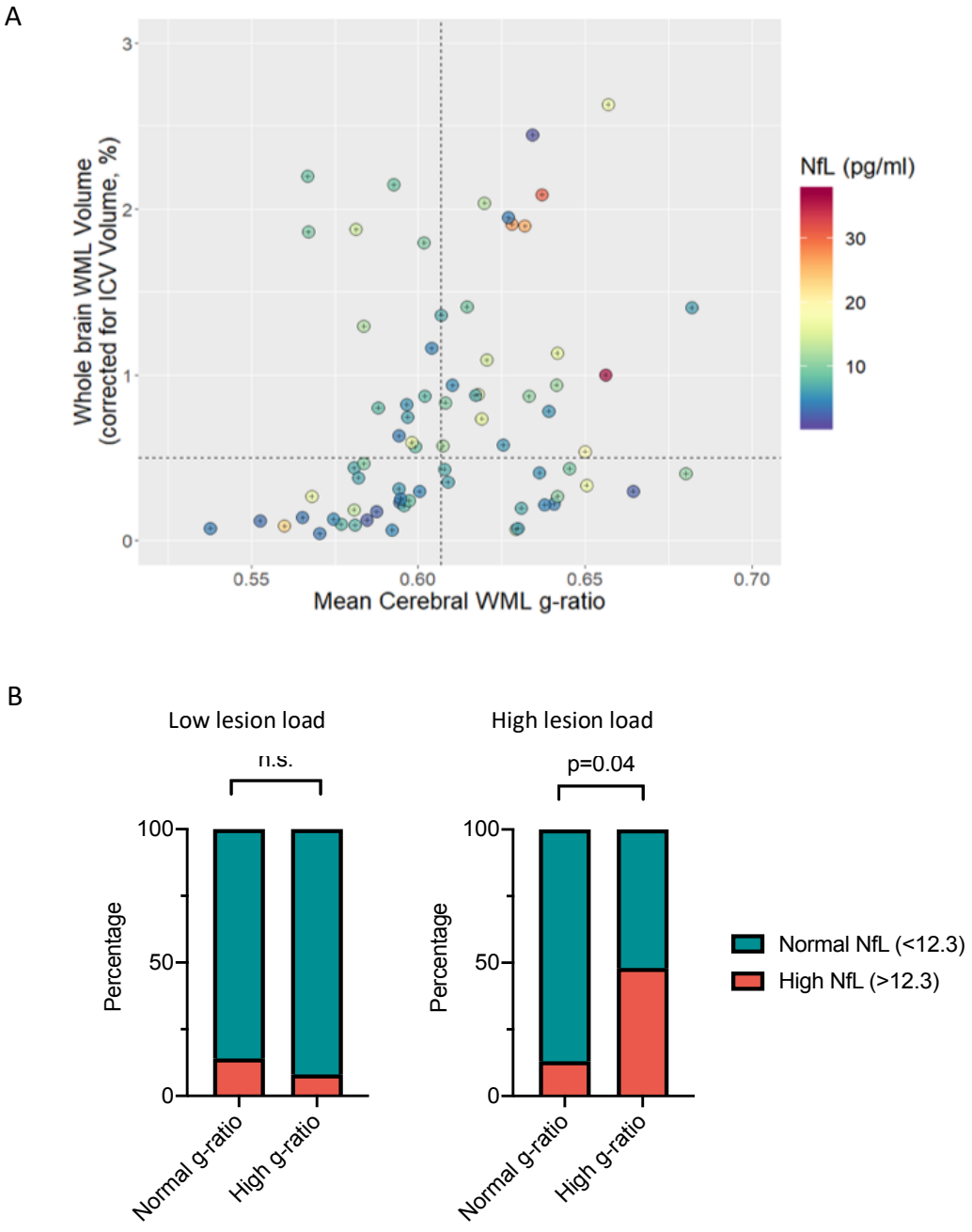


Figure 5-9 The influence of g-ratio on plasma neurofilament levels

(A) The relationship between white matter lesion volume at diagnosis, g-ratio within lesions and plasma neurofilament levels. (B) Patients with both a high white matter lesion load and a high aggregate g-ratio within lesions were significantly more likely to have high plasma neurofilament levels.

Within the high lesion load group, individuals with a high g-ratio did not differ significantly from those with a normal g-ratio in terms of the volume of white matter lesions, age, time to diagnosis, or the g-ratio of NAWM.

Average baseline disability scores were higher in those with a high g-ratio compared with those with a low g-ratio (EDSS 2.5 v 2, respectively), but were not statistically significant, and EDSS scores at 12 month follow-up did not differ between the groups.

Table 5-1 Comparison between patients with a high white matter lesion load at diagnosis

	High lesion load		Statistical significance*
	Normal g-ratio N=15	**High g-ratio N=23**	
WML volume (% of ICV)	1.2 (0.75-1.9)	1 (0.83-1.9)	p=0.71
Age of participants (years)	38 (32-53)	33 (26-49)	p=0.10
Time to diagnosis (months)	33 (13-75)	50 (7.8-105)	p=0.82
g-ratio in NAWM	0.57 (0.56-0.58)	0.57 (0.56-0.58)	p=0.88
Baseline EDSS, med (IQR)	2 (1.5-4.25)	2.5 (1.75-3)	p=0.87
Follow-up EDSS, med (IQR)	3 (2-5.75)	3 (2.4-3.6)	p=0.85

*Mann-Whitney unpaired T test used in all analyses

5.4 Discussion

Multiple lines of evidence suggest that axonal loss is an important biological substrate of disability (Tallantyre E.C. *et al*., 2010). Classic histopathological studies have demonstrated that significant axonal loss occurs in areas of acute, active, demyelination (Ferguson *et al.*, 1997; B D Trapp *et al.*, 1998). This infers that axonal loss probably occurs from disease onset. However, the extent to which demyelination drives axonal loss, and at what point in the disease course this occurs, has been debated.

The identification and quantification of axonal loss at the point of diagnosis could aid prognostication and inform treatment decisions. MR imaging is the gold standard clinical tool for identification of focal demyelination but is insensitive in quantifying axonal loss. This is particularly true in early disease when lesion loads may be minimal; or over short periods of time where atrophy changes may be tiny.

We used advanced MR imaging techniques to measure a biomarker of myelin integrity (the aggregate g-ratio) and combined this with a single molecule ELISA (Simoa) to quantify axonal loss in the blood of newly diagnosed RRMS patients.

We showed that the aggregate g-ratio detected loss of myelin integrity in white matter lesions, and that it varied between individuals. The MRI aggregate g-ratio was significantly higher in white matter lesions than in normal appearing white matter, Mann-Whitney T-test, $p<0.0001$. This suggests that it may have potential utility in quantifying myelin integrity in MS patients at the point of diagnosis. Our findings replicate other smaller studies (Stikov *et al*., 2015), but require histopathological validation by electron microscopy measurement of g-ratio within and outwith demyelinating lesions in humans. To date this has only been demonstrated, to our knowledge, in an animal (long-tailed macaque) model (Stikov *et al.,* 2015).

That the aggregate g-ratio of white matter lesions varied between individuals is in keeping with neuropathological studies which show that the extent of demyelination differs between patients (Lucchinetti *et al.*, 2000). Furthermore, the MRI aggregate g-ratio appeared to vary between individuals with similar structural MR imaging, *figure5-6.* We therefore asked whether the aggregate g-ratio might account for some of the clinical heterogeneity between patients with similar conventional radiological findings.

Using Simoa, we demonstrated that in early MS there is already evidence of axonal loss, and that plasma NfL levels were associated with both lesion volume and the aggregate g-ratio within lesions.

Published literature suggests that lesions need to be above a certain threshold volume before neurofilament levels become abnormally elevated (Schreiber *et al.*, 2018). We therefore stratified our newly diagnosed MS cohort according to lesion volume in order to study the effect of the g-ratio on axonal damage. In patients with a high white matter lesion load, we found that a high g-ratio was associated with high plasma NfL levels, whereas patients with a normal g-ratio were significantly less likely to have high NfL levels. This effect was not seen in patients with a minimal volume of lesions.

This is the first study to demonstrate the use of these particular non-invasive, advanced techniques (MRI aggregate g-ratio and blood neurofilament levels) for simultaneous evaluation of myelin and axon integrity in multiple sclerosis. This style of approach is frequently used in oncology where imaging and fluids biomarkers are combined to perform what is sometime referred to as a "liquid biopsy", whereby circulating biomarkers can give biological insight into focal imaging abnormalities and act as a proxy for invasive pathological studies. Our ability to perform these analyses *in vivo* provides us the significant advantage of linking these results with long-term clinical and radiological outcomes.

We chose to study the relationship between demyelination and axonal loss in newly diagnosed RRMS patients. This is a cohort in whom histopathological studies are rarely, if ever, performed. However, it is at this time-point - at the point of diagnosis, that prognostication is becoming critical given the increasing availability of disease modifying treatments. There is therefore a need for greater understanding of lesion biology in early disease. Our results, which suggest the demyelination influences axonal loss, are consistent with similar findings from classic neuropathological studies, and indicate that this relationship between myelin loss and axonal loss might apply throughout the disease course. Our data however only suggest an association and may be confounded by other aspects of lesion biology, such as level of inflammatory cell infiltrate and age of lesion.

There are several additional limitations to our study. Firstly, the MR aggregate g-ratio is an indirect measure of the g-ratio, and the imaging techniques employed are

dependent on certain assumptions (Campbell *et al.*, 2018). The g-ratio shows slight spatial variability (Stikov *et al.*, 2015) and therefore lesion location may be of relevance, although we did not find an obvious trend regarding lesion location in our cohort. The g-ratio also does not differentiate between demyelination and remyelination states, meaning that the same g-ratio result could occur as a result of irreversible myelin loss, or during a period of early remyelination with potential for axonal recovery.

A further limitation in the analysis of our results was that we did not correct our results for age or take into consideration time since recent relapse (clinical or radiological relapse). Blood NfL levels are age dependent, however no definitive study on age-appropriate levels has yet been published. The g-ratio may also show an age dependency, but this has not yet been extensively studied and may vary according to the metrics used to derive the aggregate g-ratio (Campbell *et al.*, 2018). Blood NfL levels increase during acute relapse and remain elevated for several weeks. Therefore, a relapse in the weeks prior to blood sampling could have resulted in a particularly elevated plasma NfL level, potentially prior to corresponding imaging findings "catching up".

No spinal imaging was performed as part of the Future MS study, and this was therefore not available to add to the analysis. Demyelinating spinal lesions could significantly elevate blood NfL levels without intra-cranial representation of significant disease burden.

In addition, we did not include analysis of grey matter lesions or correct for black holes. Grey matter lesions are not well visualised on conventional MRI sequences or without use of Gadolinium enhancement (which was also not included in the Future MS study protocol), and as a result are often under-reported compared with histological quantification (Geurts *et al*., 2005). For this reason, we did not attempt to quantify grey matter lesion volumes with this set of MRI sequences. Although a very high burden of grey matter lesions may increase neurofilament levels, previous studies have shown no correlation between MRI grey matter lesion volume and CSF neurofilament levels (Kuhle, J. *et al.,* 2016). This may be because neurofilament is predominantly expressed in large-calibre myelinated axons, and therefore more closely associated with white matter than grey matter disease.

Measurement of the MRI aggregate g-ratio is particularly susceptible to artefact at tissue-CSF interfaces (Duval, T., *et al,* 2017), making measurement of the g-ratio more difficult and less reproducible for cortical and spinal cord lesions. Finally, analysis of the MRI aggregate g-ratio in black holes would have required a second manual segmentation process with T1 mapping, which was outwith the scope of this study and is an area for future study.

Despite these limitations, this remains one of the largest studies to date of g-ratio mapping in multiple sclerosis (Hagiwara *et al.*, 2017; Kamagata *et al.*, 2019; Yu *et al.*, 2019)

In conclusion, we combined an imaging marker of myelin integrity and a blood biomarker of axonal damage to draw insights into lesion biology and prognostic relevance in early MS. We measured these biomarkers in a homogenous cohort of newly diagnosed, treatment naïve patients.

Our data suggest that measurement of the MRI aggregate g-ratio at an early stage in the disease is of potential utility in identifying individuals with greater axonal damage (as measured by blood neurofilament). Since axonal loss is the cause of permanent disability this is of clinical importance. To better understand how well blood neurofilament levels are associated with clinical metrics of disease activity requires analysis is a larger cohort.

Summary of Chapter Five

- We demonstrated that the MRI aggregate g-ratio is higher in MS lesions compared with normal appearing white matter and varies between individuals.

- There is an association between white matter lesion volume, lesion MRI g-ratio, and axonal damage in MS patients at the point of diagnosis.

Chapter Six

Multiplexed measurement of brain proteins in blood samples in early multiple sclerosis

6 Multiplexed measurement of brain proteins in blood samples in early multiple sclerosis

6.1 Introduction - Multiplex single molecule ELISA

We have shown that neurofilament levels, and thus axonal damage, differ between individuals with MS at the point of diagnosis, and may be influenced by the extent of demyelination within lesions. This demonstrates how Simoa-based detection of NfL in blood can be used to gain insights into clinically relevant biology of the disease. We therefore asked whether the additional measurement of other brain proteins such as glial fibrillary acidic protein (GFAP) could provide further insights into disease biology and clinical outcomes.

GFAP is a cytoskeletal protein of astrocytes - the glial cells that provide structural and functional support for neurons. Axonal damage associated with astrocytic scarring is frequently observed in MS lesions (Lassmann, 2019). For that reason, GFAP is of interest as a potential biomarker of disease activity.

Single molecule ELISA technology was further advanced in 2013 with the development of a multiplex assay (Rissin *et al.*, 2013). The Simoa multiplex assay employs the same basic methodology as the original Simoa assay but enables simultaneous measurement of multiple proteins. The ability to measure multiple proteins simultaneously has several advantages, namely that it can provide contemporaneous information about different cell types and that it is more efficient when testing low-volume samples, *figure 6-1*.

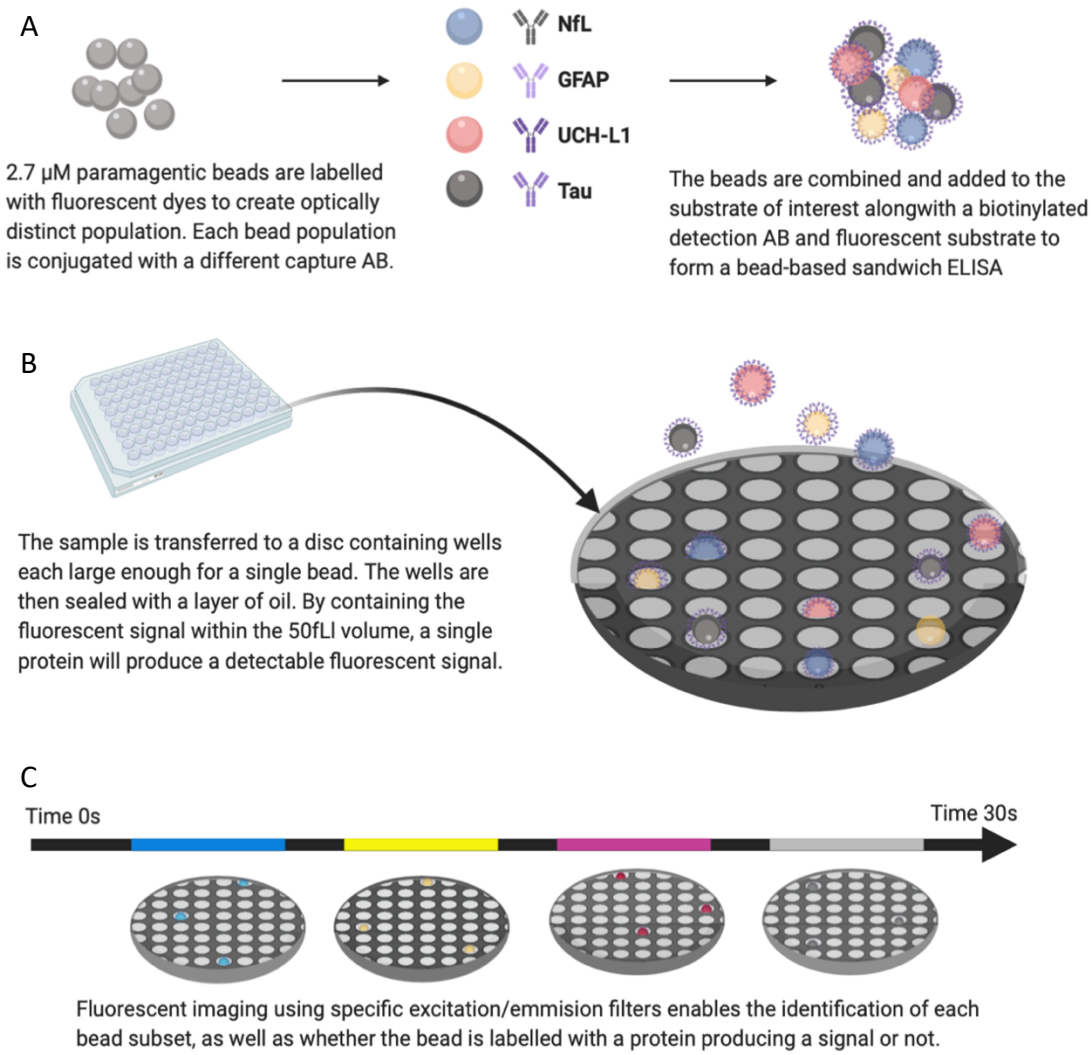


Figure 6-1 Principle behind a multiplex Simoa assay

(A) For a 4plex assay the beads are divided into four populations and each population of beads is fluorescently dyed and coupled with a specific capture antibody (e.g. anti-NfL IgG, anti-GFAP, anti-Tau and anti-UCH-L1). The different bead populations are then combined to produce a heterogenous bead population. This is added to the solution of interest, followed by the detection antibodies and a fluorescent substrate. (B) Within the Simoa analyser the beads are divided into single wells, which are then sealed. (C) A white light image determines which wells have a bead. A fluorescent image of all wells is then taken to determine which beads that have produced a fluorescent reaction ('on' beads). Finally, in order to determine which bead-type is within each well a series of fluorescent images at different wavelengths are taken. This enables the 'on beads' to be identified according to their capture antibody(Rissin *et al.*, 2013). *Image created in BioRender by SJM.*

We were particularly interested in measuring neurofilament light chain and glial fibrillary acidic protein (GFAP). NfL and GFAP were available as separate assays or included in a multiplex assay alongside ubiquitin C-terminal hydrolase-L1 (UCH-L1) and (total) tau protein. Rather than employ two separate assays, we chose to use a Simoa multiplex assay to quantify the three neuronal proteins and one glial protein in the blood of newly diagnosed, treatment naïve RRMS patients.

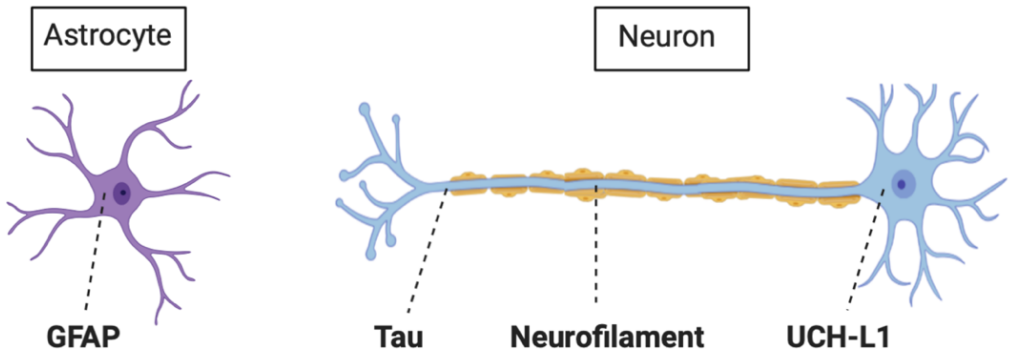


Figure 6-2 Four brain proteins measured by multiplex Simoa assay

Glial fibrillary acidic protein (GFAP) is an astrocytic protein. Tau, neurofilament and Ubiquitin C-terminal hydrolase-L1 (UCH-L1) are neuronal proteins. Image created in BioRender by SJM.

Table 6-1 Overview of the literature from individual studies analysing the potential relevance of four protein biomarkers in MS

Protein	*Main locations*	*Primary function*	*MS V HC*	*Disease subtypes*	*Relapse v Remission*	*Radiological measures*	*Clinical measures*	*Altered by DMT*
NfL	Neuronal axons in the CNS and PNS	Radial growth and structural integrity of axons	Higher in MS[1]	May be higher in RRMS[1]	Higher in clinical relapse[3] Association with Gd+[4]	Correlates with lesion volume[4]	Some correlation with clinical measures	Yes[7]
GFAP	Multiple isoforms: GFAP-α primarily found in astrocytes in CNS. GFAP-β primarily found in Schwann cells in PNS.	The cytoskeletal protein of astrocytes	Higher in MS[2]	May be higher in progressive disease[2]	Possibly association with Gd+[5]	Correlation with lesion volume[6]	Some correlation with clinical measures	No[7]
UCH-L1	Found ubiquitously throughout neurons of the CNS	Involved in the degradation of damaged or abnormal proteins	Not known	Not known	Not known	Not known	Not known	Not known
Tau	Neuronal axons and dendrites in the CNS. Also found in PNS.	Stabilizes microtubules in neuronal axons. Hyperphosphorylation results in formation of neurofibrillary tangles	Not known	Not known	No[8]	Not known	No[8]	Not known

[1](Bridel *et al.*, 2019), [2](Petzold, 2015), [3]*(Novakova, Zetterberg, et al. 2017),* [4]*(Kuhle, Kropshofer et al. 2019),* [5]*(Kuhle, Kropshofer et al. 2019)* [6](Ayrignac *et al.*, 2020), [7](Gunnarsson *et al.*, 2011), [8](Martínez *et al.*, 2015)

6.1.1 Neurofilament light chain as a biomarker in MS

See detailed discussion in chapter 1 Introduction, section 1.7.3. In short, blood neurofilament levels seem to correlate with those seen in CSF (Kuhle *et al.* 2016; Novakova, Zetterberg, *et al.* 2017) and appear too to reflect MS disease activity. Blood NfL levels are greater in MS than controls, higher in relapse than remission and decrease with disease modifying treatment (Novakova, Zetterberg, *et al.* 2017). Higher blood NfL at baseline has been linked with more significant brain atrophy at ten years, suggesting that blood NfL levels may also have a prognostic relevance (Cantó *et al.*, 2019).

6.1.2 Glial fibrillary acidic protein as a biomarker in MS

In response to cellular injury, astrocytes enlarge and develop thickened and elongated cytoplasmic processes. This process, known as reactive astrogliosis, results in an increase in GFAP expression (Eng, Ghirnikar *et al.* 2000). As activated astrocytes proliferate and migrate to areas of cellular injury, both the number of astrocytes and the concentration of GFAP within astrocytes are increased. If astrocytes are also damaged, GFAP and GFAP-breakdown products are released into extracellular fluid.

The literature on GFAP as a biomarker in MS is more mixed than for neurofilament. However, meta-analysis (Petzold, 2015) and recent studies suggest that GFAP levels are higher in MS compared with controls, and higher in in progressive disease than early or RRMS disease (Gunnarsson *et al.*, 2011; Ayrignac *et al.*, 2020)

The literature is conflicting (for both blood and CSF) as to whether GFAP levels reflect acute or chronic disease activity. One study reported that GFAP levels decreased during clinical relapse (Martínez *et al.*, 2015) whereas others have found no difference (Ayrignac *et al.*, 2020). Several studies have shown associations between GFAP levels and MRI (Gd+) lesion load (Ayrignac *et al*. 2020; Högel *et al*. 2020; Kuhle, Kropshofer *et al*. 2019); or GFAP levels and clinical metrics, such as EDSS (Axelsson *et al.*, 2011).

Several studies have combined blood GFAP and neurofilament assessment (Abdelhak *et al.*, 2018; Ayrignac *et al.*, 2020; Högel *et al.*, 2020), but in relatively small, heterogenous cohorts; and only one in early disease (Kassubek, 2017). No studies of combined blood neurofilament and GFAP levels have been performed in a large, homogenous and early disease cohort, such as Future MS.

6.1.3 Ubiquitin C-terminal hydrolase-L1 as a biomarker in MS

UCH-L1 has mainly been investigated as a potential marker of outcome in traumatic brain injury, whereby its ubiquitous expression throughout neurons may enable the identification of diffuse neuronal injury (Mondello *et al.*, 2011). UCH-L1 does not appear to have been evaluated as a biomarker in multiple sclerosis, and a lack of negative results reported suggests that accurate UCH-L1 quantification may be difficult. One study found CSF UCH-L1 was detectable in only 3 of 39 MS patients (Martínez *et al.*, 2015), although this was using a standard ELISA. A report by Quanterix showed that with a Simoa assay blood UCH-L1 levels did not differ between 16 RRMS patients and 12 healthy controls (Quanterix, 2018).

6.1.4 Tau as a biomarker in MS

The literature on tau as a biomarker in multiple sclerosis is contradictory and has mostly been conducted in CSF studies in small cohorts. One study found CSF tau to be significantly higher in MS compared with healthy controls (Abdelhak *et al.*, 2015), although that same study also reported that CSF tau levels did not appear to correlate with disease duration or clinical metrics (Martínez *et al.*, 2015). Other groups have found no difference in CSF tau levels between MS and controls, (Hein (née Maier) *et al.*, 2008), and Quanterix reported that serum tau levels were actually higher in 12 healthy controls compared with 16 RRMS patients (Quanterix, 2018).

We measured plasma NfL, GFAP, UCH-L1 and tau in the Future MS cohort and in 65 age-matched healthy controls recruited from the University of Edinburgh. We analysed differences in biomarker levels between controls and individuals with MS. Within the Future MS cohort, we then explored associations between biomarker levels and clinical and radiological measures at baseline and at 12 month follow-up, *figure 6-3*.

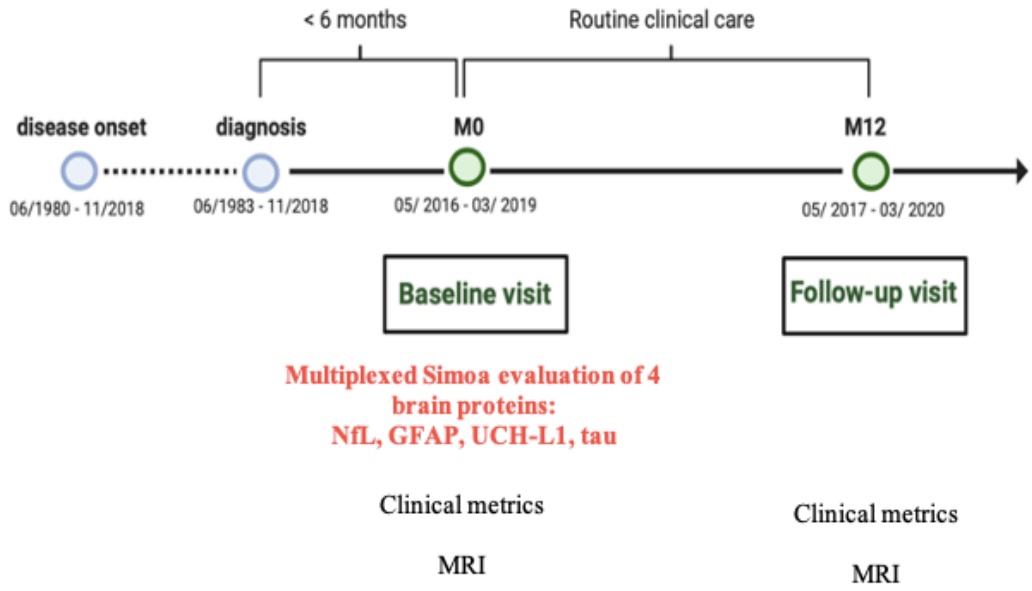


Figure 6-3 Correlation of blood biomarkers with clinical and radiological measures in the Future MS cohort

Aim of Chapter Six

To measure different brain proteins in the plasma of newly diagnosed RRMS patients and explore their prognostic value.

6.2 Methods - Using a multiplex single molecule ELISA to measure four biomarkers in newly diagnosed MS

Plasma samples were obtained at baseline study visit from Future MS participants, stored at -80ºC, and had undergone one freeze-thaw prior to use. Control samples from 65 healthy controls were handled in the same manner.

We chose to use healthy controls as the comparator in our analysis for several reasons. Samples from age-matched healthy control were available under the same ethical approvals as Future MS. The proteins measured as part of the 4plex are non-specific and may therefore be elevated to varying extents in different neurological diseases. As we were not investigating a potential *diagnostic* biomarker of multiple sclerosis this was less relevant and therefore did not require comparison with disease controls. Instead we were interested in whether any of the four proteins were reflective of disease pathology and comparison with healthy controls offered the greatest opportunity to identify that.

Quanterix Simoa Neuro 4-plex A kit was used along-with a Quanterix SR-X™ benchtop instrument. The assay was performed in accordance with the manufacturer's instructions (see methods, section 2.4.1 for lab protocol). All samples were measured in duplicate and researchers were blinded to sample identification during the experiment and in the initial data analysis. The assays were performed on non-consecutive days between December 2019 and March 2020. The mean fitted concentrations were calculated by Quanterix SR-X software and provided in the form of a run report. Data including the R2 value, accuracy of high and low control results, and the coefficient of variation between duplicates were reviewed to determine the reliability of results. Samples with a coefficient of variation between duplicates greater than 20% were repeated.

6.3 Results

Figure 6-4 shows values of the brain proteins measured in the blood. To determine whether differences were statistically significant we went on to compare MS and healthy control levels of each biomarker individually with Bonferroni correction for multiple comparisons.

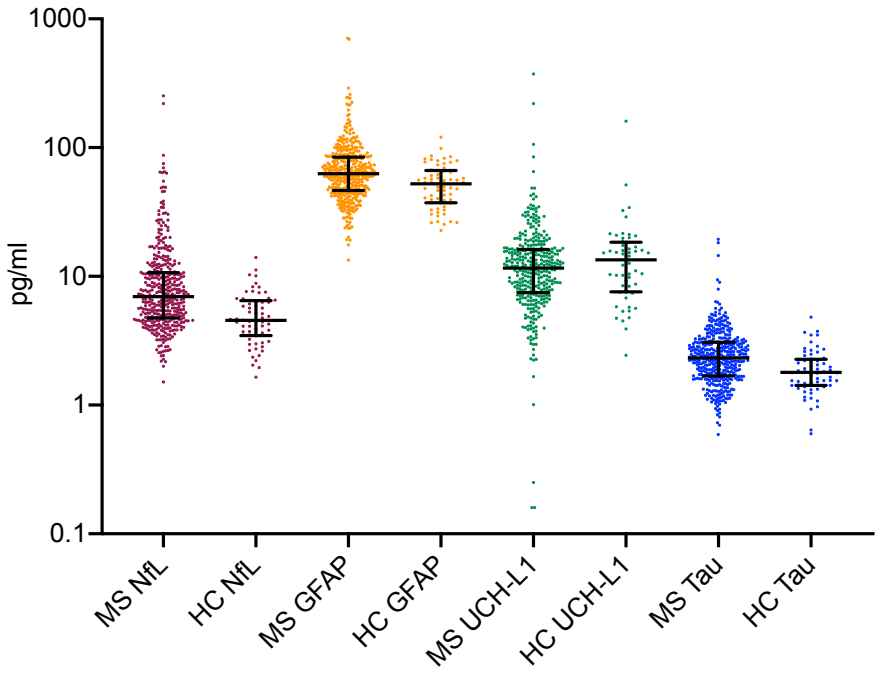


Figure 6-4 Comparison of four blood biomarkers in 440 patients with multiple sclerosis and 65 healthy controls

Neurofilament data were available for 419 MS patients and 65 healthy controls. Median NfL levels were 7.0pg/ml (IQR 4.8-11.0) in MS patients and 4.6pg/ml (3.5-6.5) in heathy controls. GFAP data were available for 416 MS patients and 65 healthy controls. Median GFAP levels were 63pg/ml (46-85) in MS patients and 52pg/ml (37-66) in healthy controls. UCH-1L data were available for 380 MS patients and 55 healthy controls. Median UCH-L1 levels were 12pg/ml (7.5-16) in MS patients and 13 (7.6-18) in healthy controls. Tau data were available for 425 MS patients and 65 healthy controls. Median tau levels were 2.3pg/ml (1.7-3.1) in MS patients and 1.8pg/ml (1.4-2.3) in healthy controls.

6.3.1 Levels of neurofilament and GFAP are significantly higher in newly diagnosed MS than healthy controls

We compared MS and healthy control levels of each biomarker individually with Bonferroni correction for multiple comparisons. Analysis showed that of the four biomarkers, only neurofilament and GFAP were significantly higher in MS than age-matched healthy controls, *figure 6-5.* However, when two (MS) outliers were removed from the analysis, GFAP levels were no longer significantly higher in the MS group compared with controls, p=0.08.

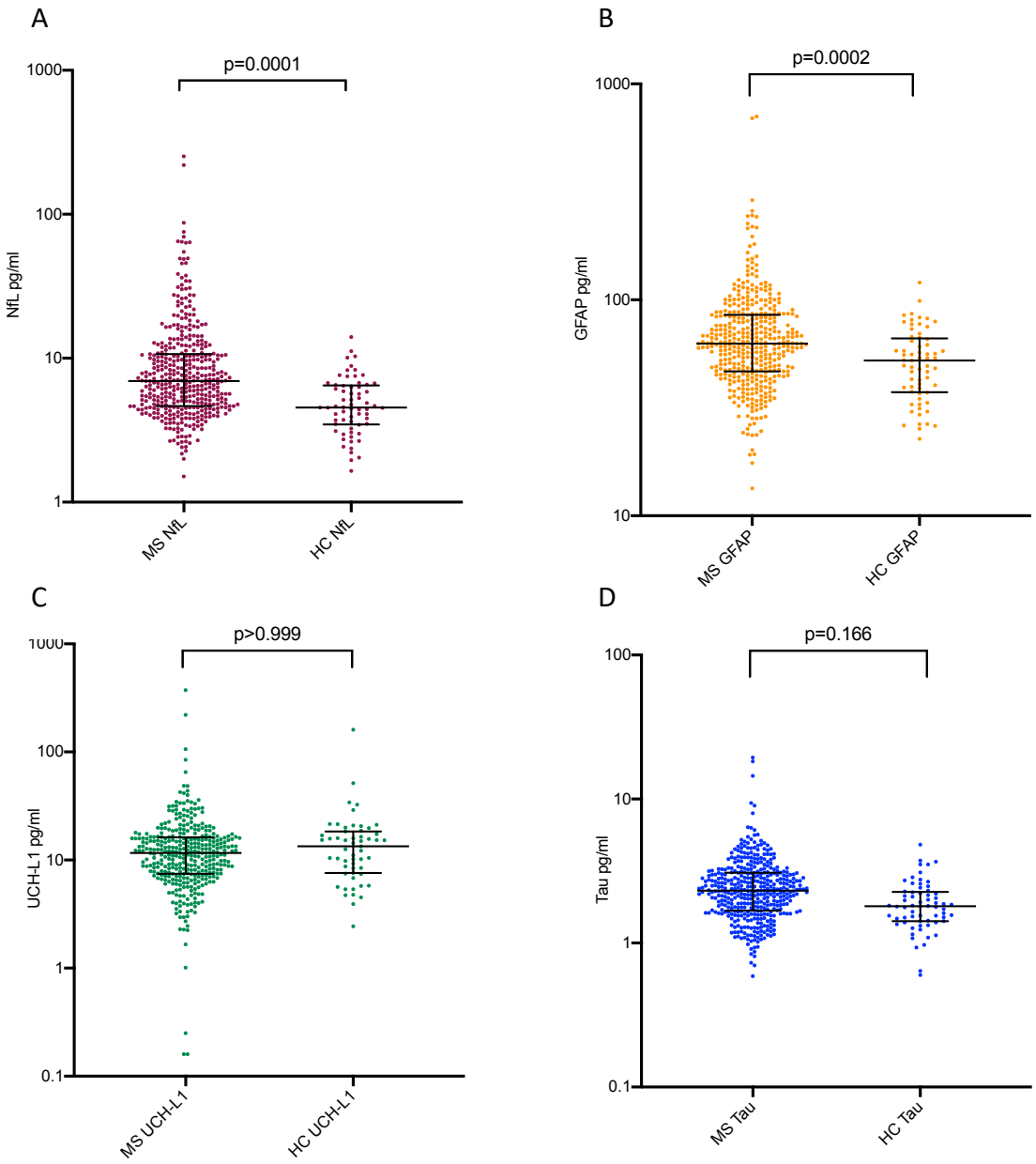


Figure 6-5 Analysis between MS and HC plasma levels of (A) neurofilament, (B) GFAP, (C) UCH-L1, (D) tau, Kruskal-Wallis ANOVA with p-values corrected for multiple comparisons.

6.3.2 Plasma neurofilament correlates weakly with plasma GFAP

We compared the correlations between blood biomarkers in (a) healthy controls and (b) multiple sclerosis. The most consistent finding was that blood levels of neurofilament and GFAP correlated weakly with each other in both healthy controls and MS.

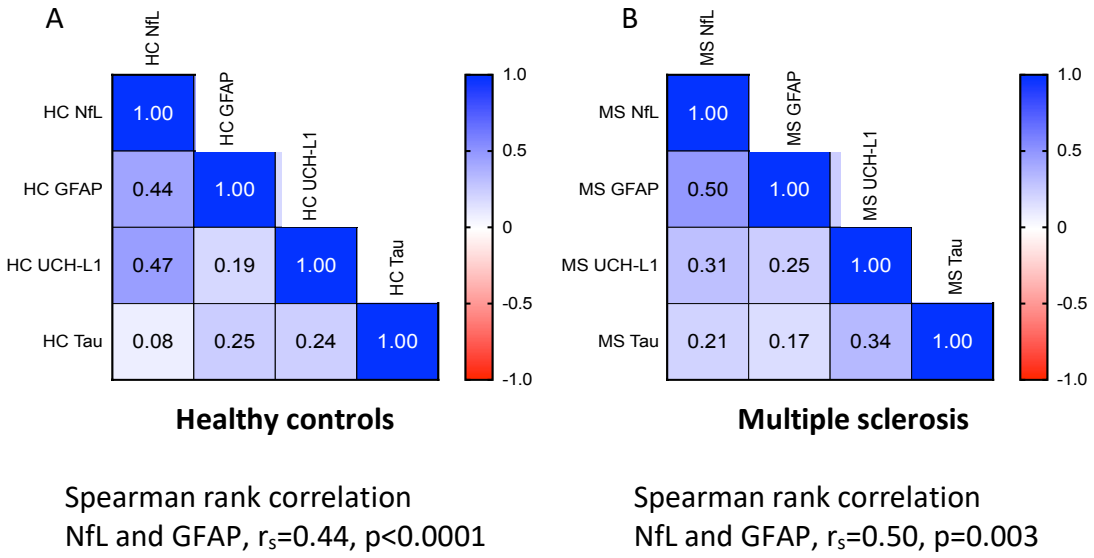


Figure 6-6 Correlation between blood biomarkers in (A) healthy controls, and (B) multiple sclerosis

The correlations between the four blood biomarkers differed slightly between healthy controls and individuals with MS. In both populations, a significant correlation was seen between neurofilament and GFAP levels. This could suggest a correlation between axonal and astrocyte damage.

6.3.3 Neurofilament and GFAP, but not other brain proteins, correlate with age

Neurofilament showed a strong positive correlation with age in healthy controls, r_s=0.52, p<0.0001. This was not observed in the context of multiple sclerosis, r_s= 0.07, p=0.16, *figure 6-7(A)*. GFAP correlated with age in healthy controls, r_s=0.41, p=0.0008, but again this association was not observed in MS patients, r= -0.003, p=0.95, *figure 6-7(B)*.

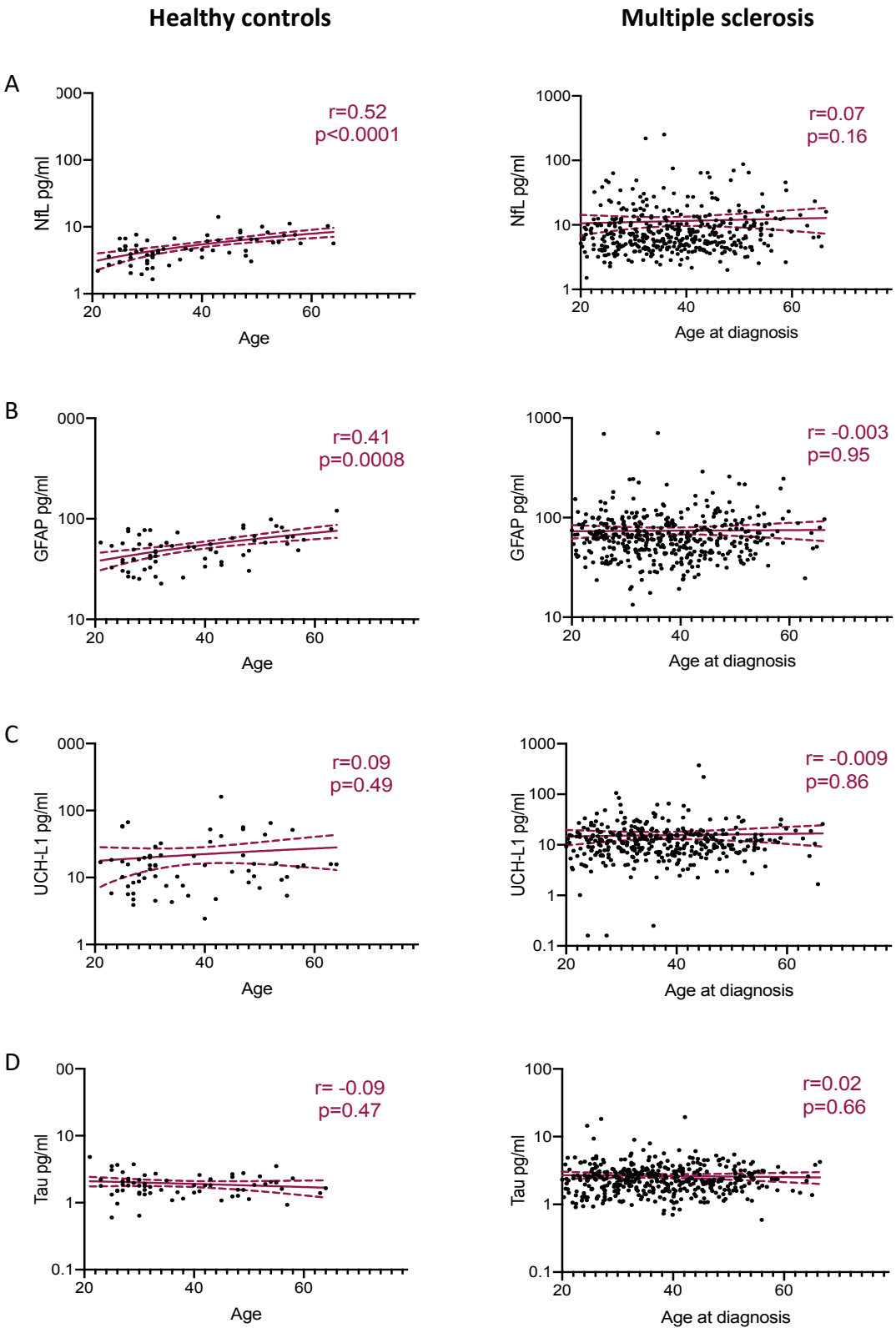


Figure 6-7 Biomarker correlation with age in healthy controls and in MS

(A) Neurofilament, (B) GFAP, (C) UCH-L1 and (D) tau, showing linear regression line and 95% CI. Neurofilament and GFP correlate with age in healthy controls, but not in MS. UCH-L1 and tau levels showed no correlation with age in either controls or individuals with MS.

6.3.3.1 Blood NFL concentrations alone correlate with CSF concentration

We next asked whether blood 4-plex Simoa concentrations correlated with CSF concentrations by analysing a subset of FMS participants who had paired CSF and serum. The correlation between CSF and serum for each biomarker were reviewed to determine how well blood levels refelect CSF levels.

The only biomarker to show a statistically significant correlation between CSF and blood concentrations was neurofilament, r_s=0.69, p<0.001, *figure 6-8(A)*.

We previously showed in chapter four (section 4.4.2, *figure 4-7*) that serum NFL concentrations were approximately 60 times less than CSF, but strongly correlated with CSF NFL. The median GFAP concentration was 25 times higher in the CSF compared with blood, and no statistically significant correlation was identified, r_s=0.06, p=0.65. UCH-L1 levels were also approximately 25 times higher in CSF than blood, and no significant correlation was found, r_s=0.23, p=0.12. The median concentration of Tau in CSF was 112 times higher than blood, and again no statistically significant correlation was identified between CSF and blood levels, r_s=0.14, p=0.31. Therefore, only NFL serum concentrations correlated with CSF levels.

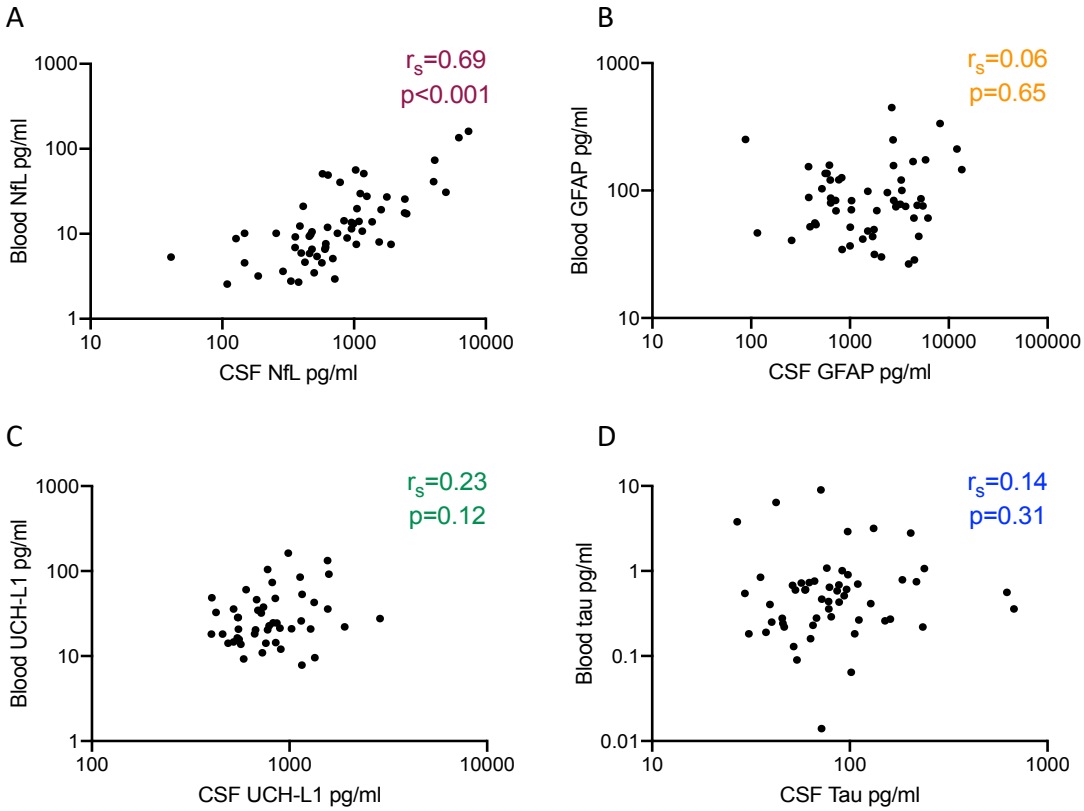


Figure 6-8 Correlation between CSF and blood concentrations of brain proteins

Correlation between CSF and serum concentrations of (A) neurofilament, (B) GFAP, (C) UCH-L1, (D) tau in 65 RRMS patients at point of diagnosis. The only correlation which was statistically significant was for neurofilament, r_s=0.69, p<0.001.

6.3.4 Blood levels of certain brain proteins are associated with MRI metrics of disease activity at baseline and at twelve months

We showed in the previous chapter in a subset (73 of 440 subjects enrolled in an advanced imaging study) a significant association between plasma NFL and WML volume (corrected for ICV) at baseline.

In this larger study we replicated this finding: Blood neurofilament correlated with baseline white matter lesion volume, r_s=0.38, p<0.0001. When we analysed GFAP we found that it too correlated with baseline white matter lesion volume, r_s=0.33, p<0.001. No such association was found for UCH-L1 and tau.

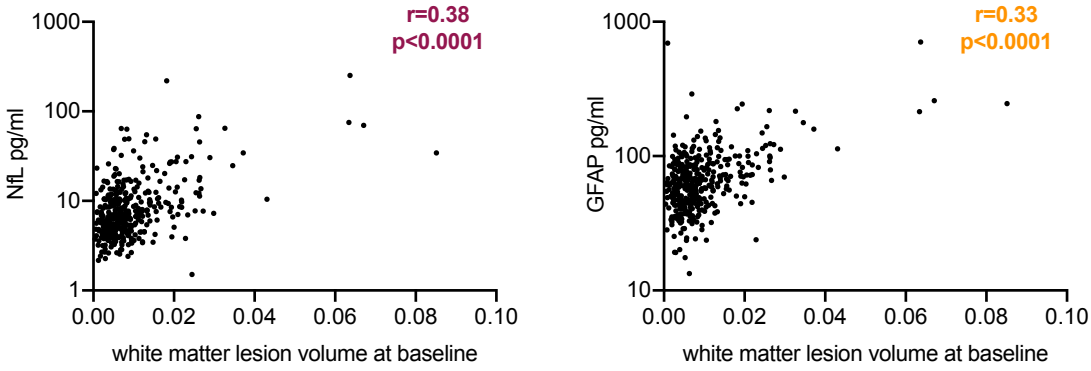


Figure 6-9 Neurofilament and GFAP levels correlate with white matter lesion volume at baseline

We then asked, if an individual has an abnormally high NfL/GFAP level, do they have an increased risk of developing new white matter lesions at one year?

We stratified the cohort into 'normal' and 'high'levels of each of the biomarkers using threshold values calculated from 65 healthy controls (mean plus three standard deviations, see also chapter five, section 5.3.3). Normal NfL levels were defined as less than or equal to 12.3pg/ml and high NfL levels as >12.3 pg/ml. Normal GFAP levels were defined as less than or equal to 113pg/ml; normal UCH-L1levels as less than or equal to 83 pg/ml; and normal tau as less than or equal to 4.2 pg/ml.

Using this approach, 86 of 420 participants (20.5%) had elevated NfL levels; 39 of 417 (9.4%) had elevated GFAP; 4/366 (1.1%) had elevated UCH-L1 and 39/426 (9.2%) had elevated tau levels, *figure 6-10.*

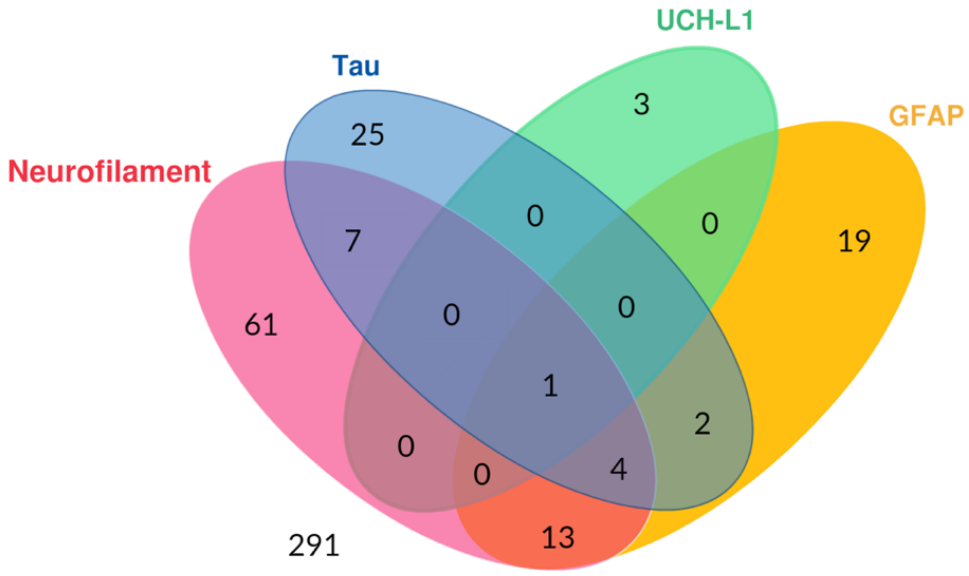


Figure 6-10 Venn diagram showing number of individuals with elevated levels for each of the four biomarkers

291 individuals had 'normal' levels for all four biomarkers and one individual had elevated levels of all four biomarkers.

A binary outcome of 'new or enlarging T2 WML lesions - yes/no' determined by visual reads was available for 83% (367 participants) of the Future MS cohort. White matter lesion volumes at 12 month follow-up were available for approximately half the cohort (225 participants).

Within the cohort for whom we had MRI visual reads data, we found that individuals within the 'high' baseline neurofilament group (>12.3pg/ml, n=73) were more likely to develop new/enlarging WMLs than those with a 'normal' baseline neurofilament level (<12.3pg/ml, n=294), 71.2% versus 34.9%, Fisher's exact test for significance p=0.00002, f*igure 6-11(A)*.

However, the absolute change in volume of WML over 12 months was not significantly different between those with normal blood NfL levels and those with high blood NfL levels. This suggests that either the changes in new/enlarging lesions were relatively small, or that larger changes in volume were masked by the simultaneous resolution of active lesions alongside the development of new lesions, *figure 6-11(B)*.

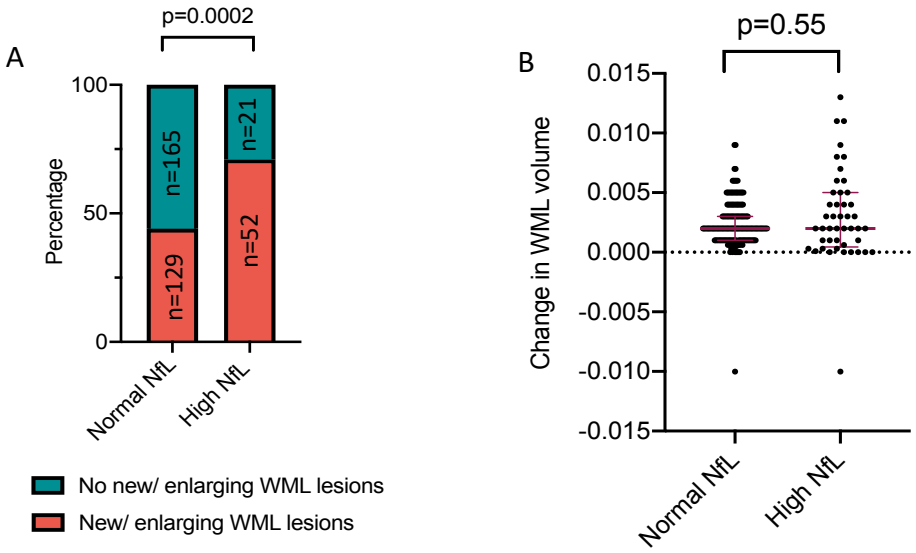


Figure 6-11 Association between baseline neurofilament levels and development of new WML lesions over 12 months

(A) High NfL levels at baseline were associated with a greater risk of new/enlarging white matter lesions at 12 months, Fisher's exact test for significance, p=0.0002. (B) Absolute changes in WML volume did not differ significantly between groups, Mann-Whitney T test.

Blood levels of GFAP at baseline did not show any prognostic significance in the risk of developing new or enlarging white matter lesions at 12 months (figure 6-12). The absolute change in WML volume was also no different between those with normal and high plasma GFAP levels.

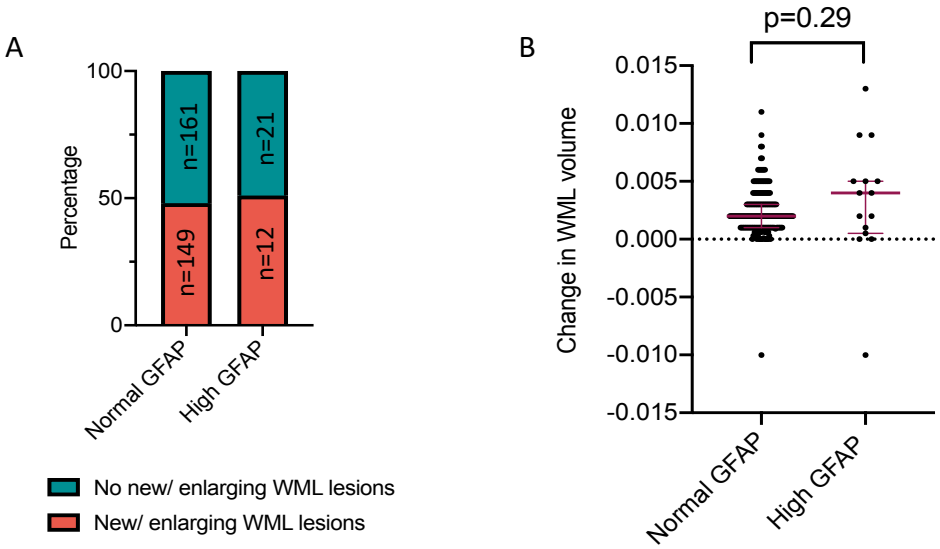


Figure 6-12 Association between baseline GFAP levels and development of new WML lesions over 12 months

(A) Baseline plasma levels of GFAP were not associated with a greater risk of new/enlarging WML at 12 months. (B) Absolute change in WML volume did not differ significantly between groups, Mann-Whitney T test.

6.3.5 Blood levels of certain brain proteins are associated with clinical metrics at baseline

We next examined the relationship between biomarker blood levels and clinical disability of the cohort at baseline and at follow-up. EDSS scores were available for 393 of 440 patients at 12 month follow-up.

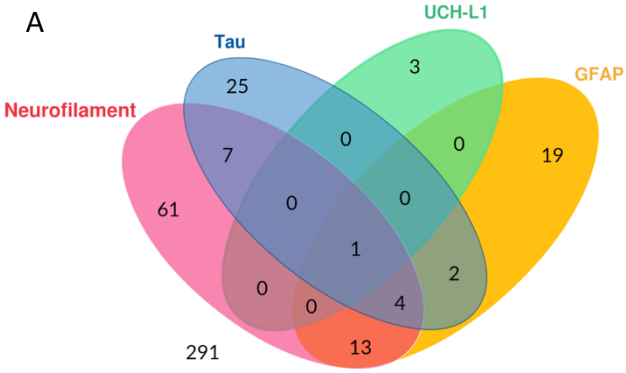
Neurofilament, GFAP and UCH-L1 all showed statistically significant positive correlations with *baseline* EDSS scores, NfL r_s=0.13, p=0.0008; GFAP r_s=0.15, p=0.003; UCH-L1 r_s=0.14, p=0.009. However none of the brain proteins were significantly associated with EDSS scores at one year.

6.3.6 Combining blood biomarkers - The Simoa brain biomarker score

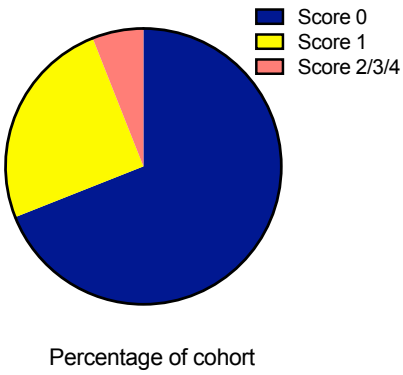
Although our results so far suggested that only neurofilament and GFAP may be of prognostic value, we investigated the combinatorial potential of measuring all four brain proteins.

Each biomarker was scored as 0 if the level was below the threshold derived from mean + 3SD of healthy controls, and 1 if the level was above the threshold (see also section 6.3.5). The scores were added together to give an overall 'Simoa brain biomarker score'. The minimum Simoa brain biomarker score was therefore 0, which equated to normal levels of all four biomarkers, and the maximum Simoa brain biomarker score was 4. Scores 2,3 and 4 were combined ('2+') for comparison given the low numbers.

A



B



C

Simoa Brain Biomarker Score	Score 0	Score 1	Score 2+
N/426*	291 68.4% of cohort	108 25.3% of cohort	27 6.3% of cohort
Median age	37 (30-45)	36 (29-45)	43 (31-50)
% female	73%	69.7%	70.4%
Months from 1st symptom to diagnosis	23 (12-65)	17 (5.2-55)	8.8 (4.1-50)
Baseline WML vol	0.006 (0.004-0.008)	0.009 (0.006-0.02)	0.016 (0.006-0.03)
Follow-up WML vol	0.007 (0.005-0.01)	0.01 (0.006-0.02)	0.008 (0.007-0.03)
New lesions over 12 months - Yes	44.9%	56.2%	70.8%
Baseline EDSS	2 (1.5-3)	2.5 (2-3)	3 (2-5.6)
Follow-up EDSS	2.5 (2-3)	2.5 (1.8-3)	3 (2-3.3)
Baseline MSFC	0.21 (-0.3-0.6)	0.09 (-0.5-0.4)	-0.36 (-1.1 - 0.1)
Follow-up MSFC	0.37 (-0.1-0.7)	0.37 (-0.2-0.7)	0.09 (-0.5-0.4)

* All four biomarker results were available for 426 of 440 individuals.

Figure 6-13 The Simoa brain biomarker score

(A) Venn diagram showing breakdown of the individuals with elevated scores in one or more of the four biomarkers. 291 individuals had a score of 0. (B) Pie chart showing breakdown for the Future MS cohort according to Simoa brain biomarker score. (C) Table showing the demographic, radiological and clinical associations with each score.

The majority of the cohort (68.4%) had a Simoa brain biomarker score of 0 out of 4. A quarter of patients had a score of 1. Of the 27 participants with a score of 2 or greater, 22 scored 2, 4 scored 3 and a single individual had a score of 4.

Age and sex were not different between individuals with different scores, *figure 6-13(C)*. However, patients with higher Simoa brain biomarker scores had shorter times from initial symptom onset to diagnosis of MS, score 0 vs score 1, p=0.03, and score 0 vs score2+, p=0.03, ANOVA Kruskal-Wallis test with Dunn's test for multiple comparisons.

A higher Simoa brain biomarker score was associated with signficantly higher white matter lesion volumes at both baseline and 12 month follow-up, *figure 6-14(A/B)*, and greater chance of developing new or enlarging T2 white matter lesions, *figure 6-14(C)*.

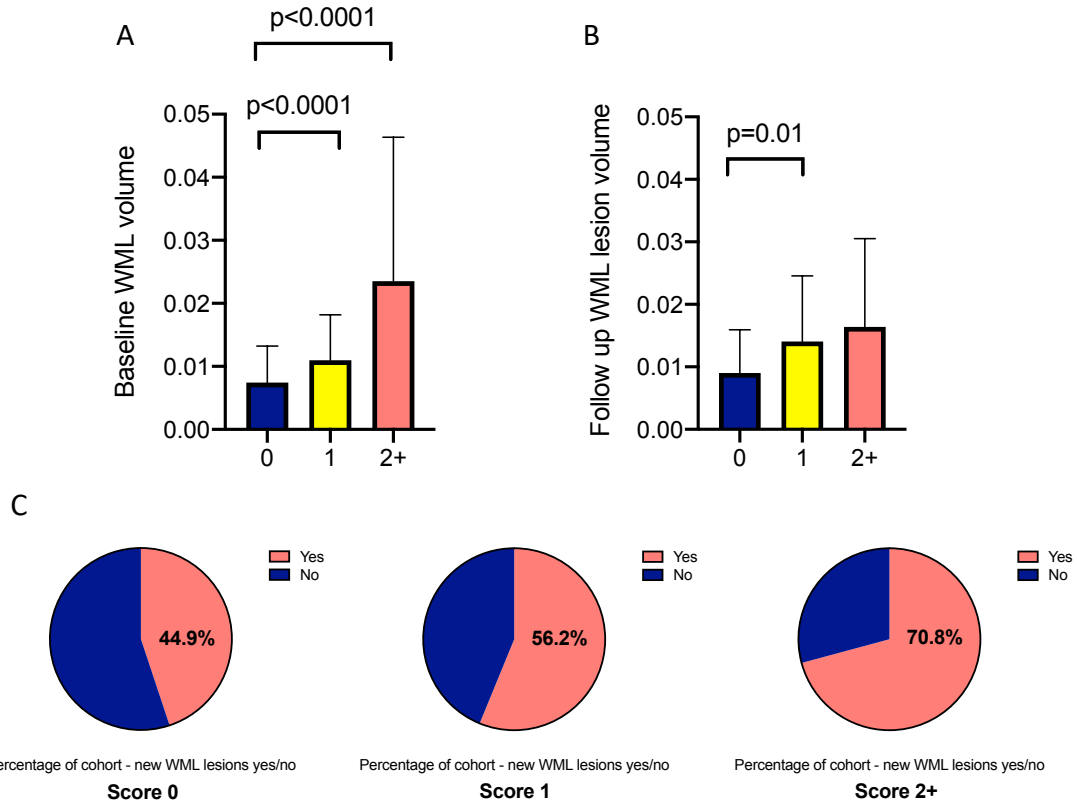


Figure 6-14 Association between Simoa brain biomarker score and radiological outcomes

At baseline (A) and 12 month follow-up (B) higher 4plex scores were associated with significantly higher WML volumes, ANOVA Kruskal-Wallis test with Dunn's test for multiple comparisons. (C) Higher 4plex scores were associated with greater risk of developing new or enlarging WML over the subsequent 12 month period, Chi squared test for trend, p=0.003.

A higher Simoa brain biomarker score was associated with greater disability (as measured by EDSS) and poorer clinical performance (as measured by MSFC) at baseline, but not at 12 month follow-up, *figures 6-13(C) and 6-15.*

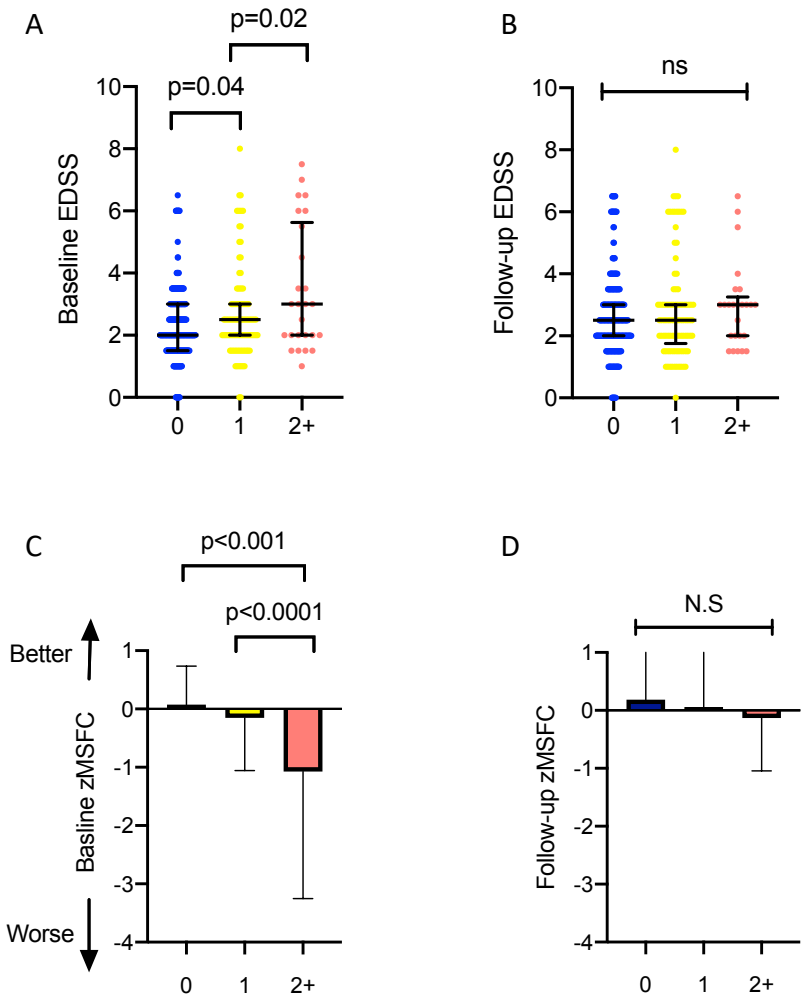


Figure 6-15 Higher 4plex scores were associated with greater disability and poorer performance

A higher Simoa brain biomarker score was associated with signficantly greater disability at baseline (A) but not at 12 month follow-up (B). A lower MSFC score represents a poorer performance across MSFC. Higher Simoa brain biomarker scores were associated with significantly lower MSFC scores at baseline (C), but not at 12 month follow-up (D). Statistical analysis using ordinary one-way ANOVA with Turkey's multiple comparisons tests.

6.4 Discussion

We have demonstrated the use of a multi-plex next generation ELISA to simultaneously quantify blood levels of four different brain proteins in a large cohort of newly diagnosed RRMS patients. This is one of the largest studies, to our knowldege, to report results using this 4plex of neurofilament, GFAP, UCH-L1 and tau in multiple sclerosis, and in a treatment naïve cohort.

Our results show that neuronal and glial proteins can be detected in the blood of both MS patients and heathy controls using Simoa. Plasma neurofilament and GFAP were increased in MS compared with healthy controls, however after the removal of outliers GFAP was no longer statistically significant. Published literature suggests higher GFAP levels occur in progressive disease states (Petzold, 2015), and are therefore unlikely to be significantly elevated in a newly diagnosed cohort. Particularly high GFAP levels have been noted in neuromyelitis optica spectrum disorder (NMOSD) (Watanabe, M *et al.,* 2019), a potential MS mimc. Misdiagnosis might explain the extremely high GFAP levels seen in our two outliers. We had planned to measure aquaporin-4 and anti-MOG antibody levels in the Future MS cohort as part of an assessment of rates of misdiagnoses, however this was delayed due to the COVID-19 pandemic and is an area for future study.

We have already demonstrated that blood levels of neurofilament correlate strongly with CSF NfL, r_s=0.69, p<0.0001. None of the additional three biomarkers showed a significant corrleation between CSF and blood levels. The lack of correlation between CSF and blood levels of GFAP, tau and UCH-L1 suggests that they do not reflect concurrent CSF concentrations *at time of sampling*. This may be the limiting factor in their use as blood biomarkers.

There are many potential reasons why a brain protein may or may not have the same dynamics in cerebrospinal fluid as in blood. Factors intrinsic to the protein should exert a constant influence on the ratio between CSF and blood concentrations, between patients. These include the molecular size of the protein; the specificity of the protein to the CNS; the speed of degradation or half-life of the protein in CSF; and the speed of degradation of the protein within the blood. Factors that vary in relation to disease activity, such as the rapidity and extent of protein release into the CSF, will however

differ between patients and within individuals during different disease states, for example relapse versus remission (Abdelhak *et al.*, 2018).

However, contrary to our results, published literature suggests that CSF and blood GFAP levels do generally correlate (Watanabe, M *et al.,* 2019; Abdelhak, A *et al*., 2018). One potential reason for the discrepancy may the assay we employed - Quanterix 4plex 'A'. When compared with other studies, our CSF GFAP levels appear lower than expected in relation to our plasma GFAP levels. Since we carried out our analysis, Quanterix have released a newer version of the same 4plex assay (Quanterix 4plex 'B'). The 4plex 'B' assay has a greater dynamic range of GFAP measurement in order to improve sensitivity at the upper range. Some of our CSF GFAP levels may therefore be falsely low due to a hook effect. Although this may negate our analysis of the correlation between CSF and blood GFAP, this does not affect the interpretation of our plasma GFAP levels.

We found that both plasma neurofilament and GFAP showed a positive correlation with age in healthy controls, but that this relationship was lost in the context of multiple sclerosis. These results are similar to previous findings (Bridel *et al.*, 2019) and suggest that the effect of age on NfL and GFAP concentrations are surpassed by disease-associated differences. A definitive study on age-appropriate levels has not yet been published and will require data from large numbers of healthy controls. This was outwith the scope of our relatively small number of healthy control subjects and was therefore a limitation in this study. We did however quantify the potential effect of age on blood neurofilament levels in our cohort of 65 healthy controls using linear regression modelling. Linear regression modelling has previously been used to quantify the effect of age on NfL levels in the CSF of healthy controls, r^2 =0.65 (Vågberg *et al.*, 2015). This statistical model requires that a linear relationship exists between the variables, which would appear to be biologically plausible given that a linear relationship has been shown between age and total intracranial grey matter volume loss (Ge *et al.*, 2002). We found that approximately a third of the variance seen in plasma NfL levels in healthy controls is due to age, R^2 =0.33, $p<0.001$, whereas the effect of age on plasma GFAP was slightly less marked, R^2=0.28, p=0.007. Together, these results suggest NfL and GFAP concentrations should be adjusted for age. This would be particularly important in older cohorts.

Our results suggest that blood neurofilament is the best biomarker in terms of reliability and reflection of concurrent CNS pathology.

Approximately 20% of the Future MS cohort had an elevated plasma neurofilament level at baseline assessment. As a biomarker of *concurrent disease activity*, neurofilament levels correlated better with radiological markers (T2 white matter lesion volume) than clinic markers (EDSS and MSFC). Plasma GFAP, elevated in approximately 10% of the cohort, also showed potential as a biomarker of *concurrent* disease activity.

We then explored the potential of these proteins as *prognostic biomarkers* by comparing baseline biomarker levels with white matter lesion volumes at 12 months, visual reads binary outcome (Yes/No) of new/enhancing WML at 12 months and clinical metrics at 12 months.

We found that individuals with a high baseline plasma neurofilament were significantly more likely to develop new/enlarging lesions than those with a normal NfL. The absolute change in WML volume over 12 months however did not differ between the cohort with normal plasma NfL and the cohort with high plasma NfL. This highlights the major limitation of using lesion volumes: that 'no change' in lesion volume over a time period may reflect a complete lack of disease activity *but could also occur if there is simultaneous resolution (involution) of lesions alongside the development of multiple new lesions*. This is of particular relevance in early stage disease, at the point of diagnosis, where lesions may be more likely to have an inflammatory component which would be expected to recede over time. None of the remaining biomarkers (GFAP, Tau or UCH-L1) showed a significant association with radiological or clinical outcomes at 12 months.

A sizeable proportion of the Future MS cohort had high baseline tau (9.2%) or GFAP (9.4%) levels. We hypothesised that the patients with the most significant disease (and thus extensive neuronal and glial cell damage) were more likely to have elevated levels of more than one glial biomarker. Based on this hypothesis, we explored the potential of combining all four biomarkers as a 'Simoa brain biomarker score'.

Individuals with higher Simoa brain biomarkers scores had a shorter time period from onset of first symptom to diagnosis, greater clinical disability and higher white matter lesion burden at baseline. This is in keeping with patients with more severe disease and more significant symptoms being diagnosed more quickly.

We also found that individuals with higher scores were at greater risk of developing new or enlarging white matter lesions over 12 months: 45% of individuals with a score of 0, 56% of individuals with a score of 1, and 71% of individuals with a score of 2 developed new/enlarging T2 WML at 12 months. This suggests that, *despite treatment*, patients with greater 'pathological burden' at the point of diagnosis are at greater risk of disease accrual that those with less evidence of neuronal or glial cell damage.

These results can be compared with our analysis of neurofilament levels alone to determine if the inclusion of additional protein biomarkers improves the ability to detect individuals at higher risk of future disease activity.

71% of individuals with a Simoa brain biomarker score of 2+ developed new/enlarging T2 WML. 71% of individuals with an elevated NfL (alone) developed new/enlarging T2 WML. These results are identical because of the 27 participants with a score of 2+, only 2 *did not* include neurofilament as part of their score (both had elevated GFAP and tau) – i.e. nearly all the individuals with a score of 2+ had an elevated neurofilament.

45% of individuals with a Simoa brain biomarker score of 0 *did not* develop new/enlarging T2 WMLs. 35% of individuals with a normal NfL (alone) *did not* develop new/enlarging T2 WMLs. The difference occurs because some individuals with elevated levels of GFAP, Tau or UCH-L1 (but not NfL) developed new/enlarging T2 WML over the 12 month follow-up period.

By incorporating the additional glial and astrocyte biomarkers we are able to improve the *negative* predictive value – i.e. determining which individuals are *less likely* to develop disease activity. The inclusion of additional protein biomarkers of disease activity may therefore improve the sensitivity in detecting individuals with milder disease who are less likely to have future disease activity. These individuals might be less likely to require higher efficacy disease modifying treatments. Knowing this, at the point of diagnosis, could aid in informing treatment decisions.

There are several limitations to our analysis. Firstly, we have not analysed 12 month data for the entire cohort. These data will be analysed once available. A further important limitation is that disease modifying treatments started between baseline and follow-up visit will have affected clinical and radiological outcomes at 12 months. Data pertaining to the specific DMTs initiated have not been finalised at time of writing and are therefore not available for analysis. Higher efficacy treatments are more likely to have an impact on both clinical and radiological measures at 12 months than less efficacious DMTs. Treatment decisions are therefore important potential confounders of our findings which require review. Finally, we have not adjusted NfL or GFAP levels for age. As previously discussed, these protein increase with age, however normative age-appropriate values are not yet available.

In summary, our findings suggest that axonal damage occurs early in the disease course, is present at the point of diagnosis and can be quantified using plasma neurofilament. These results support our findings from the previous chapter where we used the MRI aggregate g-ratio in a subset of patients to identify those at the greatest risk of axonal loss. We have demonstrated that neurofilament levels, in particular, reflect radiological measures of disease activity at baseline and that high neurofilament at baseline is associated with a greater risk of developing new/ enlarging white matter lesions over 12 months. The addition of further neuronal and glial markers into a Simoa brain biomarker score may improve the ability to identify individuals at lower risk of future disease activity, in whom highly efficacious treatments may not be required.

Longterm follow-up of the Future MS cohort will identify whether the biomarker trends found in this work continue to be relevant in the identification of patients at the greatest risk of poor prognosis.

Summary of Chapter Six

- Plasma levels of neurofilament, but not GFAP, UCH-L1 and tau, are significantly higher in MS patients at the point of diagnosis than healthy controls.

- Plasma neurofilament levels at baseline correlate with white matter lesion volumes, and higher neurofilament levels are associated with a greater risk of developing new/enlarging white matter lesions in the year following diagnosis.

Chapter Seven

Conclusions and future directions

7 Conclusions

Although multiple sclerosis is increasingly treatable, highly efficacious treatments carry serious potential risks. Current prognostic tools do not fully capture the scope of the pathology, particularly axonal loss. If irreversible disability is to be prevented, a clinical biomarker which can be employed *early in the disease* to identify individuals at greatest risk of axonal loss is required to guide treatment decisions.

To address this unmet need we applied next-generation single molecule array (Simoa) and advanced MR imaging techniques in a large, prospective, observational cohort of RRMS patients recruited from across Scotland - Future MS. We recruited 440 treatment naïve RRMS patients within an average of 2 months of diagnosis. When compared with other large cohorts, the Future MS study has achieved recruitment of a particularly *homogenous* and *early disease* cohort. This, coupled with the location of the cohort (in a country of 5 million with relatively little net migration and a national, linked health-care system) makes Future MS a powerful cohort for long-term study.

We began by evaluating the role of neurofilament as a fluid biomarker in MS through meta-analysis. Our results showed that CSF NfL reflects *acute* disease activity better than chronic disease activity. Neurofilament may therefore have a role in the detection of sub-clinical or sub-radiological disease activity. The clinical implications of this would be important for monitoring disease activity, not only in RRMS, but also for progressive MS patients (who are less likely to undergo regular MR imaging). A biomarker of sub-clinical and sub-radiological disease activity could identify individuals with suboptimal treatment response, and in fact has been proposed as a fifth element to the 'NEDA' (No Evidence of Disease Activity) score.

However, any biomarker used in the monitoring of a disease needs to be easily accessible, non-invasive and applicable at regular time intervals. CSF analysis does not meet these criteria, and as such CSF neurofilament has no real promise in this context.

Quantification of neurofilament levels in the blood of MS patients was first demonstrated in 2016 (Kuhle *et al*, 2016). Multiple studies since then (including our own) have shown that neurofilament levels in blood are reflective of CSF. As a result, over the past five years measurement in blood (using Simoa) has become an established alternative to measurement in CSF. Most studies have employed a Quanterix HD-1

fully automated Simoa analyser, however a smaller benchtop analyser is also available – Quanterix SR-1. The benchtop analyser requires manipulation of samples and reagents and is more user-intensive; but the difference in size and cost, compared with the HD-1 analyser, may make it a more feasible option for clinical laboratories. For this reason, we decided to run our samples using our Quanterix SR-1 analyser rather than send samples elsewhere for analysis.

Classical neuropathological studies have shown the axonal damage occurs in line with acute inflammation (B D Trapp *et al.*, 1998). However, histopathological studies are usually performed in individuals with chronic, end-stage or atypical, fulminant disease, and are thus not truly representitive of 'typical' newly-diagnosed MS.

Using Simoa, we demonstrated that neurofilament can be accurately quantified in blood *at the point of diagnosis* of relapsing remitting MS. This demonstrates, *in vivo,* that axonal damage is present at diagnosis and emphasises the need to control inflammatory disease activity *quickly and effectively* in order to minimise longterm axonal loss.

We went on to investigate the relationship between axonal damage and inflammation (demyelination) by combining measurement of blood neurofilament with an advanced MRI biomarker of myelin integrity - the MRI aggregate g-ratio. We demonstrated that the MRI aggregate g-ratio not only varied between individuals, but between individuals with similar structural imaging. Interestingly, this reflects what is often seen in the clinic, where patients with similar findings on routine MR imaging can vary significantly in their clinical presentation.

After finding an association between lesion volume, lesion MRI g-ratio and blood neurofilament levels, we dichotomised patients into two groups according to lesion burden. We found that individuals with a high burden of lesions *and* a high g-ratio within those lesions were significantly more likely to have elevated blood neurofilament levels when compared with those with a high burden of lesions but a normal g-ratio within lesions. Our results suggest that the extent of myelin damage within lesions varies between individuals early in the disease course, and that greater loss of myelin integrity may be associated with greater axonal damage.

We have demonstrated how simultaneous assessment of advanced imaging and liquid biomarker techniques can be used, *in vivo,* to gain insights into clinically relevant

biology of multiple sclerosis. This is, to our knowledge, one of the largest studies to date of the MRI aggregate g-ratio in MS, and the first to combine it with quantification of axonal damage.

As technologies continue to improve, we may soon be able to assess multiple different aspects of pathology, *in vivo,* at greater resolutions. Combining complimentary biomarkers could provide insights into biological heterogeneity of MS between individuals, enable identification of particular cohorts for inclusion in clinical trials and provide biomarkers for which to measure success or failure of new treatments.

We continued this theme in the final aspect of this work where we asked whether measurement of multiple brain proteins could provide further insights into disease biology or the prognosis of the Future MS cohort.

After the removal of outliers, neurofilament levels (but not GFAP, UCH-L1 or tau) were significantly higher in MS patients than healthy controls. Neurofilament levels, in particular, reflected radiological measures of disease activity at baseline and were associated with risk of radiological disease activity over the subsequent 12 months.

We did not find an association between baseline levels of GFAP, UCH-L1 or tau and radiological or clinical outcomes at 12 months. However, combination of all four biomarkers as a 'Simoa brain biomarker score' improved the *negative* predictive value over neurofilament alone by increasing the ability to identify individuals who *did not* have radiological evidence of disease activity at 12 months.

Although identification of individuals at the greatest risk of future disability is the aim of the Future MS study, identification of individuals with 'milder' disease has clinical relevance also. Such individuals may achieve remission with less efficacious disease modifying therapies and might therefore avoid the potential associated risks associated with certain DMTs.

In summary, based on our work, it is our belief that blood neurofilament levels should be incorporated into the routine clinical care of patients with multiple sclerosis. However, at present, there remain limitations which first need addressed.

Measuring blood neurofilament levels on a three-monthly basis may reveal disease activity which occurs asymptomatically, which resolves between annual MRIs, or in individuals with a very high plaque burden in whom the identification of a single new lesion can be difficult. Routine monitoring of neurofilament levels would therefore inform treatment decisions by improving the sensitivity to detect disease activity. This is increasingly relevant because of the number and scope of disease modifying treatments now available. Neurofilament is not a replacement for annual MR imaging. Neurofilament is measuring *acute/subacute disease activity* whereas MRI best demonstrates *disease accrual*. In addition, neurofilament could not replace the role of MRI in safety monitoring for individuals on DMTs, for example, in detection of PML.

At present, there are several limitations to the implementation of neurofilament into *routine* clinical practice. The first is that age-appropriate normative values are yet to be published. These are needed for individual results to be interpreted accurately. Work on this is already underway in a large Swiss cohort of healthy controls and will hopefully be published in the near future.

The second limiting factor is that the potential influence of comorbidities on the measurement of this blood biomarker is unknown. We already know that comorbidities of the central or peripheral nervous system, such as stroke or neuropathy, can elevate neurofilament levels; however other comorbidities (such as renal impairment) may impact measurement in blood. To address this would require enormous studies of healthy and disease controls and is therefore probably best answered with 'real-world' data.

Further limiting factors to the widespread introduction of this blood-based biomarker are the cost of the Simoa assay and the equipment required to analyse it. At present, a single Simoa plate analyses 32 patient samples (in duplicate) at a cost of several thousand pounds. Cost efficient, high throughput systems would be required for implementation in clinical settings where the care of several thousand MS patients could require tens of thousands of neurofilament tests each year.

Future directions

Options enabling greater personalisation of multiplex assays are likely to become increasingly available, and with this, the opportunity to study different aspects of MS disease pathology in greater depth. Although axonal damage has been the focus of this work, an important potential future direction of study is myelin biology, and in particular, combining fluid biomarker analysis of myelin damage with advanced imaging modalities.

For example, an elevated MR aggregate g-ratio does not tell us whether an axon is undergoing demyelination or remyelination. However, by combing the aggregate g-ratio with quantification of myelin proteins (such as MOG, MAG, MBP and PLP) in the blood, myelin destruction could be analysed and quantified *in vivo*, longitudinally measured, and linked with clinical outcomes. This might produce new insights into an individual's potential for remyelination. This in turn could improve understanding of the mechanisms involved in remyelination, inform the discovery of potential treatments, and identify cohorts for clinical trials.

Long-term follow-up of the Future MS cohort is planned, and we will link the findings presented in this work with outcomes at 5 years and beyond. We hope that this study will aid in improving the management of MS patients, with the aim of reducing, halting, or one day even reversing the disability associated with multiple sclerosis.

References

Abdelhak, A. *et al.* (2018) 'Serum GFAP as a biomarker for disease severity in multiple sclerosis', *Scientific Reports*, 8(1), pp. 1–7. doi: 10.1038/s41598-018-33158-8.

Aeinehband, S. *et al.* (2015) 'Complement component C3 and butyrylcholinesterase activity are associated with neurodegeneration and clinical disability in multiple sclerosis.', *PloS one*. Edited by I. Silman, 10(4), p. e0122048. doi: 10.1371/journal.pone.0122048.

Al-Chalabi, A. and Miller, C. (2003) 'Neurofilaments and neurological disease', *BioEssays*, 24, pp. 346–355.

Arrambide, G. *et al.* (2018) 'The value of oligoclonal bands in the multiple sclerosis diagnostic criteria', *Brain*, 141(4), pp. 1075–1084. doi: 10.1093/brain/awy006.

Avasarala, J. R., Cross, A. H. and Trotter, J. L. (2001) 'Oligoclonal band number as a marker for prognosis in multiple sclerosis', *JAMA Neurology*, 58(12), pp. 2044–2045. doi: 10.1001/archneur.58.12.2044.

Axelsson, M. *et al.* (2011) 'Glial fibrillary acidic protein: A potential biomarker for progression in multiple sclerosis', *Journal of Neurology*, 258(5), pp. 882–888. doi: 10.1007/s00415-010-5863-2.

Axelsson, M. *et al.* (2014) 'Immunosuppressive therapy reduces axonal damage in progressive multiple sclerosis.', *Multiple sclerosis (Houndmills, Basingstoke, England)*, 20(1), pp. 43–50. doi: 10.1177/1352458513490544.

Ayrignac, X. *et al.* (2020) 'Serum GFAP in multiple sclerosis: correlation with disease type and MRI markers of disease severity', *Scientific Reports*. Nature Publishing Group UK, 10(1), pp. 1–5. doi: 10.1038/s41598-020-67934-2.

Barkhof, F. (2002) 'The Clinico-Radiological Paradox in Multiple Sclerosis Revisited', *Current opinions Neurology*, 15(3), pp. 239–245.

Bjartmar, C. *et al.* (2000) 'Neurological disability correlates with spinal cord axonal loss and reduced N-acetyl aspartate in chronic multiple sclerosis patients', *Annals of Neurology*, 48(6), pp. 893–901. doi: 10.1002/1531-8249(200012)48:6<893::AID-ANA10>3.0.CO;2-B.

Bjartmar, C. and Trapp, B. D. (2001) 'Axonal and neuronal degeneration in multiple sclerosis: Mechanisms and functional consequences', *Current Opinion in Neurology*, 14(3), pp. 271–278. doi: 10.1097/00019052-200106000-00003.

Bloomgren, G. *et al.* (2012) 'Risk of natalizumab-associated progressive multifocal leukoencephalopathy', *New England Journal of Medicine*, 366(20), pp. 1870–1880. doi: 10.1056/NEJMoa1107829.

Bove, R. *et al.* (2018) 'SUMMIT:Creating a repository of deeply phenotyped contemporary multiple sclerosis cohorts', *Multiple Sclerosis*, 24(12), pp. 1485–1498. doi: 10.1016/j.physbeh.2017.03.040.

Breij, E. C. W. *et al.* (2008) ‘Homogeneity of active demyelinating lesions in established multiple sclerosis’, *Annals of Neurology*, 63(1), pp. 16–25. doi: 10.1002/ana.21311.

Brex, P. A. *et al.* (2002) ‘A longitudinal study of abnormalities on MRI and disability from multiple sclerosis’, *New England Journal of Medicine*, 346(3), pp. 158–164. doi: 10.1056/NEJMoa011341.

Bridel, C. *et al.* (2019) ‘Diagnostic Value of Cerebrospinal Fluid Neurofilament Light Protein in Neurology: A Systematic Review and Meta-analysis’, *JAMA Neurology*, 76(9), pp. 1035–1048. doi: 10.1001/jamaneurol.2019.1534.

Brønnum-Hansen, H., Koch-Henriksen, N. and Stenager, E. (2011) ‘The Danish multiple sclerosis registry’, *Scandinavian Journal of Public Health*, 39(7), pp. 62–64. doi: 10.1177/1403494810390729.

Brownlee, W. J. *et al.* (2019) ‘Early imaging predictors of long-term outcomes in relapse-onset multiple sclerosis’, *Brain*, 142(8), pp. 2276–2287. doi: 10.1093/brain/awz156.

Burman, J. *et al.* (2014) ‘Assessing tissue damage in multiple sclerosis: a biomarker approach’, *Acta Neurologica Scandinavica*, 130(2), pp. 81–89. doi: 10.1111/ane.12239.

Cai, L. and Huang, J. (2018) 'Neurofilament light chain as a biological marker for multiple sclerosis: a meta-analysis study', *Neuropsychiatric Disease and Treatment,* 14, pp.2241-2254.

Campbell, G. R., Worrall, J. T. and Mahad, D. J. (2014) ‘The central role of mitochondria in axonal degeneration in multiple sclerosis’, *Multiple Sclerosis Journal*, 20(14), pp. 1806–1813. doi: 10.1177/1352458514544537.

Campbell, J. S. W. *et al.* (2018) ‘Promise and pitfalls of g-ratio estimation with MRI’, *NeuroImage*, 182, pp. 80–96. doi: 10.1016/j.neuroimage.2017.08.038.

Cantó, E. *et al.* (2019) ‘Association between Serum Neurofilament Light Chain Levels and Long-term Disease Course among Patients with Multiple Sclerosis Followed up for 12 Years’, *JAMA Neurology*, 76(11), pp. 1359–1366. doi: 10.1001/jamaneurol.2019.2137.

Cerqueira, J. J. *et al.* (2018) ‘Time matters in multiple sclerosis: Can early treatment and long-term follow-up ensure everyone benefits from the latest advances in multiple sclerosis?’, *Journal of Neurology, Neurosurgery and Psychiatry*, 89(8), pp. 844–850. doi: 10.1136/jnnp-2017-317509.

Coles, A. J. *et al.* (2008) ‘Alemtuzumab vs. Interferon Beta-1a in Early Multiple Sclerosis’, *N Engl J Med*, 359, pp. 1786–1801.

Compston, A. and Coles, A. (2008) ‘Multiple sclerosis’, *The Lancet*, pp. 1502–1517. doi: 10.1016/S0140-6736(08)61620-7.

Confavreux, C. and Vukusic, S. (2006) ‘Age at disability milestones in multiple sclerosis’, *Brain*, 129(3), pp. 595–605. doi: 10.1093/brain/awh714.

Confavreux, C., Vukusic, S. and Adeleine, P. (2003) 'Early clinical predictors and progression of irreversible disability in multiple sclerosis: An amnesic process', *Brain*, 126(4), pp. 770–782. doi: 10.1093/brain/awg081.

Cree, B. A. C. *et al.* (2016) 'Long-term evolution of multiple sclerosis disability in the treatment era', *Annals of Neurology*, 80(4), pp. 499–510. doi: 10.1002/ana.24747.

Degenhardt, A. *et al.* (2009) 'Clinical prognostic factors in multiple sclerosis: A natural history review', *Nature Reviews Neurology*. Nature Publishing Group, 5(12), pp. 672–682. doi: 10.1038/nrneurol.2009.178.

Disanto, G. *et al.* (2017) 'Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis', *Annals of Neurology*, 81(6), pp. 857–870. doi: 10.1002/ana.24954.

Duval , T. *et al* (2017) 'g-ratio weighted imaging of the human spinal cord in vivo.', *Neuroimage,* 15(145), pp. 11-23.

Dziedzic, T. *et al.* (2010) 'Wallerian degeneration: A major component of early axonal pathology in multiple sclerosis', *Brain Pathology*, 20(5), pp. 976–985. doi: 10.1111/j.1750-3639.2010.00401.x.

Ebers, G. C. and PRISMS (Prevention of Relapses and Disability by Interferon-1a Subcutaneously in Multiple Sclerosis) Study Group (1998) 'Randomised double-blind placebo-controlled study of interferon 1a in relapsing / remitting multiple sclerosis', *Lancet*, 352, pp. 1498–504.

Eng, L. F., Ghirnikar, R. S. and Lee, Y. L. (2000) 'Glial Fibrillary Acidic Protein : GFAP-Thirty-One Years (1969 – 2000)*', *Neurochemical research*, 25, pp. 1439–1451.

Felts, P. A., Baker, T. A. and Smith, K. J. (1997) 'Conduction in segmentally demyelinated mammalian central axons', *Journal of Neuroscience*, 17(19), pp. 7267–7277. doi: 10.1523/jneurosci.17-19-07267.1997.

Ferguson, B. *et al.* (1997) 'Axonal damage in acute multiple sclerosis lesions', *Brain*, 120(3), pp. 393–399. doi: 10.1093/brain/120.3.393.

Fischer, J. S. *et al.* (1999) 'The multiple sclerosis functional composite measure (MSFC): An integrated approach to MS clinical outcome assessment', *Multiple Sclerosis*, 5(4), pp. 244–250. doi: 10.1177/135245859900500409.

Fisher, E. *et al.* (2000) 'Relationship between brain atrophy and disability: An 8-year follow-up study of multiple sclerosis patients', *Multiple Sclerosis*. doi: 10.1191/135245800701566331.

Fisniku, L. K. *et al.* (2008) 'Disability and T2 MRI lesions: A 20-year follow-up of patients with relapse onset of multiple sclerosis', *Brain*, 131(3), pp. 808–817. doi: 10.1093/brain/awm329.

Frischer, J. M. *et al.* (2015) 'Clinical and Pathological Insights into the Dynamic Nature of the White Matter Multiple Sclerosis Plaque', *Annals of Neurology*, 78(5), pp. 710–721.

G.A. Schumacher, G. Beebe, R.F. Kibler, L.T. Kurland, J.F. Kurtzke, F. McDowell, et al. (1965) 'Problems of experimental trials of therapy in multiple sclerosis: report by the panel on the evaluation of experimental trials of therapy in multiple sclerosis', *Ann N Y Acad Sci*, 122, pp. 552–568.

Gauthier, S. A. *et al.* (2006) 'A model for the comprehensive investigation of a chronic autoimmune disease: The multiple sclerosis CLIMB study', *Autoimmunity Reviews*, 5(8), pp. 532–536. doi: 10.1016/j.autrev.2006.02.012.

Ge, Y. *et al.* (2002) 'Age-related total gray matter and white matter changes in normal adult brain. Part I: Volumetric MR imaging analysis', *American Journal of Neuroradiology*, 23(8), pp. 1327–1333.

Geurts, J. J. G. *et al.* (2005) 'Cortical lesions in multiple sclerosis: Combined postmortem MR imaging and histopathology', *American Journal of Neuroradiology*, 26(3), pp. 572–577.

Gledhill, R. F., Harrison, B. M. and McDonald, W. I. (1973) 'Pattern of Remyelination', *Nature*, 244(1972), pp. 1972–1973.

Goldstein, M. E., Sternberger, N. H. and Sternberger, L. A. (1987) 'Phosphorylation protects neurofilaments against proteolysis', *Journal of Neuroimmunology*, 14(2), pp. 149–160. doi: 10.1016/0165-5728(87)90049-X.

Gulani, V. *et al.* (2017) 'Gadolinium deposition in the brain: summary of evidence and recommendations', *The Lancet Neurology*. Elsevier Ltd, 16(7), pp. 564–570. doi: 10.1016/S1474-4422(17)30158-8.

Gunnarsson, M. *et al.* (2011) 'Axonal damage in relapsing multiple sclerosis is markedly reduced by natalizumab.', *Annals of neurology*, 69(1), pp. 83–9. doi: 10.1002/ana.22247.

Hagiwara, A. *et al.* (2017) 'Analysis of white matter damage in patients with multiple sclerosis via a novel in Vivo MR method for measuring myelin, axons, and G-ratio', *American Journal of Neuroradiology*, 38(10), pp. 1934–1940. doi: 10.3174/ajnr.A5312.

Hakansson, I. *et al.* (2017) 'Neurofilament light chain in cerebrospinal fluid and prediction of disease activity in clinically isolated syndrome and relapsing???remitting multiple sclerosis', *European Journal of Neurology*, 24(5), pp. 703–712. doi: 10.1111/ene.13274.

Hauser, S. L. *et al.* (2017) 'Ocrelizumab versus interferon beta-1a in relapsing multiple sclerosis', *New England Journal of Medicine*, 376(3), pp. 221–234. doi: 10.1056/NEJMoa1601277.

Hein (née Maier), K. *et al.* (2008) 'Biological markers for axonal degeneration in CSF and blood of patients with the first event indicative for multiple sclerosis', *Neuroscience Letters*, 436(1), pp. 72–76. doi: 10.1016/j.neulet.2008.02.064.

Högel, H. *et al.* (2020) 'Serum glial fibrillary acidic protein correlates with multiple sclerosis disease severity', *Multiple Sclerosis Journal*, 26(2), pp. 210–219. doi: 10.1177/1352458518819380.

ISD Data dictionary, CHI number (no date). Available at: http://www.ndcdev.scot.nhs.uk/Dictionary-A-Z/Definitions/index.asp?Search=C&ID=128&Title=CHI Number (Accessed: 1 July 2020).

Jiménez-Jiḿnez, F. J. *et al.* (2005) 'Tau protein concentrations in cerebrospinal fluid of patients with amyotrophic lateral sclerosis', *Acta Neurologica Scandinavica*, 111(2), pp. 114–117. doi: 10.1111/j.1600-0404.2005.00370.x.

Jones, J. L. and Coles, A. J. (2010) 'New treatment strategies in multiple sclerosis', *Experimental Neurology*. Elsevier Inc., 225(1), pp. 34–39. doi: 10.1016/j.expneurol.2010.06.003.

Kalincik, T. *et al.* (2013) 'Sex as a determinant of relapse incidence and progressive course of multiple sclerosis', *Brain*, 136(12), pp. 3609–3617. doi: 10.1093/brain/awt281.

Kamagata, K. *et al.* (2019) 'MR g-ratio-weighted connectome analysis in patients with multiple sclerosis', *Scientific Reports*, 9(1), pp. 1–13. doi: 10.1038/s41598-019-50025-2.

Kassubek, R. (2017) 'GFAP in early multiple sclerosis: A biomarker for inflammation', *Neuroscience letters.* [East Park, Ireland] :, pp. 166–170. doi: 10.1016/j.neulet.2017.07.050.

Kearns, P. K. A. *et al.* (2019) 'Regional variation in the incidence rate and sex ratio of multiple sclerosis in Scotland 2010–2017: findings from the Scottish Multiple Sclerosis Register', *Journal of Neurology*. doi: 10.1007/s00415-019-09413-x.

Kobelt, G. *et al.* (2017) 'New insights into the burden and costs of multiple sclerosis in Europe', *Multiple Sclerosis Journal*, 23(8), pp. 1123–1136. doi: 10.1177/https.

Kornek, B. and Lassmann, H. (1999) 'Axonal pathology in multiple sclerosis. A historical note', *Brain Pathology*, 9(4), pp. 651–656. doi: 10.1111/j.1750-3639.1999.tb00547.x.

Kuhle, J. *et al.* (2016) 'Serum neurofilament light chain in early relapsing remitting MS is increased and correlates with CSF levels and with MRI measures of disease severity', *Multiple Sclerosis*, 22(12), pp. 1550–1559. doi: 10.1177/1352458515623365.

Kuhle, J., Kropshofer, H., Haering, D. A., *et al.* (2019) 'Blood neurofilament light chain as a biomarker of MS disease activity and treatment response', *Neurology*, 92(10), pp. E1007–E1015. doi: 10.1212/WNL.0000000000007032.

Kuhle, J., Kropshofer, H., Meinert, R., *et al.* (2019) 'Kuhle ECTRIMS 2019 poster GFAP in EXPAND.pdf', p. poster 1630.

Kuhlmann, T. *et al.* (2002) 'Acute axonal damage in multiple sclerosis is most extensive in early disease stages and decreases over time', *Brain*, 125(10), pp. 2202–2212. doi: 10.1093/brain/awf235.

Kuhlmann, T. *et al.* (2017) 'An updated histological classification system for multiple sclerosis lesions', *Acta Neuropathologica*. Springer Berlin Heidelberg, 133(1), pp. 13–24. doi: 10.1007/s00401-016-1653-y.

Kurtzke, J. F. (1983) 'Rating neurologic impairment in multiple sclerosis', *Neurology*, 33(11), pp. 1444 LP – 1444. doi: 10.1212/WNL.33.11.1444.

Kutzelnigg, A. *et al.* (2005) 'Cortical demyelination and diffuse white matter injury in multiple sclerosis', *Brain*, 128(11), pp. 2705–2712. doi: 10.1093/brain/awh641.

Langer-Gould, A. *et al.* (2006) 'Clinical and demographic predictors of long-term disability in patients with relapsing-remitting multiple sclerosis: A systematic review', *Archives of Neurology*, 63(12), pp. 1686–1691. doi: 10.1001/archneur.63.12.1686.

Lassmann, H. (2019) 'Pathogenic mechanisms associated with different clinical courses of multiple sclerosis', *Frontiers in Immunology*, 10(JAN), pp. 1–14. doi: 10.3389/fimmu.2018.03116.

Lebrun-Frenay, C. *et al.* (2019) 'Radiologically isolated syndrome: a 10-year follow-up study to identify factors predicting a clinical event', *ECTRIMS 2019 Scientific session presentation*.

Lee, M. K. and Cleveland, D. W. (1996) 'Neuronal Intermediate Filaments', *Annual Review of Neuroscience*, 19, pp. 187–217. doi: 10.1016/B978-0-12-378630-2.00436-9.

Lublin, F. D. *et al.* (2014) 'Defining the clinical course of multiple sclerosis: The 2013 revisions', *Neurology*, 83(3), pp. 278–286. doi: 10.1212/WNL.0000000000000560.

Lucchinetti, C. *et al.* (2000) 'Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination', *Annals of Neurology*, 47(6), pp. 707–717. doi: 10.1002/1531-8249(200006)47:6<707::AID-ANA3>3.0.CO;2-Q.

Lucchinetti, C. F. *et al.* (2011) 'Inflammatory Cortical Demyelination in Early Multiple Sclerosis Claudia', *New England Journal of Medicine*, 365(23), pp. 2133–2197. doi: 10.1038/jid.2014.371.

Luchetti, S. *et al.* (2018) 'Progressive multiple sclerosis patients show substantial lesion activity that correlates with clinical disease severity and sex: a retrospective autopsy cohort analysis', *Acta Neuropathologica*. Springer Berlin Heidelberg, 135(4), pp. 511–528. doi: 10.1007/s00401-018-1818-y.

Lycke, J. N. *et al.* (1998) 'Neurofilament protein in cerebrospinal fluid: a potential marker of activity in multiple sclerosis.', *Journal of neurology, neurosurgery, and psychiatry*, 64(3), pp. 402–4.

Mackenzie, I. S. *et al.* (2014) 'Incidence and prevalence of multiple sclerosis in the UK 1990-2010: a descriptive study in the General Practice Research Database.', *Journal of neurology, neurosurgery, and psychiatry*, 85(1), pp. 76–84. doi: 10.1136/jnnp-2013-305450.

Malmeström, C. *et al.* (2003) 'Neurofilament light protein and glial fibrillary acidic protein as biological markers in MS.', *Neurology*, 61(12), pp. 1720–5.

Martínez, M. A. M. *et al.* (2015) 'Glial and neuronal markers in cerebrospinal fluid predict progression in multiple sclerosis', *Multiple Sclerosis Journal*, 21(5), pp. 550–561. doi: 10.1177/1352458514549397.

McCrone, P. *et al.* (2008) 'Multiple sclerosis in the UK: Service use, costs, quality of life and disability', *PharmacoEconomics*, 26(10), pp. 847–860. doi: 10.2165/00019053-200826100-00005.
McDonald, W. I. (1986) 'The mystery of the origin of multiple sclerosis', *Journal of Neurology Neurosurgery and Psychiatry*, 49(2), pp. 113–123. doi: 10.1136/jnnp.49.2.113.

McDonald, W. I. *et al.* (2001) 'Recommended diagnostic criteria for multiple sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple Sclerosis', *Annals of Neurology*. doi: 10.1002/ana.1032.

Metz, I. *et al.* (2014) 'Pathologic Heterogeneity Persists in Early Active Multiple Sclerosis Lesions', *Annals of Neurology*, 75(5), pp. 728–738. doi: 10.1002/ana.24163.Pathologic.

Miller, D. H. (1994) 'Magnetic resonance in monitoring the treatment of multiple sclerosis', *Annals of Neurology*, 36, pp. S91–S94. doi: 10.1007/bf01997712.

Minneboo, A. *et al.* (2004) 'Infratentorial Lesions Predict Long-term Disability in Patients with Initial Findings Suggestive of Multiple Sclerosis', *Archives of Neurology*, 61(2), pp. 217–221. doi: 10.1001/archneur.61.2.217.

Mistry, N. *et al.* (2011) 'Focal multiple sclerosis lesions abound in normal appearing white matter', *Multiple Sclerosis Journal*, 17(11), pp. 1313–1323. doi: 10.1177/1352458511415305.

Mohammadi, S. *et al.* (2015) 'Whole-brain in-vivo measurements of the axonal G-ratio in a group of 37 healthy volunteers', *Frontiers in Neuroscience*, 9(NOV), pp. 1–13. doi: 10.3389/fnins.2015.00441.

Mondello, S. *et al.* (2011) 'Neuronal and glial markers are differently associated with computed tomography findings and outcome in patients with severe traumatic brain injury: A case control study', *Critical Care*, 15(3). doi: 10.1186/cc10286.

Murphy, A. and Gaillard, F. (2020) 'MRI sequences (overview)'.
National Institute for Health Research PROSPERO (2017). Available at: https://www.crd.york.ac.uk/prospero/.

National Institutes of Health Magnetic Resonance Imaging (2020) 'Magnetic Resonance Imaging (MRI)', *National Institute for Biomedical imaging and Bioengineering*. Available at: https://www.nibib.nih.gov/science-education/science-topics/magnetic-resonance-imaging-mri.

National Records of Scotland (2018) 'Scottish consensus results 1991-2011'. Available at: https://www.scotlandscensus.gov.uk/ethnicity-identity-language-and-religion.
Norgren, N., Rosengren, L. and Stigbrand, T. (2003) 'Elevated neurofilament levels in neurological diseases', *Brain Research*, 987(1), pp. 25–31. doi: 10.1016/S0006-8993(03)03219-0.

Norgren, N. *et al* (2004) 'Neurofilament and glial fibrillary acidic protein in multiple sclerosis', *Neurology,* 63(9), pp. 1586-1590.

Novakova, L. *et al.* (2016) 'Cerebrospinal fluid biomarkers of inflammation and degeneration as measures of fingolimod efficacy in multiple sclerosis', *Multiple Sclerosis Journal*, pp. 1–10. doi: 10.1177/1352458516639384.

Novakova, L., Axelsson, M., *et al.* (2017) 'Cerebrospinal fluid biomarkers as a measure of disease activity and treatment efficacy in relapsing-remitting multiple sclerosis', *Journal of Neurochemistry*, 141(2), pp. 296–304. doi: 10.1111/jnc.13881.

Novakova, L., Zetterberg, H., *et al.* (2017) 'Monitoring disease activity in multiple sclerosis using serum neurofilament light protein', *Neurology*, 89(22), pp. 2230–2237. doi: 10.1212/WNL.0000000000004683.

O'Connor, K. C. *et al.* (2006) 'Comprehensive phenotyping in multiple sclerosis: Discovery based proteomics and the current understanding of putative biomarkers', *Disease Markers*, 22(4), pp. 213–225. doi: 10.1155/2006/670439.

Ontaneda, D. *et al.* (2019) 'Early highly effective versus escalation treatment approaches in relapsing multiple sclerosis', *The Lancet Neurology*, 18(10), pp. 973–980.

Paty, D. W. and Li, D. K. (1993) 'Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. The IFNB Multiple Sclerosis Study Group.', *Neurology*. United States, 43(4), pp. 655–661. doi: 10.1212/wnl.43.4.655.

Petzold, A. (2005) 'Neurofilament phosphoforms: Surrogate markers for axonal injury, degeneration and loss', in *Journal of the Neurological Sciences*. doi: 10.1016/j.jns.2005.03.015.

Petzold, A. (2015) 'Glial fibrillary acidic protein is a body fluid biomarker for glial pathology in human disease', *Brain Research*. Elsevier, 1600, pp. 17–31. doi: 10.1016/j.brainres.2014.12.027.

Podbielska, M. *et al.* (2013) 'Myelin recovery in multiple sclerosis: The challenge of remyelination', *Brain Sciences*, 3(3), pp. 1282–1324. doi: 10.3390/brainsci3031282.

Polman, C. H. *et al.* (2005) 'Diagnostic criteria for multiple sclerosis: 2005 Revisions to the "McDonald Criteria"', *Annals of Neurology*, pp. 840–846. doi: 10.1002/ana.20703.

Polman, C. H. *et al.* (2011) 'Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria', *Annals of Neurology*, 69(2), pp. 292–302. doi: 10.1002/ana.22366.

Popescu, B. F. G. and Lucchinetti, C. F. (2016) 'Neuropathology of multiple sclerosis', in *Multiple Sclerosis: A Mechanistic View*, pp. 181–200.

Poser, C. M. *et al.* (1983) 'New diagnostic criteria for multiple sclerosis: Guidelines for research protocols', *Annals of Neurology*. doi: 10.1002/ana.410130302.

Public Health Scotland (2020) 'Scottish Multiple Sclerosis Register National Report 2020', *Scottish MS Register annual report*. Available at: https://www.msr.scot.nhs.uk/Reports/Dashboard-2020.html.

Quanterix (2018) 'Neurology 4-Plex A (Tau , NF-Light , GFAP , UCH-L1) A Multiple Sclerosis Study', *Quanterix applicatio note*, pp. 1–3.

Raghavan, K. *et al.* (2015) 'Progression rates and sample size estimates for PPMS based on the CLIMB study population', *Multiple Sclerosis Journal*, 21(2), pp. 180–188. doi: 10.1177/1352458514541976.

Ramanujam, R. *et al.* (2015) 'Effect of smoking cessation on multiple sclerosis prognosis', *JAMA Neurology*, 72(10), pp. 1117–1123. doi: 10.1001/jamaneurol.2015.1788.

Raphael, I. *et al.* (2014) 'Body fluid biomarkers in multiple sclerosis: How far we have come and how they could affect the clinic now and in the future', *Expert Review of Clinical Immunology*, 11(1), pp. 69–91. doi: 10.1586/1744666X.2015.991315.

Reich, D. S., Lucchinetti, C. F. and Calabresi, P. A. (2018) 'Multiple sclerosis', *N Engl J Med*, 378, pp. 169–180. doi: 10.1016/j.physbeh.2017.03.040.

Rissin, D. M. *et al.* (2010) 'Single-Molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations', *Nature Biotechnology*, 28(6), pp. 595–599.

Rissin, D. M. *et al.* (2013) 'Multiplexed single molecule immunoassays', *Lab on a Chip*, 13(15), pp. 2902–2911. doi: 10.1039/c3lc50416f.

Robinson, A. P. *et al.* (2014) 'The EAE model of MS: utility for understanding disease pathophysiology and treatment', *Hand Clin Neurol*, 122, pp. 173–189. doi: 10.1016/B978-0-444-52001-2.00008-X.The.

Rosengren, L. E. *et al.* (1996) 'Patients with Amyotrophic Lateral Sclerosis and Other Neurodegenerative Diseases Have Increased Levels of Neurofilament Protein in CSF', *Journal of Neurochemistry*, 67(5), pp. 2013–8. doi: 10.1046/j.1471-4159.1996.67052013.x.

Rouleau, G. *et al.* (1996) 'SODl Mutation Is Associated with Accumulation of Neurofilaments in Amyotrophic Lateral Sclerosis', *Annals of neurology*, 39, pp. 128–131.

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

Sawcer, S. *et al.* (2011) 'Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis', *Nature*, 476(7359), pp. 214–219. doi: 10.1038/nature10251.

Scalfari, A. *et al.* (2010) 'The natural history of multiple sclerosis, a geographically based study 10: Relapses and long-term disability', *Brain*, 133(7), pp. 1914–1929. doi: 10.1093/brain/awq118.

Scalfari, A. *et al.* (2011) 'Age and disability accumulation in multiple sclerosis', *Neurology*, 77(13), pp. 1246 LP – 1252. doi: 10.1212/WNL.0b013e318230a17d.
Schmierer, K. *et al.* (2004) 'Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain', *Annals of Neurology*, 56(3), pp. 407–415. doi: 10.1002/ana.20202.

Schreiber, S. *et al.* (2018) 'Reader response: Serum neurofilament light is sensitive to active cerebral small vessel disease', *Neurology*, 90(24), p. 1126. doi: 10.1212/WNL.0000000000005672.

Scolding, N. *et al.* (2015) 'Association of British Neurologists: Revised (2015) guidelines for prescribing disease-modifying treatments in multiple sclerosis', *Practical Neurology*, 15(4), pp. 273–279. doi: 10.1136/practneurol-2015-001139.

Scottish Medicines Consortium (2020) *NHS SMC - Ocrelizumab*. Available at: https://www.scottishmedicines.org.uk/medicines-advice/ocrelizumab-ocrevus-full-smc2223/.

Scottish MS Register (2019) 'Scottish Multiple Sclerosis Register', *National Health Services National report*, (September). Available at: http://www.msr.scot.nhs.uk/Reports/Main.html.

Serafini, B. *et al.* (2004) 'Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis', *Brain Pathology*, 14(2), pp. 164–174. doi: 10.1111/j.1750-3639.2004.tb00049.x.

Simpson, S. *et al.* (2011) 'Latitude is significantly associated with the prevalence of multiple sclerosis: A meta-analysis', *Journal of Neurology, Neurosurgery and Psychiatry*, 82(10), pp. 1132–1141. doi: 10.1136/jnnp.2011.240432.

Sinnecker, T. *et al.* (2019) 'Evaluation of the Central Vein Sign as a Diagnostic Imaging Biomarker in Multiple Sclerosis', *JAMA Neurology*, 76(12), pp. 1446–1456. doi: 10.1001/jamaneurol.2019.2478.

Soleimani, B., Murray, K. and Hunt, D. (2019) 'Established and Emerging Immunological Complications of Biological Therapeutics in Multiple Sclerosis', *Drug Safety*. Springer International Publishing, 42(8), pp. 941–956. doi: 10.1007/s40264-019-00799-1.

St-Amant, M. and O'Gorman, P. (2020) *Magnetisation Transfer*. Available at: https://radiopaedia.org/articles/magnetisation-transfer-1?lang=gb (Accessed: 29 June 2020).

Stang, A. (2010) 'Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses', *European Journal of Epidemiology*, 25(9), pp. 603–605. doi: 10.1007/s10654-010-9491-z.
Steenhof, M. *et al.* (2019) 'Familial multiple sclerosis patients have a shorter delay in diagnosis than sporadic cases', *Multiple Sclerosis and Related Disorders*. Elsevier B.V., 32(October 2018), pp. 97–102. doi: 10.1016/j.msard.2019.04.012.

De Stefano, N. and Arnold, D. L. (2015) ‘Towards a better understanding of pseudoatrophy in the brain of multiple sclerosis patients’, *Multiple Sclerosis*, 21(6), pp. 675–676. doi: 10.1177/1352458514564494.

De Stefano, N., Battaglini, M. and Smith, S. M. (2007) ‘Measuring brain atrophy in multiple sclerosis’, *Journal of Neuroimaging*. doi: 10.1111/j.1552-6569.2007.00130.x.
Stikov, N. *et al.* (2015) ‘In vivo histology of the myelin g-ratio with magnetic resonance imaging’, *NeuroImage*. Elsevier Inc., 118, pp. 397–405. doi: 10.1016/j.neuroimage.2015.05.023.

Stikov, N. *et al.* (2015) 'In vivo histology of the myelin g-ratio with magetic resonance imaging.', *Neuroimage,* 118, pp. 397-405.

Stilund, M. *et al.* (2015) ‘Biomarkers of inflammation and axonal degeneration/damage in patients with newly diagnosed multiple sclerosis: contributions of the soluble CD163 CSF/serum ratio to a biomarker panel.’, *PloS one*. Edited by M. Reindl, 10(4), p. e0119681. doi: 10.1371/journal.pone.0119681.

Swanton, J. K. *et al.* (2007) ‘MRI criteria for multiple sclerosis in patients presenting with clinically isolated syndromes: a multicentre retrospective study’, *Lancet Neurol*, 6(8), pp. 677–686. doi: S1474-4422(07)70176-X [pii]\r10.1016/S1474-4422(07)70176.

Tallantyre, E. C. *et al*. (2010) 'Clinicopathological evidence that axonal loss underlies disability in progressive multiple sclerosis.' *Multiple Sclerosis Journal* 16.4, pp 406-411.

Teunissen, C. *et al.* (2013) ‘Consensus definitions and application guidelines for control groups in cerebrospinal fluid biomarker studies in multiple sclerosis.’, *Multiple Sclerosis Journal*, 19(13), pp. 1802–9. doi: 10.1177/1352458513488232.

Teunissen, C. E. *et al.* (2009) ‘A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking’, *Neurology*. doi: 10.1212/WNL.0b013e3181c47cc2.

Thompson, A. J. *et al.* (2018) ‘Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria’, *The Lancet Neurology*. doi: 10.1016/S1474-4422(17)30470-2.

Thompson, E. J. and Freedman, M. S. (2006) ‘Cerebrospinal fluid analysis in the diagnosis of multiple sclerosis.’, *Advances in neurology*, 98(June 2005), pp. 147–160.

Tintore, M. *et al.* (2010) ‘Brainstem lesions in clinically isolated syndromes.’, *Neurology*. United States, 75(21), pp. 1933–1938. doi: 10.1212/WNL.0b013e3181feb26f.

Tintore, M. *et al.* (2015) ‘Defining high, medium and low impact prognostic factors for developing multiple sclerosis’, *Brain*, 138(7), pp. 1863–1874. doi: 10.1093/brain/awv105.

Trapp, Bruce D. *et al.* (1998) ‘Axonal transection in the lesions of multiple sclerosis’, *New England Journal of Medicine*. doi: 10.1056/NEJM199801293380502.
Trapp, B D *et al.* (1998) ‘Axonal transection in the lesions of multiple sclerosis [see comments]’, *N Engl J Med*.

Trapp, B. D. and Nave, K.-A. (2008) 'Multiple Sclerosis: An Immune or Neurodegenerative Disorder?', *Annual Review of Neuroscience*, 31(1), pp. 247–269. doi: 10.1146/annurev.neuro.30.051606.094313.

Trentini, A. *et al.* (2014) 'N-Acetylaspartate and neurofilaments as biomarkers of axonal damage in patients with progressive forms of multiple sclerosis', *Journal of Neurology*, 261(12), pp. 2338–2343. doi: 10.1007/s00415-014-7507-4.

Vågberg, M. *et al.* (2015) 'Levels and age dependency of neurofilament light and Glial Fibrillary Acidic Protein in healthy individuals and their relation to the brain parenchymal fraction', *PLoS ONE*, 10(8), pp. 1–8. doi: 10.1371/journal.pone.0135886.

Villar, L. *et al.* (2008) 'Influence of oligoclonal IgM specificity in multiple', *Multiple Sclerosis*, 14(February 2007), pp. 183–187.

Visser, E. M. *et al.* (2012) 'A new prevalence study of multiple sclerosis in Orkney, Shetland and Aberdeen city', *Journal of Neurology, Neurosurgery & Psychiatry*, 83(7), pp. 719–724. doi: 10.1136/jnnp-2011-301546.

Wallin, M. T. *et al.* (2019) 'Global, regional, and national burden of multiple sclerosis 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016', *The Lancet Neurology*, 18(3), pp. 269–285. doi: 10.1016/S1474-4422(18)30443-5.

Watanabe, M. *et al.* (2019) 'Serum GFAP and neurofilament light as biomarkers of disease activity and disability in NMOSD', *Neurology,* 93(13), pp. e1299-e1311. doi: 10.1212/WNL.0000000000008160

Wattjes, M. P. *et al.* (2015) 'Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis - Establishing disease prognosis and monitoring patients', *Nature Reviews Neurology*. Nature Publishing Group, 11(10), pp. 597–606. doi: 10.1038/nrneurol.2015.157.

Weinshenker, B. G. *et al.* (1989) 'The natural history of multiple sclerosis: A geographically based study. 8. Familial multiple sclerosis', *Brain*, 112, pp. 133–146. doi: 10.1093/brain/123.3.641.

Wijnands, J. M. A. *et al.* (2017) 'Health-care use before a first demyelinating event suggestive of a multiple sclerosis prodrome: a matched cohort study', *The Lancet Neurology*, 16(6), pp. 445–451. doi: 10.1016/S1474-4422(17)30076-5.

Yu, F. *et al.* (2019) 'Imaging G-ratio in multiple sclerosis using high-gradient diffusion MRI and macromolecular tissue volume', *American Journal of Neuroradiology*, 40(11), pp. 1871–1877. doi: 10.3174/ajnr.A6283.

Yuan, A. *et al.* (2012) 'Neurofilaments at a glance', *Journal of Cell Science*, pp. 3257–3263. doi: 10.1242/jcs.104729.

Zajicek, J. P. *et al.* (2010) 'Patient-orientated longitudinal study of multiple sclerosis in south west England (The South West Impact of Multiple Sclerosis Project, SWIMS) 1: Protocol and baseline characteristics of cohort', *BMC Neurology*. BioMed Central Ltd, 10(1), p. 88. doi: 10.1186/1471-2377-10-88.

Appendices

Appendix 1

Meta-analysis search strategy

Search Term	*Search Number*	*Number of Hits (08/09/2017)*
'MS' OR 'Multiple sclerosis'	1	Pubmed – 303,063
		Web of Science – 307,800
		Cochrane Library – 14,503
		OpenGrey database– 3,044
'**Neurite Orientation Dispersion and Density Imaging** NfL' OR 'Neurofilament light'	2	Pubmed – 1,456
		Web of Science – 1,673
		Cochrane Library - 84
		OpenGrey database- 1
'CSF' OR 'Cerebrospinal fluid'	3	Pubmed – 132,444
		Web of Science – 60,522
		Cochrane Library – 7,220
		OpenGrey database- 337
1 AND 2 AND 3	4	**Pubmed - 83**
		Web of Science - 124
		Cochrane Library - 12
		OpenGrey database - 0
Total Number of Hits		**= 219**

Appendix 2

Meta-analysis inclusion and exclusion criteria

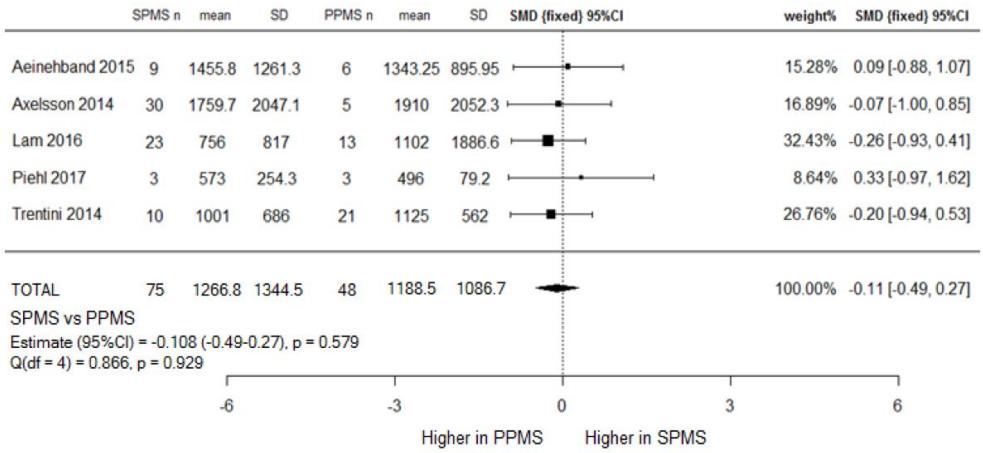
Inclusion Criteria	*Exclusion Criteria*
Diagnosis of MS meeting established diagnostic criteria	Clinically or radiologically isolated syndrome
Case control studies (retrospective or prospective) producing original work.	Animal studies, reviews or responses, or manuscripts unrelated to research topic.
Neurofilament light measured in CSF in quantitative manner.	NfL levels detectable in less than 85% of either MS or control cohort.
CSF biobanking referencing established guidelines, or if not, ELISA protocol described.	Coefficient of Variation >25%
Validated assay or, if not, description of ELISA technique and lower limit of detection	NfL data given as relative units or not absolute values.

Reasons for exclusion on abstract and on full text

	Reason for exclusion	*Number*
Excluded on Abstract	Not Multiple Sclerosis	47
	Summary or Review article	23
	Basic science study/Animal (not clinical)	14
	NfL AB or other measured	9
	Assay analytical analysis or validation	4
	Only abstract available	3
	Total	**100**
Excluded on Full text	No control group	17
	Assay sensitivity <85% for MS/ control group	7
	NfL cohort data published elsewhere	3
	Semi-quantitative data	3
	NfL measured only in serum	1
	Total	**31**
Excluded as required dataset unavailable		**6**
Raw data or dataset available for inclusion in meta-analysis		**14**

Appendix 3

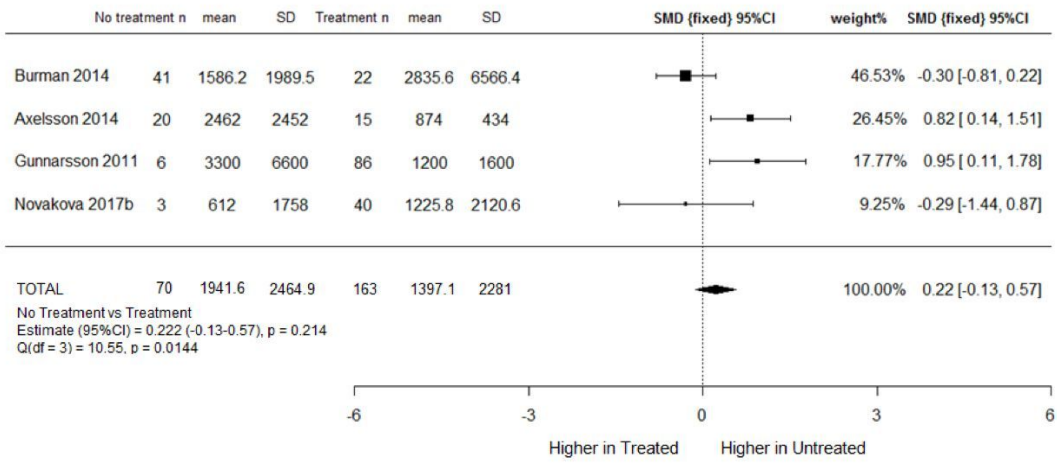
Meta-analysis of CSF neurofilament levels primary versus secondary MS



CSF NfL levels do not differ between primary progressive and secondary progressive MS

Appendix 4

Meta-analysis of CSF neurofilament levels in treated and untreated MS



CSF NfL levels do not differ between treated and untreated MS patients. The majority of disease modifying treatments used were first line therapies.