Improving longitudinal spinal cord atrophy measurements for clinical trials in multiple sclerosis by using the generalised boundary shift integral (GBSI)

Marcello Moccia, MD PhD

Queen Square MS Centre, NMR Research Unit

Department of Neuroinflammation

UCL Queen Square Institute of Neurology

Faculty of Brain Sciences

University College London, London, United Kingdom

Thesis submitted to University College London for the degree of Doctor of Medicine Research.

December 2019

Declaration

I, Marcello Moccia, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated.

Acknowledgements

I am extremely grateful to my supervisors. To Olga Ciccarelli and Frederik Barkhof, for their warm guidance and support, through the ups and downs of this journey. To Ferran Prados, for demonstrating that every problem can be solved, if you get the right code. And, last but not least, to Alan Thompson, for ensuring I stayed focused on my goals.

The support I received from the team at the Queen Square MS Centre is invaluable. Without their expertise, patience and friendship, this thesis would not have been possible, evenings in London would have been much more boring, and I would not consider, unexpectedly, London as a second home.

This research received financial support from the MAGNIMS and the UK MS Society, which I am grateful for. In particular, I would like to thank the people working at MAGNIMS Centres around Europe for sharing data for this thesis, and their thoughtful comments on this work.

Finally, I would like to thank my family and friends, who have continuously supported (and visited) me during my time in London; Claudio and Susy, for my cat moments at home; and Raffaele, my source of encouragement and endless patience.

Abstract

Spinal cord atrophy is a common and clinically relevant feature of multiple sclerosis (MS), and can be used to monitor disease progression and as an outcome measure in clinical trials. Spinal cord atrophy is conventionally estimated with segmentation-based methods (e.g., cross-sectional spinal cord area (CSA)), where spinal cord change is calculated indirectly by numerical difference between timepoints.

In this thesis, I validated the generalised boundary shift integral (GBSI), as the first registration-based method for longitudinal spinal cord atrophy measurement. The GBSI registers the baseline and follow-up spinal cord scans in a common half-way space, to directly determine atrophy on the cord edges.

First, on a test dataset (9 MS patients and 9 controls), I have found that GBSI presented with lower random measurement error, than CSA, reflected by lower standard deviation, coefficient of variation and median absolute deviation.

Then, on multi-centre, multi-manufacturer, and multi–field-strength scans (282 MS patients and 82 controls), I confirmed that GBSI provided lower measurement variability in all MS subtypes and controls, than CSA, resulting into better separation between MS patients and controls, improved statistical power, and reduced sample size estimates.

Finally, on a phase 2 clinical trial (220 primary-progressive MS patients), I demonstrated that spinal cord atrophy measurements on GBSI could be obtained from brain scans, considering their quality and association with corresponding spinal cord MRI-derived measurements. Not least, 1-year spinal

cord atrophy measurements on GBSI, but not CSA, were associated with upper and lower limb motor function.

In conclusion, spinal cord atrophy on the GBSI had higher measurement precision and stronger clinical correlates, than the segmentation method, and could be derived from high-quality brain acquisitions. Longitudinal spinal cord atrophy on GBSI could become a gold standard for clinical trials including spinal cord atrophy as an outcome measure, but should remain a secondary outcome measure, until further advancements increase the ease of acquisition and processing.

Impact statement

In this thesis, I have validated the generalised boundary shift integral (GBSI), as the first registration-based method for quantification of spinal cord atrophy in multiple sclerosis (MS). The GBSI pipeline is based on the latest iteration of the boundary shift integral, and has been specifically designed to overcome limitations of commonly used segmentation-based methods (i.e., spinal cord cross-sectional area (CSA)).

Improvements in spinal cord atrophy measurements presented in this thesis can expand research possibilities for future MS projects. In particular, I analysed a multicentre, multi-manufacturer and multi-field strength scan dataset from the Magnetic Resonance Imaging in MS (MAGNIMS) network, and showed that the GBSI has better measurement precision, leading to smaller sample size estimates and stronger clinical correlates, than CSA, suggesting the registration-based method holds promise for future collaborative studies. These results were then confirmed by re-analysing a phase 2 clinical trial, highlighting that the GBSI could become a gold standard for clinical trials including spinal cord atrophy as an outcome measure. I also demonstrated that obtaining spinal cord atrophy measurements from brain scans could represent a viable and clinically meaningful alternative to more technically-challenging spinal cord images, in particular in multi-centre settings where homogenous spinal cord acquisitions are not feasible. For instance, following the results of this thesis, the United Kingdom MS Society Efficient Clinical Trial Platform will be able to include spinal cord atrophy measurements with GBSI among MRI outcome measures to study the neuroprotective potential of different medications. Not least, GBSI could prove useful in many neurological diseases, such as amyotrophic lateral sclerosis and spinal cord injury, where spinal cord volume changes are representative of the most aggressive aspects of the diseases.

Last but not least, results of this thesis are also important to MS patients. Deriving spinal cord atrophy measurements from brain scans would significantly reduce the scan time for MRI and, thus, participants' burden. In addition, spinal cord atrophy on GBSI, but not CSA, can detect early subtle changes in motor function to the upper and lower limbs, as measured by both neurologists (on neurological examination) and MS patients (on patient reported outcome measures). In the future, identifying changes in spinal cord volume could improve monitoring the clinical course of MS and its response to treatment.

Table of contents

List of abbreviations

3D: three-dimensional

3D-FFE: three-dimensional gradient-echo fast field echo

9HPT: nine-hole peg test

95%CI: 95% confidence intervals

Ab: antibody

AQP4: aquaporin-4

ASM: active surface model

AUC: areas-under-the-curve

BBB: blood brain barrier

BOLD: blood oxygenation level-dependent

BPF: brain parenchymal fraction

BSLs: bright spotty lesions

BSI: boundary shift integral

CIS: clinically isolated syndrome

Coeff: coefficients

CNS: central nervous system

CSA: cross-sectional area

CSF: cerebro-spinal fluid

CUA: combined unique active

DIR: double inversion recovery

DIS: dissemination in space

DIT: dissemination in time

DMTs: disease-modifying treatments

DOF: degrees of freedom

DTI: diffusion tensor imaging

EBV: Epstein-Barr virus

EDSS: expanded disability status scale

FA: fractional anisotropy

fMRI: functional magnetic resonance imaging

FSS: Functional System Score

GBSI: generalised boundary shift integral

Glx: glutamate/glutamine

GM: grey matter

ICC: intraclass correlation coefficient

IL: interleukin

IVF: isotropic volume fraction

LETM: longitudinally extensive transverse myelitis

MAGNIMS: magnetic resonance imaging in multiple sclerosis

MD: mean diffusivity

MOG: myelin oligodendrocyte glycoprotein

MPRAGE: magnetisation-prepared rapid gradient echo

MRI: magnetic resonance imaging

MRS: magnetic resonance spectroscopy

MS: multiple sclerosis

MSWS: multiple sclerosis walking scale

MTR: magnetisation transfer ratio

MTSat: magnetisation transfer saturation

MUCCA: mean upper cervical cord area

NAA: N-acetyl-aspartate

NAA/Cr: N-acetyl-aspartate/ creatinine

NDI: neurite density index

NMOSD: neuromyelitis optica spectrum disorders

NODDI: neurite orientation dispersion and density imaging

OCT: optical coherence tomography

ODI: orientation dispersion index

OR: odds ratios

PBVC: percent brain volume change

PD: proton density

PET: Positron emission tomography

PP: primary progressive

PSIR: phase-sensitive inversion recovery

qMT: quantitative magnetisation transfer

QSI: Q-space imaging

RD: radial diffusivity

RIS: radiologically isolated syndrome

RNFL: retinal nerve fibre layer

ROC: receiver operating characteristics

RRMS: relapsing-remitting multiple sclerosis

SCT: spinal cord toolbox

SMT: spherical mean technique

SP: secondary progressive

SIENA: structural image evaluation using normalisation of atrophy

STIR: short-tau inversion recovery

T25FWT: timed 25-foot walking test

UCCA: upper cord cross-sectional area

WM: white matter

List of tables

List of figures

Figure 6.6. Spinal cord atrophy rates by type of scan and analysis technique.

Figure 6.7. Spinal cord atrophy rates in patients with all measurements and in high-quality scans.

Figure 6.8. Spinal cord atrophy longitudinal changes from brain and spinal cord MRI.

Publications associated with this thesis

Moccia M, Valsecchi N, Ciccarelli O, Van Schijndel R, Barkhof F, Prados F. Spinal cord atrophy in a primary progressive multiple sclerosis trial: improved sample size using GBSI. *Submitted*.

Prados F, Moccia M, Johnson A, Yiannakas M, Grussu F, Cardoso MJ, Ciccarelli O, Ourselin S, Barkhof F, Wheeler-Kingshott C. Generalised boundary shift integral for longitudinal assessment of spinal cord atrophy. *Neuroimage*. 2020;209:116489.

Moccia M, Prados F, Filippi M, Rocca MA, Valsasina P, Brownlee WJ, Zecca C, Gallo A, Rovira A, Gass A, Palace J, Lukas C, Vrenken H, Ourselin S, Gandini Wheeler-Kingshott CAM, Ciccarelli O, Barkhof F, on the behalf of the MAGNIMS Study Group. Longitudinal spinal cord atrophy in multiple sclerosis using the generalized boundary shift integral. *Annals of Neurology*. 2019;86(5):704-713.

Moccia M, Ruggieri S, Ianniello A, Toosy A, Pozzilli C, Ciccarelli O. Advances in spinal cord imaging in multiple sclerosis. *Therapeutic Advances in Neurological Disorders*. 2019 22;12:1756286419840593.

Moccia M, de Stefano N, Barkhof F. Imaging outcome measures for progressive multiple sclerosis trials. *Multiple Sclerosis Journal*. 2017;23(12):1614-1626.

1. Introduction to multiple sclerosis

This thesis concerns a novel method for spinal cord atrophy measurement in multiple sclerosis (MS). In the Chapter 1, I have provided a general background to MS. After briefly mentioning epidemiology and aetiopathogenesis, I will review current MS classification, clinical assessments, diagnostic criteria and prognosis. Finally, I will discuss available treatments, highlighting limitations and future perspectives for development of new drugs.

1.1. Introduction

MS is a chronic inflammatory demyelinating and neurodegenerative disease of the central nervous system (CNS), potentially causing any neurological deficit [1, 2]. Previous studies have implicated a combination of genetic and environmental factors in the aetio-pathogenesis of MS, with chronic inflammation and neurodegeneration mediated by the patient's immune system [3, 4].

1.1.1. Epidemiology

According to the most recent Global Burden of Disease Study estimates (2016), MS is the most common immune-mediated disease of the CNS, with over 2.2 million cases world-wide, corresponding to 10% increased prevalence from 1990 [5]. North America, Western Europe and Australasia hold higher prevalence rate (91-164 cases per 100,000), compared with Africa (2-3 cases per 100,000) [5]. On the contrary, incidence of MS has been relatively stable or slightly increased over the past four to five decades [5]. As such, the rising prevalence mostly reflects improved survival, with a global mortality rate for MS being decreased by 11% between 1990 and 2016 [5].

Clinical onset is generally in early adult life, though there is increased awareness of presentation in childhood [1]. Prevalence of MS is similar in preteen boys and girls, but, then, progressively increases through lifetime among women, with a 2:1 sex ratio in favour of women in the sixth decade of life [5]. As discussed above, the life expectancy for a person with MS is relatively unimpeded by the disease, with a 5 to 10-year reduction versus non-affected individuals [6]. However, MS is one of the leading causes of disability from CNS disease among young adults, and has a severe impact on quality of life, made further strenuous by a relatively young average age of onset [1, 5].

1.1.2. Aetiology

1.1.2.1. Genetic risk

The possibility of a genetic background to MS was originally explored in family studies, which show up to 33% recurrence rate of MS within affected families. In particular, first-degree relatives have a ten-fold increased risk of MS, whilst second-degree family members have a three-fold increased risk [1, 2, 7, 8]. Sibling and twin studies display similar patterns of increased risk [9, 10].

The first allele to be accurately associated with risk of MS development was the human leukocyte antigen class II haplotype HLA-DRB*1501 [1, 2, 11, 12]. More recently, the development of genome-wide association study technology has allowed the identification of more than 200 genetic loci implicated in the risk of MS. These loci include genes encoding for the immune system (such as interleukin (IL)-2 and IL-7 receptors), cytokine pathways, and co-stimulatory molecules [11, 13, 14]. However, genetics alone cannot explain MS aetiology fully.

1.1.2.1. Environmental risk

Multiple environmental factors have been implicated in MS susceptibility. Prevalence and incidence rates of MS have been shown to increase with latitude [2, 15]. Individuals who move to high-risk latitudes during childhood (e.g., from Africa to Northern Europe) have the MS risk profile of their new high-risk area, rather than the risk levels seen in their area of birth, suggesting that environmental risk factors for MS play a role during the developmental process [16]. The association between latitude and MS risk could be at least in part mediated by vitamin D. Indeed, there is an increased risk of MS in relation to early-life deficiency of vitamin D, whose production is primarily enhanced by sunlight exposure and, thus, is generally reduced in high latitude areas [2, 17]. Still, the exact effect of vitamin D supplementation on healthy individuals (or on patients with MS) is currently unknown [13].

A number of infectious agents have been associated with MS susceptibility. Among them, the Epstein-Barr virus (EBV) has the strongest correlates. Individuals who contract EBV during childhood have a 15-fold increased risk of MS, which further soars to 30-fold for those who are infected during or after adolescence [1, 2, 17]. On the contrary, individuals who are seronegative to EBV have almost zero risk of developing MS. However, the exact relationship between EBV and MS pathogenesis remains unclear [18, 19]. Other herpetic infections have also been studied (e.g., Human Herpes virus 6, Cytomegalovirus, Herpes simplex virus), but with less definite results [1, 2, 20].

Other risk factors include tobacco smoking, reproductive factors (e.g., breastfeeding, menopause, contraceptives), dietary consumption of polyunsaturated fatty acid, obesity, and socioeconomic status, but the exact role of these factors in MS pathogenesis remains poorly understood [1, 2, 20].

Overall, multiple environmental interactions with predisposing genotype could be considered as the initial trigger to MS development [2].

1.1.3. Pathology

The most typical pathology sign of MS is the presence of demyelinating lesions in the white matter (WM) and grey matter (GM), in the brain and in the spinal cord [21–25]. Demyelinating lesions generally originate around venules, where accumulation of inflammatory lymphocytes can be observed [26, 27], along with astrocytic response and macrophages/microglia infiltrates to the active injury site, eventually resulting into gliotic scars [28]. Acute inflammatory demyelination is clinically associated with the acute onset of new neurological symptoms (i.e., clinical relapse) [29–31].

Neuro-axonal loss is another prominent hallmark of MS, and is a key factor of irreversible disability accrual [32]. In the early stages of the disease, axonal loss is generally seen in areas of pathological demyelination, in association with inflammatory infiltrates consisting of macrophages/microglia and lymphocytes [33]. During the course of the disease, axonal loss can occur in areas of prolonged demyelination without active inflammation, suggesting that axonal survival is related to the presence of myelin support [34–38]. Notwithstanding this, the presence of chronically demyelinated axons suggests that demyelination does not necessarily leads to neuro-axonal loss [35, 39, 40]. In advanced MS, axonal loss results into shrinking of the brain parenchyma (i.e., atrophy), and is associated with impaired function of macrophages/microglia and astrocytes, and with increased oxidative stress and mitochondrial damage. In particular, demyelination and subsequently impaired axonal dysfunction increase the energy demand, further contributing to altered metabolism, neuronal dysfunction and, ultimately, axonal loss [41–43]

The mechanisms by which demyelination and axonal loss occur are profoundly heterogeneous and involve a variety of cellular subsets [1, 44, 45]. The sequence of pathological events might include perivenular infiltration of macrophages, CD8+ T lymphocytes, and, CD4+ T lymphocytes and B lymphocytes, with profound blood brain barrier (BBB) leakage, giving rise to classical active demyelinated plaques [45–50]. As the disease progresses, infiltrates of T and B lymphocytes, activation of microglia/macrophages and astrocytes, and mitochondrial dysfunction become obvious throughout the brain parenchyma, also in the absence of major BBB damage [51, 52], with formation of aggregates of inflammatory cells in the form of meningeal follicle-like structures, and expansion of previously-existing WM and GM lesions [42, 43, 45, 46, 53–57]. These changes ultimately lead to progressive demyelination, axonal loss and neurodegeneration in the brain and the spinal cord [42, 43, 45, 54, 57].

Inflammation, demyelination and axonal loss can be measured *in vivo* by using magnetic resonance imaging (MRI) [58]. This will be further discussed in Chapter 2.

1.1.4. Clinical features

At the onset of the disease, the clinical course in most patients is characterised by recurring episodes of neurological deficits (i.e., relapses). A relapse is defined as the acute-subacute occurrence of neurological symptoms lasting at least 24 hours, in absence of infections, fever or other symptoms of systemic disease [59]. On the first relapse, the condition is named clinically isolated syndrome (CIS) or relapsing-remitting MS (RRMS), depending on whether MS pathology is isolated or disseminated in time (DIT) and space (DIS), as further detailed in paragraph 1.1.5 [59–61]. On the occasion of a second clinical relapse, the condition inevitably falls within RRMS [59–61]. In the early

phases of MS, patients generally recover from relapses fully, in absence of any treatment, but, later on, disability from relapses can accumulate, resulting in irreversible neurological disability. Following this initial relapsing-remitting course, patients can present with progressive disability accrual, independently from relapses, namely secondary progressive MS (SPMS) [60, 62]. For other patients, gradual disability worsening can occur from the start of the disease, in this case termed primary progressive MS (PPMS) [60]. In recent years, PPMS and SPMS have been reclassified as different parts of the same progressive spectrum (**Table 1.1**) [63, 64]. Of note, patients diagnosed with the progressive form of the disease may also experience periods of symptom remission (and sometimes improvement), within an overall declining trajectory, and, still, can present with inflammatory activity (e.g., relapses or new lesions). As such, progressive MS can be further classified into four categories: active or not active (based on the presence or absence of relapses or new lesions), and with or without symptom progression (based on disability accrual independently from relapses) (**Table 1.1**) [60].

Initial symptoms can consist of nearly any neurological sign, based on the location of the inflammatory demyelinating lesions. In RRMS, most common symptoms at onset are sensorimotor disturbances, visual defects or a combination of multiple symptoms. In PPMS, symptoms at onset generally consist of sensorimotor disturbances, bladder dysfunction, unbalance or a combination of multiple symptoms [65]

Clinical features of MS can be measured with a wide variety of clinical assessments, tests and scales. Functional systems scores (FSS) and the extended disability status scale (EDSS) are the most widely used clinical measures [66]. The FSS measure disability level in various neurological systems on an ordinal scale; the full FSS battery examines disability related to the pyramidal, cerebellar, brainstem,

Table 1.1. Classification of MS clinical phenotypes.

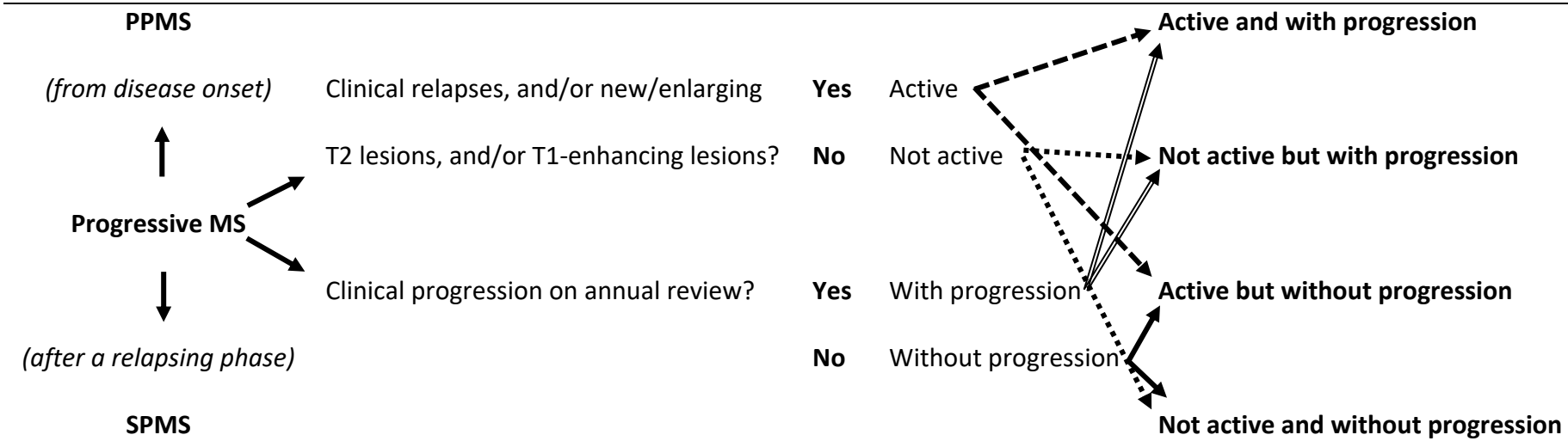
Table shows clinical phenotype classification for relapsing and progressive patients. Adapted from Lublin FD et al. Neurology 2014 [60].

Phenotype classification for relapsing patients

CIS →	Clinical relapses, and/or new/enlarging	**Yes**	Active	→	**Active CIS***
	T2 lesions, and/or T1-enhancing lesions?	**No**	Not active	→	**Not active CIS**
RRMS →	Clinical relapses, and/or new/enlarging	**Yes**	Active	→	**Active RRMS**
	T2 lesions, and/or T1-enhancing lesions?	**No**	Not active	→	**Not active RRMS**

*Active CIS classifies as RRMS

Phenotype classification for progressive patients



sensory, bowel/bladder, visual, and cerebral systems. Based on these measurements, an overall disability level can be scored with the EDSS, which uses an ordinal scale of increasing severity from 0 (absence of neurological symptoms and signs) to 10 (death related to MS), with 0.5 unit increments [66]. EDSS steps 1.0 to 3.5 indicate unrestricted walking, whilst EDSS steps 4.0 to 7.5 indicate progressive reduction in walking ability [66]; as such, most of the EDSS is highly driven by ambulatory function. Overall, the EDSS is still the internationally accepted primary endpoint in clinical trials, but, considering its documented weaknesses, it is generally combined with other more specific clinical measures [63, 67].

The timed 25-foot walking test (T25FWT) is an objective assessment of ambulatory impairment in MS, wherein patients are measured by the average time they take to walk 25 feet (7.62m), at maximum speed, on two different trials, with a walking aid if necessary [63]. Ambulatory function can also be assessed by MS patients within the so-called patient-reported outcome measures (PROMs) [63]. In particular, the MS walking scale (MSWS) is a 12-item questionnaire showing good correlates with lower limb motor impairment and functional independence [68, 69].

The 9-hole peg test (9HPT) assesses disability of the upper limbs, by measuring the average time for placing nine pegs into nine holes, with dominant and non-dominant hand, in two different trials [63].

More recently, there has been growing attention to cognitive symptoms of MS [70–74]. Among neuropsychological tests, the symbol digit modalities test (SDMT) has shown the best correlates with overall cognitive functioning [75–77]. The SDMT measures information processing speed

through the number of correct digits (1-9) corresponding to a symbol, the patient is able to write or to orally report during a 90-second trial [77].

1.1.5. Diagnosis

Diagnosis of MS requires that clinical and radiological signs of MS are disseminated in both time and space (DIT and DIS), as originally described in the 2001 McDonald criteria and subsequently revised to achieve an earlier and more accurate diagnosis [59, 61, 78–80]. MS diagnosis is currently based on a combination of clinical features (relapses and clinical progression), MRI findings (T2 and T1-Gadolinum enhancing lesions in periventricular, cortical/juxtacortical, infratentorial and/or spinal cord areas), and laboratory measures (oligoclonal bands in the cerebrospinal fluid (CSF)) [61]. In the present thesis, I included populations where diagnosis was performed in accordance with 2010 revision of McDonald criteria, since 2017 revision was presented and published while this work was already ongoing. Possible implications are mainly related to the diagnosis of CIS and will be further discussed in Chapter 5. For comparison purposes, 2010 and 2017 revisions of McDonald criteria are presented in **Table 1.2**.

1.1.6. Prognosis

The prognosis of MS is unpredictable at patient level. However, there are a number of clinical, MRI and laboratory measures that have been shown to be associated with MS prognosis [1, 65].

Age is a main determinant of MS progression across all disease phenotypes [1]. In natural history cohort studies, PPMS and SPMS patients present with similar age at onset of progression, suggesting that disease progression occurs under the dominant influence of age-related pathological

processes, irrespectively of previous relapses [6, 81]. Other demographic factors include female sex, which is associated with higher risk of relapses [1, 82, 83], and genes [84, 85].

At the time of diagnosis, topographic characteristics (non-optic neuritis presentations) and residual disability of relapses are poor prognostic factors [83, 86, 87]. The presence and the number of CSF un-matched oligoclonal bands is another factor associated with poor prognosis and, as such, has been also included in the most recent revision of MS diagnostic criteria [61, 83, 88]. Brain and spinal cord MRI predictors of MS progression will be fully detailed in Chapter 2.

1.1.7. Treatment

No cure is available for MS, and the current therapeutic strategy is aimed at reducing the risk of relapses and disability progression [89]. The rapid advent of disease-modifying treatments (DMTs) in the mid-1990s was heralded by a period of rapid progress in the understanding of MS [90]. In particular, with the support of MRI, new possible treatments have been quickly screened on well selected populations [91], as discussed in Chapter 2.

DMTs can prevent relapses, new brain and spinal cord lesions, and, at least in part, worsening neurological disability [92]. Historically, DMTs have been immunosuppressant or immunomodulatory, meaning that continuous treatment is required to maintain suppression of inflammation (and disease activity). Some more recent DMTs can be given as short courses, with the aim of producing enduring immunological actions (immune reconstitution therapies) [93]. Generally speaking, all DMTs target neuroinflammation and, as such, have a direct effect on the inflammatory component of MS; on the contrary, an indirect (and much less obvious) effect could be expected on neurodegeneration and disability progression [90]. Still, the long-term exposure to

Table 1.2. 2010 and 2017 revisions of McDonald Criteria for MS diagnosis.

Table shows clinical, radiological and laboratory requirements for MS diagnosis in different clinical presentations, according to 2010 and 2017 revisions of McDonald Criteria. Main novelties for the 2017 revision of McDonald Criteria, when compared with 2010 criteria, are inclusion of CSF-specific oligoclonal bands for DIT, evaluation of both cortical and juxtacortical lesions, removal of the differentiation between symptomatic and asymptomatic lesions. Adapted from Thompson AJ et al. Lancet Neurology 2018, and from Polman CH et al. Ann Neurol 2011 [59] [61].

Clinical Presentation	**2010 revision of McDonald Criteria**	**2017 revision of McDonald Criteria**
≥2 relapses and objective clinical evidence of ≥2 lesions or of 1 lesion with reasonable historical evidence of a prior relapse.	None	None
≥2 relapses and objective clinical evidence of 1 lesion.	DIS demonstrated by: ≥1 T2 lesion in at least 2 of 4 MS-typical regions, or further relapse implicating a different CNS site.	DIS demonstrated by: ≥1 T2 lesion in at least 2 of 4 MS-typical regions, or further relapse implicating a different CNS site.
1 relapse and objective clinical evidence of ≥2 lesions.	DIT demonstrated by: simultaneous presence of symptomatic T1-enhancing and T2 lesions, or a new T1-enhancing or T2 lesions on follow-up, or further relapse.	DIT demonstrated by: simultaneous presence of T1-enhancing and T2 lesions, or a new T1-enhancing or T2 lesions on follow-up, or further relapse, or CSF-specific oligoclonal bands.

1 relapse and objective clinical evidence of 1 lesion (CIS).	DIS demonstrated by: ≥1 T2 lesion in at least 2 of 4 MS-typical regions, or further relapse implicating a different CNS site. DIT demonstrated by: simultaneous presence of symptomatic T1-enhancing and T2 lesions, or a new T1-enhancing or T2 lesions on follow-up, or further relapse.	DIS demonstrated by: ≥1 T2 lesion in at least 2 of 4 MS-typical regions, or further relapse implicating a different CNS site. DIT demonstrated by: simultaneous presence of T1-enhancing and T2 lesions, or a new T1-enhancing or T2 lesions on follow-up, or further relapse, or CSF-specific oligoclonal bands.
Insidious neurological progression suggestive of MS (PPMS).	One-year disease progression and 2 of the following 3 criteria: 1. ≥1 T2 lesions in periventricular, juxtacortical or infratentorial areas; 2. ≥2 T2 lesions in the spinal cord; 3. Positive CSF (CSF-specific oligoclonal bands and/or elevated IgG index).	One-year disease progression and 2 of the following 3 criteria: 1. ≥1 T2 lesions in periventricular, cortical/juxtacortical or infratentorial areas; 2. ≥2 T2 lesions in the spinal cord; 3. Positive CSF (CSF-specific oligoclonal bands).

DMTs has significantly reduced the risk of severe ambulatory impairment (EDSS 6.0), SP conversion and death in people with relapsing MS, compared with no treatment or placebo [94–96]. Most evidence is coming from Interferon Beta and Glatiramer Acetate studies, whilst few has recently come from newer DMTs [95, 97].

Over ten DMTs have been developed and approved for RRMS and CIS over the last 20 years, whilst only three for progressive MS (Interferon Beta 1b and Siponimod for SPMS, and Ocrelizumab for PPMS) (**Table 1.3**) [98–100]. Indeed, translating medications originally developed for RRMS, to progressive MS has been unsuccessful so far, with several negative trials, including those investigating Fingolimod (INFORMS) [101], Natalizumab (ASCEND) [102], and Rituximab (OLYMPYUS) [103]. This is because clinical and radiological outcome measures for RRMS do not necessarily apply to progressive MS [63, 91]. Also, the poor understanding of the pathological mechanisms in progressive MS makes it challenging to develop targeted therapeutic agents. Thus, the understanding of MS progression and the development of biomarkers could enhance the development of new medications.

Despite the successful development of treatments for MS, a number of clinical trials have failed to show any significant clinical effect [43, 90, 104]. For instance, laquinimod is an immunomodulatory compound downregulating pro-inflammatory T cell responses in the peripheral blood and within the CNS. Furthermore, animal models have suggested neuroprotective effects of laquinimod, exerted through modulation of pathological microglia, with reduced demyelination and axonal damage [104]. Laquinimod consistently showed a positive effect on new or enlarging T2 lesions in both RRMS and PPMS, but did not reach primary clinical endpoints in clinical trials [104, 105],

suggesting that clinical effects might require longer time to disclose, when compared with MRI [87], and further highlighting the need of accurate and clinically meaningful MRI outcome measures.

Thus, after providing the reader with a general background to MS and showing its multifaceted pathology and clinical features, in the following Chapter 2, I will review brain and spinal cord imaging outcome measures that have been used in clinical trials and observational studies, specifically mentioning current limitations and possible future perspectives.

Table 1.3. Disease modifying treatments.

Table shows the list of approved DMTs in the United Kingdom, with recommendations for clinical use (adapted from the National Institute for Health and Care Excellence, the National Healthcare System England, and the European Medicine Agency) [98–100]. MRI activity is defined as ≥1 T1 gadolinium-enhancing lesion or significant increase in T2 lesion load. Inadequate response to previous DMT is defined as ≥1 relapse in the last year and MRI evidence of disease activity. Rapidly evolving severe RRMS is defined as ≥2 relapses in the last year and, on brain MRI, ≥1 T1 gadolinium-enhancing lesion or significant increase in T2 lesion load.

DMT	Administration	Subtype	Additional recommendations
Alemtuzumab	Infusion	RRMS	≥2 relapses in last 1 year and MRI activity; RRMS with inadequate response to previous DMT.
Beta Interferon 1a	Injection	CIS, RRMS	≥2 relapses in last 2 years; ≥1 relapse in last 2 years and MRI activity.
Beta Interferon 1b	Injection	CIS, RRMS	≥2 relapses in last 2 years;

			≥1 relapse in last 2 years and MRI activity.
		SPMS	Continuing relapses.
Cladribine	Oral	RRMS	Rapidly evolving severe RRMS; RRMS with inadequate response to previous DMT.
Dimethyl fumarate	Oral	RRMS	≥2 relapses in last 2 years.
Fingolimod	Oral	RRMS	RRMS with inadequate response to previous DMT; Alternative to Natalizumab (de-escalation strategy in patients at risk of progressive multifocal leukoencephalopathy).
Glatiramer Acetate	Injection	CIS, RRMS	≥2 relapses in last 2 years; ≥1 relapse in last 2 years and MRI activity.
Natalizumab	Infusion	RRMS	Rapidly evolving severe RRMS.
Ocrelizumab	Infusion	RRMS	≥2 relapses in last 2 years; ≥1 relapse in last 2 years and MRI activity.
		PPMS	Early PPMS with MRI activity.
Siponimod	Oral	SPMS	SPMS with MRI activity.
Teriflunomide	Oral	RRMS	≥2 relapses in last 2 years.

2. Introduction to multiple sclerosis imaging outcome measures

2.1. Introduction

As discussed in Chapter 1, during the last 20 years, over a dozen DMTs received the approval for the treatment of RRMS, being facilitated by screening the anti-inflammatory activity of putative treatments using active MRI lesions as outcomes in phase 2 trials [92, 106]. On the contrary, the paucity of active medications for both PPMS and SPMS is striking [1, 107]. In view of this, the Progressive MS Alliance recently suggested to develop and validate biomarkers of progression that could make clinical trials for progressive MS less time and resource-consuming, when compared with conventional clinical measures [108]. This could be achieved with the identification of reliable, repeatable and sensitive-to change imaging outcomes [91, 109].

Several brain MRI measures are able to reflect the inflammatory and neurodegenerative pathology of MS [110, 111]. Brain lesion count and volume are robust markers of inflammation and demyelination, and are important outcomes in both RRMS and progressive MS trials [63]. Following recent improvements in analysis methods, brain atrophy has gained relevance, in light of its strong association with disability accrual [63, 91]. Advanced brain MRI techniques, such as magnetisation transfer ratio (MTR), diffusion tensor imaging (DTI) and magnetic resonance spectroscopy (MRS), have been included in few trials so far, and hold promise for the future, as they can reflect specific pathological changes targeted by neuroprotective treatments, such as improved myelination measures within lesional tissue, following treatment [91, 112]. Positron emission tomography (PET) and optical coherence tomography (OCT) are also emerging as candidate imaging outcomes of MS progression [91].

More recently, improvements in MRI acquisition protocols and post-processing have overcome some of the limitations associated with imaging the spinal cord, a small and mobile structure at risk of motion artefacts from breathing, cardiac movement, CSF pulsation and blood flow [109, 113]. Conventional spinal cord MRI provides information on focal lesions, which are necessary for the diagnosis and prognosis of MS, and is commonly used in the clinical setting [63, 114]. Spinal cord volume loss is the result of demyelination, neuro-axonal loss, oligodendrocyte damage, and gliosis, ultimately resulting in chronic motor, sensory and autonomic dysfunction [115, 116], and will be at the very centre of this thesis. Advanced spinal cord MRI techniques assess the type and extent of spinal cord abnormalities, but, as discussed below, their use is currently limited to specialised centres for research purposes [109].

Overall, looking at the paradigm of treatment development for PPMS and SPMS, the number of imaging outcomes included in clinical trials has almost doubled from 2.3±1.5 in the decade 1996-2006, to 4.1±2.6 in most recent years (2007 to current) (**Figure 2.1**) [91].

In the present chapter, I will discuss clinical correlates, applications, limitations and future perspectives of brain and spinal cord outcome measures for the study of MS. Methodological and statistical drawbacks will be also discussed.

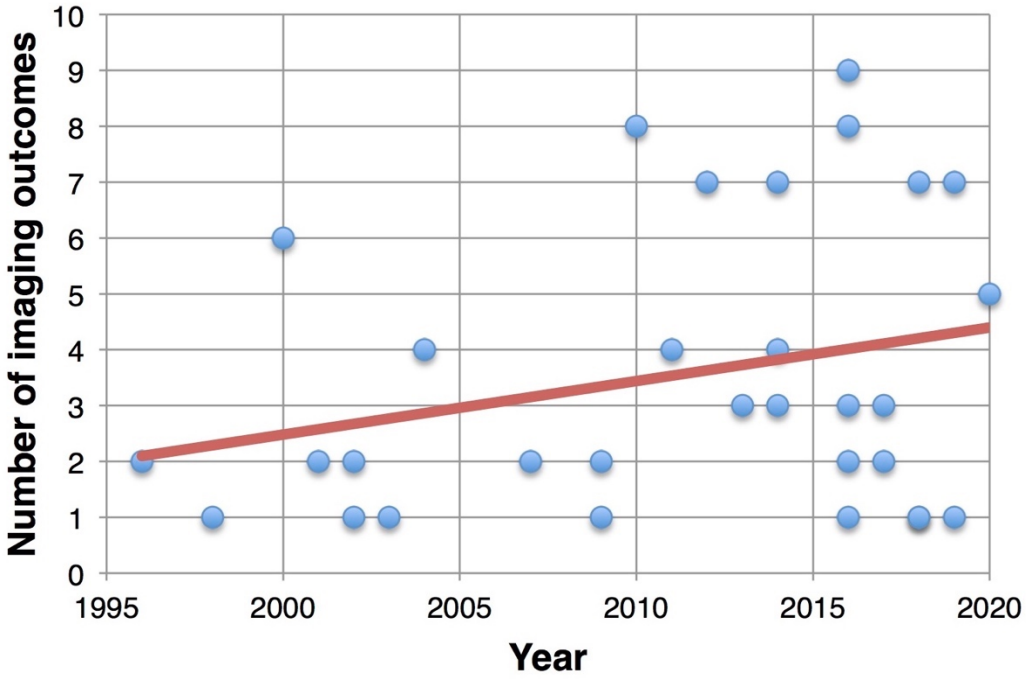
2.2. Brain outcome measures

2.2.1. Brain lesions

Brain lesions were the first MRI-derived outcome measure for MS clinical trials, and include number of gadolinium-enhancing, new/enlarging T2 lesions, and T1 lesions (or their combination into

Figure 2.1. Clinical trials and imaging outcomes.

Scatter plot shows the number of imaging outcomes used in clinical trials conducted from 1996 up to recent years (the expected conclusion date has been used for on-going clinical trials), independently of trial results (i.e., medication approval or not). Adapted from Moccia M et al. Mult Scler 2017 [91].



combined unique active lesions (CUA)), and their related volumes. Lesion measures are the best biomarker of active inflammation in MS, allowing the screen for early disease activity in phase 2 clinical trials in RRMS [106]. On the contrary, lesion-derived measures play a secondary – but not negligible role in the study of progressive MS. In PPMS, the burden of T2-visible lesion load and of gadolinium-enhancing lesions is low, despite clinical severity, and seems to have only a minimal impact on the disability accrual over time [107, 117]. Still, MRI measures of focal brain lesions are the most common imaging metric in phase 2 and 3 clinical trials in progressive MS [103, 118, 127–136, 119, 137–143, 120–126].

Future clinical trials on progressive MS might include these outcomes if the presence of inflammation is expected and targeted. Indeed, trials might select populations with relatively high inflammatory activity, depending on inclusion/exclusion criteria (e.g., 24.7-27.5% of PPMS patients presented with gadolinium-enhancing lesions at baseline visit of the ORATORIO trial, as expected in the earliest phases of the disease) [64, 120]. However, clinical outcomes might be difficult to predict based on results derived from brain lesions. Indeed, the use of DMTs specifically designed for RRMS in clinical trials in progressive MS can result in a positive effect on lesion count and volume measures, but not on neurodegenerative clinical (e.g. disability progression) and imaging outcomes (e.g. brain and spinal cord atrophy), as occurred in the INFORMS and the ARPEGGIO trials [105, 118, 144]. Similarly, the use of interferon-beta in SPMS was associated with fewer active lesions, but no effect was established on clinical disability [145].

2.2.2. Brain atrophy

Brain atrophy is detectable on MRI scans from the earliest clinical stages of MS, including radiologically isolated syndrome (RIS) and CIS, and is a biomarker of irreversible neurodegenerative processes [146]. Global brain atrophy has been associated with the degree of disability in large cohorts of both RR and progressive MS [147–149]. Besides, improvements in MRI post-processing have allowed segmentation of WM and GM (both cortical and deep) separately, allowing an improved association with clinical features [149–152]. Regional volumes might show a greater change over time, resulting in higher sensitivity and smaller sample size when compared with global measures [124, 153]. Intriguingly, brain atrophy has not been associated with relapse risk in RRMS, suggesting that atrophy is probably driven more by (possibly independent) neurodegenerative changes than inflammatory lesions, which further support the use of this measure in progressive MS [146, 154].

There are several methods to quantify whole brain atrophy. In general, brain tissue volume needs to be normalised for head size, and longitudinal changes can be detected by using registration and segmentation-based techniques. Registration-based methods compare longitudinally acquired images and measure changes in brain surface after registration in a common space; structural image evaluation using normalisation of atrophy (SIENA) is the most popular example. Segmentation-based techniques measure brain volume on a single scan, and, then, determine change over time indirectly, and include brain parenchymal fraction (BPF) (which is the ratio of brain parenchymal volume to the total volume within the brain surface contour) [146, 155]. In comparative analyses, brain atrophy measured with registration-based techniques shows better repeatability, and higher power to detect treatment effect, when compared with segmentation-based [156–158].

Whole brain atrophy has been included in several phase 2 and 3 clinical trials in progressive MS as primary [105, 124, 159–163], or secondary outcome (**Table 2.1**) [103, 118, 136, 138, 142, 143, 164–167, 119, 125, 126, 128–130, 133, 135]. The first trial demonstrating a beneficial effect on global brain atrophy (using simvastatin) was a phase 2 trial study in SPMS [121]. Positive results have been reported also in the phase 3 ORATORIO study in PPMS [120], and the phase 3 EXPAND study in SPMS [143]; future positive results will certainly be available in next years, as a result of the number of on-going trials measuring global brain atrophy (**Table 2.1**).

Regional brain atrophy has been used as secondary outcome in a few trials, where measures were obtained from cortical GM, deep GM, WM [124], putamen, thalamus, and optic nerve [126]. Considering that MS does not affect the brain uniformly, the detection of regional pathology might be predictive of more specific clinical features, when compared with whole-brain measures

Table 2.1. Clinical trials in progressive MS using imaging outcomes.

Table shows phase 2 and 3 clinical trials conducted in PP, SP and mixed populations (PP and SP) of MS patients, with study design, intervention and results on different MRI outcomes (O= ongoing; P= positive treatment effect; N= no effect detected). Adapted from Moccia M et al. Mult Scler 2017 [91].

	Study	*Design*	*Intervention*	*Brain atrophy*	*Regional brain atrophy*	*T2 lesions*	*T2 lesion volume*	*T1 lesions*	*Gadolinum+ lesions*	*CUA lesions*	*MTR*	*DTI*	*fMRI*	*Spinal cord atrophy*	*OCT*
PPMS	**FUMAPMS**, 2019 [125]	Phase 2 N= 90	Dimethyl fumarate vs Placebo	**O**	**O**	**O**	**O**		**O**		**O**	**O**		**O**	
	IPPoMS, 2018 [168]	Phase 2 N=85	Idebenone	**O**											
	ARPEGGIO, 2017 [105]	Phase 2 N=374	Laquinimod vs Placebo	**N**	**N**	**P**								**N**	
	ORATORIO, Montalban et al. *New Eng J Med* 2017 [164]	Phase 3 N= 732 Duration= 120 weeks	Ocrelizumab 600mg vs Placebo (2:1)	**P**		**P**	**P**								

	INFORMS, Lublin et al. *Lancet* 2016 [118]	Phase 3 N= 823 Duration= 36 months	Fingolimod 0.5mg vs Placebo (1:1.5)	**N**			**P**			**P**		**P**						**N**
	OLYMPUS, Hawker et al. *Ann Neurol* 2009 [103]	Phase 2/3 N= 439 Duration= 96 weeks	Rituximab 1000mg vs Placebo	**N**					**N**									
	Montalban et al. *Mult Scler* 2009 [165]	Phase 2 N= 71 Duration= 24 months	Interferon beta-1b (250µg on alternate days) vs. Placebo															**N**
	PROMISE, Wolinsky et al., *Ann Neurol* 2007 [139]	Phase 3 N= 943 Duration= 36 months	Glatiramer acetate vs Placebo						**N**			**N**						
	Leary et al. *Neurology* 2003 [166]	Phase 2 N=50 Duration= 24 months	Interferon beta-1a (30µg vs. 60µg per week) vs. Placebo															**N**
	Kalkers et al. *Mult Scler* 2002 [167]	Phase 2 N=16 Duration= 24 months	Placebo for 12 months vs. Riluzole for following 12 months (2x50mg per day)															**N**

SPMS	**EXPAND**, 2018 [143]	Phase 3, N= 1651	Siponimod vs Placebo	**P**		**P**	**P**	**P**	**P**		
	NCT02057159, 2019 [142]	Phase 2/3	NeuroVax vs Placebo						**O**		
	MS-SMART, 2018 [163]	Phase 2 N= 445	Amiloride vs Riluzole vs Fluoxetine vs Placebo (1:1:1:1)	**N**							
	Abili-T, 2017 [161]	Phase 2 N=183	Tcelna vs Placebo	**O**							
	B7493-W, 2016 [162]	Phase 2/3 N=54	Lipoic acid vs Placebo	**O**							
	ASCEND, 2016 [102]	Phase 3 N= 889	Natalizumab 300mg vs Placebo	**N**	**N**	**N**					
	MS-STAT, Chataway et al. *Lancet* 2014 [159]	Phase 2 N=140 Duration= 24 months	Simvastatin 80mg vs Placebo (1:1)	**P**	**P**	**N**					
	NCT00395200, Connick et al. *Lancet Neurol* 2012 [126]	Phase 2 N= 10 Duration= 20+10 months	Autologous mesenchymal stem cells transplantation, open	**N**	**P**		**N**	**N**		**N**	**N**

		label (before vs after treatment)										
MAESTRO, Freedman et al. *Neurology* 2011 [133]	Phase 3 N= 612 Duration= 2 years	MBP8298 500mg vs Placebo	**N**			**N**	**N**		**N**			
Lamotrigine trial, Kapoor et al. *Lancet Neurol* 2010 [124]	Phase 2 N=120 Duration=2 years	Lamotrigine 400mg vs Placebo (1:1)	**N**	**N**			**N**	**N**				**N**
ESIMS, Hommes et al. *Lancet* 2004 [119]	Phase 3 N=318 Duration= 27 months	IVIG vs Placebo (1:1)	**P**				**N**					
NA-SPMS, The North American Study Group on Interferon beta-1b in Secondary Progressive MS, *Neurology* 2004 [131]	Phase 3 N=939 Duration= 3 years	Interferon beta-1b 250 µg and 160 µg/m^2 vs Placebo (1:1:1)					**P**					

IMPACT, Cohen et al., *Neurology* 2002 [140]	Phase 3 N= 436 Duration= 24 months	IFN beta-1a 60mcg/week IM vs Placebo	**P**	**P**			
SPECTRIMS, Li et al. *Neurology* 2001 [132]	Phase 3 N= 618 Duration= 3 years	Interferon beta-1a 44µg and 22µg vs Placebo	**P**	**P**			**P**
Cladribine MRI Study Group, Rice et al. Neurology 2000 [137]	Phase 2 N= 159 Duration: 24 months	Cladribine 0.7mg/kg and 2.1 mg/kg vs Placebo	**N**	**P**		**P**	
EU-SPMS, European Study Group on Interferon beta-1b in Secondary Progressive MS *Lancet* 1998 [123]	Phase 3 N=718 Duration= 36 months	Interferon beta-1b vs Placebo (1:1)		**P**			
Karussis et al. *Neurology* 1996 [141]	Phase 2 N= 30 Duration= 6 months	Linomide 2.5mg vs Placebo				**P**	**P**

PPMS and SPMS	**SPRINT-MS**, Fox et al. *Cont Clin Trials* 2016 [160]	Phase 2 N= 255 Duration= 96 weeks	Ibudilast 100mg vs Placebo (1:1)	**P**														**N**
	NCT01144117, Schreiber et al. *Mult Scler* 2016 [129]	Phase 2 N= 50 Duration= 24 weeks	Erythropoietin 48000 UI vs Placebo	**N**				**N**										
	ACTiMuS, Rice et al. *Trials* 2015 [169]	Phase 2 N= 80	Early vs Late autologous bone marrow cellular therapy	**O**			**O**								**O**	**O**	**O**	**O**
	FLUOX-PMS, Cambron et al. *Trials* 2014 [128]	Phase 2 N= 120 Duration= 108 weeks	Fluoxetine 40mg vs Placebo	**O**	**O**	**O**									**O**			**O**
	NAPMS, Romme et al. *Neurology* 2014 [130]	Phase 2 N= 24 Duration= 60 weeks	Natalizumab 300mg open label (before vs after treatment)	**P**	**P**		**N**		**P**		**P**	**P**						
	CUPID, Zajicek et al., *Lancet Neurol* 2013 [138]	Phase 3 N=498 Duration= 3 years	Dronabinol vs. Placebo (1:1)	**N**		**N**		**N**										

[170, 171]. However, standardisation of software for analysis is needed to made widespread application in clinical trials possible [172].

Overall, measures of global and regional brain atrophy are gaining relevance in clinical trials on progressive MS, reflective of improvements in measurement techniques allowing good repeatability and sensitivity to change. Nevertheless, there are several possible limitations, including changes in magnet, gradients, coils, distortion corrections and image-contrast changes. Patients treated with anti-inflammatory treatments have a slight decrease in the brain volume in the first six to twelve months (pseudoatrophy), compared with placebo, due to the resolution of inflammation and oedema [173, 174]. A possible solution is to re-baseline subjects after 6 months [175, 176], although longer periods may be required for more toxic types of treatment (e.g. chemotherapy during bone marrow transplantation) [177]. However, re-baselining implies reduced time of observation on treatment, with subsequently reduced number of patients reaching study outcomes, and, then, loss of statistical power. In the OPERA II trial (one of the two phase 3 trials for Ocrelizumab in RRMS), statistical significance in brain volume change was lost when analysing data from week 24 to 96, instead of baseline to week 96 [178].

A reversible fluctuation of brain volumes can also occur because of variations in hydration status, which could be affected by disability status and time-of-the-day of MRI acquisition [179, 180].

2.2.3. Advanced brain MRI techniques

Conventional neuroimaging techniques lack specificity with regard to different pathophysiological substrates of MS, and are not able to explain the heterogeneous and long-term clinical evolution of the disease [176, 181–183]. Advanced MRI techniques, such as MTR, DTI and MRS, may provide

higher pathological specificity for the more destructive aspects of the disease (i.e., demyelination and neuroaxonal loss), and be more closely associated with clinical correlates [173, 184]. Moreover, functional MRI (fMRI) is contributing to the definition of the role of cortical reorganisation after MS tissue damage [151].

MTR values reflect the efficiency of the magnetisation exchange between mobile protons in tissue water and those bound to the macromolecules, such as myelin. MTR has been associated with disease progression in PPMS [148, 185]. In view of this, MTR has been included in several clinical trials in progressive MS and has been measured in GM (cortical and deep), WM, T2 lesions, putamen, thalamus and optic nerve [125–127, 130].

DTI measures brain tissue microstructure by the exploitation of the properties of water diffusion. From the tensor, it is possible to calculate the magnitude of diffusion, reflected by mean diffusivity (MD), and diffusion anisotropy, which is a measure of tissue organisation, generally expressed as fractional anisotropy (FA). In line with this, MD is increased and FA is decreased in T2 lesions, WM and GM from MS patients [36, 186]. DTI has been assessed across multiple scanners/platforms and is suitable for multi-centre studies [187–189]. DTI is the most frequently-used advanced MRI metric in phase 2 and 3 clinical trials in progressive MS, so far with demonstrable beneficial effects. MD and FA have been measured in pyramidal tracts, WM, GM and lesions in different trials in progressive MS [125, 127, 128, 130, 136]. More specific measures such as axial and radial diffusivity (RD) can be calculated as measures of the mobility of water along and perpendicular to axons (reflecting axonal density and demyelination respectively) [184]; however, they have not been included in clinical trials in progressive MS so far, due to difficulties in standardised acquisitions between centres.

The fMRI provides signal related to brain activation based on blood oxygen consumption and blood flow in the brain, and has only been included in two clinical trials on progressive MS [136, 163].

MRS can measure brain levels of several metabolites [190, 191]. The most commonly measured is total N-acetyl-aspartate (NAA), a marker of axonal loss and metabolic dysfunction [192]. NAA has been included in a few clinical trials in RRMS [193], and one in PPMS [194].

2.2.4. PET

PET is a quantitative imaging technique, which investigates cellular and molecular processes in vivo using positron-emitting molecules, ideally binding a selective target [191, 195, 196]. As MS is a complex and multifactorial disease, various radioligands have been tested. Amyloid tracers, measuring myelin loss and repair, and ^{11}C-Flumazenil, reflecting neuronal integrity, might be of interest for clinical trials on neuroprotective compounds [182, 196–199].

To date, no MS clinical trials have included PET, reflecting its invasive nature and high costs. In the future, the development of standardised procedures might represent a trigger for the application of this technique in phase 1 and 2a clinical trials [119].

2.2.5. OCT

OCT is a non-invasive method to obtain high spatial resolution images of the retina, measuring retinal nerve fibre layer (RNFL) thickness and macular volume.

RNFL is thinner in patients with MS than in healthy controls, even in patients with MS who have not experienced episodes of optic neuritis [200]. Therefore, OCT measures a more diffuse pathological process which better corresponds to overall CNS damage [201, 202].

RNFL and macular volume have been included in a few clinical trials on progressive MS (**Table 2.1**) [126, 128, 136, 160], so far without demonstrable neuroprotective effects.

OCT is a fast, non-invasive, easy-to-use imaging method producing quantitative measures reliably, with great potential in MS for testing neuroprotective strategies over a short time frame [146]. Like brain volume, RNFL is sensitive to biological variations. However, there is the need for high-quality acquisitions and appropriate image processing, performed by trained examiners following specific consensus criteria [203].

2.3. Spinal cord outcome measures

2.3.1. Spinal cord lesions

Spinal cord lesions on MRI correspond to areas of demyelination, oedema/inflammation, neuro-axonal loss and gliosis, affecting spinal cord structure and function [25, 116]. Post-mortem spinal cord studies have described a larger proportion of demyelination in the GM (33%), than in the WM (20%), with lesions involving either both GM and WM, or GM alone [204]. No difference in the extent of GM demyelination was seen between different cord levels, whilst most of WM lesions occur in the cervical cord [204, 205].

2.3.1.1. Characteristics of MS lesions on spinal cord MRI

Spinal cord lesions are visualised as areas of T2 hyperintensity (**Figure 2.2**) and, less commonly, as areas of T1 hypointensity on conventional spin echo sequences. Although T1 hypointensity in the spinal cord is thought to be rare in MS, a recent study using inversion-recovery prepared fast field echo sequence (e.g., heavily T1 weighted sequence) at 3T demonstrated that 87% of patients with MS show T1 hypointense lesions in the spinal cord, and most of the lesions seen on the short-tau inversion recovery (STIR) T2 weighted sequence are hypointense on T1 [206].

After administration of gadolinium, new inflammatory activity of MS lesions, with associated BBB breakdown, allow the MS lesions to appear as areas of T1 hyperintensity; gadolinium enhancement in acute spinal cord lesions is generally nodular and, in 20% of the cases, may have a ring-shape [207–209].

MS lesions often occur in the cervical region (59%), and, less frequently, in the lower thoracic spinal cord (T7-12) (20%) [210]. On sagittal scans, they appear as cylindrical lesions, while on axial images they appear as wedge-shape lesions. In sagittal views, they rarely exceed two vertebral segments in length. On axial scans, MS lesions involve less than 50% of the cross-sectional area, occupy preferentially the lateral and posterior WM columns and do not spare the GM. However, spinal cord involvement can be diffuse, as shown by diffuse signal abnormalities on PD images, especially in the progressive forms of MS; diffuse signal abnormalities in RRMS are associated with a poor prognosis [211, 212].

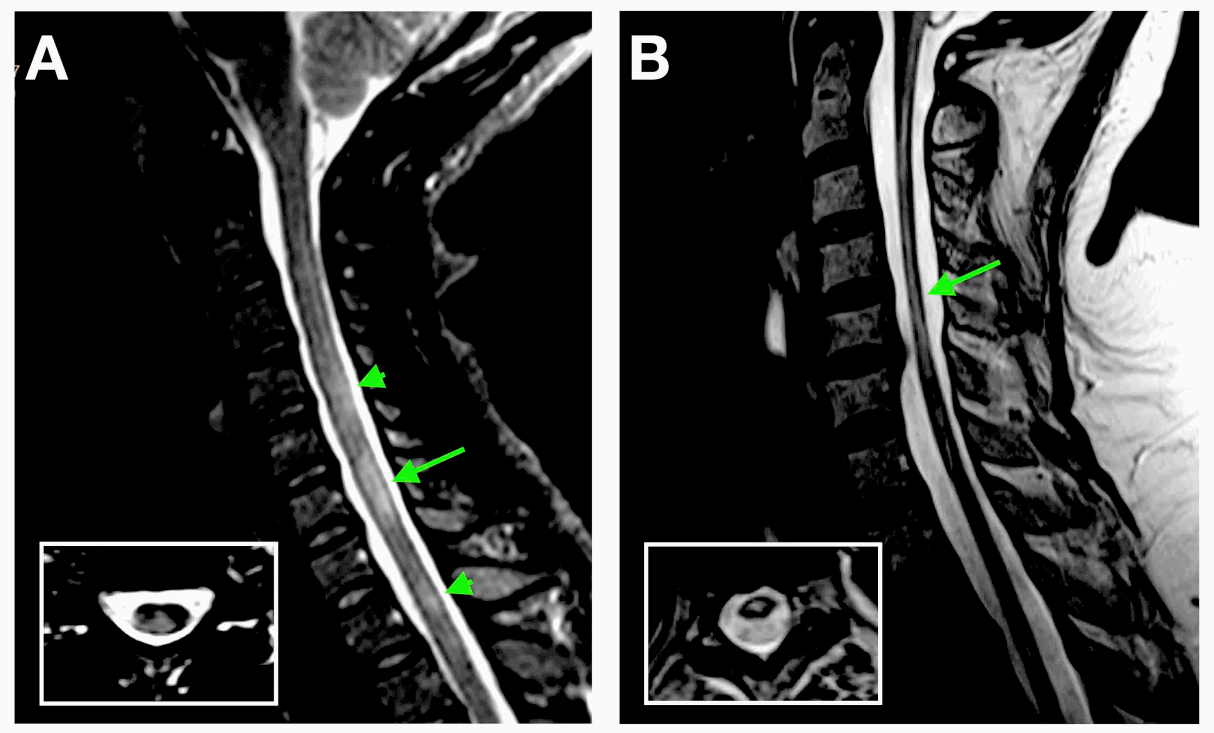
2.3.1.2. Recommended spinal cord MRI protocols

The MRI protocols recommended for spinal cord lesion detection in the clinical setting include both sagittal and axial scans [213]. For sagittal imaging, conventional or fast dual-echo (proton density

(PD) and T2-weighted, either in combination or independently) spin-echo sequences are usually considered the gold standard. A recent study at 1.5T suggested that PD fast spin-echo sequence detects cord lesions in patients who have normal T2 fast spin-echo MRI, and should therefore be used as a core sequence at 1.5T [214]. Either the T2 or PD spin-echo sequence can be substituted with a STIR T2-weighted sequence, which improves the visibility of MS lesions [215]. In general, it is

Figure 2.2. Lesions in MS and NMO.

T2 sagittal and axial (inset) spinal cord MRI of a patient with MS and a patient with AQP4-Ab positive NMOSD. In MS (**A**), spinal cord MRI shows a single area of T2 hyperintensity which involves both GM and WM in the lateral-posterior part of the cord, and has cylindric shape on sagittal view and wedge shape on axial view. In AQP4 NMOSD (**B**), there is a longitudinally extensive transverse myelitis from C1 to C5, with preferential involvement of the central area of the spinal cord. Arrows on the sagittal plane indicate the level at which the lesion is presented on the axial plane; arrow-heads indicate additional lesions. Adapted from Moccia M et al. Ther Adv Neurol Disord 2019 [109].



not recommended to use the STIR sequence on its own because of its susceptibility to flow-related artefacts and possible lower observer concordance [216]. An alternative to STIR sequence or to one of the two dual-echo T2-weighted sequences, is a heavily T1-weighted sequence [217], such as phase-sensitive inversion recovery (PSIR) or magnetisation-prepared rapid gradient echo (MPRAGE), which improves the detection of MS lesions in the cervical cord [218]. The three-dimensional (3D) acquisition of the MPRAGE permits multiplanar reconstruction that facilitates the delineation of lesions [219]. A recent 3T study has reported that a 3D double inversion recovery (DIR) sequence for cervical spinal cord imaging is more sensitive at detecting inflammatory lesions than conventional 2D T2-weighted TSE sequence [220]. However, the DIR sequence of the spinal cord is not widely used in clinical practice because it is strongly affected by artefacts, especially in obese patients, and by magnetic field inhomogeneities, and its coverage capability is currently limited to the cervical spine [220].

For axial imaging, possible sequences are 2D and/or 3D T2-weighted fast spin echo sequences. A full cervical cord axial coverage detects more lesions (9-22%) than sagittal scans alone [210, 221], and can also exclude lesions in cases of equivocal abnormalities on sagittal scans [210, 221, 222].

No significant improvement in lesion detection was found when using 3T field strength compared with 1.5T [223]. Improvements in lesion detection are expected at 7T, although its application and relevance require further studies, especially in the context of new coil designs and optimised acquisition times [224–226].

Although pathological involvement of the spinal cord GM contributes significantly to disability in RRMS and SPMS [227], its assessment with conventional MRI techniques is not achievable because

of insufficient contrast between tissue compartments and low spatial resolution. In the research setting, improved delineation of cervical cord lesions and their involvement of the WM and GM are obtained by using 3D-PSIR in combination with axial 3D gradient-echo fast field echo (3D-FFE) [228], although this MRI protocol requires a long acquisition time and has limited coverage of the cervical cord.

2.3.1.3. Diagnosis of MS supported by spinal cord MRI

The 2017 revised McDonald criteria confirmed that MRI is the most useful paraclinical test to aid the diagnosis of MS, and can be used to establish dissemination of lesions in space (DIS) and time (DIT) [61]. The spinal cord is one of the four areas of the CNS where lesions with characteristics typical of MS are scored to confirm DIS. Prior to the 2017 McDonald criteria, only asymptomatic spinal cord lesions were scored for DIS, which led to the high specificity of the DIS criteria; in order to facilitate the scoring of the criteria, and avoid discussing which lesion is the symptomatic one in cases of multiple lesions occurring in the same CNS location, the 2017 revised criteria do not distinguish anymore between symptomatic and asymptomatic lesions when testing the DIS criteria, as already discussed in Chapter 1 [61]. In particular, the inclusion of spinal cord symptomatic lesions for DIS or DIT increases diagnostic sensitivity, with little or no reduction in specificity [80, 229, 230].

Whilst brain MRI is recommended in all patients who are undergoing investigations for the diagnosis of MS, spinal cord MRI is advisable when: (1) The clinical presentation suggests a spinal cord lesion; (2) The clinical presentation is suggestive of PPMS; (3) Brain MRI is normal, but there is a strong clinical suspicion of MS; (4) Brain MRI findings are inconclusive (e.g., age-related vascular changes) [61, 217, 231]. Therefore, spinal cord MRI is generally recommended in patients with spinal cord CIS and in those with non-spinal MS not fulfilling the DIS criteria. It is debated whether all the remaining

CIS patients, who have non-spinal MS and fulfil DIS criteria on brain MRI brain, should undergo spinal cord MRI [232].

More recently, patients with clinical features typical of MS, but showing evidence of pathology exclusively in the spinal cord, even with a single lesion, and whose MRI does not fulfil the DIS criteria, have been described as two novel clinical entities: (1) Progressive solitary sclerosis, when insidiously progressive upper motor neuron impairment can be attributed to an isolated demyelinating lesion within the CNS (within the spinal cord in 90% cases) [233]; and (2) Pure spinal MS, when relapsing episodes of short-segment myelitis occur over time, in the absence of typical brain or optic nerve lesions [234]. Progressive solitary sclerosis and pure spinal MS are proposed novel MS phenotypes, characterised by a predominant spinal cord pathology. Myelocortical MS is another suggested pathology subtype of MS where axonal loss in the WM occurs in absence of obvious demyelination, but is still lacking further pathology and clinical validation [39].

2.3.1.4. Differential diagnosis facilitated by spinal cord MRI

MS could be responsible for up to 50% cases of inflammatory myelopathies and, thus, a number of conditions should be considered in the differential diagnosis. These include neuromyelitis optica spectrum disorders (NMOSD), myelin oligodendrocyte glycoprotein-antibody (MOG-Ab) associated disease, sarcoidosis, paraneoplastic syndromes, infectious and post-infectious diseases, and require different treatment and management strategies [235, 236]. Certain lesion characteristics on spinal cord MRI may aid the clinicians to navigate through the differential diagnosis of spinal cord inflammatory pathology [213].

NMOSD is responsible for up 50% cases of longitudinally extensive transverse myelitis (LETM), defined as T2 hyperintense spinal cord lesion extending ≥3 vertebral levels (**Figure 1.2**) [236]. However, the length of spinal cord lesions in NMOSD depends on timing of MRI with respect to clinical onset [237]. NMOSD can also present with involvement of <3 vertebral segments [238]. Additionally, LETM is not a pathognomonic feature of NMOSD, and other inflammatory demyelinating conditions can cause a LETM.

One of the most important spinal cord MRI features differentiating NMOSD from MS, and other LETM aetiologies, is the presence of bright spotty lesions (BSLs) [239, 240], defined as lesions with signal intensities at least as high as, but not higher than, that of the surrounding CSF on a T2-weighted image, and not as low as that of the surrounding CSF on a T1-weighted image. Bright spotty lesions are seen in the majority of patients without LETM, and it is thought that they indicate severe damage to the spinal cord. Other spinal cord distinctive features of NMOSD are lesions occupying ≥50% axial cross-sectional area (transversally-extensive lesion), T1 hypointense lesions, and centrally- or both centrally- and peripherally-located lesions [236]. Gadolinium enhancement is common in NMOSD, but variable in its appearances (frequently irregular and punctuate); ring-enhancement is seen in one-third of NMOSD myelitis episodes and distinguishes NMOSD from other causes of longitudinally extensive myelopathies, but not from MS [209]. Additionally, NMOSD lesions are more frequently located in the cervical and/or dorsal spinal cord, when compared with the lumbar cord [236].

20-40% of NMOSD patients negative for aquaporin-4 antibody (AQP4-Ab), are instead MOG-Ab positive [241, 242]. The LETM of MOG-Ab-associated disease is virtually indistinguishable from that of NMOSD AQP4 positive disease.

Spinal cord sarcoidosis is an under-recognised cause of LETM, and can precede symptoms of systemic and pulmonary sarcoidosis. Linear dorsal leptomeningeal enhancement extending ≥2 vertebral segments and persisting >2 months differentiates spinal cord sarcoidosis from NMOSD and MS, where gadolinium enhancement is patchy, diffuse or ring-like [243, 244]. When linear dorsal subpial enhancement is combined with central canal enhancement in cases of sarcoidosis, a "trident" sign on axial post-gadolinium sequences has been described [244, 245].

Other important causes of spinal cord lesions are post-infectious myelitis (e.g., cytomegalovirus, herpes simplex virus, varicella zoster, enterovirus, etc), which often present with LETM, and are associated with variable radiological appearances on T2, T1 and post-contrast T1 weighted images.

Non-inflammatory myelopathies include vascular aetiologies (e.g., acute spinal cord infarction), spinal dural arteriovenous fistula, tumours, nutritional deficiencies, infections, and compressive myelopathies. In these cases, timely diagnosis and management is crucial to improve clinical outcomes [235, 246].

Additional clinical (e.g., hyper-acute or gradually progressive onset), radiological (e.g., presence of lesions on brain MRI, abnormalities on chest PET imaging), laboratory (e.g., presence of AQP4 and MOG antibodies) features might be necessary to establish the exact diagnosis of myelopathy [236, 243, 245]. The most striking consequence of a more appropriate and widespread use of spinal cord MRI and additional tests in cases of spinal cord myelopathy is that the recognition of an "idiopathic" transverse myelitis is reducing over time [235].

2.3.1.5. Prognosis of MS using spinal cord MRI

In patients with radiologically isolated syndrome (RIS), the presence of asymptomatic spinal cord lesions is seen in 64% of patients who later develop CIS or MS, and in 100% of patients who later develop PPMS [247].

In patients with CIS, the presence and the number of spinal cord lesions are associated with increased risk of clinical conversion to MS and disability progression, regardless of demographics, clinical features and brain MRI [248–250]. In contrast, the probability of disability progression is very low in the absence of spinal cord lesions [248].

In established MS, spinal cord lesions are associated with a higher risk of relapse [251], disability progression [252, 253], and switching of DMT due to poor treatment response [254]. Also, upper cervical cord lesion load, quantified on PSIR sequences, is greater in progressive forms of MS than in RRMS, and is associated with disability [252]. In SPMS, spinal cord lesions frequently involve at least two spinal cord WM columns and extend to the GM [212]. The main limitation of these studies is that spinal cord coverage was confined to the upper cervical cord, in order to minimise physiologic artefacts and enable high-resolution acquisitions within an acceptable time frame, thus limiting generalizability to clinical practice [212].

2.3.1.6. Monitoring MS with spinal cord MRI

Spinal cord lesions are more likely to be symptomatic and leave residual neurological impairment, due to poor compensatory capacity of the spinal cord, than brain lesions [250, 251]. Despite this, 58% of new spinal cord lesions were reported to be asymptomatic and 25% of RRMS patients develop at least one asymptomatic spinal cord lesion over 1.5 years [255]. When only stable RRMS

patients are considered, 10% of them show subclinical spinal cord lesion activity alone [251]. Interestingly, asymptomatic spinal cord lesions predict clinical relapses when combined with asymptomatic brain lesions [251]. Thus, spinal cord MRI could disclose subclinical disease activity in otherwise clinically stable MS, and could enhance a more thorough understanding of the course of MS [114]. Asymptomatic spinal cord lesions may not be restricted to MS patients, as they have also been observed in NMOSD patients [256], but more data in NMOSD are needed.

2.3.1.7. Spinal cord lesion mapping

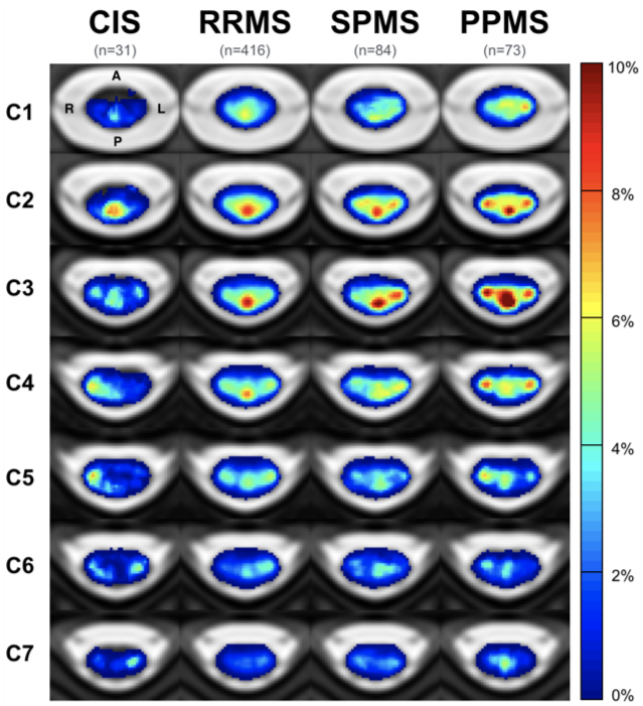
Recent developments in imaging post-processing have allowed the creation of probability maps of spinal cord lesions, which show probability of each voxel being "lesional". Single-centre studies combining 3D-T1 weighted FFE and the active surface model (ASM), a semi-automated voxel-based analysis of the spinal cord within JIM, showed that SPMS and especially PPMS patients have higher lesion counts and volumes, when compared with RRMS, and that lesions are more frequently located in the posterior cord than in the anterior cord, and in the upper cervical cord than in the lower cord [206, 257]. A larger, multi-centre study, employing fully automated methods based on the Spinal Cord Toolbox (SCT) confirmed that lesions are more frequently located in the posterior columns in all MS subtypes, and that lesion mapping at C3 clearly distinguishes between MS subtypes [258]. In particular, high lesion probability was found in the posterior columns in RRMS, posterior and lateral cord in SPMS and posterior, lateral and central regions in PPMS (**Figure 2.3**) [258]. High disability levels were associated with lateral and central cord involvement [258].

2.3.2. Spinal cord atrophy

Spinal cord atrophy is a common and clinically relevant aspect of MS. An increasing number of

Figure 2.3. Lesion probability maps in the spinal cord.

Lesion probability maps at the cervical level are shown for different disease subtypes. Adapted from Eden et al. Brain 2019 [258].



studies have focused on the importance of spinal cord atrophy as a biomarker of disability progression and as an outcome measure in clinical trials.

2.3.2.1. Pathological correlates

Spinal cord atrophy is the consequence of different pathological processes, including axonal transection and associated neuro-axonal loss, demyelination, loss of oligodendrocytes, gliosis, and, ultimately diffuse tissue injury [23, 24, 226, 259–262]. Although these pathological abnormalities occur within focal lesions, extensive tissue abnormalities are also present in the normal-appearing spinal cord of MS patients, and this finding may explain why spinal cord atrophy occurs independently of spinal cord lesions [23–25, 35, 263–266]. Additionally, spinal cord atrophy also occurs, at least in part, independently of brain pathology [225, 264, 267].

2.3.2.2. Advances in spinal cord atrophy measurements

Spinal cord atrophy is generally measured as cross-sectional area at the cervical level, which is least affected by movement artefacts, yields the most reproducible results, and provides the best clinical correlates [268–272]. The most common levels are C1-C2 and C2-C3, but measurements can be also made between C1 and C7 [273].

Atrophy assessment can be done on a variety of sequences, mainly 3D T1-weighted and T2*-weighted gradient echo sequences on different MRI scanners (e.g. Philips, Siemens, GE) [115, 274, 275].

In the present thesis, we will use two different methods for spinal cord image segmentation and atrophy calculation: surface-based and image-based methods [276]. The older methods were fully manual, whilst the most recent methods have been semi- or fully-automated. For example, JIM is a surface-based method that semi-automatically outlines the cord, after marking the centre of the spinal cord [277]. Within the JIM tool, the ASM has provided more prompt and reproducible measures of the spinal cord volume, compared with manual methods [278]. The ASM offers a considerable reduction in user interaction time, and can be performed over long spinal tracts. The user needs to identify landmarks at the extremes of the region to study, and, then, mark the centreline of the cord. Sagittally-acquired images are then reformatted to the axial plane to obtain five contiguous 3 mm slices; the program automatically calculates the radius and the centre of each axial slice and, finally, the cross-sectional area is obtained by averaging these contiguous slices [274, 279, 280]. Other semi-automated method is NeuroQLab (an image-based method that segments

the upper cervical cord from surrounding non-spinal cord tissue by using a Gaussian mixture modelling method) [281, 282].

Recent efforts have attempted to develop fully-automated methods, such as the SCT, which is an open-source comprehensive software dedicated to the processing of spinal cord MRI. SCT is built on previously-validated methods and includes motion correction tools, templates and algorithms to segment the spinal cord, allowing standardisation and automation of the processing pipeline [283]. The segmentation tool (*PropSeg*) contained in the first version of SCT (version 3.0) has already been tested on a large cohort of MS patients and healthy controls, and has been used in Chapter 5 of the present thesis. This fully-automated surface-based image segmentation method has the same sensitivity as the ASM tool within JIM, but has higher inter-rater reproducibility and is more time-efficient [284]. A newer version of SCT (version 4.0) includes a fully-automated segmentation method based on convolutional neural networks (*DeepSeg*) [285], that has been used in Chapter 6 of the present thesis; this version also has a fully-automated framework for intramedullary lesion segmentation, presenting with higher efficiency and reproducibility in lesion count and volume, when compared with manual detection [286].

A recent study has demonstrated that there is a systematic difference in the values of the cross-sectional area between methods, with lower values provided by fully-automated methods (SCT) than semi-automated methods (NeuroQLab and JIM) [287, 288]; a good agreement between these two semi-automated techniques was observed [287].

When using these methods for longitudinal spinal cord atrophy calculation, the rate of atrophy is estimated indirectly by numerical subtraction of spinal cord cross-sectional area measurements calculated at different time-points [289].

There has been a recent shift towards calculating spinal cord atrophy using brain volumetric images [282, 290]. A recent collaborative study has confirmed that the spinal cord cross-sectional area, calculated at C1-C2 level using dedicated volumetric MRI of the spinal cord, is similar to that obtained using volumetric brain MRI [287, 290, 291]. Chapter 6 will aim to validate this new approach, which has the potential to allow calculation of spinal cord atrophy without acquiring a dedicated cord sequence, thereby saving scanning time, in both clinical trials and observational cohorts.

Few clinical trials in progressive MS have included spinal cord atrophy as outcome measure (**Table 2.1**) [63, 124, 125, 136, 144, 165–167, 213], but yielded inconclusive or negative results. Those disappointing results may be, at least in part, due to the relatively high noise and low reproducibility of the segmentation-based methods [91, 109, 292]. A more widespread use of spinal cord atrophy is currently hampered by challenges to obtain high reproducibility and responsiveness to changes when measuring such a small structure. Small absolute changes in spinal cord area are difficult to detect in a multi-centre setting, where there may be a great variability of imaging protocols and scanners [156]. The acquisition of high-quality spinal cord MRI can be affected by artefacts (e.g. breathing, pulsation of blood and CSF), and this may limit the precision of spinal cord atrophy measurements. As a consequence, sample size estimates obtained for current measurement techniques are fairly large and generally prohibitive, when compared with brain atrophy [144].

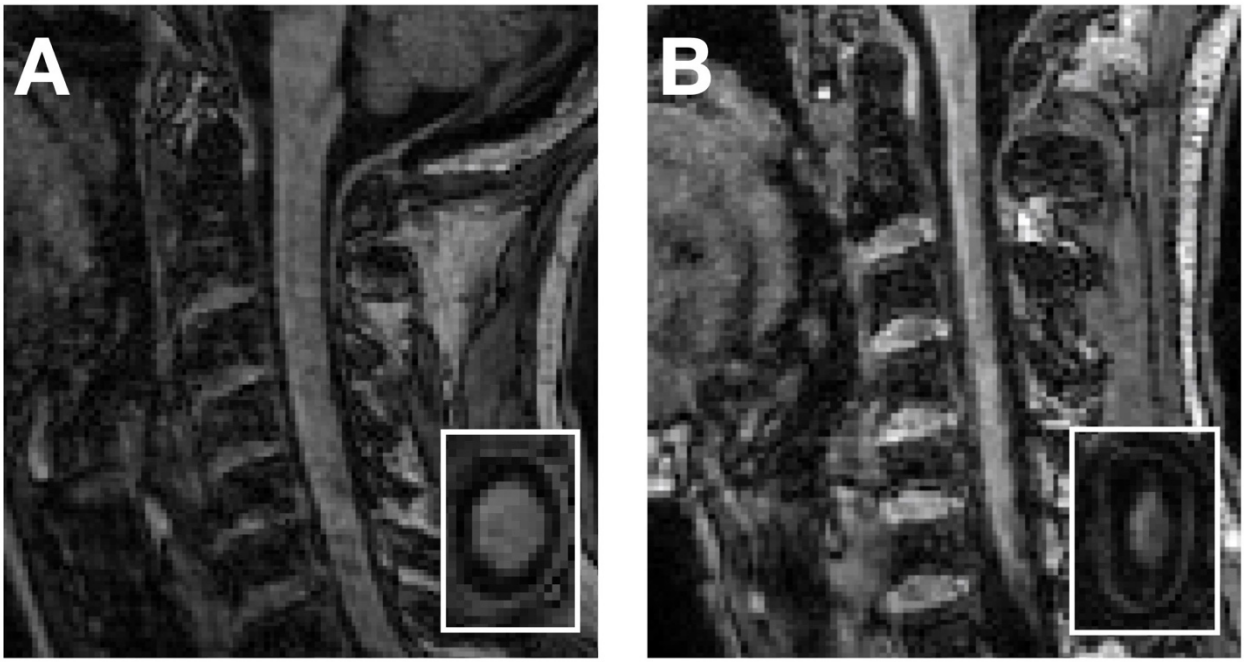
2.3.2.3. Spinal cord atrophy in disease phenotypes

Spinal cord atrophy occurs even in early stages of MS, and has been detected in patients with CIS [275, 293–295]. In CIS patients who were followed-up for 5 years after onset, the lowest rate of spinal cord atrophy (-0.1% a year) was observed in those who remained CIS, whilst the highest rate (-1.4% a year) was detected in patients who developed MS and had an EDSS at the last time point equal or greater than 3 [250]. In general, a high rate of spinal cord atrophy is observed in the progressive forms of MS, especially SPMS (-2.2% per year) (**Figure 2.4**) [226, 273, 294–296]. Overall, in clinically-definite MS, the rate of cord atrophy has been reported to vary between 1 and 5% per year [115, 226, 297–299]. A multicentre study has detected a rate of -1.22% per year in patients with stable MS and -2.01% in patients who deteriorated over time [273]. Interestingly, there is a significant development of spinal cord atrophy in early PPMS patients when compared with healthy controls over only 1-year follow-up, but not in patients with established SPMS, who had a higher disability and more atrophic cord than early PPMS patients [300]. Although the rate of atrophy may vary slightly between studies, because of different cohorts and different methods, it is consistently higher than the rate of brain atrophy, which is known to be around -0.5% per year in MS patients [149]. A recent meta-analysis of twenty-two longitudinal studies assessing spinal cord atrophy in all MS subtypes revealed a pooled rate of spinal cord atrophy of -1.78% per year, that increased to -2.08% per year when considering progressive patients alone [268].

The segmentation of GM areas on PSIR images at 3T allows the evaluation of GM atrophy in MS. Relapsing MS patients show smaller spinal cord GM areas (i.e., higher atrophy) than age- and sex-matched controls, without significant differences in spinal cord WM areas [301]; the GM of progressive MS patients shows the highest degree of atrophy [301].

Figure 2.4. Spinal cord atrophy visible on conventional MRI.

Cervical cord MRI with sagittal and axial views at C2 level (inset, used for spinal cord cross-sectional area measurements) in CIS (**A**) and PPMS (**B**) of same age (40 y.o.) and sex (males). Adapted from Moccia et al. Ther Adv Neurol Disord 2019 [109].



Only few studies have examined cervical cord atrophy in NMOSD, and reported conflicting results. Some studies found more pronounced spinal cord atrophy in AQP4 positive patients than MOG patients [302], and in MS than NMOSD [303], whereas another study found similar reductions of cross-sectional areas in NMOSD and MS [304].

2.3.2.4. Spinal cord atrophy and MS disability

A number of studies have shown associations between: (1) the extent of spinal cord atrophy at a single time point and concurrent disability [305], and (2) the rate of spinal cord atrophy over time and disability progression [250, 275, 295, 306–308]. A recent study has reported that every 1% increase in the annual rate of spinal cord volume loss is associated with a 28% risk of developing

disability progression in the subsequent year [296]. In a longitudinal cohort of non-spinal CIS, upper cord cross-sectional area (UCCA) decrease was associated with 5-year increased disability, measured by EDSS [250]. Overall, spinal cord atrophy can account for 77% of disability progression after 5 years [227, 305, 309]. Within EDSS, the sub-scores that reflect the neurological functions mediated by spinal cord pathways, such as the pyramidal, sensory, bowel and bladder functional scores, correlated with spinal cord atrophy [310]. Higher spinal cord atrophy rate is associated with worsening of more specific measures of motor disability, such as the 9HPT and the T25FWT [296, 305]. Associations between the development of spinal cord atrophy and disability progression are particularly strong in PPMS [311].

2.3.2.5. Spinal cord atrophy in clinical trials

Since spinal cord atrophy rates are two-to-three times higher than brain atrophy (-1.78% vs -0.5% per year), in particular in progressive MS [268, 312], and the spinal cord is a very eloquent site of pathology in MS, spinal cord atrophy has been considered as an exploratory outcome measure in phase 2 and phase 3 clinical trials, especially in patients with progressive MS, although much less frequently than brain atrophy [63]. However, clinical therapeutic trials that incorporated spinal cord atrophy as an outcome measure did not demonstrate beneficial drug effects on this metric [124, 144, 313, 314]. In addition to the possibility that the medications tested were not effective, there may be other reasons for these negative results, related to methodological difficulties of calculating spinal cord atrophy; these include: movement artefacts and subsequent image noise; the limited spatial resolution of MRI scanners, which is an important issue, given the small cord size; multicentre design, with inter-site variability related to the use of different scanners with different acquisition settings; and inter-study variability related to the use of different methods to calculate spinal cord area [156, 315]. Also, spinal cord normalisation using the intracranial volume, which aims to reduce

the effect of biological conditions unrelated to the disease, has been suggested [269, 270, 316], but it is not always performed.

There have been encouraging results from a recent, single-centre, study employing spinal cord atrophy [300, 315]. If patients at the early stage of PPMS, with mild disability and a non-atrophic cord are selected, the sample size necessary to run a trial over only 1 year is achievable [300].

2.2.3. Advanced spinal cord MRI techniques

Advanced spinal cord imaging techniques are currently used in exploratory studies to investigate microstructural abnormalities which reflect neurodegeneration, and to develop new targets for therapeutic intervention [91, 191]. These techniques include methods that study neuro-axonal integrity (DTI and new models of diffusion), myelin content (MTR and myelin water imaging), metabolic changes (MRS and functional connectivity on fMRI) (**Table 2.2**). However, advanced MRI techniques remain technically challenging, and results from studies using different acquisition protocols are difficult to compare [317–319]. Hereby, I will focus on the most recent advances in these techniques and their latest applications to MS patients, and refer to other manuscripts for more technical and comprehensive reviews [184, 320, 321].

2.2.4.1. Diffusion based techniques

DTI provides quantitative measures of microstructural abnormalities, which have been found to be abnormal in MS when compared with healthy subjects [322]. Recent studies have reported increased magnitude of diffusion in the direction perpendicular to the main direction of fibre bundles (i.e., RD), and reduced diffusion anisotropy (i.e., FA) in RRMS with acute spinal cord involvement, when compared with healthy controls, and in SPMS, when compared with clinically

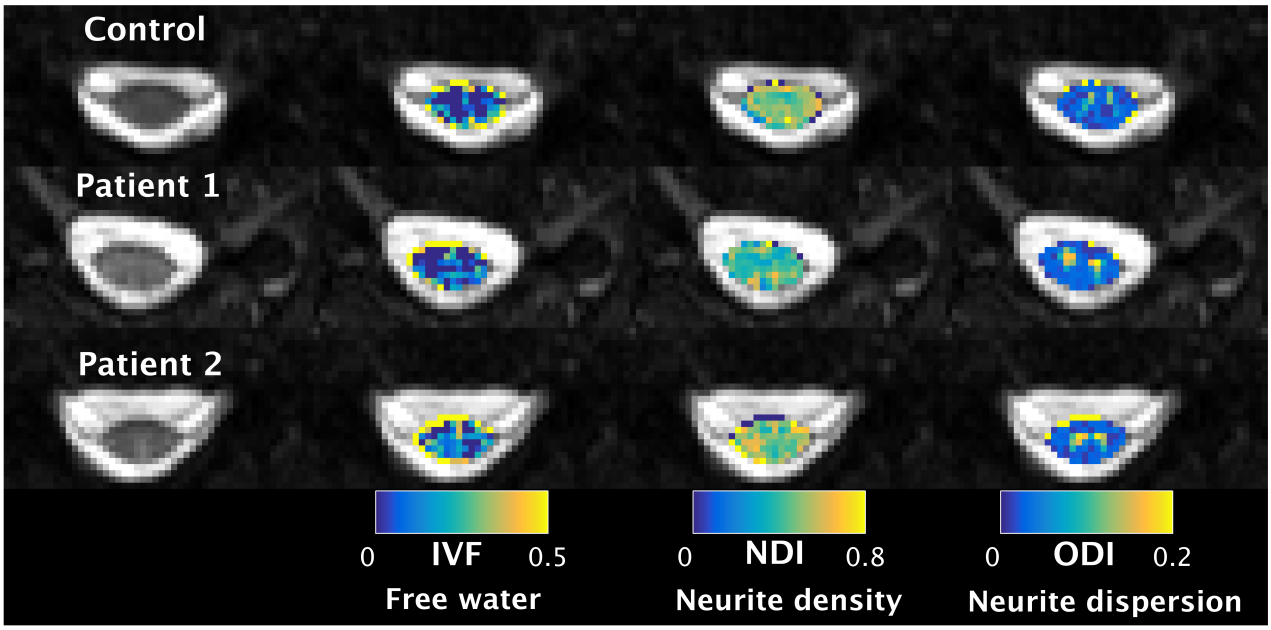
stable RRMS [227, 318], suggesting reduced myelin and axonal integrity and impaired neuronal fibre coherence [323]. A combination of DTI indices could explain up to 77% of the EDSS variability, suggesting a strong contribution of spinal cord microstructural changes to irreversible disability (**Table 2.2**) [227]. A recent study, which investigated the reproducibility of DTI-derived measures at C1-C6 between different sites, has shown the feasibility of multi-centre spinal cord DTI, with adequate matching of the sequence design across sites, in particular for different manufacturers [324]. The main advantages of spinal cord DTI are that it is simple to acquire and easy to interpret; its main limitation is that the DTI-derived measures are based on model approximations that the biological substrate is likely to violate and have low pathological specificity.

Q-space imaging (QSI) is a model-free technique that determines the voxel wise probability density function of fibre orientation (using apparent diffusion coefficient (ADC), full-width half-maximum (FWHM), and zero displacement probability (P0)), and seems to be more sensitive than conventional DTI measures at detecting MS-related abnormalities [325]. QSI-derived indices of perpendicular diffusivity are increased and indices of parallel diffusivity are decreased in the spinal cord of early PPMS, when compared with controls, possibly reflecting increased movement of water in the direction perpendicular to the long axis of the cord, due to the breakdown of myelin and axonal membranes, even in the absence of a significant degree of spinal cord atrophy [326, 327]. Changes in QSI-derived measures are associated with different measures of clinical disability, suggesting that they reflect pathological abnormalities that contribute to neurological deficits (**Table 2.2**) [326, 327]. The main limitations of QSI include the need to acquire a large number of data points, therefore necessitating long acquisition times, limited directional resolution, and difficulty in interpreting the probability density function.

Neurite orientation dispersion and density imaging (NODDI) is a recently developed multi-compartmental diffusion model, providing microstructural indices related to geometrical complexity of neurite architecture (**Figure 2.5**) [328]. This technique applied to the spinal cord has been recently validated by comparison with histology, and a trend toward lower neurite complexity in demyelinated lesions, has been demonstrated [329, 330]. In a pilot study, neurite dispersion index was reduced in the spinal cord of RRMS patients when compared with healthy controls [331], and a recent study has described reduced orientation index in the normal-appearing WM and lesions of the spinal cord from 6 MS patients, when compared with 8 healthy subjects (**Table 2.2**) [332]. The

Figure 2.5. Neurite orientation dispersion and density imaging in the spinal cord.

NODDI provides tissue-specific indices related to geometrical complexity of neurite architecture. Cervical spinal cord NODDI maps (at C2 level) of isotropic volume fraction (IVF) (estimating the amount of free water), neurite density index (NDI) (estimating the amount of neurites), and orientation dispersion index (ODI) (estimating the variability of neurite orientations) are shown from healthy control and MS patients. Courtesy Dr Francesco Grussu, University College London, United Kingdom.



main findings of brain and spinal cord NODDI studies is that for similar value of FA there are different combinations of orientation dispersion index (ODI) and neurite dispersion index, so NODDI is expected to be more pathologically specific than DTI.

Finally, an exploratory study has assessed the feasibility of the spherical mean technique (SMT), which is another multi-compartmental diffusion model, in the spinal cord in 6 MS patients and 8 controls (**Table 2.2**) [333]. SMT, which is feasible on standard MRI scanners, enables mapping of the neurite density and compartment diffusivities, and is sensitive in identifying abnormal changes in MS lesions when compared with healthy WM [333].

An *in vivo* study of the spinal cord, which fits, studies and compares several biophysical models, similar to what has been done in the brain [334], would be important to establish the limitations and the advantages of each model and the clinical potential of the latest models. Reducing the acquisition times, without sacrificing the accuracy of the derived indices, may be possible with the latest techniques [333]. The development of more advanced hardware (high-field MRI scanners), software (localisation, gating, and motion compensation), and coils (multi-channel phased-array coils) will contribute to expand the use of diffusion-derived metrics in MS clinical practice and trials [320].

2.2.4.2. Techniques reflecting myelin content

MTR is a quantitative technique measuring the magnetisation exchange between freely mobile protons and those associated with macromolecules such as myelin, providing an indirect estimate of myelin content, in addition to neuroaxonal integrity and water content. A large study carried out in patients with MS reported lower MTR values in patients with higher EDSS, than those with lower

EDSS, independent of lesion load [335], suggesting that this measure can detect clinically relevant differences beyond conventional imaging. Reduced MTR values were found in the cervical spinal cord of 60 patients with early RRMS when compared with 34 age-matched controls, in the absence of spinal cord atrophy [336]. This study also showed that the main contribution to low MTR levels is from the normal-appearing spinal cord tissue, since the effect of the lesions is minimal [336]. In patients, there was a correlation between lower MTR and higher lesion load [336]. Evidence for reduced MTR values was also found in the cervical cord of patients with RIS [317], although this finding requires further confirmation. Of note, the distribution of MTR reduction in the spinal cord supports a spatial pattern of microstructural damage that resembles that in the brain [336, 337], and suggests that MTR abnormalities in a region involving the pia mater and subpial cord occur early in the course of MS and are more marked in those with a progressive course [338]. Clinical correlates of MTR are reported in **Table 2.2**.

Myelin water imaging has been validated as a myelin marker [339, 340], but it has been applied to the spinal cord in only a few studies. Myelin water fraction varies along the spinal cord proportionally to the WM area fraction [341]. In patients with cervical spondylotic myelopathy, myelin water imaging shows high specificity in detecting impaired spinal cord conduction, when compared with conventional imaging (e.g., T2 signal intensity), which only provides a measurement of the extent of spinal cord compression [342]. In MS, spinal cord myelin water fraction decreased by 11% in PPMS, but not in healthy subjects [343], and was associated with disability scores (**Table 2.2**) [344], suggesting progressive demyelination and neuro-axonal loss in this disease subtype, that is related to progressive disability accrual. Cervical spinal cord myelin water volume fraction progressively decreases in MS, but not in NMOSD, in the absence of clinical relapses, suggesting that

neurodegenerative and demyelinating processes occur continuously in MS, but not in NMOSD where inflammation might dominate [345].

MTR and myelin water imaging appear to provide complementary information. Although MTR is not pathologically specific, it is commonly available and fast to acquire, whilst myelin water imaging is more specific as a myelin marker, but requires more complicated post-processing and is not a sequence product.

A recent paper investigated the role of spinal cord quantitative MT saturation technique (MTSat), minimally affected by T1 relaxation and field inhomogeneity, and demonstrated that MTSat correlates with disability more strongly than MTR, suggesting higher sensitivity to tissue damage for future clinical applications (**Table 2.2**) [346].

Finally, quantitative magnetisation transfer (qMT) applied at 3T with reasonable acquisition time showed excellent grey/white contrast and sensitivity to MS pathology (lesions) [347].

2.2.4.3. Metabolic imaging techniques

^{1}H-MRS estimates levels of metabolites, as long as they are available in relatively high concentrations [191, 348, 349]. The most commonly estimated metabolites are total NAA (a marker of neuronal metabolism and, more in general, of neuronal integrity); NAA/Cr (NAA values normalised by voxel creatinine (Cr)); glutamate and its precursor glutamine (Glx) (whose reduction indicates chronic neuro-axonal degeneration); and *myo*-inositol (a marker of glial cell activation and proliferation) [191, 349]. A recent investigation of MRS at 3T using *in vivo* and *post-mortem*

experiments reported an extended metabolic profile of the spinal cord, thereby indicating the rich information which can be provided by this technique [350].

When MRS was applied to early PPMS patients, lower NAA, lower Glx, and higher *myo*-inositol, in particular within lesions, were detected in patients compared with normal-appearing tissue in healthy controls [326]. Metabolic changes within the spinal cord occurred in the absence of significant spinal cord atrophy, pointing towards early neuro-axonal loss and tissue remodelling, and were associated with disability measures [326]. When including patients at different disease stages, lower NAA/Cr was associated with spinal cord atrophy and with disability progression during follow-up (**Table 2.2**) [351]. Also, diffuse lesions were characterised by lower NAA/Cr when compared with focal lesions [351].

MRS may assist with the differential diagnosis of myelopathies. It has been used to define the metabolic profile of different spinal cord tumours (e.g., strongly reduced NAA and strongly increased myo-inositol in the ependymoma, or absence of significant metabolic changes in extradural tumours, such as the schwannoma), and traumatic spinal cord injury (reduced NAA) [352].

^{23}Na-MRS has been investigated in the brains of MS patients in several studies [353, 354], that have shown increased total sodium concentration in the MS lesions and normal appearing tissue in patients when compared with controls, suggesting either an expansion of the extracellular compartment as a consequence of neuro-axonal loss, or an accumulation of sodium in the swollen axonal terminals with ongoing degeneration [191]. Advances in sodium imaging acquisition and

Table 2.2. Pathological specificity of advanced spinal cord MRI and clinical correlates.

Table shows pathophysiologic mechanism of MS that can be studied with different advanced MRI techniques. Changes occurring in different MS subtypes and clinical correlates are presented. Adapted from Moccia M et al. Ther Adv Neurol Disord 2019 [109].

Pathophysiologic mechanisms	**Advanced MRI technique**	**Changes in MS** *(compared with controls)*	**Clinical correlates** *(if abnormal)*	**Reference**
Neuro-axonal integrity	**DTI** *(FA, RD, MD)*	MD, FA = in RIS RD ↑↑↑ in RRMS and SPMS MD ↑↑↑ in RRMS and SPMS FA ↓↓↓ in RRMS and SPMS	EDSS Upper limb function Lower limb function	[227, 317, 318]
	QSI	ADCxy, FWHMxy, P0xy ↑↑↑ in PPMS ADCz, FWHMz, P0z ↓↓↓ in PPMS	Spasticity Postural instability Sensory dysfunction	[326, 327]
	NODDI *(v_{in}, ODI)*	Intra-neurite volume fraction ↓↓↓ in RRMS lesions ODI ↑↑↑ in RRMS normal-appearing WM ODI ↑↑↑ in RRMS lesions		[331, 332]

	SMT *(V_{ax})*	Axonal volume fraction ↓↓↓ in RRMS		[333]
	MRS *(NAA/Cr)*	↓↓↓ in PPMS, RRMS and SPMS	EDSS Upper limb function Lower limb function	[351]
	MRS *(NAA)*	↓↓↓ in PPMS	EDSS Spasticity Postural instability Sensory dysfunction	[326]
	MRS *(Glx)*	↓↓↓ in PPMS	Postural instability	[326]
Myelin content	**MTR**	↓ in RIS ↓↓↓ in RRMS	EDSS	[317, 335, 336]
	Myelin water fraction	↓↓↓ in PPMS ↓↓↓ in RRMS	EDSS Lower limb function	[343, 345]
	MTSat	n/a	EDSS Lower limb function	[346]

Astrocytic activation and proliferation	**MRS** *(Myo-inositol)*	↑↑↑ in PPMS lesions	Postural instability	[326]
Functional connectivity	**BOLD fMRI**	↓↓↓ in RRMS lesions ↑↑↑ in RRMS peri-lesional area		[355]

analysis have allowed application of this metabolic technique to the spinal cord of healthy subjects [356], and patients with MS [357]; preliminary findings mirror brain results, with higher total sodium concentration in MS patients than healthy subjects.

2.2.4.4. Functional MRI

Very few studies have investigated the resting state blood oxygenation level-dependent (BOLD) signal in the spinal cord on fMRI mainly because its sensitivity and reliability are still suboptimal and technical limitations are significant [358]. In a 7T fMRI study measuring the BOLD signal, spinal cord functional networks were generally intact in RRMS (**Table 2.2**). However, increased connectivity was found at the boundaries of lesions, possibly indicating compensatory changes to demyelination/axonal loss, and/or disruption of inhibitory spinal interneurons [355].

2.3 Design issues

2.3.1. Sample size

Sample size calculation is a pivotal aspect of planning clinical trials and is based on the primary outcome of the study, generally being imaging for phase 2 and clinical for phase 3 trials [359, 360]. Imaging outcome measures are often included as secondary or exploratory variables in all patients in phase 3 clinical trials, even though they might require a smaller sample size to detect significant difference. A caveat though is that the size of the treatment effect may differ between clinical and imaging outcomes; i.e. 30% reduction in rate of brain volume may not equate in 30% reduction in disability progression.

In order to further explore this issue, I have estimated the treatment effect on brain atrophy which could have been detected in populations recruited in phase 3 clinical trials in progressive MS (**Table**

2.3), based on the actual sample size and the measured rates of brain atrophy in the placebo arm (accepting a power of 80% and an α-error of 5%). Most recent studies would have been able to detect 15-30% treatment effects on brain atrophy [118, 120, 138, 143], in line with actual detected statistical effect (percent brain volume change (PBVC) relative difference was 17% in the ORATORIO trial and 15% in the EXPAND trial).

Inclusion criteria can also impact on the sample size. For instance, in RRMS populations, the rate of inflammatory activity is high, and measures of inflammatory activity (new or enlarging T2 lesions, new T1 lesions and Gadolinium-enhancing lesions) can lead towards relatively lower sample sizes, compared with progressive MS [361, 362]. When considering the number of enhancing lesions, the detection of 50% treatment effect for interferon-beta treatment requires about 120 patients per arm in RRMS trials, and a three-fold number in SPMS [363]. By contrast, use of imaging markers more specific for progressive features (e.g., brain atrophy) will reduce the sample size needed in clinical trials in PPMS and SPMS. Advanced MRI techniques, such as MTR, might also require smaller sample sizes [364, 365], in particular when trials with neuroprotective agents are conducted in selected populations.

Sample size can be affected also by variability of imaging outcomes, which is directly related to the standard deviation of the measure, a major determinant of the sample size [158]. Increasing the number of scans performed in clinical trials can reduce measurement variability and, accordingly, sample size. However, at least for brain atrophy, the effect of increasing the number of observations is modest, when compared with the effect of increasing the duration of follow-up [289].

Overall, DMTs can have a specific effect on each MRI outcome and thus the sample size should be estimated depending on the expected efficacy profile in the selected population. As such, MRI may be particularly useful in early-phase clinical studies on novel therapeutic agents, where drugs can be easily screened before they are taken forward to larger-scale studies [366], as is common practice for anti-inflammatory drugs in phase 2 RRMS studies.

2.3.2. Measurement sensitivity

Quantitative MRI measures are strongly dependent not only on acquisition parameters, but also on processing methods, presenting with different sensitivity to change, repeatability and measurement error.

Brain atrophy measures have been a cornerstone in the study of interventions with putative neuroprotective effects [63, 120, 143], because of the application of registration-based methods that provide *direct* estimates of brain atrophy, such as the SIENA [367], and the boundary shift integral (BSI) method [368–370]. Both SIENA and BSI have reduced sample size requirements to detect significant differences between groups or over time, and are nowadays well-established methods to measure longitudinal brain atrophy in clinical trials and in observational studies for neurodegenerative diseases [289]. In particular, clinical trial results can be affected by different analyses [157]. For instance, in a clinical trial of teriflunomide in RRMS, changes were measured by BPF, a segmentation-based technique, and no significant effect was found initially [202]. However, in a post-hoc analysis, the use of a registration-based automated technique (SIENA) revealed that teriflunomide was associated with significant reductions in brain volume loss [371].

Table 2.3. Phase 3 clinical trials in progressive MS evaluating brain atrophy.

Table shows phase 3 clinical trials in progressive MS which included brain atrophy as outcome measure. Characteristics of trials (sample size, duration) and of MRI measures (results and technique applied) are reported. The potentially-detectable effect size has been calculated based on placebo arm results and sample. Adapted from Moccia M et al. Mult Scler 2017 [91].

Clinical trials	**Sample size recruited** *(treatment* vs *placebo)*	**Volume change** *(treatment* vs *placebo)*	**Duration**	**Methods**	**Effect size potentially detectable**
EXPAND, Kappos et al. *Lancet* 2018 [143]	470 vs 239 (SPMS)	PBVC -0.71%±0.77 vs -0.84%±0.71 (p=0.02)	From baseline to month 24	SIENA	21.6%
ORATORIO, Montalban et al. *New Eng J Med* 2017 [164]	488 vs 244 (PPMS)	PBVC -0.90%±1.12 vs -1.09%±1.15 (p=0.02)	From week 24 to 120	SIENA	26.7%
INFORMS, Lublin et al. *Lancet* 2016 [118]	336 vs 487 (PPMS)	PBVC -1.49%±1.35 vs -1.53%±1.35 (p=0.673)	From baseline to month 36	SIENA	15.8%

CUPID, Zajicek et al., *Lancet Neurol* 2013 [138]	329 vs 164 (182 vs 91 in the MRI sub-study population) (PPMS and SPMS)	PBVC -1.95%±1.51 vs -1.82%±1.47 (p=0.94)	From baseline to year 3	SIENA	33.5%
OLYMPUS, Hawker et al. *Ann Neurol* 2009 [103]	292 vs 147 (PPMS)	Volume change -10.8cm^3±40.3 vs -9.9cm^3±37.0 (p=0.62)	From baseline to week 96	BPF	>99%
ESIMS, Fazekas et al. *Mult Scler* 2005 [372]	159 vs 159 (SPMS)	PCV -0.62%±0.88 vs -0.88%±0.91 (p=0.0093)	From baseline to month 27	Six-slice volume	32.5%

On the contrary, the current status of spinal cord atrophy measurement is suboptimal. Spinal cord atrophy is conventionally determined indirectly, by numerical subtraction of volume (based on 3D surface fitting) or cross-sectional area (CSA) (based on 2D edge detection on serial images) obtained separately at each time-point, providing *indirect* estimates of atrophy rates [278, 310, 321, 358, 373], that could introduce a lot of noise due to the miss-segmentations. Thus, the development and optimisation of registration-based techniques for spinal cord atrophy may reduce measurement noise and improve its precision [374], and is the main goal of this thesis.

3. Rationale and aims of thesis

3.1. Rationale

As fully discussed in Chapter 1, MS is a common neurological disorder in young adults, with a highly variable clinical course, but potentially leading to chronic disability and poor quality of life [1]. Thus, as mentioned in Chapter 2, identifying reliable imaging outcome measures is a cornerstone for improving the understanding of the disease mechanisms and for monitoring the clinical course of MS and its response to treatment [63, 109]. Progressive MS represents a unique opportunity for studying imaging markers of neurodegeneration, with equal bearing on relapsing forms of the disease. Several imaging candidates hold promise for filling the unmet need of biomarkers in progressive MS, by capturing the effect on neurodegeneration [91].

The study of atrophy looks particularly promising, by the *in vivo* characterisation of most aggressive aspects of MS pathology (e.g., demyelination and neuro-axonal loss). So far, brain atrophy is the best examined and most robust outcome with attainable sample sizes and first positive results, though treatment effects tend to be more modest than those seen for inflammatory MRI markers (i.e., lesions) [91]. Brain volume has already being used as primary outcome measure in phase 2 trials, and as secondary exploratory measure in phase 3 trials in progressive MS [105]. However, spinal cord MRI holds great promise for future trials due to higher rates of atrophy and better sensitivity to change compared with brain volume changes (-1.7%/year spinal cord volume loss vs -0.6%/year brain volume loss) [149, 268]. As discussed in Chapter 2, robust application of spinal cord atrophy in clinical trials requires implementation of registration-based techniques with lower measurement noise and higher sensitivity to change, when compared with currently available segmentation-based methods.

3.2. Aims

In keep with this background, this thesis will primarily aim to:

(1) Validate a registration-based method for the quantification of spinal cord atrophy (generalised boundary shift integral (GBSI));

(2) Evaluate clinical correlates of this novel registration-based method;

(3) Evaluate possible future implication for the design of clinical trials and observational studies.

The pipeline of the GBSI, along with detailed methods to overcome the limitations posed by current segmentation-based atrophy measurements, and the observational cohorts and clinical trial used for testing and implementing this technique, are fully described in Chapter 4.

In Chapter 5, I preliminarily applied this novel registration-based method to a monocentric test dataset and, then, to a multi-centre cohort acquired across different European sites with specific MRI expertise, within the Magnetic Resonance Imaging in Multiple Sclerosis (MAGNIMS) network. Hereby, I have compared precision and repeatability of spinal cord atrophy measurements obtained with different segmentation methods (e.g., JIM and SCT), and with the novel registration-based method, and, then, explored possible clinical correlates.

In Chapter 6, I re-analysed a phase 2 clinical trial on PPMS to further evaluate clinical correlates and treatment effect of this novel registration-based method, and possible implications for image acquisition and clinical trial design (e.g., measurement precision, image quality, deriving spinal cord atrophy measurements from brain scans).

Finally, in Chapter 7, I have summarised the results, highlighting novel contributions of this thesis to the field and possible future directions in MS research.

4. Methods

In Chapter 4, I report on general methods for the present thesis, including study design, populations and techniques for cross-sectional and longitudinal spinal cord atrophy measurements. In particular, I am hereby presenting general methods, but will leave more specific methods to Chapter 5 and Chapter 6 (e.g., inclusion and exclusion criteria, demographic and clinical features). Also, I present detailed methods for cross-sectional and longitudinal spinal cord atrophy measurements, but will then discuss specific applications, in relation to different research questions, in Chapter 5 and Chapter 6.

Looking at specific contributions to this thesis, I need to acknowledge that the first GBSI pipeline for spinal cord atrophy measurement was developed by Prados and colleagues, as the adaptation of a method used for brain atrophy to measure longitudinal spinal cord atrophy [374, 375]. I have then applied this preliminary pipeline to different cohorts of MS patients and healthy controls (as described in Chapter 5 and Chapter 6). While performing spinal cord segmentation, as detailed below, and GBSI pipeline, as from Prados and colleagues [374, 375], I regularly extracted data, and run statistical models and graphical presentations, in order to evaluate overall GBSI performance, and to define possible areas of improvement. Based on this, Dr Prados and I further implemented the pipeline (e.g, we improved definition of the probabilistic XOR masks, we included denoising within the pipeline). As such, the GBSI pipeline presented in Chapter 4 is the result of the analyses I run, and subsequent changes in collaboration with Dr Prados.

4.1. Study design

The present thesis is based on retrospective analysis of prospectively collected data from observational studies conducted at the Queen Square MS Centre ([228, 250, 278]MAGNIMS

Collaboration network (www.magnims.eu) [113, 273, 287], and from a phase 2 clinical trial (A Randomized Placebo-controlled trial Evaluating Laquinimod in primary progressive multiple sclerosis, Gauging Gradations In MRI and clinical Outcomes (ARPEGGIO)) [105].

The study was conducted in accordance with the Declaration of Helsinki. Each participant had provided written consent for research within each centre, for the main study and its subsequent analyses, with local ethics committees approval. All scans were pseudo-anonymised, transferred, stored and analysed in accordance with current regulations (EU GDPR 2016/679).

4.2. Populations

4.2.1. London cohorts

London cohorts were composed of three different cohorts with longitudinal spinal cord MRI acquisitions from the Queen Square MS Centre (The London cohorts have been fully reported in previous publications [228, 250, 278], and will be briefly presented below and in subsequent Chapter 5.

4.2.1.1. Test dataset

The test dataset includes spinal cord images from 9 healthy controls and 9 MS patients, who were scanned twice within the same session, with repositioning of the subjects in-between acquisitions [284]. This dataset has been previously used to compare a fully automated spinal cord segmentation method (*PropSeg*), with the semi-automated ASM tool within JIM [284].

Spinal cord MRI was done using a 3T Philips Achieva MRI system with RF dual-transmit technology (Philips Medical Systems, Best, Netherlands) and a 16-channel neurovascular coil. Acquisition

protocol is reported in Chapter 5, and further details are available from previous publications [113, 284].

4.2.1.2. CIS cohort

The CIS cohort has been studied at the Queen Square MS Centre in the past 20 years. This cohort includes 131 non-spinal CIS patients who underwent spinal cord MRI at diagnosis (between 1995 and 2004), and after an average period of 5 years; demographic and clinical data were recorded at the two time points [250]. Twenty healthy controls were also included. This cohort has been previously used to study MRI predictors of subsequent MS progression [250].

Spinal cord MRI was done at baseline and follow-up on the same 1.5T Signa scanner (General Electric, WI, USA); there was a scanner upgrade during the study period that, however, in the previous studies, did not seem to affect statistical analyses [250]. Acquisition protocol is reported in Chapter 5, and further details are available from previous publications [113, 250].

4.2.1.3. Spinal cord cohort

The spinal cord cohort includes spinal cord images from 41 MS patients (CIS, RRMS, SPMS, PPMS) and 17 healthy controls, who underwent spinal cord MRI at baseline and after one year; demographic and clinical data were recorded at the two time points [228, 278]. This cohort has been previously used to validate the ASM tool within JIM, and to study possible associations between spinal cord atrophy and other clinical and MRI variables [228, 278].

Spinal cord MRI was done at baseline and follow-up on the same 3T Philips Achieva MRI system with RF multi-transmit technology (Philips Healthcare, Best, the Netherlands), and a 16-channel

neurovascular coil. Acquisition protocol is reported in Chapter 5, and further details are available from previous publications [228, 278].

4.2.2. MAGNIMS cohort

The MAGNIMS is a European network of academics that share a common interest in the study of MS using MRI techniques (www.magnims.eu). Eight MAGNIMS centres have collaborated to acquire spinal cord images from 155 MS patients (CIS, RRMS, SPMS, PPMS) and 59 healthy controls (collected by Professor Mara Rocca and Professor Massimo Filippi, Hospital San Raffaele, Milan); MRIs were acquired at baseline and after one year; demographic and clinical data were recorded at the two time points [113, 273, 287].

MRIs were acquired on multi-producer 3T scans; acquisition protocols are reported in Chapter 5. Inclusion and exclusion criteria have been reported previously [113, 273, 287], and will be better discussed in Chapter 5.

4.2.3. ARPEGGIO clinical trial

The ARPEGGIO trial was a randomised, double-blind, parallel-group, placebo-controlled study. PPMS patients were randomised in a 1:1:1 ratio to receive oral laquinimod 0.6 mg or 1.5 mg, or placebo once daily from January 2015 to April 2016 at 85 sites in 10 countries. Duration of the core study was 48 weeks. The laquinimod 1.5 dose arm was discontinued as of January 1, 2016, due to cardiovascular side effects (patients were followed-up, but no further treatment was given) [105].

All patients underwent brain and cervical spinal cord MRI at baseline, week 24 and week 48, including 3D T1-weighted isotropic images of the brain and spinal cord (1x1x1mm^3), within 14 days

of the scheduled clinical visit. MRI scans were collected centrally at the VUmc in Amsterdam and, for the purposes of the present thesis, were then transferred to the NMR Research Unit (. Inclusion and exclusion criteria have been reported previously [105], and will be better discussed in Chapter 6.

In the original clinical trial dataset, a number of clinical tests (EDSS, T25FW, 9HPT, SDMT and MSWS), and MRI variables (e.g., number of new T2 lesions, T2 lesion volume change, T1 lesion volume change and PBVC) were included. Full details of acquisitions and processing have been previously reported [105].

4.3. Cross-sectional spinal cord measurements

In the present thesis, we used two different methods for cross-sectional spinal cord image segmentation and atrophy calculation: surface-based and image-based methods [276]. The older methods are fully manual, whilst the most recent methods are semi- or fully-automated.

4.3.1. Spinal cord toolbox

When the work related to this thesis started (Chapter 5), *PropSeg* was the standard segmentation tool within the SCT (SCT version 3.0), after being tested on a large cohort of MS patients and healthy controls [283, 284, 376]. This fully-automated surface-based image segmentation method has the same sensitivity as the ASM within JIM, but has higher inter-rater repeatability and is more time-efficient [284]. The most recent release of the SCT (version 4.0) includes the *DeepSeg* tool, a fully-automated segmentation method based on convolutional neural networks [283, 285], which was used in Chapter 6. The *DeepSeg* also allows automated spinal cord segmentation from brain MRI images, which was not validated on the *PropSeg* [283, 285].

4.3.2. JIM

Within the JIM tool, the ASM is a surface-based method that semi-automatically outlines the cord, after marking the centre of the spinal cord [274, 277]. The ASM has provided more prompt and repeatable measures of the spinal cord volume, compared with manual methods [278]. The ASM offers a considerable reduction in user interaction time, and can be performed over long spinal segments. The user needs to identify landmarks at the extremes of the region to study, and, then, mark the centreline of the cord. Sagittally-acquired images are then reformatted to the axial plane to obtain five contiguous 3 mm slices; the program automatically calculates the radius and the centre of each axial slice and, finally, the cross-sectional area is obtained by averaging these contiguous slices [274].

4.3.3. NeuroQLab

The NeuroQLab is an image-based method that segments the upper cervical cord from surrounding non-spinal cord tissue by using a Gaussian mixture modelling [275, 281, 282]. After definition of the spinal cord subsection to be segmented, watershed segmentation of the spinal cavity and surrounding cerebrospinal fluid is performed, and mean upper cervical cord area (MUCCA) is computed [275, 281, 282].

4.4. Longitudinal spinal cord measurements

4.4.1. Segmentation methods

When using segmentation methods for longitudinal spinal cord atrophy calculation, the rate of atrophy is estimated by numerical subtraction of spinal cord cross-sectional area measurements calculated at different time-points. For instance, percent change of cord area is calculated using the

following formula: Atrophy = 100*(follow-up area - baseline area)/baseline area. Years between baseline and follow-up scans can also be included in the denominator of the formula, if there is variability of the interval between scans [109, 113].

4.4.2. Registration method (GBSI)

The GBSI is applicable to datasets with T1-weighted images, with identical acquisition parameters, ideally using 1 mm isometric voxels, at two time-points for each subject. A graphic overview of the pipeline is presented in **Figure 4.1** and further detailed below.

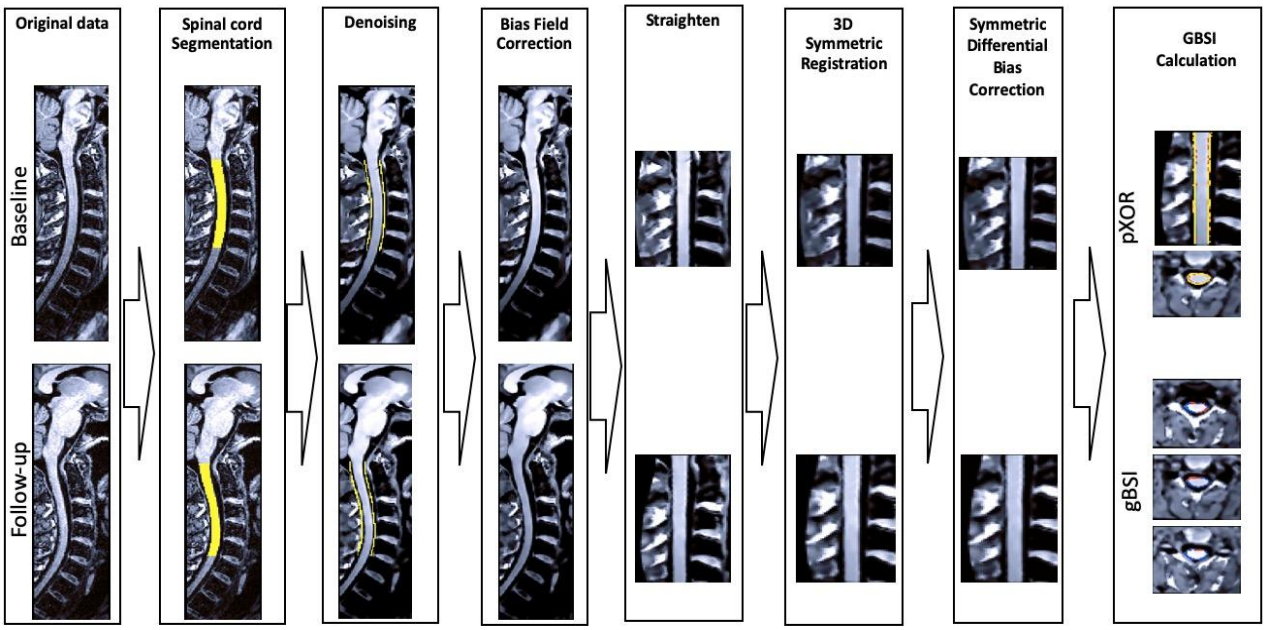
4.4.2.1. Spinal cord segmentation

The first step is the segmentation of the spinal cord from T1-weighted images [113, 283, 373]. The whole cord is segmented (i.e., WM and GM together), defining the spinal cord boundaries and the cranio-caudal extension of tissue over which the GBSI estimates are required. This segmentation is computed separately and independently for each time-point, over the same spinal cord segments; for instance, in the present thesis I will work on C2-C5 segmentation in Chapter 5, and, then, on C1-2 and C2-5 segmentations in Chapter 6. The segmentation can be obtained manually, semi- or fully-automatically using the wide range of techniques, as detailed above. The extracted spinal cord segmentation can be represented with a hard (binary) or a soft (probabilistic) mask. Considering that the acquired spinal cord extension can vary by a few slices between time-points (e.g., due to positioning of the subject in the scanner), the longitudinal atrophy is computed over the

Figure 4.1. Pipeline for spinal cord longitudinal atrophy measurement using the registration-based technique (GBSI).

After spinal cord segmentation, baseline and follow-up spinal cord images are denoised, bias-field corrected, straightened and ultimately registered to the halfway space. Intensity changes in the

vicinity of the cord boundaries are estimated for GBSI calculation. Adapted from Prados F, Moccia M, et al. Submitted [374, 375].



intersection of these regions of interest, discarding the pixels that are not covered at both time-points.

4.4.2.2. Image denoising

The extracted masks are used to compute a ring surrounding the spinal cord to scale the signal intensity of the images accounting for the presence of the noise floor [377]. For this step, the signal intensities in the original T1-weighted images are denoised using a fast version of the adaptive non-local means filter algorithm, using the mask from the segmented spinal cord [378]. To be more precise, the standard deviation of the signal in a ring within the CSF was calculated to compute the root power of the noise, and, then, used to account for the presence of a noise floor (this approach is standard on T1-weighted images, where this can be obtained from regions where the signal from the CSF is suppressed) [377]. Any value of voxels within the extracted ring that were more than 2 standard deviations above the mean were discarded, in order to avoid the inclusion of values from

nerve roots or other spurious signal intensities. Finally, the ring within the CSF was derived by dilation of the spinal cord mask by 2-pixel layers, and, then, by subsequent subtraction of the mask once.

4.4.2.3. Inhomogeneity correction

An intensity inhomogeneity correction is applied to the 3D data using the N4 algorithm [379], within the region determined by the two-time dilatation of spinal cord masks. The following parameters are used: full width at half maximum (FWHM)=0.05, convergence threshold=0.0001, and maximum number of iterations=1000 [370].

4.4.2.4. Spinal cord straightening

To remove any difference in the cord curvature between time-points resulting from positioning in the scanner, straightening of the MRI images is obtained from the previously computed spinal cord segmentation, using a robust and accurate method, freely available as part of the SCT software package [283, 380] This method preserves spinal cord topology, which is essential for measuring subtle changes in spinal cord edges when using GBSI.

4.4.2.5. Half-way space registration

Images are registered to the half-way space using an affine transformation in order to avoid biases that would be introduced if registering one time-point to the other [381–383]. This step is achieved using an inverse-consistent and symmetric algorithm [384]. Once the transformations are obtained, images and corresponding masks for each time-point are linearly resampled to the common half-way space.

4.4.2.6. Differential bias correction

A longitudinal differential bias correction method is used to remove the residual intensity inhomogeneity-derived differences between the baseline and the repeated images (radius=5 voxels) [385]. This step is needed, despite the previous cross-sectional inhomogeneity correction, to avoid artificial atrophy values by the remaining intensity differences from the cross-sectional inhomogeneity correction method.

4.4.2.7. Intensity normalisation

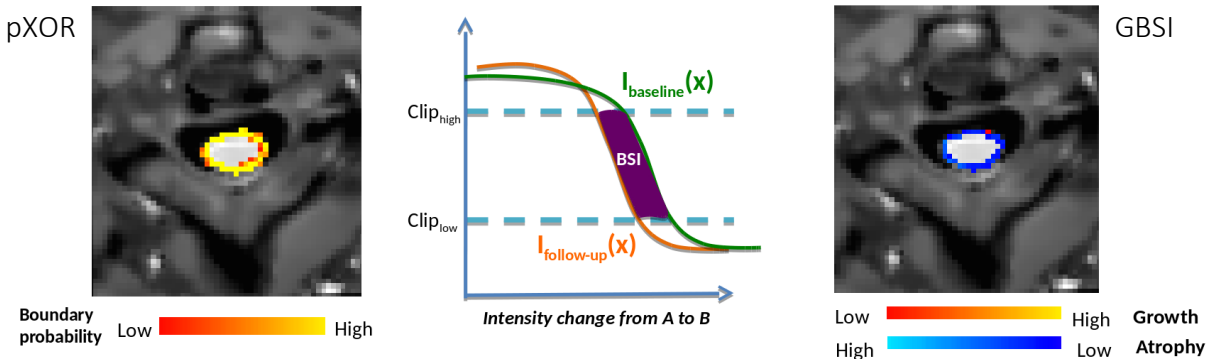
Image intensities are normalised so that the probabilistic area from which GBSI will be computed can be extracted [369, 370]. The intensity normalisation of the baseline and follow-up images is obtained from a linear regression between the average tissue intensity inside the cord and inside the CSF. The tissue intensity values are computed using a k-means algorithm (k=2), which is limited to a region of interest obtained from each input mask (after 2 dilations).

4.4.2.8. Probabilistic XOR

The probabilistic XOR mask is obtained from the half-way linearly-resampled segmentation masks following the same steps already introduced for the brain GBSI calculations (**Figure 4.2**) [370]. This mask identifies the voxels with high probability to be tissue at the edge of the cord.

Figure 4.2. GBSI calculation.

pXOR mask is shown on the left, the intensity profile function that computes the GBSI is shown in the middle and the GBSI results are displayed on the right. Adapted from Prados F, et al. ISMRM 2016 [375].



4.4.2.9. Atrophy computation

The GBSI is computed on a voxel-by-voxel basis as the difference in intensity between baseline and follow-up images within a clipped window [368–370]. The clipped window aims to catch the difference between tissue intensities at the two time-points, by reducing the background influence. The intensity differences are weighted by the probabilistic XOR mask voxel-wise. For the spinal cord GBSI, a predetermined clipping window was used [374, 375]. To increase robustness, the "forward" and "backward" BSI is calculated for each pair of images (i.e., swapping baseline and follow-up images, and repeating the intensity normalisation, probabilistic XOR and atrophy computation steps), and the mean of the results is included. Finally, percent spinal cord volume change is calculated by dividing the GBSI value by the binarised, straightened and registered baseline cord mask volume.

4.4.3. Software

The N4, denoising, differential bias correction, registration and GBSI methods are all freely available as part of NifTK software package at https://cmiclab.cs.ucl.ac.uk/CMIC/NifTK. SCT has been used for spinal cord straightening, and it can be found at https://github.com/neuropoly/spinalcordtoolbox. In this thesis, the GBSI has been implemented using these software packages; however, other software packages offer similar processing options and could be used alternatively.

5. Longitudinal spinal cord atrophy in multiple sclerosis using the generalised boundary shift integral

5.1. Background

As previously discussed in Chapter 2, spinal cord atrophy on MRI is a marker of neurodegeneration in MS [226, 260], and is one of the main substrates of long-term disease progression [109, 213, 227, 250, 296, 305, 308, 311, 386]. Spinal cord atrophy progresses faster than brain atrophy (1.7%/year vs 0.4-0.6%/year), is greater in progressive MS than in the relapsing forms of MS, and predicts disability [268, 305, 312]. It is crucial to obtain an accurate and precise longitudinal measurement of spinal cord atrophy, because it could be used to monitor disease progression and become a primary outcome measure in phase 2 clinical trials with neuroprotective therapies, not only in MS, but also in other neurodegenerative disorders [91, 108, 300, 387].

Spinal cord atrophy is conventionally estimated with segmentation-based methods (e.g., CSA), applied to volumetric spinal cord images [278], that measure cord characteristics at each time point; indirect longitudinal atrophy measurements are obtained by numerical subtraction, with relatively-low repeatability and responsiveness to change [109, 292]. On the contrary, longitudinal brain atrophy measurements are nowadays based on registration-based techniques that significantly reduce measurement noise [289].

As such, in Chapter 5, I have applied the GBSI pipeline fully described in Chapter 4 to measure longitudinal spinal cord atrophy in cohorts of MS patients and healthy controls. I have performed spinal cord segmentation, GBSI calculation, and, then, statistical analyses.

5.2. Aims

The present chapter reports on a multicentre, multi-manufacturer and multi-field strength scan analysis, and aims to:

(1) Evaluate precision and repeatability of spinal cord atrophy measurements obtained with different segmentation methods (e.g., JIM and SCT), and with the registration-based GBSI;

(2) Compare measurements of spinal cord atrophy obtained using GBSI with those obtained with conventional CSA (semi-automatic segmentation with the SCT);

(3) Explore associations between GBSI- and CSA-derived spinal cord measurements and MS clinical features;

(4) Estimate the sample size needed to detect changes in spinal cord atrophy over one year using GBSI and CSA.

5.3. Methods

5.3.1. Study design and population

In this chapter, I performed a retrospective analysis of prospectively collected data from Queen Square MS Centre (MAGNIMS centres. Overall, I included a test dataset composed of 9 healthy controls and 9 MS patients, and a validation dataset composed of 327 MS patients and 96 healthy controls. The London and MAGNIMS cohorts have been reported in previous publications and in Chapter 4 [228, 250, 273, 278, 287].

Eligibility criteria were:

(1) Diagnosis of CIS or MS according to the 2010 McDonald Criteria [59], or healthy controls without history of neurological or psychiatric disorders;

(2) The presence of at least two volumetric MRI scans (interval between scans was collected), acquired with isotropic voxel of 1x1x1mm^3;

(3) Information on: EDSS score at each time point [66], subtype of MS (CIS, RRMS and progressive MS (including both PPMS and SPMS)) [59], age, gender and disease duration (time from clinical onset to baseline MRI).

Each participant had provided a written consent for research within each centre. The final protocol for the analysis of pseudo-anonymised scans, acquired independently and prospectively in each centre, was approved by the European MAGNIMS collaboration and by the local ethics committees.

5.3.2. MRI acquisition and processing

Dedicated cervical spinal cord 3DT1-weighted images (1x1x1mm^3) were analysed. Images were acquired in 8 MAGNIMS sites on 1.5T and 3T scanners, from different manufacturers and with different MRI parameters (**Table 5.1**).

In the whole population, for calculating CSA and GBSI, masks of C2-5 spinal cord level were obtained for images acquired at each time point with SCT, using the routine method known as *PropSeg* (version 3.1.1) (**Figure 5.1**), that has been fully explained in Chapter 2 and Chapter 4 [283]. C2-5 CSA was obtained by averaging all cord cross-sectional areas. For GBSI, I followed the previously described pipeline, which is applicable to datasets with identical image acquisitions (i.e., T1-weighted), and similar voxel size, with acquisitions at two time-points for each subject [370, 374, 375]. Briefly, the extracted segmentation masks were used to compute a ring surrounding the spinal cord to scale the signal intensity of the images accounting for the presence of the noise floor [377]. For this step, the whole 3D volume was corrected for signal intensities, using a fast version of the

Table 5.1. London and MAGNIMS cohorts and included patients.

Site	*Number of included individuals*	*Manufacturer, model and field strength*	*Acquisition parameters (TR/TE/TI msec)*
Barcelona	21	Siemens Magnetom Trio 3T	2300/3.2/900
Bochum	4	Philips Achieva 3T	8/"shortest"/"shortest"
London	123	GE Signa 1.5T	15.6/4.2/450
	55	Philips Achieva 3T	8/3.7/860
	18	Philips Achieva 3T	8/3.7/860
Lugano	43	Siemens Magnetom Skyra 3T	2300/5.1/1140
Mannheim	14	Siemens Magnetom Skyra 3T	1900/2.6/1000
Milan	49	Philips Achieva 3T	8/"shortest"/"shortest"
Naples	28	General Electric Signa HDtx 3T	7.9/3.2/450
Oxford	15	Siemens Magnetom Prisma 3T	2300/3.59/900

adaptive non-local means filter algorithm [378], and, for intensity inhomogeneities, using the N4 algorithm [379]. Spinal cord images were then straightened using a specific software available within the SCT [388], to remove the difference in curvature between time-points due to subject positioning in the scanner. A 3D symmetric and inverse-consistent rigid-only (9 degrees of freedom (DOF)) registration to the half-way space between baseline and follow-up images was performed [384]; masks were resampled to the same space using linear interpolation and registered to the halfway space. To reduce the residual bias field and homogenise the grey scale between both registered time-points, a symmetric differential bias correction was applied [385]. Overall, this method does not generate any bias between the baseline and follow-up images as the exact same

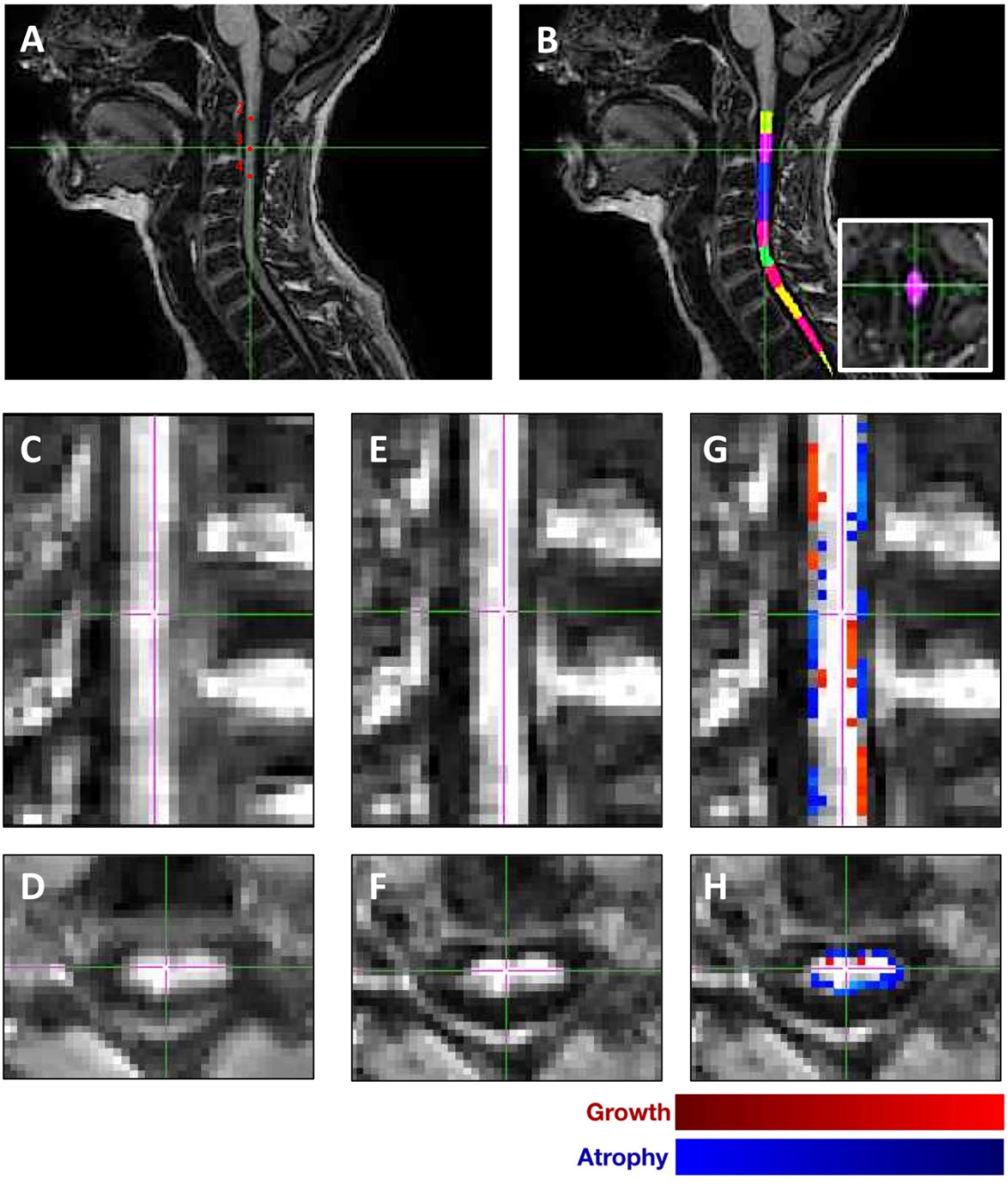
image processing pipeline is applied to both time points. The probabilistic XOR (pXOR) region was adaptively estimated from baseline and follow-up cord segmentations, to calculate the BSI over specific tissue boundaries, excluding voxels in areas that might reduce its sensitivity (e.g., voxels with partial volume with CSF or the centre voxels of the spinal cord). The GBSI was finally computed (**Figure 5.1**).

Percent CSA annual change between time points was calculated using the following formula: 100*[(CSA at follow-up – CSA at baseline)/CSA at baseline)/years between baseline and follow-up scans]. Similarly, percent GBSI annual change was calculated using the following formula: GBSI/ years between baseline and follow-up scans. Of note, spinal cord atrophy based on CSA is measured as the average of the slice-wise 2D edge detection over a fixed cord section/length, hence effectively it is a 3D volume divided by its height; whilst GBSI measures the voxel-by-voxel difference in intensities between the baseline and the follow-up images, as a 3D volume change [113]. I segmented the spinal cord over a fixed cord length (C2-5), and I used the same masks to compute CSA and GBSI, thus making the atrophy measurement effectively computed over the same boundary/region for both measurements [292].

To determine measurement precision and repeatability, I analysed test dataset from 9 healthy controls and 9 MS patients that were re-scanned after having been removed from the scanner and repositioned between the scans during the same visit at the Queen Square MS Centre, [284]. Images of the spinal cord were acquired with 1 mm isotropic voxel using a 3T Philips Achieva MRI system with RF dual-transmit technology (Philips Medical Systems, Best, Netherlands) (**Table 5.1**). I segmented the spinal cord using two different well-established CSA segmentation techniques: a semi-automatic delineation of the CSA using the

Figure 5.1. Spinal cord segmentation and GBSI.

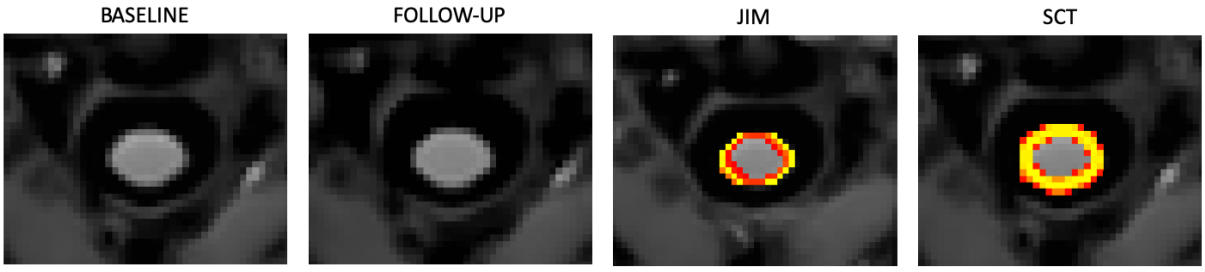
SCT was used for spinal cord segmentation. C2-3, C3-4, and C4-5 reference points were set manually (**A**). Representative images of semi-automatic spinal cord segmentation output are shown (sagittal and, in the inset, axial views) (**B**). Afterwards, baseline (**C**/**D**) and follow-up (**E**/**F)** spinal cord images were straightened, and, ultimately, registered to the halfway space. Intensity changes in the vicinity of the cord boundaries were estimated for generalised boundary shift integral (GBSI) calculation (**G**/**H**). Adapted from Moccia M, Prados F et al. Ann Neurol 2019 [113].



ASM within JIM (version 6) [277, 284], and a semi-automatic segmentation method using the *PropSeg* within the SCT (version 3.1.1) (refer to Chapter 2 and Chapter 4 for further details on these methods). Then, different GBSI measurements were obtained using masks of spinal cord segmentation obtained with JIM and SCT at both time-points (**Figure 5.2**).

Figure 5.2. pXOR masks.

Figure shows pXOR masks obtained from baseline and follow-up segmentations using ASM within JIM, and *PropSeg* within SCT. The semi-automatic delineation using ASM within JIM generated a slightly thinner pXOR mask than the *PropSeg* within the SCT. Adpated from Prados F, Moccia M, et al. Submitted.



5.3.3. Statistical analyses

On the test dataset, to evaluate repeatability of GBSI measurements obtained with different segmentation methods (aim 1), I calculated the intraclass correlation coefficient (ICC) for GBSI obtained from JIM and SCT segmentations. Then, assuming spinal cord atrophy would be expected to be zero in healthy controls and MS patients scanned twice during the same session, measurement error between scans was quantified by the coefficient of variation and the median absolute deviation on GBSI and CSA (SCT segmentation).

On the London and MAGNIMS validation datasets, means, medians and proportions of demographics, clinical features and MRI measures (percentage spinal cord changes obtained with GBSI and CSA) were calculated for patients (and their subgroups) and healthy controls (aim 2). Differences were evaluated with *t*-test, Mann–Whitney test, χ^2 test or Fisher's exact test, as appropriate. GBSI and CSA atrophy progression measurements were compared using a paired t-test.

Linear regression models were employed to estimate spinal cord atrophy changes (with GBSI and CSA) in different disease phenotypes, when compared with healthy controls (used as reference group); age, gender, site of acquisition and disease duration were included as covariates (results are presented as adjusted coefficients (Coeff), 95% confidence intervals (95%CI), and p-values). Then, I used different disease phenotypes as reference group in the linear regression models, to perform direct comparisons between different disease phenotypes (e.g., CIS, RRMS, PMS – including both PPMS and SPMS).

To compare GBSI and CSA in their ability to predict different clinical variables, I employed logistic regression models to estimate associations between percentage spinal cord atrophy change obtained using GBSI and CSA (independent continuous variable), and different binary clinical variables (dependent variable) (e.g., disease subtypes/healthy controls, EDSS progression (which was defined as 1 point change if baseline EDSS≤5.5, and 0.5 point if ≥6.0); age, gender, site of acquisition and disease duration were included as covariates. Results are presented as odds ratios (OR), 95%CI and p-values. Based on this, I obtained areas-under-the-curve (AUC), using GBSI and CSA, in turn, as the main explanatory variables; I used bootstrap resampling (1000 repetitions) to calculate pointwise confidence intervals for the receiver operating characteristics (ROC) curve (aim 3).

Sample sizes required for a hypothetical clinical trial evaluating a neuroprotective medication over one year were estimated using CSA and GBSI. Sample size was computed using the formula $n = \frac{2(Z_\alpha + Z_{1-\beta})^2 \sigma^2}{\Delta^2}$, where *n* is the required sample size per treatment arm in 1:1 controlled trials, Z_α and $Z_{1-\beta}$ are constant (set at 5% alpha-error and 80% power, respectively), σ is the standard deviation (from each disease phenotype), and Δ is the estimated effect size (aim 4) [289, 300]. Effect size was derived from adjusted beta-coefficients at linear regression models, estimating spinal cord loss in MS patients, when compared with physiologic loss in controls. As such, I assumed that 100% treatment effect is theoretically reached when the spinal cord atrophy change in patients is equal to that observed in healthy controls. From there, with a conservative approach, I hypothesised a number of effect sizes (e.g., 30%, 60% and 90%), that were smaller than the observed difference between MS cases and physiologic spinal cord loss in controls.

Stata 15.0 was used for data processing and analysis. Results were considered statistically significant when associated with p values <0.05.

5.4. Results

5.4.1. Measurement repeatability and error

On the test dataset, I observed similar spinal cord atrophy measurements for GBSI (ICC=0.66), independently of the segmentation technique used as input (GBSI from JIM segmentation=0.38±3.48%; GBSI from SCT segmentation=0.28±3.62%)

Among healthy controls and MS patients scanned twice during the same session, spinal cord atrophy would be expected to be zero. Both GBSI and CSA indeed had overall atrophy rates tending towards

zero (0.28±3.62% and -0.17±5.39%, respectively), indicating no systematic bias. However, GBSI presented with lower random measurement error, when compared with CSA, reflected by a clearly lower coefficient of variation (12.92% and 31.70%) and slightly lower median absolute deviation (0.21% and 0.25%).

5.4.2. Study population

Longitudinal spinal cord scans from 327 MS patients and 96 healthy controls were collected. Forty-five patients' scans and 14 healthy controls' scans were excluded because of either SCT failure, likely as a consequence of poor contrast (n=13), wrong voxel size (n=26), wrong acquisition parameters (n=10), and artefacts (n=8), or GBSI failure (n=2), mostly from the eldest cohorts acquired using 1.5T scanner (e.g., CIS cohort) (see **Figure 5.3** for the study flow diagram). Therefore, scans from 282 patients and 82 healthy controls acquired in 8 MAGNIMS centres were included in the analysis (see **Table 5.1** for the number of patients per centre and acquisition parameters). The demographic and clinical features of patients and healthy controls are given in **Table 5.2**.

5.4.3. Spinal cord atrophy obtained with CSA and GBSI

On paired t-test, the percentage spinal cord changes obtained with CSA were similar to those obtained with GBSI (p=0.55).

On linear regression models adjusted for age, gender, site of acquisition and disease duration, using CSA as main variable of interest, there was a significant decrease in the percentage spinal cord change over one year between RRMS (-1.74±4.02%) and healthy controls (-0.56±3.77%) (Coeff=-1.45;95%CI=-2.81, -0.10; p=0.03), but not between CIS (-1.19±3.67%) and healthy controls (Coeff=-0.84; 95%CI=-2.49, 0.80; p=0.31), nor between PMS (-1.29±3.20%) and healthy controls (Coeff=-

1.45; 95%CI=-3.70, 0.80; p=0.21) (**Figure 5.4A**). There were no differences in percentage CSA change over one year between CIS and RRMS (Coeff=-0.61; 95%CI=-1.95, 0.72; p=0.37), nor between CIS and PMS (Coeff=-0.60; 95%CI=-2.73, 1.53; p=0.58), nor between RRMS and PMS (Coeff=-0.00; 95%CI=-1.75, 1.73; p=0.99), when adjusted for the same covariates.

When using GBSI, overall, the rates of spinal cord decline were similar (or even higher among PMS) to those obtained with CSA, but the standard deviations of the measurements were smaller. On linear regression models adjusted for age, gender, site of acquisition and disease duration, using

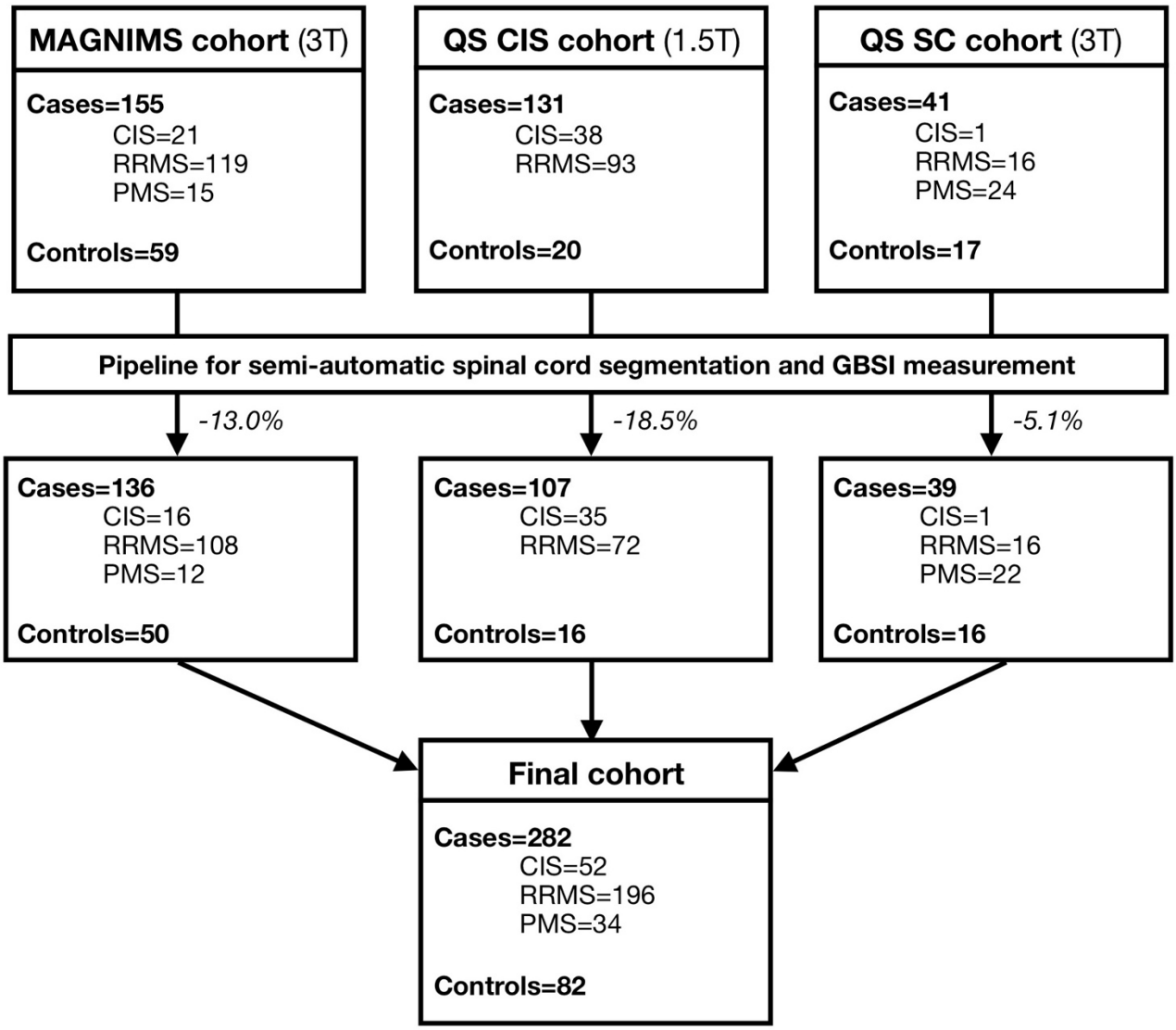
Table 5.2. Demographic and clinical features.

Table shows demographic and clinical features of MS and controls. P-values are shown from *t*-test, Mann–Whitney test, χ^2 test or Fisher's exact test, as appropriate (*p<0.05).

		MS *(n=282)*	**Controls** *(n=82)*	*p-values*
Age*, years*		38.2 ± 11.2	36.6 ± 12.5	*0.339*
Gender*, female*		169 (59.8%)	46 (55.7%)	*0.590*
Interval between scans*, years*		1.6 ± 1.1	1.2 ± 0.7	*0.001**
Disease duration*, years*		5.9±8.2		
Disease subtype	*CIS*	52 (18.4%)		
	RRMS	196 (69.5%)		
	PMS	34 (12.1%)		
EDSS at baseline		1.5 (0-7.5)		
EDSS at follow-up		2.0 (0-8.0)		
Patients with EDSS progression		74 (26.2%)		

Figure 5.3. Patient disposition.

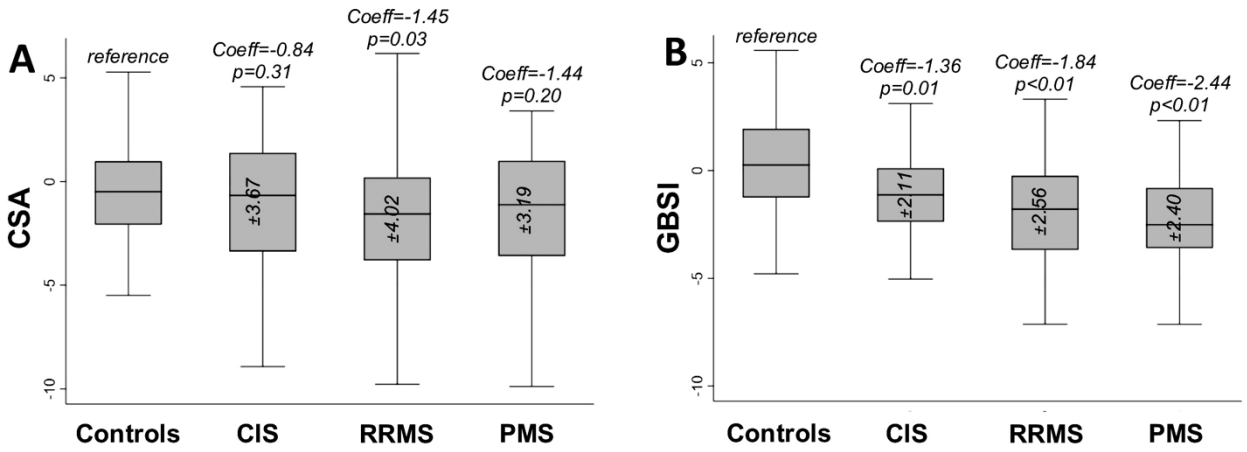
Figure shows MS cases and controls from MAGNIMS and UCL Queen Square Institute of Neurology cohorts; scanner filed strength is reported. Exclusion rate from the original cohort is shown, as a consequence of poor contrast, wrong voxel size, wrong acquisition parameters, and artefacts, mostly present in the eldest cohorts acquired using 1.5T scanners (e.g., CIS cohort). Adapted from Moccia M, Prados F et al. Ann Neurol 2019 [113].



GBSI as main variable of interest, there was a significant decrease in the percentage spinal cord change obtained with GBSI between healthy controls (0.02±2.39%) when compared to: CIS (-0.95±2.11%) (Coeff=-1.37; 95%CI=-2.40, -0.33; p=0.01), RRMS (-1.74±2.57%) (Coeff=-1.84; 95%CI=-2.70, -0.99; p<0.01), and PMS (-2.29±2.40%) (Coeff=-2.44; 95%CI=-3.87, -1.02; p<0.01) (**Figure 5.4B**). Similarly to the findings obtained with CSA, no differences were detected in GBSI percent annual reduction between CIS and RRMS (Coeff=-0.48; 95%CI=-1.32/0.37; p=0.27), nor between CIS and PMS (Coeff=-1.08; 95%CI=-2.43, 0.27; p=0.12), nor between RRMS and PMS (Coeff=0.60; 95%CI=-0.60, 1.70; p=0.28), when adjusted for the same covariates.

Figure 5.4. Box-and-Whisker plot of CSA and GBSI measurements

Box-and-Whisker plot for 1-year percentage change with CSA (**A**) and GBSI (**B**) in healthy controls, CIS, RRMS and progressive MS (PMS) (including both PPMS and SPMS). Coeff and p-values are reported from linear regression models using healthy controls as reference group (age, gender, disease duration and site of MRI acquisition were used as covariates). Standard deviation is also reported for each group. Adapted from Moccia M, Prados F et al. Ann Neurol 2019 [113].



5.4.4. Clinical correlates of CSA and GBSI

On logistic regression models adjusted by age, gender, site of acquisition and disease duration, RRMS had higher probability of spinal cord atrophy progression on CSA, when compared with healthy controls. On GBSI, all MS subtypes (CIS, RRMS and PMS patients) had higher probability of spinal cord atrophy progression, when compared with controls. Also, on GBSI, MS patients with EDSS progression had higher probability of spinal cord atrophy progression, than those without (**Table 5.3**).

Table 5.3. Clinical correlates of spinal cord atrophy with CSA and GBSI.

Table shows associations between spinal cord atrophy (with CSA and GBSI), and different clinical variables. OR, 95%CI and p-values are shown from logistic regression models; age, gender, site of acquisition and disease duration were included as covariates (*p<0.05).

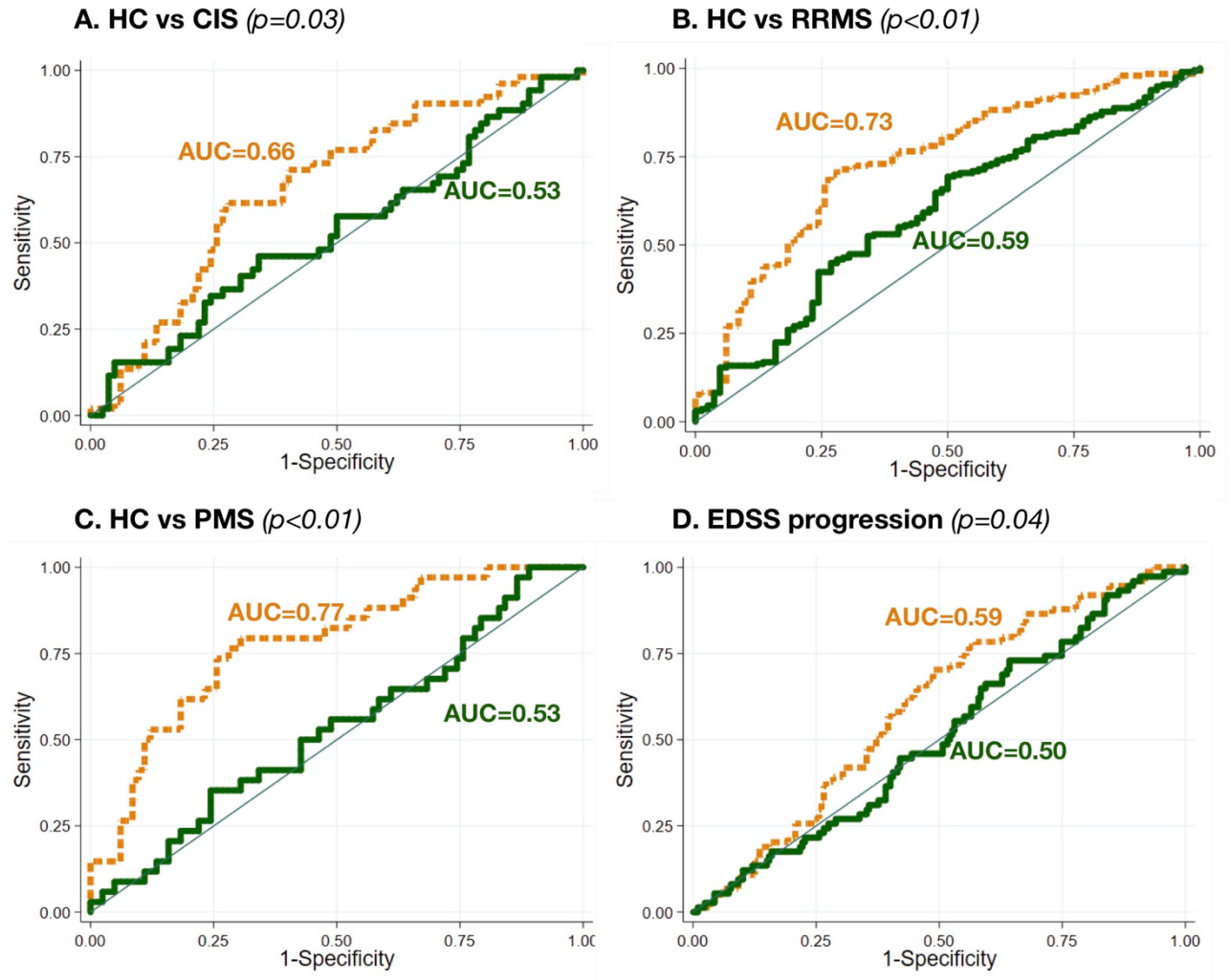
	CSA				**GBSI**			
	OR	**95%CI**		**p-values**	**OR**	**95%CI**		**p-values**
		Lower	*Upper*			*Lower*	*Upper*	
CIS vs HC	-0.04	-0.14	0.04	0.33	-0.23	-0.40	-0.07	<0.01*
RRMS vs HC	-0.07	-0.14	-0.00	0.02*	-0.31	-0.42	-0.19	<0.01*
PMS vs HC	-0.05	-0.17	0.05	0.32	-0.43	-0.64	-0.23	<0.01*
EDSS progression	-0.00	-0.07	0.06	0.83	-0.13	0.25	-0.02	0.02*

CIS patients were better differentiated from controls using GBSI (AUC=0.66; 95%CI=0.57, 0.75), than CSA (AUC=0.53; 95%CI=0.43, 0.63) (p=0.03; **Figure 5.5A**). RRMS patients were better differentiated from controls using GBSI (AUC=0.73; 95%CI=0.66, 0.80), than CSA (AUC=0.59; 95%CI=0.52, 0.66) (p<0.01; **Figure 5.5B**). PMS patients were better differentiated from controls using GBSI (AUC=0.77;

95%CI=0.68, 0.86), than CSA (AUC=0.53; 95%CI=0.45, 0.64) (p<0.01; **Figure 5.5C**). Patients with EDSS progression (n=76) were better differentiated from those without EDSS progression (n=206) using GBSI (AUC=0.59; 95%CI=0.52-0.66) than CSA (AUC=0.50; 95%CI=0.43-0.58) (p=0.04; **Figure 5.5D**).

Figure 5.5. ROC curves for CSA and GBSI in relation to clinical variables.

ROC curves for CSA (solid green line) and GBSI (dashed orange line) in relation to differentiating healthy controls (n=82) from CIS (n=52) (**A**), RRMS (n=196) (**B**), and PMS (n=34) (including both PPMS and SPMS) (**C**), and patients with EDSS progression (n=74), from those without EDSS progression (n=208) (**D**). AUC and p-value are reported. Adapted from Moccia M, Prados F et al. Ann Neurol 2019 [113].



5.4.5. Sample size estimates using GBSI and CSA

The minimum sample sizes per arm required to detect a 60% treatment effect in one year clinical trial (i.e., a 60% reduction in percentage spinal cord change in MS cases when compared with physiologic spinal cord loss in controls, adjusted for age, gender, site of acquisition and disease duration) were lower for GBSI compared with CSA (CIS: 106 vs. 830; RRMS: 95 vs. 335; PMS: 44 vs. 215) (power=80%, alpha=5%). Similar results were obtained when estimating the sample size required to detect different treatment effects (**Figure 5.6**).

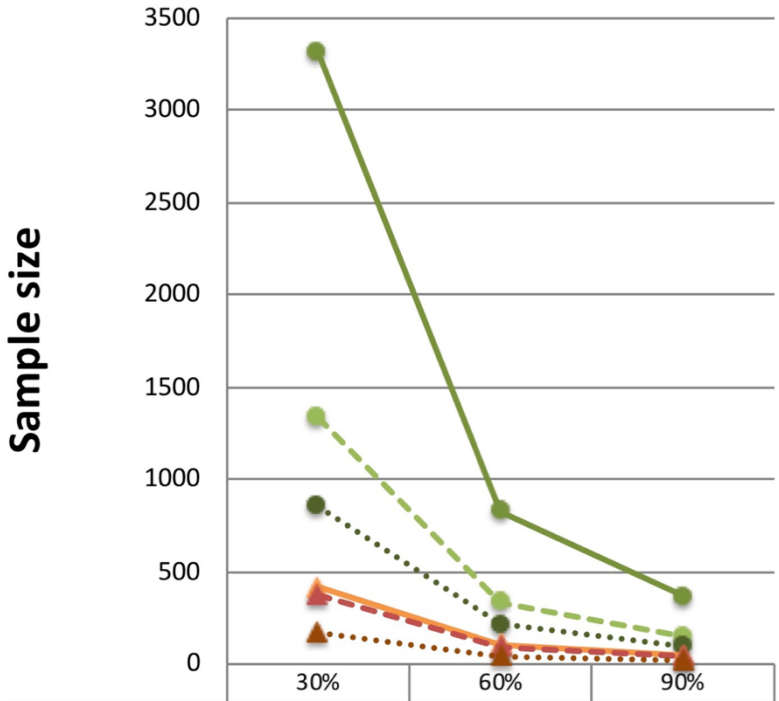
5.5. Discussion

5.5.1. Main findings

There have been few clinical trials and observational studies in MS that used spinal cord atrophy as an outcome measure, because of the large sample size required when using the available CSA method [63, 91, 144, 292]. In the present study, I applied a standard semi-automatic pipeline for spinal cord segmentation using the SCT, and, then, a fully automated registration-based technique (GBSI) for spinal cord atrophy to a large, multicentre, multi-manufacturer and multi-field strength scan cohort, derived from longitudinal observational studies. The rates of spinal cord loss over one year obtained with GBSI were similar to those obtained with CSA, but they were associated with lower variability, greater ability to distinguish between MS patients and controls, and more robust clinical correlates, thereby holding promise for future MS research on spinal cord imaging. Use of GBSI yielded increased statistical power to detect treatment changes, suggesting that future treatment trials –particularly those testing neuroprotective agents– could include spinal cord atrophy as a primary outcome measure.

Figure 5.6. Sample size estimates for CSA and GBSI.

Profile plot shows sample size estimates for CSA and GBSI in different disease phenotypes (CIS, RRMS and PMS). For sample size calculation, I included adjusted beta-coefficients from linear regression models, estimating spinal cord loss in MS patients, when compared with physiologic loss in controls, and standard deviation from each disease phenotype. Power was set at 80% and alpha at 5%. Different treatment effects were hypothesised (e.g., 30%, 60% and 90%), that were smaller than the observed difference between MS cases and physiologic spinal cord loss in controls. Adapted from Moccia M, Prados F et al. Ann Neurol 2019 [113].



	30%	60%	90%
CSA in CIS	3314	830	370
CSA in RRMS	1336	335	150
CSA in PMS	854	215	96
GBSI in CIS	419	106	48
GBSI in RRMS	374	95	43
GBSI in PMS	170	44	20

5.5.2. Interpretation

CSA and GBSI provided similar rates of spinal cord atrophy in each MS subtype, but CSA yielded a larger variability (standard deviation), when compared with GBSI (e.g., in RRMS ±4.02% vs. ±2.57%, respectively), implying that GBSI measurements are more precise. Similarly, when images were acquired with a scan-rescan fashion, GBSI presented with lower coefficient of variation, when compared with CSA, indicating higher precision; the median absolute deviation (0.21%) was similar to what has been previously described in the validation of a registration-based method for brain atrophy (0.15-0.20% for the SIENA) [367]. Spinal cord measurements with GBSI can be repeated using different segmentation methods, with high agreement between techniques (e.g., JIM and SCT). Of note, a higher percentage of spinal cord decline (and a lower standard deviation) was found in PMS when using GBSI than when using CSA (-2.29±2.40% vs -1.29±3.20%), and this was probably due to the relatively small sample in this MS subtype, further highlighting the higher measurement precision of GBSI. A more detailed description of spinal cord atrophy measurements in progressive MS will be done in the following Chapter 6.

The smaller variability in GBSI measurements may be due to the ability of GBSI to derive spinal cord atrophy rates directly from small intensity changes between images at the cord boundaries, accounting for partial volume effects in these regions which are critical for measuring changes. Indeed, partial volume effects can lead to the inclusion of tissue outside of the area of interest with subsequent segmentation errors, and to the variability when calculating the absolute cross-sectional areas [368–370]. The use of the probabilistic XOR (pXOR) region for weighting the BSI over specific tissue boundaries, is particularly relevant in the spinal cord where there are extremely close-fitting surfaces that can be affected by changes even in a small number of voxels. On the contrary, conventional segmentation-based methods (e.g., CSA) rely on the numerical difference between

areas obtained from the hard segmentation at each time-point; this approach provides an indirect estimate of atrophy and can introduce greater variability, especially in acquisitions with large voxel sizes or between scans with different intensity scales.

I found that GBSI allowed a four-fold smaller (and therefore achievable) sample size than that obtained with CSA. In a recent review of previous clinical trials in MS [213], differences between treated and untreated patients ranged from 0.4% to 1.8%/year spinal cord loss on CSA, in line with a roughly 60% treatment effect (corresponding to 0.7%/year spinal cord loss). Overall, the sample size estimates for spinal cord atrophy measurements with GBSI are of the same order of magnitude as those for brain atrophy obtained with registration-based methods [153, 158, 289, 360]. However, monitoring spinal cord atrophy could be more clinically relevant than brain atrophy, due to its robust clinical correlates [227, 250, 305]. Spinal cord atrophy occurs in the early stages of MS (e.g., CIS), is more obvious in progressive patients than relapsing types of MS, and progresses faster than brain atrophy [227, 250, 305]. The risk of disability progression increases with the rate of spinal cord atrophy [296], accounting for up to 77% of motor disability, as measured by the EDSS [227, 250, 305]. In line with this, the percentage spinal cord decline measured with GBSI was associated with EDSS progression and with disease subtype, and performed better than CSA-derived spinal cord atrophy in detecting more disabled patients. Of note, differentiation between healthy controls and MS patients (especially progressive MS) revealed better performance of GBSI when compared with CSA, whilst differentiation between patients with and without EDSS progression showed similar GBSI and CSA measurements, with CSA tending to random prediction of results (AUC=0.50) and GBSI presenting with statistically higher but not particularly better predictive value (AUC=0.59). The lack of separation in this chapter might be due to the limited proportion of patients presenting with 3-

month confirmed disability progression over 1-year observation time, and these findings should be confirmed with a longer follow-up in order to avoid false positive and negative results.

5.5.3. Limitations

Limitations of the present study include the short follow-up duration (1.6 years in MS patients), meaning that I could not assess the association between spinal cord atrophy and long-term disability progression. Previous studies have shown that spinal cord atrophy predicts disease progression and conversion from CIS to RRMS over a long follow-up period [307]. However, study duration was *a priori* set the at 1 year in order to obtain estimates for clinical trials and short-term observational studies; of note, I obtained annualised atrophy rates in line with pooled estimates from 94 studies (1.78%/year) [268], and the number of participants per arm was consistent with that obtained by other similar studies [153, 158, 289]. The present cohort was recruited before presentation and publication of the 2017 revision of McDonald criteria [61], and, thus, diagnosis was performed with the 2010 revision [59]. As such, I cannot exclude that at least some CIS patients would nowadays classify as RRMS [61], possibly determining an (almost) physiological rate of spinal cord atrophy in the remaining CIS population. I segmented the spinal cord using *PropSeg* within the SCT. The *PropSeg* tool has already been demonstrated as a robust, accurate and fast segmentation method of the spinal cord in both MS patients and controls, compared with other segmentation methods (e.g., semi-manual ASM within JIM) [284, 376, 388]. Also, the possibility to derive spinal cord GBSI measurements from brain scans needs to be explored [282, 287, 292, 296].

An important caveat of this chapter, is that good quality images (e.g., dedicated spinal cord 3DT1 images with 1mm isotropic voxel) are needed for spinal cord atrophy measurements. I was only able to use 86% of the patients whose scans were originally collected in eight experienced imaging

centres, using slight variations in MRI protocols and field strength. However, there is an initiative to mitigate the problem of setting-up a standard high-quality image protocol for spinal cord MRI (http://www.spinalcordmri.org - Protocols). In this project, a consensus acquisition protocol has been developed and tested in about 30 different sites across the world, for different vendors (GE, Siemens and Philips). This protocol is easy to apply and facilitates the adoption curve of spinal cord MRI acquisitions. Also, latest software improvements, such as machine-learning based segmentation methods (e.g., *DeepSeg* within the SCT), could facilitate the spinal cord image processing and will be further explored in Chapter 6.

5.5.4. Conclusions

In this chapter, I presented results from a longitudinal multicentre, multi-manufacturer and multi-field strength scan analysis, and showed that GBSI and CSA provided similar rates of spinal cord atrophy, but the registration-based method (GBSI) was associated with lower variability, providing smaller sample size estimates as well as a higher significance at differentiating between different MS subtypes and patients with disability progression, compared with segmentation-based method (CSA), though still far from perfect. This study provides evidence that GBSI should be considered as a precise and reliable tool for calculating MS-related spinal cord atrophy in clinical trials and in observational datasets.

6. Spinal cord atrophy in a primary progressive multiple sclerosis trial

6.1. Background

As already discussed in Chapter 1 and Chapter 2, spinal cord atrophy is a common feature of MS, can be detected *in vivo* using MRI, and is one of the main substrates of disease progression [226, 294, 297, 298]; as such, spinal cord atrophy can be used to monitor disease progression and has been included in clinical trials evaluating medications with putative neuroprotective effects in MS [227, 300, 305, 309].

In Chapter 5 [113, 389], I have preliminary applied the GBSI pipeline to cohorts of MS patients and healthy controls. This technique significantly reduced measurement error by providing a direct measure of atrophy based on registration of all images in a common space [113, 289], when compared with indirect measures obtained from numerical subtraction between segmentations at different timepoints [113], being at risk of differences in cord coverage, miss-segmentations and/or changes in cord curvature [109, 292]. Improvements in measurement precision could be particularly relevant to clinical trials, using gold-standard segmentation methods (e.g., CSA), that, so far, have failed to show any significant treatment effect on spinal cord atrophy, especially in progressive MS patients [105, 213], where this outcome measure holds the strongest clinical correlates [109, 213].

Thus, following previous results, I have hereby further implemented the GBSI pipeline in collaboration with Dr Prados. In particular, I have run spinal cord segmentation and GBSI calculation on a large cohort of progressive MS, and evaluated possible correlates to clinical and patient-reported outcomes. Additionally, I obtained registration-based spinal cord atrophy measurements for different levels and also from brain scans.

6.2. Aims

In the present study, I re-analysed a phase 2 clinical trial on PPMS to:

(1) Compare spinal cord atrophy measurements using segmentation- and registration-based methods, with possible implications for image acquisition and clinical trial design (e.g., measurement precision, image quality);

(2) Compare spinal cord atrophy measurements obtained from brain (C1-2) and spinal cord MRI (C1-2 and C2-5), using segmentation- and registration-based methods;

(3) Explore possible clinical correlates, also in relation to conventional brain MRI measures;

(4) Explore possible treatment effect.

6.3. Methods

6.3.1. Study design

This is a secondary analysis on PPMS patients who participated in the ARPEGGIO phase 2 clinical trial. As already discussed in Chapter 4, the ARPEGGIO trial was a randomised, double-blind, parallel-group, placebo-controlled study. Patients were randomised in a 1:1:1 ratio to receive oral laquinimod 0.6 mg or 1.5 mg, or placebo once daily from January 2015 to April 2016 at 85 sites in 10 countries. Duration of the core study was 48 weeks [105].

6.3.2. Population

Inclusion and exclusion criteria have been reported previously [105]. Briefly, inclusion criteria were:

(1) Aged 25–55 years;

(2) Diagnosis of PPMS [59];

(3) EDSS at baseline from 3.0 to 6.5;

(4) Documented worsening of clinical disability in the 2 years prior to screening;

(5) A FSS ≥2 for the pyramidal system or gait impairment due to lower limb dysfunction.

Exclusion criteria were:

(1) Clinical history of any MS exacerbations or relapses;

(2) Any other progressive neurological disorder (e.g., cervical cord compression, vitamin B12 deficiency);

(3) Previous use of immunosuppressive/cytotoxic agents, experimental/investigational drugs and/or MS-specific treatments (e.g., fingolimod, dimethyl fumarate, glatiramer acetate, interferon-β, laquinimod).

6.3.3. MRI acquisition and processing

All patients underwent brain and cervical spinal cord MRI at baseline, week 24 and week 48, including 3D T1-weighted isotropic images of the brain and spinal cord (1x1x1mm^3), within 14 days of the scheduled clinical visit. MRI scans were collected centrally at the VUmc in Amsterdam. For the purposes of the present study, I included baseline and week-48 MRI.

For the primary analysis of the ARPEGGIO trial, cord area at C1-2 level was determined using the NeuroQLab, a segmentation-based method using a Gaussian mixture modelling [275, 281, 282]. After definition of the spinal cord subsection to be segmented, watershed segmentation of the spinal cavity and surrounding cerebrospinal fluid was performed, and MUCCA was computed [275]. Percent change of cord area was calculated using the following formula: MUCCA = 100*(week 48 MUCCA - baseline MUCCA)/baseline MUCCA.

In the present study, masks of C1-2 and C2-5 levels were obtained from spinal cord MRI acquired at each time point, using the *DeepSeg* tool within the SCT (version 4.0), a fully-automated segmentation method based on convolutional neural networks (**Figure 6.1**) [283, 285]. Similarly, masks of C1-2 level were obtained from brain MRI (**Figure 6.2**). Percent change of CSA was calculated using the following formula: CSA = 100*(week 48 CSA - baseline CSA)/baseline CSA. For GBSI, I followed the pipeline previously described in Chapter 4 [113, 389]. The GBSI was finally computed and the percent variation was measured (**Figure 6.1; Figure 6.2**).

From the original clinical trial dataset, I extracted the number of new T2-hyperintense lesions, T2-hyperintense lesion volume change, T1-hypointense lesion volume change and PBVC, at baseline and week 48 visits. Full details of acquisitions and processing have been previously reported [105].

6.3.4. MRI quality

To classify scans based on their quality, I used the σ (standard deviation of the MR signal), calculated with the following formula: $\sigma = \frac{\sigma_\eta}{\sqrt{2-\frac{\pi}{2}}}$, where σ_η is the standard deviation of the magnitude–reconstructed cerebrospinal fluid signal [377]. For the purpose of this chapter, I classified scans based on the median σ for each spinal cord segment.

6.3.5. Clinical variables

Baseline clinical variables were age, gender, disease duration and EDSS. From the original clinical trial dataset, I extracted following clinical variables corresponding to MRI acquisitions: EDSS, T25FW, 9HPT, SDMT, and MSWS. EDSS progression was defined as ≥1 point from baseline EDSS if EDSS at entry was ≤5.5 or increase of ≥0.5 point if EDSS at entry was >5.5. T25FW, 9HPT, SDMT and MSWS progression was defined as ≥20% increase from baseline score.

Figure 6.1. Spinal cord MRI processing using SCT and GBSI.

Reference points were preliminarily set at C1, C2 and C5 on sagittal scans (**a**). Then, masks of C1-2 (**b**) and C2-5 (**c**) spinal cord levels were obtained from spinal cord images acquired at each time point using the *DeepSeg* tool within the SCT (version 4.0). The probabilistic boundary-shift region-of-interest was then adaptively estimated from baseline and follow-up cord segmentations and the GBSI integral was finally computed for C1-2 (**d** for sagittal plane, **e** for axial plane) and C2-5 (**f** for sagittal plane, **g** for axial plane).

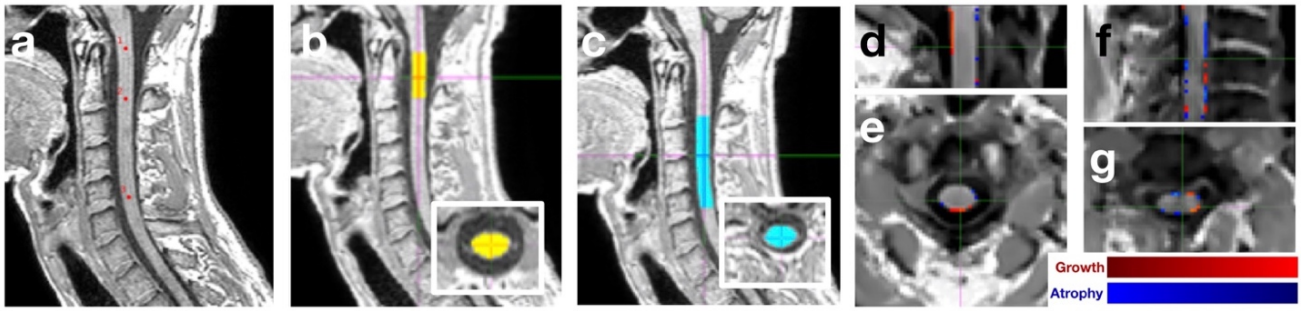
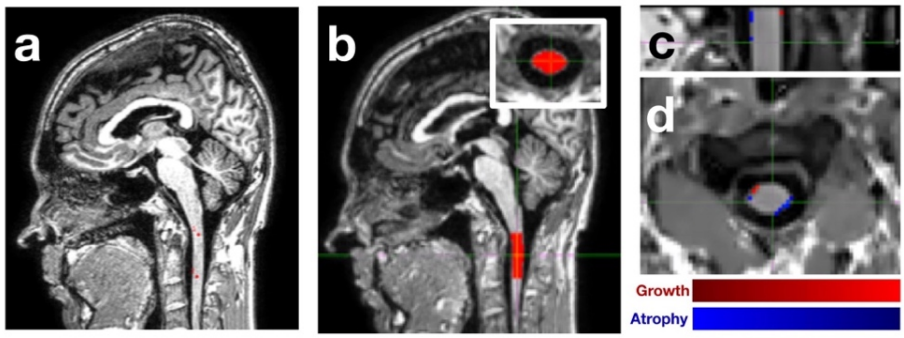


Figure 6.2. Brain MRI processing using SCT and GBSI.

Reference points were preliminarily set at C1 and C2 on sagittal scans (**a**). Then, masks of C1-2 (**b**) levels were obtained from spinal cord images acquired at each time point using the *DeepSeg* tool within the SCT (version 4.0). The probabilistic boundary-shift region-of-interest was then adaptively estimated from baseline and follow-up cord segmentations and the GBSI integral was finally computed for C1-2 (**c** for sagittal plane, **d** for axial plane).



6.3.6. Treatment exposure

Patients were randomised in a 1:1:1 ratio to receive oral laquinimod in a dose of 0.6 mg or 1.5 mg or placebo (once daily). The laquinimod 1.5 dose arm was discontinued as of January 1, 2016, due to cardiovascular side effects (patients were followed-up, but no further treatment was given).

6.3.7. Statistical analyses

Baseline demographics and clinical characteristics are presented as mean (and standard deviation), number (and percent), or median (and range), as appropriate.

To evaluate possible implications of different spinal cord atrophy measurements on image acquisition and clinical trial design (aim 1), I computed the sample size required for a hypothetical clinical trial evaluating a neuroprotective medication over one year. Sample size was computed using the formula $n = \frac{2(Z_{\alpha}+Z_{1-\beta})^2\sigma^2}{\Delta^2}$, where *n* is the required sample size per treatment arm in 1:1 controlled trials, Z_{α} and $Z_{1-\beta}$ are constant (set at 5% alpha-error and 80% power, respectively), σ is the standard deviation (from each spinal cord atrophy measurement), and Δ is the estimated effect size [289, 300]. Effect size was derived from the spinal cord change on each atrophy measurement (MUCCA, CSA at C1-2 level from brain MRI, GBSI at C1-2 level from brain MRI, CSA at C1-2 level from spinal cord MRI, GBSI at C1-2 level from spinal cord MRI, CSA at C2-5 level from spinal cord MRI, and GBSI at C2-5 level from spinal cord MRI); different treatment effects were simulated (e.g., 30%, 60% and 90%). Sample size estimates were performed on the whole population first and, then, on the subset of scans above the median quality (based on the standard deviation of the MR signal (σ)). As additional estimates of measurement precision, I also computed coefficients of variation and median absolute deviations.

To evaluate the possibility of deriving longitudinal spinal cord atrophy measurements from brain MRI (aim 2), I used linear regression models including spinal cord atrophy measurements from brain MRI (CSA at C1-2 level and GBSI at C1-2 level in turn) and corresponding spinal cord atrophy measurements from spinal cord MRI (CSA at C1-2 and C2-5 levels, and GBSI at C1-2 and C2-5 levels, respectively). Results were presented as Coeff, 95%CI and p-values.

To evaluate possible clinical correlates of each spinal cord atrophy measurement (aim 3), I used different stepwise linear regression models (one for each MRI variable) with backward selection for p=0.20 as the critical value for entering clinical variables in the model (EDSS progression, T25FWT progression, 9HPT progression, SDMT progression, and MSWS progression). For the purposes of this analysis, I considered both spinal cord MRI variables (MUCCA, CSA at C1-2 level from brain MRI, GBSI at C1-2 level from brain MRI, CSA at C1-2 level from spinal cord MRI, GBSI at C1-2 level from spinal cord MRI, CSA at C2-5 level from spinal cord MRI, and GBSI at C2-5 level from spinal cord MRI), and brain MRI variables (number of new T2 lesions, T2 lesion volume change, T1 lesion volume change and PBVC). Results were presented as Coeff, 95%CI and p-values.

To evaluate differences in spinal cord atrophy between laquinimod 0.6mg and placebo (aim 4), I used linear regression models. Covariates were age, gender, height, country and baseline CSA [270]. Results were presented as Coeff, 95%CI and p-values.

Stata 15.0 was used for data processing and analysis. Results were considered statistically significant when associated with p-values <0.05.

6.4. Results

6.4.1. Patient disposition

Patient disposition and reasons for exclusion are reported in **Figure 6.3**. Overall, among patients with baseline and week 48 visits and with MRI acquisitions suitable for analyses, I was able to include 98.1% of patients for CSA at C1-2 level from brain MRI (211/215), 84.2% of patients for GBSI at C1-2 level from brain MRI (181/215), 67.4% of patients for CSA at C1-2 and at C2-5 level from spinal cord MRI (114/169), and 66.9% of patients for GBSI at C1-2 and at C2-5 level from spinal cord MRI (113/169). Examples of MRI acquisitions not suitable for analyses or resulted into SCT failure, GBSI failure and atrophy rate above what reasonably expected (±5%) are presented in **Figure 6.4**. For comparison, I included 220 patients with MUCCA, number of new T2 lesions, T2 lesion volume change, T1 lesion volume change and PBVC from the original clinical trial.

Demographic and clinical features of included patients are reported in **Table 6.1**.

Table 6.1. Demographics and clinical features.

Baseline demographics and clinical features are reported for included patients.

Population (n=220)	
Age*, years*	46.5±6.8
Gender*, male*	118 (53.6%)
Baseline EDSS	4.5 (3.0-6.5)
Disease duration*, years*	3.4±3.2
Height*, cm*	172.0±9.8
Country*, n*	10

Figure 6.3. Patient disposition.

Patient disposition flow diagram shows number of included and excluded patients. Reasons for exclusion from the original trial population were early termination, lack of upper spinal cord segments included in brain MRI, lack of dedicated spinal cord MRI, SCT failure, GBSI failure, and >5% atrophy rate on both SCT and GBSI. Total number of patients with each MRI measure is presented.

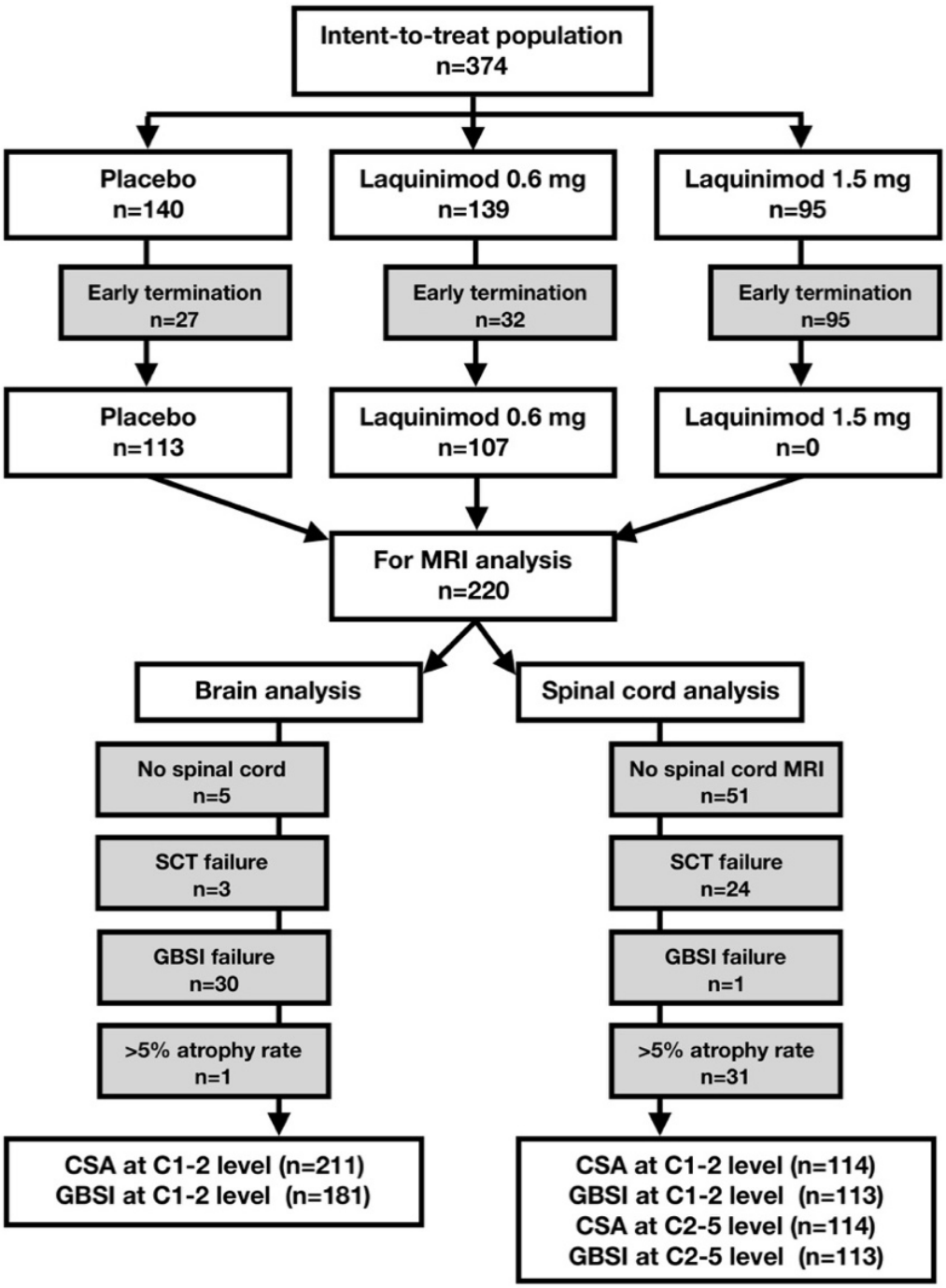
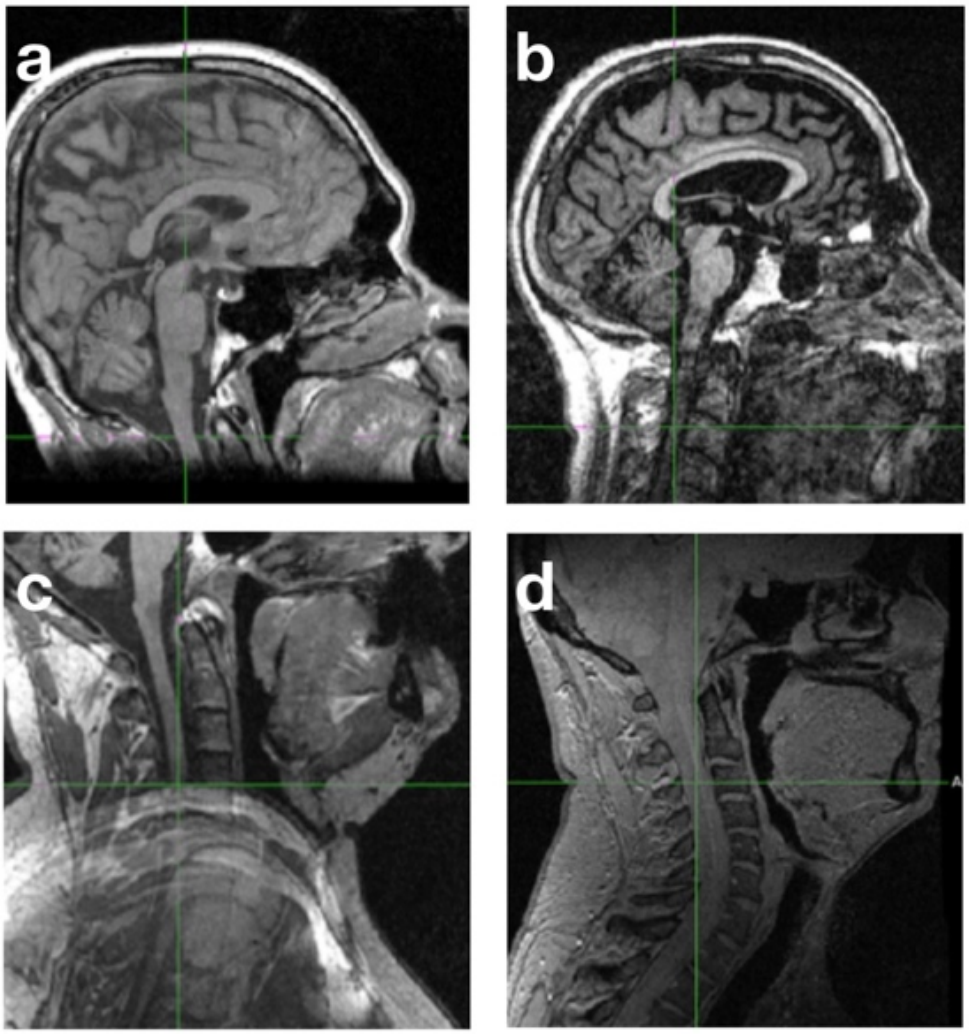


Figure 6.4. Examples for exclusion.

I excluded brain MRI if acquisition did not include C1-2 spinal cord level (**a**). SCT failure, GBSI failure and atrophy rate above what expected (±5%) were related to noisy images (**b**), artefacts (**c**), and poor contrast (**d**) for both brain and spinal cord MRI.

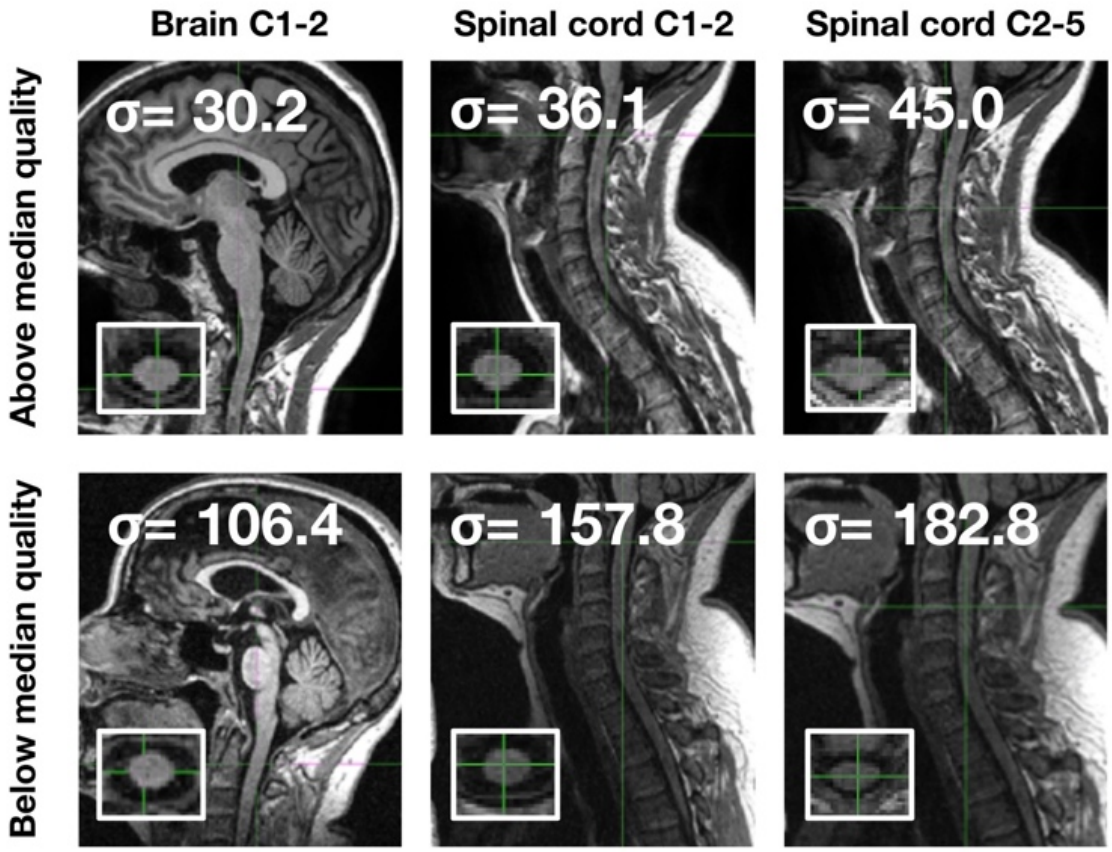


6.4.2. MRI quality

Scans were classified on the median value of σ (standard deviation of the MR signal) at each spinal cord level. The lowest median standard deviation was found on brain MRI at C1-2 level (σ=43.3), followed by spinal cord MRI at C1-2 level (σ=57.0), and spinal cord MRI at C2-5 level (σ=76.5). Representative images are presented in **Figure 6.5**.

Figure 6.5. MRI quality.

Figure shows sagittal plane of brain and spinal cord images above (upper row) and below (lower row) median quality (axial plane in the inset). Different spinal cord levels are reported (C1-2 from brain MRI, C1-2 from spinal cord MRI, and C2-5 from spinal cord MRI). Standard deviation of the MR signal (σ) is reported for each image.



6.4.3. Spinal cord atrophy

Spinal cord atrophy measurements obtained with GBSI (-1.5±3.4% at C1-2 level from brain MRI; -1.8±3.1% at C1-2 level from spinal cord MRI; -1.5±3.6% at C2-5 level from spinal cord MRI) had a larger effect size and smaller standard deviation when compared with corresponding CSA (-0.9±4.2% at C1-2 level from brain MRI; -1.1±4.1% at C1-2 level from spinal cord MRI; -1.5±4.7% at C2-5 level from spinal cord MRI) and MUCCA (-0.9±3.1%) (**Figure 6.6**), resulting into smaller sample size estimates, smaller coefficients of variation and smaller median absolute deviations (**Table 6.2**).

Table 6.2. Sample size estimates and measurement error.

Table shows sample size estimates using 1-year atrophy rate and standard deviation from each MRI measure for the whole population; power was set at 80% and alpha-error at 5%. Different treatment effects were simulated (30%, 60% and 90%). Coefficients of variation and median absolute deviations are also reported.

	Sample size *Treatment effect*			**Coefficient of variation**	**Median absolute deviation**
	30%	*60%*	*90%*		
MUCCA	2073	518	230	-3.56	3.06
CSA at C1-2 level from brain MRI	3454	864	384	-4.57	4.13
GBSI at C1-2 level from brain MRI	832	208	92	-2.13	3.26
CSA at C1-2 level from spinal cord MRI	2528	632	281	-3.42	3.86
GBSI at C1-2 level from spinal cord MRI	523	131	58	-1.83	2.83
CSA at C2-5 level from spinal cord MRI	1581	395	176	-3.09	4.86
GBSI at C2-5 level from spinal cord MRI	1027	257	114	-2.37	3.33

Results were confirmed when considering the subset of patients with availability of all measurements (MUCCA, CSA and GBSI at different cord levels from both brain and spinal cord MRI), avoiding the effect of different sample size for computing mean and standard deviation (**Figure 6.7a**), and the MRIs with quality above the median (**Figure 6.7b**).

Figure 6.6. Spinal cord atrophy rates by type of scan and analysis technique.

Box-and-Whisker plots show 1-year spinal cord percent volume change for different measurements (MUCCA, CSA and GBSI at different cord levels from brain and spinal cord MRI). Mean, standard deviation and number of included patients are reported.

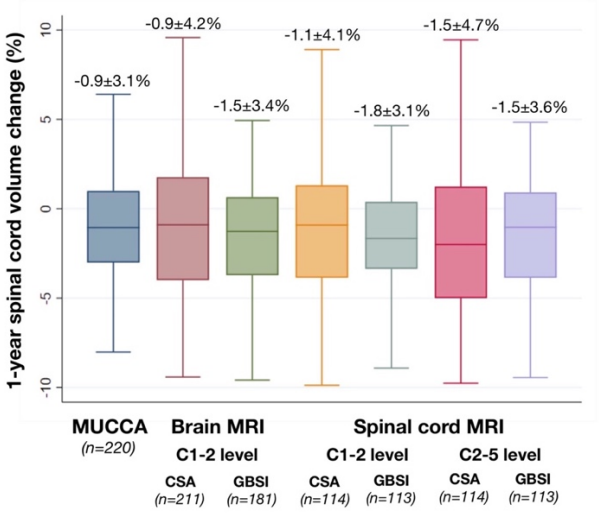
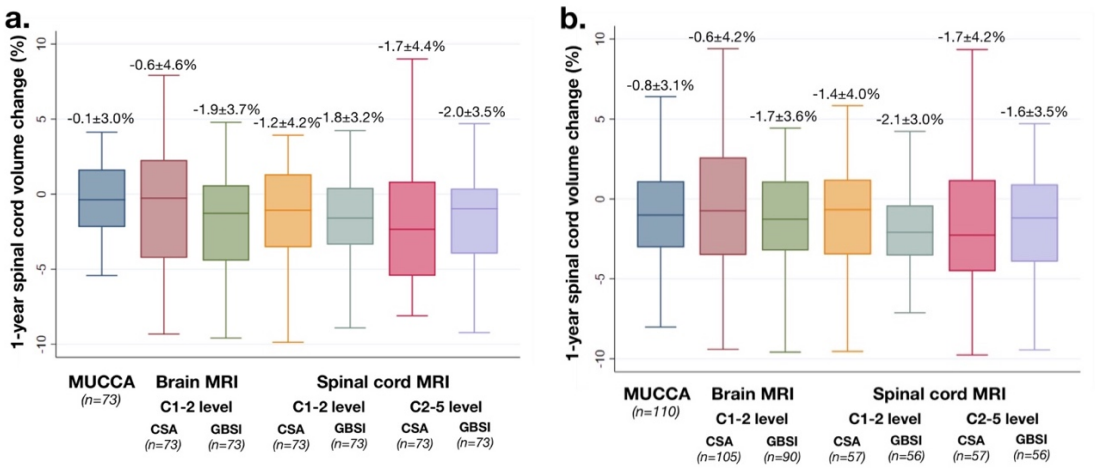


Figure 6.7. Spinal cord atrophy rates in patients with all measurements and in high-quality scans.

Box-and-Whisker plots show 1-year spinal cord percent volume change for different measurements (MUCCA, CSA and GBSI at different cord levels from brain and spinal cord MRI) in a subset of patients with all measurements (n=73) (**a**), and in scans above median quality (50% of the original sample, as presented in **Figure 6.6**), estimated as the standard deviation of the MR signal (σ) (**b**). Mean, standard deviation, and number of included patients are reported.

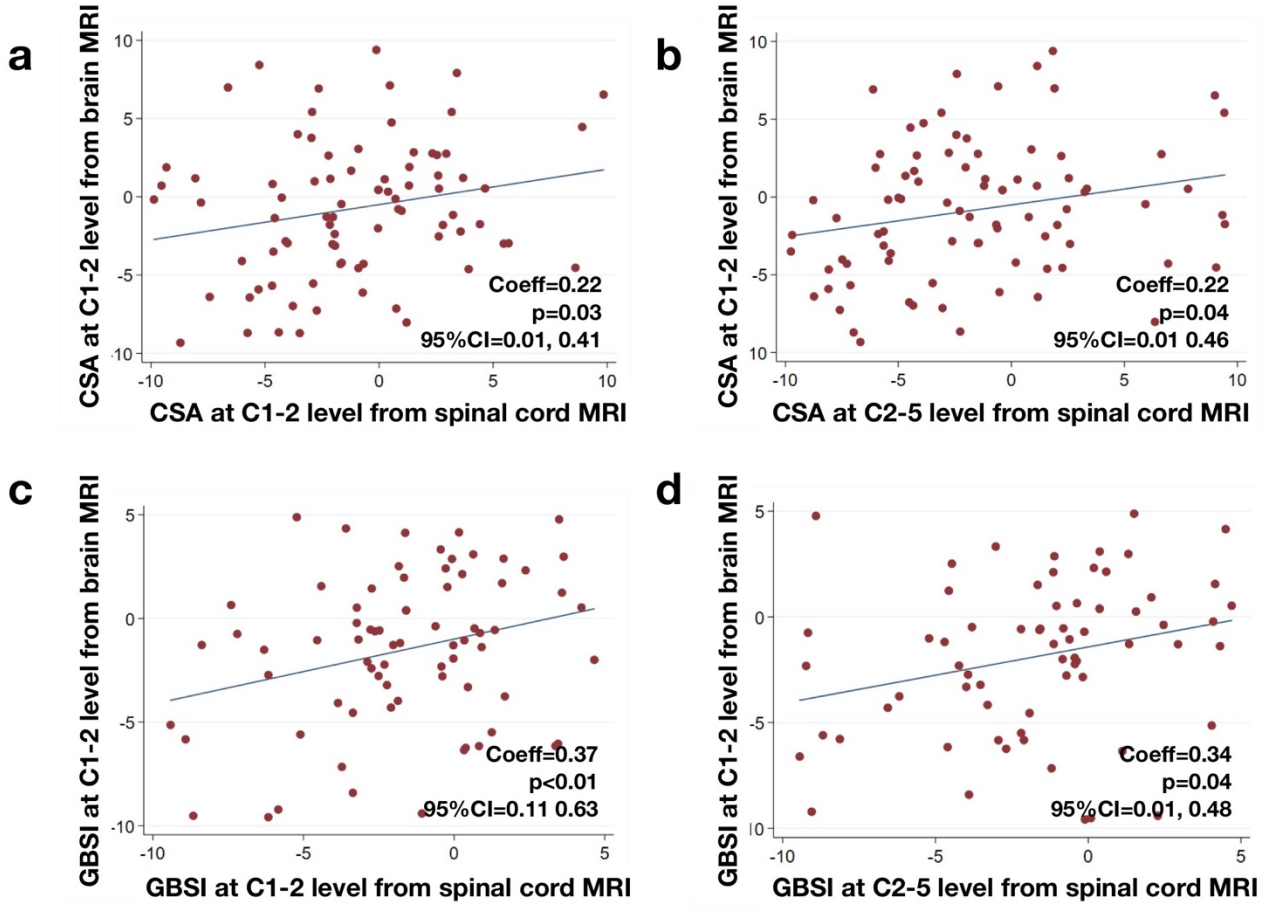


6.4.4. Spinal cord atrophy from brain scans

Longitudinal spinal cord atrophy measurements obtained from brain MRI were related to corresponding measurements from dedicated spinal cord MRI for both CSA (**Figure 6.8a-b**) and GBSI (**Figure 6.8c-d**).

Figure 6.8. Spinal cord atrophy longitudinal changes from brain and spinal cord MRI.

Scatter plots show 1-year spinal cord atrophy longitudinal changes obtained from brain scans (CSA at C1-2 level from brain MRI in **a** and **b**; GBSI at C1-2 level from brain MRI in **c** and **d**) in relation to measurements obtained with the same technique but from spinal cord MRI (CSA at C1-2 level and CSA at C2-5 level from spinal cord MRI in **a** and **b**; GBSI at C1-2 level and GBSI at C2-5 level from spinal cord MRI in **c** and **d**). Coeff, p-values and 95%CI are reported from linear regression models.



6.4.5. Clinical correlates

Looking at spinal cord atrophy measurements, patients with T25FWT progression presented with more pronounced spinal cord atrophy on GBSI at C1-2 level from brain MRI (Coeff=-1.24; 95%CI=-2.30, -0.19; p=0.02), and GBSI at C2-5 level from spinal cord MRI (Coeff=-1.47; 95%CI=-2.18, -0.75; p<0.01). Patients with 9HPT progression presented with more pronounced spinal cord atrophy on GBSI at C1-2 level from brain MRI (Coeff=-0.92; 95%CI=-1.33, -0.52; p<0.01). Patients with MSWS progression presented with more pronounced spinal cord atrophy on GBSI at C1-2 level from brain MRI (Coeff=-1.86; 95%CI=-2.44, -1.27; p<0.01), and GBSI at C1-2 level from spinal cord MRI (Coeff=-1.73; 95%CI=-2.38, -1.07; p<0.01). No significant clinical correlates were detected for MUCCA, CSA at C1-2 level from brain MRI, CSA at C1-2 level from spinal cord MRI, and CSA at C2-5 level from spinal cord MRI. Spinal cord atrophy measurements in relation to progression of EDSS, T25FWT, 9HPT, MSWS and SDMT are reported in **Table 6.3**.

Looking at brain measurements, patients with EDSS progression presented with increased number of new T2 lesions (Coeff=2.95; 95%CI=0.61, 5.29; p=0.01), higher T2 lesion volume change (Coeff=1075.53; 95%CI=479.09, 1671.97; p<0.01), and more pronounced PBVC (Coeff=-0.40; 95%CI=-0.54, -0.26; p<0.01). Patients with T25FWT progression presented with higher T1 lesion volume change (Coeff=142.88; 95%CI=23.78, 261.97; p=0.01). Patients with SDMT progression presented with increased number of new T2 lesions (Coeff=2.66; 95%CI=0.49, 4.84; p=0.01), higher T2 lesion volume change (Coeff=266.56; 95%CI=21.54, 511.58; p=0.03), and more pronounced PBVC (Coeff=-0.46; 95%CI=-0.58, -0.33; p<0.01).

Table 6.3. Spinal cord atrophy longitudinal changes from brain and spinal cord MRI.

Table shows spinal cord atrophy in relation to progression of EDSS, T25FWT, 9HPT, MSWS and SDMT, for different measurements. Results in bold indicate p<0.05 on stepwise linear regression models.

		MUCCA	**Brain MRI**		**Spinal cord MRI**			
			C1-2		**C1-2**		**C2-5**	
			CSA	*GBSI*	*CSA*	*GBSI*	*CSA*	*GBSI*
EDSS	Yes	-1.03±3.59	-0.93±4.22	-1.71±3.51	-1.39±4.32	-1.89±3.17	-1.63±4.64	-1.62±3.56
	No	-0.82±2.94	-0.89±4.31	-1.25±3.18	-0.53±3.56	-0.96±2.72	-1.20±5.27	-1.23±3.83
T25FWT	Yes	-1.33±3.02	-1.04±4.31	**-1.77±3.48**	-1.55±4.28	-1.84±3.42	-2.06±4.66	**-2.03±4.01**
	No	-0.75±3.10	-0.86±4.22	**-1.26±3.36**	-1.11±4.16	-1.63±3.00	-1.35±4.82	**-1.36±3.50**
9HPT	Yes	-1.16±2.35	-1.65±5.52	**-1.68±3.42**	-1.27±4.18	-2.15±4.22	-1.52±4.65	-2.69±2.73
	No	-0.98±3.05	-0.85±4.11	**-0.88±3.78**	-0.71±5.31	-1.61±3.07	-0.05±7.75	-1.45±3.64
MSWS	Yes	-0.95±3.19	-1.41±4.06	**-1.96±3.46**	-1.24±4.37	**-2.21±2.93**	-1.50±5.58	-2.15±3.42
	No	-0.51±2.74	-0.55±4.50	**-0.72±3.25**	-1.20±4.15	**-0.51±3.24**	-1.50±4.00	-1.22±3.58
SDMT	Yes	-1.38±3.09	-0.96±4.36	-1.87±3.45	-1.23±4.07	-1.80±2.86	-2.11±5.63	-1.84±2.98
	No	-0.59±3.07	-0.72±4.02	-1.59±3.47	-1.03±4.74	-1.68±3.21	-1.29±4.58	-1.14±3.75

6.4.6. Treatment effect

In the original ARPEGGIO analysis, no effect of treatment was seen on clinical or atrophy measures, including brain atrophy (using SIENA) and spinal cord atrophy (using MUCCA) [105]. In this post-hoc analysis, again, no treatment effect was detected with any of spinal cord atrophy measurements (**Table 6.4**).

6.5. Discussion

6.5.1. Main findings

In this chapter, I explored clinical correlates and implications for clinical trials of segmentation- and registration-based measurements of spinal cord atrophy in progressive MS, obtained at different levels from both brain and spinal cord MRI. The registration-based method (GBSI) is a candidate secondary outcome measure for clinical trials in MS, providing clinically meaningful measurements of spinal cord atrophy at relatively-low sample size. In particular, deriving spinal cord atrophy measurements from volumetric brain MRI using the GBSI would allow the use of spinal cord atrophy in multi-centre studies, where acquiring high quality MRI of the spinal cord is difficult to achieve.

6.5.2. Interpretation

Overall, I confirmed that GBSI measurements provide similar spinal cord volume change (1.5-1.8%/year), when compared with CSA (0.9-1.5%/year) [113, 268, 389], but are more precise, with smaller standard deviations, coefficients of variation and median absolute deviations. GBSI boundary contours are indeed less affected by partial volume effects (i.e., inclusion of tissue outside of the area of interest) with subsequent segmentation errors [113, 369, 389]. Also, I confirmed the strong clinical correlates of spinal cord atrophy [213, 309]. In this chapter, 1-year spinal cord atrophy

Table 6.4. Treatment effect on spinal cord atrophy.

Table shows spinal cord atrophy in relation to treatment with placebo or laquinimod 0.6 mg. Coeff, 95%CI and p-values are reported from linear regression models adjusted by age, gender, height, country and baseline CSA.

	Placebo	**Laquinimod**	**Coeff**	**95%CI**		**P-values**
		0.6 mg		*Lower*	*Upper*	
MUCCA	-0.81±3.20	-1.38±3.04	-0.25	-1.13	0.62	0.56
CSA at C1-2 level from brain MRI	-1.08±4.10	-1.14±4.12	-0.19	-1.48	1.09	0.76
GBSI at C1-2 level from brain MRI	-1.28±3.40	-2.14±3.51	-0.78	-1.99	0.41	0.19
CSA at C1-2 level from spinal cord MRI	-0.67±4.56	-1.51±4.25	-0.58	-2.55	1.38	0.55
GBSI at C1-2 level from spinal cord MRI	-1.47±2.56	-1.97±3.39	-0.56	-1.97	0.84	0.42
CSA at C2-5 level from spinal cord MRI	-0.80±4.77	-1.80±4.73	-0.42	-2.51	1.66	0.68
GBSI at C2-5 level from spinal cord MRI	-1.31±4.05	-1.51±2.94	-0.38	-2.09	1.32	0.65

measurements obtained with GBSI, but not with CSA, at different levels (C1-2 and C2-5) and from both brain and spinal cord MRI were associated with upper and lower limb motor function, as measured by neurologists (e.g., walking test, and hand test), and by people with MS (e.g., self-reported scale on walking difficulties). By comparison, brain MRI variables were associated with clinical measures of upper and lower limb motor function, and cognitive disability (e.g., EDSS and SDMT), suggesting they better depict the overall patients' status. Looking at previous longitudinal studies measuring spinal cord atrophy in PPMS, early changes in spinal cord area are associated with clinical changes in the long term (e.g., six to fifteen years) [307, 311, 390], making spinal cord atrophy a useful short term surrogate marker of disease progression, with even stronger clinical correlates in the long term.

This chapter has further highlighted the difficulties in acquiring good quality spinal cord images. Unexpectedly, the noise in the dedicated spinal cord MRI was higher than in spinal cord images derived from brain MRI. By applying an semi-automatic pipeline for spinal cord atrophy calculation, I was able to include from 65 to 95% of scans, depending on the spinal cord level. In the previous chapter on combined MAGNIMS and Queen Square spinal cord cohorts, I was able to include 85% of scans [113]. Overall, failures in SCT and GBSI were related to noise, poor contrast and/or artefacts, that are difficult to account for when operating with semi-automatic methods for spinal cord segmentation and registration. Of note, acquiring good quality images is easier for brain than spinal cord, as demonstrated by the variability of the MR signal (standard deviation increased from 43 on C1-2 level from brain MRI, to 57 and 76 on C1-2 and C2-5 levels on spinal cord MRI, respectively). As such, obtaining spinal cord atrophy measurements from brain scans could represent a viable, efficient and clinically meaningful alternative to more technically-challenging spinal cord images, in particular in multi-centre settings where homogenous spinal cord acquisitions are not feasible. So

far, the possibility to derive spinal cord measurements from brain scans has been successfully explored only with a cross-sectional design [282, 292, 296]. In this chapter, notwithstanding the large variability of centres acquiring images with different expertise, protocols and field strength, correlation coefficients were statistically, but minimally, higher for GBSI than CSA (0.3 vs 0.2), suggesting that longitudinal spinal cord atrophy could be better derived from brain scans using GBSI, than conventional segmentation-based methods (e.g., CSA). Differences in quality between brain and spinal cord MRI could explain the partial concordance of measurements, and could be due to the use of different coils, that can affect image quality at C1-2 level. Indeed, more advanced coils, optimised for both brain and spinal cord, can improve image quality, especially at upper cervical cord level, when compared with conventional coils [391]. Unfortunately, the ARPEGGIO dataset did not include coils at different sites to be accounted for in the statistical models.

Sample size estimates with spinal cord GBSI (e.g., 130-250 per treatment arm, for a 60% effect) are two-to-four fold smaller than CSA (e.g., 400-800 per treatment arm, for a 60% effect), and are on the same order of magnitude of brain atrophy (the pooled rate of 1-year brain atrophy in placebo- and laquinimod-treated patients from the ARPEGGIO trial (-0.45±1.00%) corresponds to 215 patients per treatment arm for a 60% treatment effect) [91, 105]. In Chapter 5 [113], I obtained even smaller estimates for both GBSI (50-100), and CSA (200-800), considering a similar treatment effect and also accounting for physiological spinal cord loss in healthy controls. However, that study was conducted on a much more selected population, using MRIs acquired within the MAGNIMS network [113]. In the present study, I tried to simulate similar conditions by classifying scans on the median quality; however, when including top quality scans, atrophy variability only partly improved, with increased effect size and reduced standard deviation for some measurements. Thus, future clinical trials aiming to use spinal cord atrophy as an outcome measure, rather than simply increase

the sample size, could only include sites that apply appropriate acquisition protocols, and/or restrict MRI acquisition to a limited number of sites with possibility to acquire high quality images. A consensus spinal cord acquisition protocol has been recently developed and could be used in future multicentre MRI studies on the spinal cord (http://www.spinalcordmri.org - Protocols), as already discussed in Chapter 5. Alternatively, longer follow-up could also be considered [289]; unfortunately, in the ARPEGGIO trial, no MRI scans were available beyond week 48.

6.5.3. Limitations

Limitations of the present study include the relatively short follow-up duration (48 weeks). On the one hand, this is in line with phase 2 clinical trial requirements in MS where brain atrophy is the primary outcome measure. However, one-year follow-up is relatively short to acquire sufficiently reliable clinical outcomes, that could have been better studied in the long term. Also, I used *DeepSeg* for spinal cord segmentation, which is an automated method with high repeatability, but providing smaller volumes than other methods (e.g., NeuroQLab and JIM) [285, 288]; I previously ruled out the impact of different segmentation methods on longitudinal spinal cord atrophy measurements on GBSI, but cannot exclude results using the segmentation-based method (CSA) were affected [389]. Also, I used the standard deviation of the MR signal to classify scans based on their quality [377], whilst other measures should be considered for future studies; for instance the contrast-to-noise ratio provides information on both functional performance (contrast) and data quality (noise caused by the system and physiology) [392, 393]. Additional limitations are related to the study design, including patients within a specific disability range, and, thus, limiting the possibilities of clinical correlates.

6.5.4. Conclusions

Imaging the spinal cord is challenging, but clinically relevant. Spinal cord atrophy can detect early subtle changes in motor function to the upper and lower limbs, with possible long-term more obvious clinical correlates, as a consequence of tract-specificity of spinal cord anatomy. Improvements in spinal cord acquisition, processing and analysis (e.g., SCT and GBSI), along with the possibility of deriving spinal cord atrophy measurements from brain MRI, can enhance the application of this clinically meaningful imaging outcome measure in multi-centre longitudinal observational studies and clinical trials. Rigorous quality control will be required though, as spinal cord scans (and the resulting measurements) remain more variable than those for brain atrophy. Despite power calculations not being far off from those for brain atrophy in the ARPEGGIO study, spinal cord atrophy is destined to remain a secondary outcome for the moment, given the issues around controlling image quality in the spinal cord, especially when people with more advanced levels of disability are to be included.

7. Conclusions and future directions

7.1. Introduction

In this thesis, I have validated the first registration-based method (GBSI) for quantification of spinal cord atrophy, expanding research possibilities for future MS projects, especially for collaborative multicentre studies and clinical trials. In particular, I have achieved thesis objectives by:

(1) Developing a registration-based method (GBSI) for quantification of spinal cord atrophy, which proved more precise, when compared with segmentation-based methods (i.e., CSA);
(2) Demonstrating stronger clinical correlates for this novel registration-based method (GBSI), when compared with segmentation-based methods (CSA), from the perspective of both neurologists and MS patients;
(3) Evaluating implication for the design and the conduction of future clinical trials and observational studies using the registration-based method (GBSI), especially in terms of sample size estimates and image acquisition.

The major findings presented in this thesis, and relating to each of these objectives, are summarised below, along with future directions.

7.2. Improving longitudinal spinal cord atrophy measurements with GBSI

In Chapter 4, I introduced a new pipeline based on the latest iteration of BSI for computing longitudinal spinal cord atrophy, to overcome limitations of commonly used segmentation-based methods. In following Chapter 5 and Chapter 6, I demonstrated that longitudinal spinal cord changes with GBSI hold better measurement precision than CSA.

Overall, GBSI and CSA provided similar rates of spinal cord atrophy, but the registration-based method (GBSI) was associated with lower variability and increased statistical power, providing smaller sample size estimates and stronger clinical correlates. In particular, in Chapter 5, in a longitudinal multicentre, multi-manufacturer and multi-field strength scan setting, I showed better performance of GBSI at differentiating between different MS subtypes and patients with disability progression, compared with segmentation-based method (CSA). Then, in Chapter 6, I confirmed that spinal cord atrophy on GBSI can detect early subtle changes in motor function of the upper and lower limbs. Also, I showed the feasibility of deriving spinal cord atrophy measurements from volumetric brain MRI using the GBSI, allowing the use of spinal cord atrophy in multi-centre studies, where acquiring high quality MRI of the spinal cord is difficult to achieve.

An important caveat of this thesis is that good quality images (e.g., dedicated spinal cord 3DT1 images with 1mm isotropic voxel) are needed for measuring spinal cord atrophy accurately. Depending on the spinal cord segment, the dataset and the segmentation method, I was able to process from 65% to 95% of scans, suggesting that rigorous quality control would be required for future acquisitions. Thus, for the moment, spinal cord atrophy measurements with GBSI should remain a secondary, but clinically relevant, outcome measure, until issues around controlling image quality are solved, especially when people with more advanced levels of disability are to be included.

7.3. Future directions

Future research will focus on characterising spinal cord pathology in a clinical setting (i.e., measurement of spinal cord atrophy with GBSI on clinical MRI acquisitions). In particular, spinal cord atrophy could be combined to other clinical and brain MRI outcome measures to further improve the surveillance of the clinical course of MS and its response to treatment. Indeed, combined

endpoints have become increasingly common in MS, and are conventionally based on relapses, disability progression, brain lesions, and brain atrophy [394, 395]; as such, the inclusion of spinal cord atrophy could further improve statistical power and clinical correlates.

BSI is a registration method which measures intensity changes and, hence, is currently applicable over T1-weighted images, as mentioned in Chapter 4. Further improvements to the GBSI technique will require the application to other sequences (i.e., PSIR or T2*), pending changes in the integral and in the intensity normalisation steps [396, 397]. The possibility of using PSIR and T2* lesions will allow spinal cord lesion masking and filling, which can make atrophy estimates more precise, as it happened for brain atrophy [398–400].

GBSI will need to be tested on images acquired on ultra-high-field (7T) scanners, which have appeared in recent years. 7T MRI of the spinal cord can potentially overcome limitations of 1.5 and 3T scanners, by improving spatial resolution, increasing the contrast-to-noise ratio, and allowing better characterisation of WM and GM [401, 402]. Preliminary reports have shown increased sensitivity and spatial accuracy in characterizing pathology in the spinal cord than lower field MRI [403]. However, 7T spinal cord MRI remains technically challenging due to motion artefacts and field inhomogeneities, and requires time-consuming acquisition and complex post-processing. New coils that reduce field inhomogeneities and allow whole spine coverage will help to overcome these limitations and develop this exciting tool in MS.

The ultimate validation of the GBSI would require pathology samples of the spinal cord, that could be acquired with a scan-rescan fashion, where atrophy is expected to be zero. As such, we could

evaluate the measurement which actually performs the best, and, not least, the pathology background of spinal cord atrophy [23, 24]

Finally, the GBSI should be considered for measuring spinal cord atrophy in other neurological diseases. Indeed, spinal cord atrophy does not only occur in MS, but also in other inflammatory diseases of the central nervous system, such as NMO [302, 303, 404], and human T- lymphotropic virus 1 (HTLV-1)-associated myelopathy/tropical spastic paraparesis [405]. Not least, spinal cord atrophy is associated with more severe clinical features in amyotrophic lateral sclerosis [406–408], spinal cord injury [409], and Friedreich's ataxia [410], and with worse recovery from spinal cord surgery [411]. There is also preliminary evidence of spinal cord atrophy in Alzheimer's disease [412], and in Huntington disease [413], suggesting that spinal cord atrophy can be part of more widespread neurodegenerative diseases.

7.4. Conclusions

Progressive MS represents a unique opportunity for studying imaging markers of neurodegeneration. Several imaging candidates hold promise for filling the unmet need of biomarkers in progressive MS, and brain atrophy is currently the best examined and most robust outcome, with attainable sample sizes and first positive results in phase 2 and 3 trials.

Imaging the spinal cord is challenging, but clinically relevant. Following the present thesis, improvements in spinal cord processing, along with the possibility of deriving spinal cord atrophy measurements from brain MRI, can enhance the application of this clinically meaningful imaging outcome to multi-centre longitudinal observational studies and clinical trials [109, 414]. For instance, results from this thesis will be particularly helpful to the United Kingdom MS Society

Efficient Clinical Trial Platform [415], to study the impact of different medications on the spinal cord atrophy. Not least, GBSI could prove useful in many neurological diseases where spinal cord volume changes are representative of the most aggressive aspects of the diseases (e.g., spinal cord injury, amyotrophic lateral sclerosis). Thus, accurate and sensitive-to-change monitoring of spinal cord atrophy over time can shed light on the most aggressive aspects of many neurological diseases, and the suggested GBSI pipeline is a suitable endpoint in studies of neuroprotection, also combined with other MRI measures.

References

1. Thompson AJ, Baranzini SE, Geurts J, et al (2018) Multiple sclerosis. Lancet 391:1622–1636 . doi: 10.1016/S0140-6736(18)30481-1
2. Waubant E, Lucas R, Mowry E, et al (2019) Environmental and genetic risk factors for MS: an integrated review. Ann Clin Transl Neurol 6:1905–1922 . doi: 10.1002/acn3.50862
3. Compston A, Coles A (2008) Multiple sclerosis. Lancet 372:1502–1517 . doi: 10.1016/S0140-6736(08)61620-7
4. Trapp BD, Nave K-A (2008) Multiple Sclerosis: An Immune or Neurodegenerative Disorder? Annu Rev Neurosci 31:247–269 . doi: 10.1146/annurev.neuro.30.051606.094313
5. Wallin MT, Culpepper WJ, Nichols E, et al (2019) Global, regional, and national burden of multiple sclerosis 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 18:269–285 . doi: 10.1016/S1474-4422(18)30443-5
6. Scalfari A (2019) MS progression is predominantly driven by age-related mechanisms – YES. Mult Scler 25:902–904 . doi: 10.1177/1352458518820633
7. Robertson N, Clayton D, Fraser M, et al (1996) Clinical concordance in sibling pairs with multiple sclerosis. Neurology 47:347–352
8. Carton H, Vlietinck R, Debruyne J, et al (1997) Risks of multiple sclerosis in relatives of patients in Flanders, Belgium. J Neurol Neurosurg Psychiatry 62:329–33
9. Mumford C, Wood N, Kellar-Wood H, et al (1994) The British Isles survey of multiple sclerosis in twins. Neurology 44:11–15
10. Willer CJ, Dyment DA, Risch NJ, et al (2003) Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Natl Acad Sci 100:12877–12882 . doi: 10.1073/pnas.1932604100
11. Svejgaard A (2008) The immunogenetics of multiple sclerosis. Immunogenetics 60:275–286 . doi: 10.1007/s00251-008-0295-1

12. DeLuca G, Ramagopalan S, Herrera B, et al (2007) An extremes of outcome strategy provides evidence that multiple sclerosis severity is determined by alleles at the HLA-DRB1 locus. Proc Natl Acad Sci U S A 104:20896–20901

13. Høglund RA, Maghazachi AA (2014) Multiple sclerosis and the role of immune cells. World J Exp Med 4:27–37 . doi: 10.5493/wjem.v4.i3.27

14. Steri M, Orrù V, Idda ML, et al (2017) Overexpression of the Cytokine BAFF and Autoimmunity Risk. N Engl J Med 376:1615–1626 . doi: 10.1056/NEJMoa1610528

15. Simpson S, Blizzard L, Otahal P, et al (2011) Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. J Neurol Neurosurg Psychiatry 82:1132–1141 . doi: 10.1136/jnnp.2011.240432

16. Compston A, Coles A (2008) Multiple sclerosis. Lancet (London, England) 372:1502–17 . doi: 10.1016/S0140-6736(08)61620-7

17. Ascherio A, Munger KL (2007) Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. Ann Neurol 61:504–513 . doi: 10.1002/ana.21141

18. Ascherio A, Munger KL (2007) Environmental risk factors for multiple sclerosis. Part I: The role of infection. Ann Neurol 61:288–299 . doi: 10.1002/ana.21117

19. Lang HLE, Jacobsen H, Ikemizu S, et al (2002) A functional and structural basis for TCR cross-reactivity in multiple sclerosis. Nat Immunol 3:940–943 . doi: 10.1038/ni835

20. Ascherio A, Munger KL, Lennette ET, et al (2001) Epstein-Barr virus antibodies and risk of multiple sclerosis: a prospective study. JAMA 286:3083–8

21. Peterson JW, Bö L, Mörk S, et al (2001) Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. Ann Neurol 50:389–400 . doi: 10.1002/ana.1123

22. Kutzelnigg A, Lucchinetti CF, Stadelmann C, et al (2005) Cortical demyelination and diffuse

white matter injury in multiple sclerosis. Brain 128:2705–2712 . doi: 10.1093/brain/awh641

23. Petrova N, Carassiti D, Altmann DR, et al (2018) Axonal loss in the multiple sclerosis spinal cord revisited. Brain Pathol 28:334–348 . doi: 10.1111/bpa.12516

24. Schmierer K, McDowell A, Petrova N, et al (2018) Quantifying multiple sclerosis pathology in post mortem spinal cord using MRI. Neuroimage 182:251–258 . doi: 10.1016/j.neuroimage.2018.01.052

25. Lassmann H (2015) Spinal cord pathology in multiple sclerosis. Lancet Neurol 14:348–349 . doi: 10.1016/S1474-4422(15)70037-2

26. Bramow S, Frischer JM, Lassmann H, et al (2010) Demyelination versus remyelination in progressive multiple sclerosis. Brain 133:2983–2998 . doi: 10.1093/brain/awq250

27. Frischer JM, Weigand SD, Guo Y, et al (2015) Clinical and pathological insights into the dynamic nature of the white matter multiple sclerosis plaque. Ann Neurol 78:710–721 . doi: 10.1002/ana.24497

28. Lassmann H (2018) Multiple sclerosis pathology. Cold Spring Harb Perspect Med 8:1–16 . doi: 10.1101/cshperspect.a028936

29. Barkhof F, Scheltens P, Frequin S, et al (1992) Relapsing-remitting multiple sclerosis: sequential enhanced MR imaging vs clinical findings in determining disease activity. AJR Am J Roentgenol 159:1041–1047

30. Lassmann H (1998) Neuropathology in multiple sclerosis: new concepts. Mult Scler J 4:93–98 . doi: 10.1177/135245859800400301

31. Lucchinetti CF, Brück W, Parisi JE, et al (2000) Heterogeneity of Multiple Sclerosis Lesions: Implications for the Pathogenesis of Demyelinatio. Ann Neurol 47:707–717

32. Nave K-A, Trapp BD (2008) Axon-Glial Signaling and the Glial Support of Axon Function. Annu Rev Neurosci 31:535–561 . doi: 10.1146/annurev.neuro.30.051606.094309

33. Trapp BD, Peterson J, Ransohoff RM, et al (1998) Axonal Transection in the Lesions of Multiple Sclerosis. N Engl J Med 338:278–285 . doi: 10.1056/NEJM199801293380502

34. Schmierer K, Scaravilli F, Altmann DR, et al (2004) Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain. Ann Neurol 56:407–415 . doi: 10.1002/ana.20202

35. DeLuca GC, Williams K, Evangelou N, et al (2006) The contribution of demyelination to axonal loss in multiple sclerosis. Brain 129:1507–1516 . doi: 10.1093/brain/awl074

36. Mottershead J, Schmierer K, Clemence M, et al (2003) High field MRI correlates of myelin content and axonal density in multiple sclerosis--a post-mortem study of the spinal cord. J Neurol 11:1301

37. Criste G, Trapp B, Dutta R (2014) Axonal loss in multiple sclerosis. Causes and mechanisms

38. Trapp BD, Ransohoff RM, Fisher E, Rudick RA (1999) Neurodegeneration in Multiple Sclerosis: Relationship to Neurological Disability. Neurosci 5:48–57 . doi: 10.1177/107385849900500107

39. Trapp BD, Vignos M, Dudman J, et al (2018) Cortical neuronal densities and cerebral white matter demyelination in multiple sclerosis: a retrospective study. Lancet Neurol 17:870–884 . doi: 10.1016/S1474-4422(18)30245-X

40. Haider L, Zrzavy T, Hametner S, et al (2016) The topograpy of demyelination and neurodegeneration in the multiple sclerosis brain. Brain 139:807–815 . doi: 10.1093/brain/awv398

41. Haider L (2015) Inflammation, Iron, Energy Failure, and Oxidative Stress in the Pathogenesis of Multiple Sclerosis. Oxid Med Cell Longev 725370 . doi: 10.1155/2015/725370

42. Haider L, Fischer MT, Frischer JM, et al (2011) Oxidative damage in multiple sclerosis lesions. Brain 134:1914–1924 . doi: 10.1093/brain/awr128

43. Faissner S, Plemel JR, Gold R, Yong VW (2019) Progressive multiple sclerosis: from

pathophysiology to therapeutic strategies. Nat Rev Drug Discov. doi: 10.1038/s41573-019-0035-2

44. Moccia M, Ciccarelli O (2017) Molecular and Metabolic Imaging in Multiple Sclerosis. Neuroimaging Clin N Am 27:343–356 . doi: 10.1016/j.nic.2016.12.005

45. Lassmann H (2019) Pathogenic Mechanisms Associated With Different Clinical Courses of Multiple Sclerosis. Front Immunol 9:1–14 . doi: 10.3389/fimmu.2018.03116

46. Dendrou C, Fugger L, Friese M (2015) Immunopathology of multiple sclerosis. Nat Rev Immunol 15:545–558 . doi: 10.1038/nri3871

47. Kasper LH, Shoemaker J (2010) Multiple sclerosis immunology: The healthy immune system vs the MS immune system. Neurology 74:S2–S8 . doi: 10.1212/WNL.0b013e3181c97c8f

48. Gandhi R, Laroni A, Weiner HL (2010) Role of the innate immune system in the pathogenesis of multiple sclerosis. J Neuroimmunol 221:7–14 . doi: 10.1016/j.jneuroim.2009.10.015

49. Bell J, Spencer J, Yates R, DeLuca G (2018) The cortical blood-brain barrier in multiple sclerosis: a gateway to progression? J Neurol 265:966–967 . doi: 10.1007/s00415-017-8727-1

50. Spencer J, Bell J, DeLuca G (2018) Vascular pathology in multiple sclerosis: reframing pathogenesis around the blood-brain barrier. J Neurol Neurosurg Psychiatry 89:42–52 . doi: 10.1136/jnnp-2017-316011

51. Dendrou CA, Fugger L, Friese MA (2015) Immunopathology of multiple sclerosis. Nat Rev Immunol 15:545–558 . doi: 10.1038/nri3871

52. Machado-Santos J, Saji E, Tröscher AR, et al (2018) The compartmentalized inflammatory response in the multiple sclerosis brain is composed of tissue-resident CD8+ T lymphocytes and B cells. Brain 141:2066–2082 . doi: 10.1093/brain/awy151

53. Kasper LH, Shoemaker J (2010) Multiple sclerosis immunology: The healthy immune system

vs the MS immune system. Neurology 74:S2 . doi: 10.1212/WNL.0b013e3181c97c8f

54. Howell OW, Reeves CA, Nicholas R, et al (2011) Meningeal inflammation is widespread and linked to cortical pathology in multiple sclerosis. Brain 134:2755–2771 . doi: 10.1093/brain/awr182

55. Magliozzi R, Howell OW, Reeves C, et al (2010) A Gradient of neuronal loss and meningeal inflammation in multiple sclerosis. Ann Neurol 68:477–493 . doi: 10.1002/ana.22230

56. Magliozzi R, Serafini B, Rosicarelli B, et al (2013) B-cell enrichment and epstein-barr virus infection in inflammatory cortical lesions in secondary progressive multiple sclerosis. J Neuropathol Exp Neurol 72:29–41 . doi: 10.1097/NEN.0b013e31827bfc62

57. Frischer JM, Bramow S, Dal-Bianco A, et al (2009) The relation between inflammation and neurodegeneration in multiple sclerosis brains. Brain 132:1175–1189 . doi: 10.1093/brain/awp070

58. Lassmann H (2018) Multiple Sclerosis Pathology and its Reflection by Imaging Technologies: Introduction. Brain Pathol 28:721–722 . doi: 10.1111/bpa.12649

59. Polman CH, Reingold SC, Banwell B, et al (2011) Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. Ann Neurol 69:292–302 . doi: 10.1002/ana.22366

60. Lublin FD, Reingold SC, Cohen JA, et al (2014) Defining the clinical course of multiple sclerosis : The 2013 revisions. Neurology 83:278–286 . doi: 10.1212/WNL.0000000000000560

61. Thompson AJ, Banwell BL, Barkhof F, et al (2018) Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 17:162–173 . doi: 10.1016/S1474-4422(17)30470-2

62. Plantone D, De Angelis F, Doshi A, Chataway J (2016) Secondary Progressive Multiple Sclerosis: Definition and Measurement. CNS Drugs 30:517–526 . doi: 10.1007/s40263-016-0340-9

63. Tur C, Moccia M, Barkhof F, et al (2018) Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. Nat Rev Neurol 14:75–93 . doi: 10.1038/nrneurol.2017.171

64. Lublin FD (2014) New multiple sclerosis phenotypic classification. Eur Neurol 72:1–5 . doi: 10.1159/000367614

65. Confavreux C, Vukusic S (2014) The clinical course of multiple sclerosis

66. Kurtzke JF (1983) Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 33:1444–1452 . doi: 10.1212/WNL.33.11.1444

67. Sorensen PS, Blinkenberg M (2015) The potential role for ocrelizumab in the treatment of multiple sclerosis: current evidence and future prospects. Ther Adv Neurol Disord 9:44–52 . doi: 10.1177/1756285615601933

68. Goldman M, Ward M, Motl R, et al (2017) Identification and validation of clinically meaningful benchmarks in the 12-item Multiple Sclerosis Walking Scale. Mult Scler 23:1405–1414 . doi: 10.1177/1352458516680749

69. Motl R, Sandroff B, McAuley E (2018) Naturally occurring change in Multiple Sclerosis Walking Scale-12 scores over time in multiple sclerosis. Neurodegener Dis Manag 8:315–322 . doi: 10.2217/nmt-2018-0016

70. Saccà F, Costabile T, Carotenuto A, et al (2017) The EDSS integration with the Brief International Cognitive Assessment for Multiple Sclerosis and orientation tests. Mult Scler 23: . doi: 10.1177/1352458516677592

71. Moccia M, Lanzillo R, Palladino R, et al (2016) Cognitive impairment at diagnosis predicts 10-year multiple sclerosis progression. Mult Scler 22:659–667 . doi: 10.1177/1352458515599075

72. DeLuca GC, Yates RL, Beale H, Morrow SA (2015) Cognitive impairment in multiple sclerosis: Clinical, radiologic and pathologic insights. Brain Pathol 25:79–98 . doi: 10.1111/bpa.12220

73. Muhlert N, Sethi V, Schneider T, et al (2013) Diffusion MRI-based cortical complexity

alterations associated with executive function in multiple sclerosis. J Magn Reson Imaging 38:54–63 . doi: 10.1002/jmri.23970

74. Bozzali M, Spanò B, Parker G, et al (2013) Anatomical brain connectivity can assess cognitive dysfunction in multiple sclerosis. Mult Scler 19:1161–1168 . doi: 10.1177/1352458512474088

75. Van Schependom J, D'hooghe MB, Cleynhens K, et al (2014) The Symbol Digit Modalities Test as sentinel test for cognitive impairment in multiple sclerosis. Eur J Neurol 21:1219–1225 . doi: 10.1111/ene.12463

76. Parmenter BA, Weinstock-Guttman B, Garg N, et al (2007) Screening for cognitive impairment in multiple sclerosis using the Symbol digit Modalities Test. Mult Scler 13:52–57 . doi: 10.1177/1352458506070750

77. Benedict RHB, Deluca J, Phillips G, et al (2017) Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. Mult Scler 23:721–733 . doi: 10.1177/1352458517690821

78. McDonald WI, Compston A, Edan G, et al (2001) Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis. Ann Neurol 50:121–127

79. Brownlee W, Swanton J, Altmann D, et al (2015) Earlier and more frequent diagnosis of multiple sclerosis using the McDonald criteria. J Neurol Neurosurg Psychiatry 86:584–585 . doi: 10.1136/jnnp-2014-308675

80. Brownlee WJ, Swanton JK, Miszkiel KA, et al (2016) Should the symptomatic region be included in dissemination in space in MRI criteria for MS? Neurology 87:680–683 . doi: 10.1212/WNL.0000000000002975

81. Scalfari A, Neuhaus A, Daumer M, et al (2013) Early relapses, onset of progression, and late outcome in multiple sclerosis. JAMA Neurol 70:214–22 . doi: 10.1001/jamaneurol.2013.599

82. Moccia M, Palladino R, Carotenuto A, et al (2016) Predictors of long-term interferon discontinuation in newly diagnosed relapsing multiple sclerosis. Mult Scler Relat Disord 10:90–96 . doi: 10.1016/j.msard.2016.09.011

83. Tintore M, Rovira a., Rio J, et al (2015) Defining high, medium and low impact prognostic factors for developing multiple sclerosis. Brain 138:1863–1874 . doi: 10.1093/brain/awv105

84. Yates R, Esiri M, Palace J, et al (2015) The influence of HLA-DRB1*15 on motor cortical pathology in multiple sclerosis. Neuropathol Appl Neurobiol 41:371–384 . doi: 10.1111/nan.12165

85. DeLuca G, Alterman R, Martin J, et al (2013) Casting light on multiple sclerosis heterogeneity: the role of HLA-DRB1 on spinal cord pathology. Brain 136:1025–1034 . doi: 10.1093/brain/awt031

86. Weinshenker BG, Bass B, Rice GP, et al (1989) The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. Brain 112:133–46

87. Giovannoni G, Cutter G, Pia-Sormani M, et al (2017) Is multiple sclerosis a length-dependent central axonopathy? The case for therapeutic lag and the asynchronous progressive MS hypotheses. Mult Scler Relat Disord 12:70–78 . doi: 10.1016/j.msard.2017.01.007

88. Arrambide G, Tintore M, Espejo C, et al (2018) The value of oligoclonal bands in the multiple sclerosis diagnostic criteria. Brain 141:1075–1084 . doi: 10.1093/brain/awy006

89. Montalban X, Gold R, Thompson A, et al (2018) ECTRIMS/EAN guideline on the pharmacological treatment of patients with multiple sclerosis. Eur J Neurol 25:215–237 . doi: 10.1111/ene.13536

90. Filippini G, Del Giovane C, Clerico M, et al (2017) Treatment with disease-modifying drugs for people with a first clinical attack suggestive ofmultiple sclerosis (Review). Cochrane Database Syst Rev 4:CD012200 . doi: 10.1002/14651858.CD012200.pub2

91. Moccia M, de Stefano N, Barkhof F (2017) Imaging outcomes measures for progressive multiple sclerosis trials. Mult Scler 23:1614–1626 . doi: 10.1177/1352458517729456

92. De Angelis F, John N, Brownlee W (2018) Disease-modifying therapies for multiple sclerosis. BMJ 363:k4674 . doi: 10.1136/bmj.k4674

93. Dobson R, Giovannoni G (2019) Multiple sclerosis – a review. Eur J Neurol 26:27–40 . doi: 10.1111/ene.13819

94. Signori A, Gallo F, Bovis F, et al (2016) Long-term impact of interferon or Glatiramer acetate in multiple sclerosis: A systematic review and meta-analysis. Mult Scler Relat Disord 6:57–63 . doi: 10.1016/j.msard.2016.01.007

95. Claflin SB, Tan B, Taylor B V. (2019) The long-term effects of disease modifying therapies on disability in people living with multiple sclerosis: A systematic review and meta-analysis. Mult Scler Relat Disord 36:101374 . doi: 10.1016/j.msard.2019.08.016

96. Kingwell E, Leray E, Zhu F, et al (2019) Multiple sclerosis: effect of beta interferon treatment on survival. Brain 142:1324–1333 . doi: https://doi.org/10.1093/brain/awz055

97. Brown JWL, Coles A, Horakova D, et al (2019) Association of Initial Disease-Modifying Therapy with Later Conversion to Secondary Progressive Multiple Sclerosis. JAMA 321:175–187 . doi: 10.1001/jama.2018.20588

98. NICE (2019) Disease-modifying therapies for multiple sclerosis

99. England N Treatment Algorithm for Multiple Sclerosis Disease-modifying Therapies

100. EMA (2019) Lemtrada for multiple sclerosis: measures to minimise risk of serious side effects

101. Lubiln F, Miller D, Freedman M, et al (2016) Oral fingolimod versus placebo in primary progressive multiple sclerosis: results of a large phase III, randomised trial. Lancet 387:1075–1084

102. Steiner D, Dl A, Ms F, et al (2016) Natalizumab Versus Placebo in Patients with Secondary

Progressive Multiple Sclerosis (SPMS): Results from ASCEND, a Multicenter, Double-blind, Placebo-Controlled, Randomized Phase 3 Clinical Trial. AAN 375

103. Hawker K, O'Connor P, Freedman MS, et al (2009) Rituximab in patients with primary progressive multiple sclerosis: Results of a randomized double-blind placebo-controlled multicenter trial. Ann Neurol 66:460–471 . doi: 10.1002/ana.21867
104. Faissner S, Gold R (2018) Efficacy and Safety of the Newer Multiple Sclerosis Drugs Approved Since 2010. CNS Drugs 32:269–287 . doi: 10.1007/s40263-018-0488-6
105. Barkhof F, Giovannoni G, Hartung H, et al (2015) ARPEGGIO: A randomized, placebo-controlled study to evaluate oral laquinimod in patients with primary progressive multiple sclerosis (PPMS). Neurology 84:P7.210
106. Sormani MP, Bruzzi P (2013) MRI lesions as a surrogate for relapses in multiple sclerosis: A meta-analysis of randomised trials. Lancet Neurol 12:669–676 . doi: 10.1016/S1474-4422(13)70103-0
107. Miller D, Leary S (2007) Primary-progressive multiple sclerosis. Lancet Neurol 6:903–912
108. Zaratin P, Comi G, Coetzee T, et al (2016) Progressive MS Alliance Industry Forum: Maximizing Collective Impact To Enable Drug Development. Trends Pharmacol Sci 37:808–810 . doi: 10.1016/j.tips.2016.07.005
109. Moccia M, Ruggieri S, Ianniello A, et al (2019) Advances in spinal cord imaging in multiple sclerosis. Ther Adv Neurol Disord 12:1756286419840593 . doi: 10.1177/1756286419840593
110. Bö L, Geurts J, Ravid R, Barkhof F (2004) Magnetic resonance imaging as a tool to examine the neuropathology of multiple sclerosis. Neuropath Appl Neuro 30:106–117 . doi: 10.1111/j.1365-2990.2004.00521.x
111. Filippi M, Rocca MA, Barkhof F, et al (2012) Association between pathological and MRI findings in multiple sclerosis. Lancet Neurol 11:349–360 . doi: 10.1016/S1474-

4422(12)70003-0

112. Hubbard P, Zhou F, Eichhorn S, Parker G (2015) Biomimetic phantom for the validation of diffusion magnetic resonance imaging. Magn Reson Med 73:299–305 . doi: 10.1002/mrm.25107

113. Moccia M, Prados F, Filippi M, et al (2019) Longitudinal spinal cord atrophy in multiple sclerosis using the generalised boundary shift integral. Ann Neurol 86:704–713 . doi: 10.1002/ana.25571

114. Cortese R, Ciccarelli O (2018) Clinical monitoring of multiple sclerosis should routinely include spinal cord imaging – Yes. Mult Scler 24:1536–1537

115. Losseff N, Webb S, O'Riordan J, et al (1996) Spinal cord atrophy and disability in multiple sclerosis. A new reproducible and sensitive MRI method with potential to monitor disease progression. Brain 119:701–708

116. Schmierer K, Miquel ME (2018) Magnetic resonance imaging correlates of neuro-axonal pathology in the MS spinal cord. Brain Pathol 28:765–772 . doi: 10.1111/bpa.12648

117. Khaleeli Z, Ciccarelli O, Mizskiel K, et al (2010) Lesion enhancement diminishes with time in primary progressive multiple sclerosis. Mult Scler 16:317–324 . doi: 10.1177/1352458509358090

118. Lublin F, Miller DH, Freedman MS, et al (2016) Oral fingolimod in primary progressive multiple sclerosis (INFORMS): A phase 3, randomised, double-blind, placebo-controlled trial. Lancet 387:1075–1084 . doi: 10.1016/S0140-6736(15)01314-8

119. Hommes OR, Sørensen PS, Fazekas F, et al (2004) Intravenous immunoglobulin in secondary progressive multiple sclerosis: randomised placebo-controlled trial. Lancet (London, England) 364:1149–56 . doi: 10.1016/S0140-6736(04)17101-8

120. Montalban X, Hauser SL, Kappos L, et al (2017) Ocrelizumab versus Placebo in Primary

Progressive Multiple Sclerosis. N Engl J Med 376:209–220 . doi: 10.1056/NEJMoa1606468

121. Chataway J, Schuerer N, Alsanousi A, et al (2014) Eff ect of high-dose simvastatin on brain atrophy and disability in secondary progressive multiple sclerosis (MS-STAT): A randomised, placebo-controlled, phase 2 trial. Lancet 383:2213–2221 . doi: 10.1016/S0140-6736(13)62242-4

122. ClinicalTrials.gov (2016) A Phase 2 Clinical Study in Subjects With Primary Progressive Multiple Sclerosis to Assess the Efficacy, Safety and Tolerability of Two Oral Doses of Laquinimod Either of 0.6 mg/Day or 1.5mg/Day (Experimental Drug) as Compared to Placebo. https://clinicaltrials.gov/ct2/show/NCT02284568

123. European Study Group on Interferon beta-1b in Secondary Progressive MS (1998) Placebo-controlled multicentre randomised trial of interferon β-1b in treatment of secondary progressive multiple sclerosis. Lancet 352:1491–1497 . doi: 10.1016/S0140-6736(98)10039-9

124. Kapoor R, Furby J, Hayton T, et al (2010) Lamotrigine for neuroprotection in secondary progressive multiple sclerosis: a randomised, double-blind, placebo-controlled, parallel-group trial. Lancet Neurol 9:681–688 . doi: 10.1016/S1474-4422(10)70131-9

125. ClinicalTrials.gov (2016) Dimethyl Fumarate Treatment of Primary Progressive Multiple Sclerosis (FUMAPMS). https://clinicaltrials.gov/ct2/show/NCT02959658

126. Connick P, Kolappan M, Crawley C, et al (2012) Autologous mesenchymal stem cells for the treatment of secondary progressive multiple sclerosis: An open-label phase 2a proof-of-concept study. Lancet Neurol 11:150–156 . doi: 10.1016/S1474-4422(11)70305-2

127. Fox RJ, Coffey CS, Cudkowicz ME, et al (2016) Design, rationale, and baseline characteristics of the randomized double-blind phase II clinical trial of ibudilast in progressive multiple sclerosis. Contemp Clin Trials 50:166–177 . doi: 10.1016/j.cct.2016.08.009

128. Cambron M, Mostert J, Haentjens P, et al (2014) Fluoxetine in progressive multiple sclerosis

(FLUOX-PMS): study protocol for a randomized controlled trial. Trials 15:37 . doi: 10.1186/1745-6215-15-37

129. Schreiber K, Magyari M, Sellebjerg F, et al (2016) High-dose erythropoietin in patients with progressive multiple sclerosis: A randomized, placebo-controlled, phase 2 trial. Mult Scler 1–11 . doi: 10.1177/1352458516661048

130. Romme Christensen J, Ratzner R, Bornsen L, et al (2014) Natalizumab in progressive MS. Neurology 82:1499–1507 . doi: 10.1212/WNL.0000000000000361

131. The North American Study Group on Interferon beta-1b in Secondary Progressive MS (2004) Interferon beta-1b in secondary progressive MS. Neurology 63:1788–1795

132. Li DKB, Zhao GJ, Paty DW (2001) Randomized controlled trial of interferon- beta-1a in secondary progressive MS MRI results. Neurology 56:1505–1513

133. Freedman MS, Traboulsee A, Bar-Or A, et al (2011) A phase III study evaluating the efficacy and safety of MBP8298 in secondary progressive MS. Neurology 77:1551–60

134. ClinicalTrials.gov (2016) Exploring the Efficacy and Safety of Siponimod in Patients With Secondary Progressive Multiple Sclerosis (EXPAND)

135. ClinicalTrials.gov (2016) A Clinical Study of the Efficacy of Natalizumab on Reducing Disability Progression in Participants With Secondary Progressive Multiple Sclerosis (ASCEND in SPMS)

136. Rice C, Marks D, Ben-Shlomo Y, et al (2015) Assessment of bone marrow-derived Cellular Therapy in progressive Multiple Sclerosis (ACTiMuS): study protocol for a randomised controlled trial. Trials 16:463 . doi: 10.1186/s13063-015-0953-1

137. Rice GP, Filippi M, Comi G (2000) Cladribine and progressive MS: clinical and MRI outcomes of a multicenter controlled trial. Cladribine MRI Study Group. Neurology 54:1145–1155 . doi: 10.1212/WNL.54.5.1145

138. Zajicek J, Ball S, Wright D, et al (2013) Effect of dronabinol on progression in progressive

multiple sclerosis (CUPID): a randomised, placebo-controlled trial. Lancet Neurol 12:857–865 . doi: 10.1016/S1474-4422(13)70159-5

139. Wolinsky J, Narayana P, O'Connor P, et al (2007) Glatiramer acetate in primary progressive multiple sclerosis: results of a multinational, multicenter, double-blind, placebo-controlled trial. Ann Neurol 61:14–24

140. Cohen J, Cutter G, Fischer J, et al (2002) Benefit of interferon beta-1a on MSFC progression in secondary progressive MS. Neurology 59:679–687

141. Karussis DM, Lehmann D, Linde A (1996) Treatment of secondary progressive multiple sclerosis with the immunomodulator linornide : Neurology 47:341–346

142. ClinicalTrials.gov (2016) A Study of NeuroVax^{TM}, a Therapeutic TCR Peptide Vaccine for SPMS of Multiple Sclerosis

143. Kappos L, Bar-Or A, Cree BAC, et al (2018) Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. Lancet. doi: 10.1016/S0140-6736(18)30475-6

144. Yaldizli Ö, MacManus D, Stutters J, et al (2015) Brain and cervical spinal cord atrophy in primary progressive multiple sclerosis: results from a placebo-controlled phase III trial (INFORMS). Mult Scler 22:30–31

145. La Mantia L, Vacchi L, Rovaris M, et al (2013) Interferon β for secondary progressive multiple sclerosis: a systematic review. J Neurol Neurosurg Psychiatry 84:420–426 . doi: 10.1136/jnnp-2012-303291

146. Ontaneda D, Fox RJ (2017) Imaging as an Outcome Measure in Multiple Sclerosis. Neurotherapeutics 14:24–34 . doi: 10.1007/s13311-016-0479-6

147. Rovaris M, Judica E, Sastre-Garriga J, et al (2008) Large-scale, multicentre, quantitative MRI study of brain and cord damage in primary progressive multiple sclerosis. Mult Scler 14:455–

464 . doi: 10.1177/1352458507085129

148. Khaleeli Z, Ciccarelli O, Manfredonia F, et al (2008) Predicting progression in primary progressive multiple sclerosis: a 10-year multicenter study. Ann Neurol 63:790–793 . doi: 10.1002/ana.21375

149. Eshaghi A, Prados F, Brownlee WJ, et al (2018) Deep gray matter volume loss drives disability worsening in multiple sclerosis. Ann Neurol 83:210–222 . doi: 10.1002/ana.25145

150. Fisniku LK, Chard DT, Jackson JS, et al (2008) Gray matter atrophy is related to long-term disability in multiple sclerosis. Ann Neurol 64:247–254 . doi: 10.1002/ana.21423

151. Rocca MA, Absinta M, Filippi M (2012) The role of advanced magnetic resonance imaging techniques in primary progressive MS. J Neurol 259:611–621 . doi: 10.1007/s00415-011-6195-6

152. Eshaghi A, Marinescu R V., Young AL, et al (2018) Progression of regional grey matter atrophy in multiple sclerosis. Brain. doi: 10.1093/brain/awy088

153. Healy B, Valsasina P, Filippi M, Bakshi R (2009) Sample size requirements for treatment effects using gray matter, white matter and whole brain volume in relapsing-remitting multiple sclerosis. J Neurol Neurosurg Psychiatry 80:1218–23 . doi: 10.1136/jnnp.2008.154732

154. Moccia M, Quarantelli M, Lanzillo R, et al (2017) Grey:white matter ratio at diagnosis and the risk of 10-year multiple sclerosis progression. Eur J Neurol 24: . doi: 10.1111/ene.13183

155. Rudick R, Fisher E, Lee J, et al (1999) Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing-remitting MS. Multiple Sclerosis Collaborative Research Group. Neurology 53:1698–1704

156. Durand-Dubief F, Belaroussi B, Armspach JP, et al (2012) Reliability of longitudinal brain volume loss measurements between 2 sites in patients with multiple sclerosis: Comparison of 7 quantification techniques. AJNR Am J Neuroradiol 33:1918–1924 . doi:

10.3174/ajnr.A3107

157. Sormani M, Rovaris M, Valsasina P, et al (2004) Measurement error of two different techniques for brain atrophy assessment in multiple sclerosis. Neurology 62:1432–1434

158. Anderson VM, Fernando KTM, Davies GR, et al (2007) Cerebral atrophy measurement in clinically isolated syndromes and relapsing remitting multiple sclerosis: A comparison of registration-based methods. J Neuroimaging 17:61–68 . doi: 10.1111/j.1552-6569.2006.00081.x

159. Chataway J, Schuerer N, Alsanousi A, et al (2014) Effect of high-dose simvastatin on brain atrophy and disability in secondary progressive multiple sclerosis (MS-STAT): A randomised, placebo-controlled, phase 2 trial. Lancet 383:2213–2221 . doi: 10.1016/S0140-6736(13)62242-4

160. Fox RRJ, Coffey CS, Conwit R, et al (2018) Phase 2 Trial of Ibudilast in Progressive Multiple Sclerosis. N Engl J Med 379:846–855 . doi: 10.1056/NEJMoa1803583

161. ClinicalTrials.gov (2017) Study of Tcelna (Imilecleucel-T) in Secondary Progressive Multiple Sclerosis (Abili-T)

162. ClinicalTrials.gov (2016) Lipoic Acid for Secondary Progressive Multiple Sclerosis (MS)

163. Connick P, De Angelis F, Parker RA, et al (2018) Multiple Sclerosis-Secondary Progressive Multi-Arm Randomisation Trial (MS-SMART): A multiarm phase IIb randomised, double-blind, placebo-controlled clinical trial comparing the efficacy of three neuroprotective drugs in secondary progressive multiple scl. BMJ 8:1–16 . doi: 10.1136/bmjopen-2018-021944

164. Montalban X, Hauser SL, Kappos L, et al (2017) Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. N Engl J Med 376:209–220 . doi: 10.1056/NEJMoa1606468

165. Montalban X, Sastre-Garriga J, Tintoré M, et al (2009) A single-center, randomized, double-blind, placebo-controlled study of interferon beta-1b on primary progressive and transitional

multiple sclerosis. Mult Scler 15:1195–1205 . doi: 10.1177/1352458509106937

166. Leary S, Miller D, Stevenson V, et al (2003) Interferon beta-1a in primary progressive MS: an exploratory, randomized, controlled trial. Neurology 60:44–51

167. Kalkers N, Barkhof F, Bergers E, et al (2002) The effect of the neuroprotective agent riluzole on MRI parameters in primary progressive multiple sclerosis: a pilot study. Mult Scler 8:532–533

168. ClinicalTrials.gov (2016) Idebenone for Primary Progressive Multiple Sclerosis. https://clinicaltrials.gov/ct2/show/NCT01854359

169. O'Connor PW, Li D, Freedman MS, Bar-Or A RG, Confavreux C, Paty DW, Stewart JA SR (2006) A phase II study of the safety and efficacy of teriflunomide in multiple sclerosis with relapses. Neurology 66:894–900 . doi: 10.1212/01.wnl.0000203121.04509.31

170. Jasperse B, Vrenken H, Sanz-Arigita E, et al (2007) Regional brain atrophy development is related to specific aspects of clinical dysfunction in multiple sclerosis. Neuroimage 38:529–537 . doi: 10.1016/j.neuroimage.2007.07.056

171. Muhlert N, Sethi V, Cipolotti L, et al (2015) The grey matter correlates of impaired decision-making in multiple sclerosis. J Neurol Neurosurg Psychiatry 86:530–536 . doi: 10.1136/jnnp-2014-308169

172. Popescu V, Schoonheim MM, Versteeg A, et al (2016) Grey matter atrophy in multiple sclerosis: Clinical interpretation depends on choice of analysis method. PLoS One 11:e0143942 . doi: 10.1371/journal.pone.0143942

173. Barkhof F, Calabresi PA, Miller DH, Reingold SC (2009) Imaging outcomes for neuroprotection and repair in multiple sclerosis trials. Nat Rev Neurol 5:256–266 . doi: 10.1038/nrneurol.2009.41

174. Wattjes M, Rovira A, Miller D, et al (2015) Evidence-based guidelines: MAGNIMS consensus

guidelines on the use of MRI in multiple sclerosis—establishing disease prognosis and monitoring patients. Nat Rev Neurol 11:597–606 . doi: 10.1038/nrneurol.2015.157

175. De Stefano N, Giorgio A, Battaglini M, et al (2017) Reduced brain atrophy rates are associated with lower risk of disability progression in patients with relapsing multiple sclerosis treated with cladribine tablets. Mult Scler. doi: 10.1177/1352458517690269.

176. Sormani MP, Arnold DL, De Stefano N (2014) Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. Ann Neurol 75:43–49 . doi: 10.1002/ana.24018

177. Lee H, Narayanan S, Brown RA, et al (2017) Brain atrophy after bone marrow transplantation for treatment of multiple sclerosis. Mult Scler 23:420–431 . doi: 10.1177/1352458516650992

178. Hauser SL, Bar-Or A, Comi G, et al (2017) Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis. N Engl J Med 376:221–234 . doi: 10.1056/NEJMoa1601277

179. Nakamura K, Brown RA, Araujo D, et al (2014) Correlation between brain volume change and T2 relaxation time induced by dehydration and rehydration: Implications for monitoring atrophy in clinical studies. NeuroImage Clin 6:166–170 . doi: 10.1016/j.nicl.2014.08.014

180. Cincotta MC, Engelhard MM, Stankey M, Goldman MD (2016) Fatigue and fluid hydration status in multiple sclerosis: A hypothesis. Mult Scler 22:1438–1443 . doi: 10.1177/1352458516663854

181. Friese MA, Schattling B, Fugger L (2014) Mechanisms of neurodegeneration and axonal dysfunction in multiple sclerosis. Nat Rev Neurol 10:225–38 . doi: 10.1038/nrneurol.2014.37

182. Bodini B, Louapre C, Stankoff B (2015) Advanced imaging tools to investigate multiple sclerosis pathology. Press Medicale 44:e159–e167 . doi: 10.1016/j.lpm.2015.02.011

183. Popescu V, Agosta F, Hulst HE, et al (2013) Brain atrophy and lesion load predict long-term disability in multiple sclerosis. J Neurol Neurosurg Psychiatry 84:1082–1091

184. Enzinger C, Barkhof F, Ciccarelli O, et al (2015) Nonconventional MRI and microstructural cerebral changes in multiple sclerosis. Nat Rev Neurol 11:676–686 . doi: 10.1038/nrneurol.2015.194

185. Khaleeli Z, Sastre-Garriga J, Ciccarelli O, et al (2007) Magnetisation transfer ratio in the normal appearing white matter predicts progression of disability over 1 year in early primary progressive multiple sclerosis. J Neurol Neurosurg Psychiatry 78:1076–1082

186. Rovaris M, Gass A, Bammer R, et al (2005) Diffusion MRI in multiple sclerosis. Neurology 65:1526–153

187. Fox R, McColl R, Lee J, et al (2008) A preliminary validation study of diffusion tensor imaging as a measure of functional brain injury. Arch Neurol 65:1179–1184 . doi: 10.1001/archneur.65.9.1179

188. Grech-Sollars M, Hales P, Miyazaki K, et al (2015) Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. NMR Biomed 28:468–485 . doi: 10.1002/nbm.3269

189. McHugh D, Hubbard Cristinacce P, Naish J, Parker G (2019) Towards a “resolution limit” for DW-MRI tumor microstructural models: A simulation study investigating the feasibility of distinguishing between microstructural changes. Magn Reson Med 81:2288–2301 . doi: 10.1002/mrm.27551

190. Gülin Ö (2016) Magnetic Resonance Spectroscopy of Degenerative Brain Diseases, Contempora. Springer

191. Moccia M, Ciccarelli O (2017) Molecular and Metabolic Imaging in Multiple Sclerosis. Neuroimaging Clin N Am 27:343–356 . doi: 10.1016/j.nic.2016.12.005

192. De Stefano N, Filippi M, Miller D, et al (2007) Guidelines for using proton MR spectroscopy in multicenter clinical MS studies. Neurology 69:1942–52 . doi:

10.1212/01.wnl.0000291557.62706.d3

193. Filippi M, Rocca MA, Pagani E, et al (2014) Placebo-controlled trial of oral laquinimod in multiple sclerosis: MRI evidence of an effect on brain tissue damage. J Neurol Neurosurg Psychiatry 85:852–859 . doi: 10.1136/jnnp-2013-306132

194. Sajja B, Narayana P, Wolinsky J, Ahn C (2008) Longitudinal magnetic resonance spectroscopic imaging of primary progressive multiple sclerosis patients treated with glatiramer acetate: multicenter study. Mult Scler 14:73–80 . doi: 10.1177/1352458507079907

195. De Paula Faria D, Copray S, Buchpiguel C, et al (2014) PET imaging in multiple sclerosis. J Neuroimmune Pharmacol 9:468–482 . doi: 10.1007/s11481-014-9544-2

196. Matthews PM, Datta G (2015) Positron-emission tomography molecular imaging of glia and myelin in drug discovery for multiple sclerosis. Expert Opin Drug Discov 10:557–570 . doi: 10.1517/17460441.2015.1032240

197. Matthews PM, Comley R (2009) Advances in the molecular imaging of multiple sclerosis. Expert Rev Clin Immunol 5:765–77 . doi: 10.1586/eci.09.66

198. Freeman L, Garcia-Lorenzo D, Bottin L, et al (2015) The neuronal component of gray matter damage in multiple sclerosis: A [(11) C]flumazenil positron emission tomography study. Ann Neurol 78:554–67 . doi: 10.1002/ana.24468

199. Mallik S, Samson RS, Wheeler-Kingshott CAM, Miller DH (2014) Imaging outcomes for trials of remyelination in multiple sclerosis. J Neurol Neurosurg Psychiatry 85:1396–1404 . doi: 10.1136/jnnp-2014-307650

200. Costello F, Hodge W, Pan Y, et al (2008) Tracking retinal nerve fiber layer loss after optic neuritis: a prospective study using optical coherence tomography. Mult Scler 14:893–905 . doi: 10.1177/1352458508091367

201. Herrero R, Garcia-Martin E, Almarcegui C, et al (2012) Progressive degeneration of the retinal

nerve fiber layer in patients with multiple sclerosis. Invest Ophtalmol Vis Sci 53:8344–8349

202. Wolinsky JS, Narayana PA, Nelson F, et al (2013) Magnetic resonance imaging outcomes from a phase III trial of teriflunomide. Mult Scler 19:1310–9 . doi: 10.1177/1352458513475723

203. Cruz-Herranz A, Balk L, Oberwahrenbrock T, et al (2016) The APOSTEL recommendations for reporting quantitative optical coherence tomography studies. Neurology 86:2303–2309 . doi: 10.1212/WNL.0000000000002774

204. Gilmore CP, Bö L, Owens T, et al (2006) Spinal Cord Gray Matter Demyelination in Multiple Sclerosis — A Novel Patternv of Residual Plaque Morphology. Brain Pathol 16:202–208

205. Eden D, Gros C, Badji A, et al (2018) Spatial Distribution of Multiple Sclerosis lesions in the Cervical Cord. ISMRM

206. Valsasina P, Aboulwafa M, Preziosa P, et al (2018) Cervical Cord T1-weighted Hypointense Lesions at MR Imaging in Multiple Sclerosis: Relationship to Cord Atrophy and Disability. Radiology 288:234–244 . doi: 10.1148/radiol.2018172311

207. Trop I, Bourgouin PM, Lapierre Y, et al (1998) Multiple Sclerosis of the Spinal Cord : Diagnosis and Follow-up with Contrast-Enhanced MR and Correlation with Clinical Activity. AJNR Am J Neuroradiol 19:1025–1033

208. Klawiter E, Benzinger T, Roy A, et al (2010) Spinal Cord Ring Enhancement in Multiple Sclerosis. Arch Neurol 67:1395–1398

209. Zalewski NL, Morris PP, Weinshenker BG, et al (2017) Ring-enhancing spinal cord lesions in neuromyelitis optica spectrum disorders. J Neurol Neurosurg Psychiatry 88:218–225 . doi: 10.1136/jnnp-2016-314738

210. Weier K, Mazraeh J, Naegelin Y, et al (2012) Biplanar MRI for the assessment of the spinal cord in multiple sclerosis. Mult Scler 18:1560–1569 . doi: 10.1177/1352458512442754

211. Lycklama à Nijeholt GJ, Barkhof F, Scheltens P, et al (1997) MR of the Spinal Cord in Multiple

Sclerosis : Relation to Clinical Subtype and Disability. AJNR Am J Neuroradiol 18:1041–1048

212. Kearney H, Miszkiel KA, Yiannakas MC, et al (2016) Grey matter involvement by focal cervical spinal cord lesions is associated with progressive multiple sclerosis. Mult Scler 22:910–920 . doi: 10.1177/1352458515604905

213. Ciccarelli O, Cohen J, Reingold S, et al (2019) Spinal Cord Involvement in Multiple Sclerosis and Neuromyelitis Optica Spectrum Disorders. Lancet Neurol 18:185–197 . doi: 10.1016/S1474-4422(18)30460-5

214. Chong AL, Chandra R V., Chuah KC, et al (2016) Proton density MRI increases detection of cervical spinal cord multiple sclerosis lesions compared with T2-weighted fast spin-echo. AJNR Am J Neuroradiol 37:180–184 . doi: 10.3174/ajnr.A4476

215. Philpott C, Brotchie P (2011) Comparison of MRI sequences for evaluation of multiple sclerosis of the cervical spinal cord at 3 T. Eur J Radiol 80:780–785 . doi: 10.1016/j.ejrad.2010.09.031

216. Bot JCJ, Barkhof F, Lycklama à Nijeholt GJ, et al (2000) Comparison of a conventional cardiac-triggered dual spin-echo and a fast STIR sequence in detection of spinal cord lesions in multiple sclerosis. Eur Radiol 10:753–758

217. Rovira A, De Stefano N (2016) MRI monitoring of spinal cord changes in patients with multiple sclerosis. Curr Opin Neurol 29:445–452 . doi: 10.1097/WCO.0000000000000343

218. Alcaide-Leon P, Pauranik A, Alshafai L, et al (2016) Comparison of Sagittal FSE T2, STIR, and T1-weighted phase-sensitive inversion recovery in the detection of spinal cord lesions in MS at 3T. AJNR Am J Neuroradiol 37:970–975 . doi: 10.3174/ajnr.A4656

219. Nair G, Absinta M, Reich DS (2013) Optimized T1-MPRAGE Sequence for Better Visualization of Spinal Cord Multiple Sclerosis Lesions at 3T. AJNR Am J Neuroradiol 34:2215–2222

220. Riederer I, Karampinos DC, Settles M, et al (2015) Double inversion recovery sequence of the

cervical spinal cord in multiple sclerosis and related inflammatory diseases. AJNR Am J Neuroradiol 36:219–225 . doi: 10.3174/ajnr.A4093

221. Galler S, Stellmann JP, Young KL, et al (2016) Improved lesion detection by using axial T2-weighted MRI with full spinal cord coverage in multiple sclerosis. AJNR Am J Neuroradiol 37:963–969 . doi: 10.3174/ajnr.A4638

222. Breckwoldt MO, Gradl J, Hähnel S, et al (2017) Increasing the sensitivity of MRI for the detection of multiple sclerosis lesions by long axial coverage of the spinal cord: a prospective study in 119 patients. J Neurol 264:341–349 . doi: 10.1007/s00415-016-8353-3

223. Hagens MHJ, Burggraaff J, Kilsdonk ID, et al (2018) Three-Tesla MRI does not improve the diagnosis of multiple sclerosis. Neurology 91:e249–e257 . doi: 10.1212/WNL.0000000000005825

224. Dula AN, Pawate S, Dortch RD, et al (2017) Magnetic Resonance Imaging of the Cervical Spinal Cord in Multiple Sclerosis at 7T. Mult Scler 22:320–328 . doi: 10.1177/1352458515591070

225. Kearney H, Miller DH, Ciccarelli O (2015) Spinal cord MRI in multiple sclerosis - diagnostic, prognostic and clinical value. Nat Rev Neurol 11:327–338 . doi: 10.1038/nrneurol.2015.80

226. Gass A, Rocca MA, Agosta F, et al (2015) MRI monitoring of pathological changes in the spinal cord in patients with multiple sclerosis. Lancet Neurol 14:443–454 . doi: 10.1016/S1474-4422(14)70294-7

227. Kearney H, Schneider T, Yiannakas MC, et al (2015) Spinal cord grey matter abnormalities are associated with secondary progression and Physical disability in multiple sclerosis. J Neurol Neurosurg Psychiatry 86:608–614 . doi: 10.1136/jnnp-2014-308241

228. Kearney H, Miszkiel KA, Yiannakas MC, et al (2013) A pilot MRI study of white and grey matter involvement by multiple sclerosis spinal cord lesions. Mult Scler Relat Disord 2:103–108 . doi: 10.1016/j.msard.2012.09.005

229. Filippi M, Rocca MA, Ciccarelli O, et al (2016) MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. Lancet Neurol 15:292–303 . doi: 10.1016/S1474-4422(15)00393-2

230. Tintore M, Otero-Romero S, Rio J, et al (2016) Contribution of the symptomatic lesion in establishing MS diagnosis and prognosis. Neurology 87:1368–1374 . doi: 10.1212/WNL.0000000000003144

231. Bot JCJ, Barkhof F (2009) Spinal-Cord MRI in Multiple Sclerosis: Conventional and Nonconventional MR Techniques. Neuroimaging Clin N Am 19:81–99 . doi: 10.1016/j.nic.2008.09.005

232. Barkhof F (2014) Spinal cord MRI should always be performed in clinically isolated syndrome patients: Yes. Mult Scler 20:1688–1689 . doi: 10.1177/1352458514546518

233. Keegan BM, Kaufmann TJ, Weinshenker BG, et al (2016) Progressive solitary sclerosis: Gradual motor impairment from a singleCNSdemyelinating lesion. Neurology 87:1713–1719 . doi: 10.1212/WNL.0000000000003235

234. Schee JP, Viswanathan S (2018) Pure spinal multiple sclerosis: A possible novel entity within the multiple sclerosis disease spectrum. Mult Scler 1352458518775912 . doi: 10.1177/1352458518775912

235. Zalewski NL, Flanagan EP, Keegan BM (2018) Evaluation of idiopathic transverse myelitis revealing specific myelopathy diagnoses. Neurology 90:e96–e102 . doi: 10.1212/WNL.0000000000004796

236. Pekcevik Y, Mitchell CH, Mealy MA, et al (2016) Differentiating neuromyelitis optica from other causes of longitudinally extensive transverse myelitis on spinal magnetic resonance imaging. Mult Scler 22:302–311 . doi: 10.1177/1352458515591069

237. Asgari N, Pernille H, Skejoe B, Lennon VA (2013) Evolution of longitudinally extensive

transverse myelitis in an aquaporin-4 IgG-positive patient. Neurology 81:95–96

238. Flanagan EP, Weinshenker BG, Krecke KN, et al (2015) Short myelitis lesions in aquaporin-4-IgG-positive neuromyelitis optica spectrum disorders. JAMA Neurol 72:81–87 . doi: 10.1001/jamaneurol.2014.2137

239. Yonezu T, Ito S, Mori M, et al (2014) “Bright spotty lesions” on spinal magnetic resonance imaging differentiate neuromyelitis optica from multiple sclerosis. Mult Scler 20:331–337 . doi: 10.1177/1352458513495581

240. Hyun J-W, Kim S-H, Jeong I, et al (2015) Bright spotty lesions on the spinal cord : an additional MRI indicator of neuromyelitis optica spectrum disorder ? J Neurol Neurosurg Psychiatry 86:1280–1282

241. Sato DK, Callegaro D, Lana-Peixoto M, et al (2014) Distinction between MOG antibody-positive and AQP4 antibody-positive NMO spectrum disorders. Neurology 82:474–481

242. Hamid SHM, Whittam D, Mutch K, et al (2017) What proportion of AQP4 - IgG - negative NMO spectrum disorder patients are MOG - IgG positive? A cross sectional study of 132 patients. J Neurol 264:2088–2094 . doi: 10.1007/s00415-017-8596-7

243. Flanagan EP, Kaufmann TJ, Krecke KN, et al (2016) Discriminating long myelitis of neuromyelitis optica from sarcoidosis. Ann Neurol 79:437–447 . doi: 10.1002/ana.24582

244. Jolliffe EA, Keegan BM, Flanagan EP (2018) Trident sign trumps Aquaporin-4-IgG ELISA in diagnostic value in a case of longitudinally extensive transverse myelitis. Mult Scler Relat Disord 23:7–8 . doi: 10.1016/j.msard.2018.04.012

245. Zalewski NL, Krecke KN, Weinshenker BG, et al (2016) Central canal enhancement and the trident sign in spinal cord sarcoidosis. Neurology 87:743–4 . doi: 10.1212/WNL.0000000000002992

246. Zalewski N, Rabinstein A, Brinjikji W, et al (2018) Unique Gadolinium Enhancement Pattern

in Spinal Dural Arteriovenous Fistulas. JAMA Neurol 75:1542–1545 . doi: 10.1001/jamaneurol.2018.2605

247. Kantarci OH, Lebrun C, Siva A, et al (2016) Primary Progressive Multiple Sclerosis Evolving from Radiologically Isolated Syndrome. Ann Neurol 79:288–294 . doi: 10.1002/ana.24564
248. Arrambide G, Rovira A, Sastre-Garriga J, et al (2018) Spinal cord lesions: A modest contributor to diagnosis in clinically isolated syndromes but a relevant prognostic factor. Mult Scler 24:301–312 . doi: 10.1177/1352458517697830
249. Sombekke MH, Wattjes MP, Balk LJ, et al (2013) Spinal cord lesions in patients with clinically isolated syndrome: A powerful tool in diagnosis and prognosis. Neurology 80:69–75 . doi: 10.1212/WNL.0b013e31827b1a67
250. Brownlee WJ, Altmann DR, Da Mota PA, et al (2017) Association of asymptomatic spinal cord lesions and atrophy with disability 5 years after a clinically isolated syndrome. Mult Scler 23:665–674 . doi: 10.1177/1352458516663034
251. Zecca C, Disanto G, Sormani MP, et al (2016) Relevance of asymptomatic spinal MRI lesions in patients with multiple sclerosis. Mult Scler 22:782–791 . doi: 10.1177/1352458515599246
252. Kearney H, Altmann DR, Samson RS, et al (2015) Cervical cord lesion load is associated with disability independently from atrophy in MS. Neurology 84:367–373 . doi: 10.1212/WNL.0000000000001186
253. D'Amico E, Patti F, Leone C, et al (2016) Negative prognostic impact of MRI spinal lesions in the early stages of relapsing–remitting multiple sclerosis. Mult Scler – Exp Transl Clin 2:1–7 . doi: 10.1177/2055217316631565
254. Saccà F, Lanzillo R, Signori A, et al (2019) Determinants of therapy switch in multiple sclerosis treatment-naïve patients: A real-life study. Mult Scler 25:1263–1272 . doi: 10.1177/1352458518790390

255. Ruggieri S, Logoteta A, Tinelli E, et al (2018) Measuring disease activity in multiple sclerosis: the essential role of spinal cord MRI monitoring. Mult Scler 24:308

256. Flanagan E, Weinshenker M, Krecke K, Pittock S (2015) Asymptomatic myelitis in neuromyelitis optica and autoimmune aquaporin-4 channelopathy. Neurol Clin Pract 5:175–177

257. Rocca MA, Valsasina P, Damjanovic D, et al (2013) Voxel-wise mapping of cervical cord damage in multiple sclerosis patients with different clinical phenotypes. J Neurol Neurosurg Psychiatry 84:35–41 . doi: 10.1136/jnnp-2012-303821

258. Eden D, Gros C, Badji A, et al (2019) Spatial distribution of multiple sclerosis lesions in the cervical spinal cord. Brain 1–14 . doi: 10.1093/brain/awy352

259. Trapp B, Peterson J, Ransohoff R, et al (1998) Axonal Transection in the Lesions of Multiple Sclerosis. N Engl J Med 338:278–285 . doi: 10.1056/NEJM199801293380502

260. Brex P, Leary S, O'Riordan J, et al (2001) Measurement of spinal cord area in clinically isolated syndromes suggestive of multiple sclerosis. J Neurol Neurosurg Psychiatry 70:544–547 . doi: 10.1136/jnnp.70.4.544

261. Gilmore CP, Deluca GC, Bö L, et al (2009) Spinal cord neuronal pathology in multiple sclerosis. Brain Pathol 19:642–649 . doi: 10.1111/j.1750-3639.2008.00228.x

262. DeLuca G, Ebers G, Esiri M (2004) Axonal loss in multiple sclerosis: a pathological survey of the corticospinal and sensory tracts. Brain 127:1009–1018

263. Agosta F, Filippi M (2007) MRI of spinal cord in multiple sclerosis. J Neuroimaging 17:46S-49S . doi: 10.1111/j.1552-6569.2007.00137.x

264. Evangelou N, DeLuca G, Owens T, Esiri M (2005) Pathological study of spinal cord atrophy in multiple sclerosis suggests limited role of local lesions. Brain 128:29–34

265. Gilmore C, DeLuca G, Bö L, et al (2005) Spinal cord atrophy in multiple sclerosis caused by

white matter volume loss. Arch Neurol 62:1859–1862

266. Chard D, Brex P, Ciccarelli O, et al (2003) The longitudinal relation between brain lesion load and atrophy in multiple sclerosis: a 14 year follow up study. J Neurol Neurosurg Psychiatry 74:1551–1554

267. Ruggieri S, Petracca M, Miller A, et al (2015) Association of deep gray matter damage with cortical and spinal cord degeneration in primary progressive multiple sclerosis. JAMA Neurol 72:1466–1474 . doi: 10.1001/jamaneurol.2015.1897

268. Casserly C, Seyman EE, Alcaide-Leon P, et al (2018) Spinal Cord Atrophy in Multiple Sclerosis: A Systematic Review and Meta-Analysis. J Neuroimaging 28:556–586 . doi: 10.1111/jon.12553

269. Healy BC, Arora A, Hayden DL, et al (2012) Approaches to normalization of spinal cord volume: Application to multiple sclerosis. J Neuroimaging 22:e12-19 . doi: 10.1111/j.1552-6569.2011.00629.x

270. Oh J, Seigo M, Saidha S, et al (2014) Spinal Cord Normalization in Multiple Sclerosis. J Neuroimaging 24:577–584 . doi: 10.1111/jon.12097

271. Schlaeger R, Papinutto ND, Zhu AH, et al (2015) The association between thoracic spinal cord gray matter atrophy and disability in Multiple Sclerosis. JAMA Neurol 72:897–904 . doi: 10.1001/jamaneurol.2015.0993.The

272. Hua LH, Donlon SL, Sobhanian MJ, et al (2015) Thoracic spinal cord lesions are influenced by the degree of cervical spine involvement in multiple sclerosis. Spinal Cord 53:520–525 . doi: 10.1038/sc.2014.238

273. Rocca M, Valsasina P, Meani A, et al (2018) Cranio-caudal patterns of cervical cord atrophy progression in MS according to disease phenotype and clinical worsening: a multicenter study. Mult Scler 24:102–103

274. Horsfield MA, Sala S, Neema M, et al (2010) Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: Application in multiple sclerosis. Neuroimage 50:446–455 . doi: 10.1016/j.neuroimage.2009.12.121

275. Daams M, Weiler F, Steenwijk MMD, et al (2014) Mean upper cervical cord area (MUCCA) measurement in long-standing multiple sclerosis: relation to brain findings and clinical disability. Mult Scler 20:1860–1865 . doi: 10.1177/1352458514533399

276. De Leener B, Taso M, Adad JC, Callot V (2016) Segmentation of the human spinal cord. MAGMA 29:125–153 . doi: 10.1007/s10334-015-0507-2

277. Xinapse (2018) Cord Finder - Introduction. http://www.xinapse.com/Manual/. Accessed 28 Dec 2018

278. Kearney H, Yiannakas M, Abdel-Aziz K, et al (2014) Improved MRI quantification of spinal cord atrophy in multiple sclerosis. J Magn Reson Imaging 39:617–623

279. Hickman S, Coulon O, Parker G, et al (2003) Application of a B-spline active surface technique to the measurement of cervical cord volume in multiple sclerosis from three-dimensional MR images. J Magn Reson Imaging 18:368–371

280. Coulon O, Hickman S, Parker G, et al (2002) Quantification of spinal cord atrophy from magnetic resonance images via a B-spline active surface model. Magn Reson Med 47:1176–1185

281. Lukas C, Hahn HK, Bellenberg B, et al (2004) Sensitivity and reproducibility of a new fast 3D segmentation technique for clinical MR-based brain volumetry in multiple sclerosis. Neuroradiology 46:906–915 . doi: 10.1007/s00234-004-1282-3

282. Liu XY, Lukas XC, Steenwijk XMD, et al (2016) Multicenter Validation of Mean Upper Cervical Cord Area Measurements from Head 3D T1-Weighted MR Imaging in Patients with Multiple Sclerosis. AJNR Am J Neuroradiol 37:749–754

283. De Leener B, Lévy S, Dupont SM, et al (2017) SCT: Spinal Cord Toolbox, an open-source software for processing spinal cord MRI data. Neuroimage 145:24–43 . doi: 10.1016/j.neuroimage.2016.10.009

284. Yiannakas MC, Mustafa AM, De Leener B, et al (2015) Fully automated segmentation of the cervical cord from T1-weighted MRI using PropSeg: Application to multiple sclerosis. NeuroImage Clin 10:71–77 . doi: 10.1016/j.nicl.2015.11.001

285. McCoy DB, Dupont SM, Gros C, et al (2019) Convolutional neural network–based automated segmentation of the spinal cord and contusion injury: Deep learning biomarker correlates of motor impairment in acute spinal cord injury. AJNR Am J Neuroradiol 40:737–744 . doi: 10.3174/ajnr.A6020

286. Gros C, De Leener B, Badji A, et al (2019) Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. Neuroimage 184:901–915 . doi: 10.1016/j.neuroimage.2018.09.081

287. Lukas C, Prados F, Valsasina P, et al (2018) Quantification of spinal cord atrophy in MS: which software, which vertebral level, spinal cord or brain MRI? A multi-centric, longitudinal comparison of three different volumetric approaches. Mult Scler 24:88–89

288. Weeda MM, Middelkoop SM, Steenwijk MD, et al (2019) Validation of mean upper cervical cord area (MUCCA) measurement techniques in multiple sclerosis (MS): High reproducibility and robustness to lesions, but large software and scanner effects. NeuroImage Clin 24:101962 . doi: 10.1016/j.nicl.2019.101962

289. Altmann DR, Jasperse B, Barkhof F, et al (2009) Sample sizes for brain atrophy outcomes in trials for secondary progressive multiple sclerosis. Neurology 72:595–601 . doi: 10.1212/01.wnl.0000335765.55346.fc

290. Liu Z, Yaldizli Ö, Pardini M, et al (2015) Cervical cord area measurement using volumetric brain

magnetic resonance imaging in multiple sclerosis. Mult Scler Relat Disord 4:52–57 . doi: 10.1016/j.msard.2014.11.004

291. Freund P, Dalton C, Wheeler-Kingshott C, et al (2010) Method for simultaneous voxel-based morphometry of the brain and cervical spinal cord area measurements using 3D-MDEFT. J Magn Reson Imaging 32:1242–1247 . doi: 10.1002/jmri.22340

292. Prados F, Barkhof F (2018) Spinal cord atrophy rates. Ready for prime time in multiple sclerosis clinical trials? Neurology 91:157–158 . doi: 10.1212/WNL.0000000000005873

293. Biberacher V, Boucard CC, Schmidt P, et al (2015) Atrophy and structural variability of the upper cervical cord in early multiple sclerosis. Mult Scler 21:875–884 . doi: 10.1177/1352458514546514

294. Rocca MA, Horsfield MA, Sala S, et al (2011) A multicenter assessment of cervical cord atrophy among MS clinical phenotypes. Neurology 76:2096–2102 . doi: 10.1212/WNL.0b013e31821f46b8

295. Hagström IT, Schneider R, Bellenberg B, et al (2017) Relevance of early cervical cord volume loss in the disease evolution of clinically isolated syndrome and early multiple sclerosis: a 2-year follow-up study. J Neurol 264:1402–1412 . doi: 10.1007/s00415-017-8537-5

296. Tsagkas C, Magon S, Gaetano L, et al (2018) Spinal cord volume loss. A marker of disease progression in multiple sclerosis. Neurology 91:e349–e358 . doi: 10.1212/WNL.0000000000005853

297. Furby J, Hayton T, Altmann D, et al (2010) A longitudinal study of MRI-detected atrophy in secondary progressive multiple sclerosis. J Neurol 257:1508–1516

298. Rashid W, Davies G, Chard D, et al (2006) Upper cervical cord area in early relapsing-remitting multiple sclerosis: cross-sectional study of factors influencing cord size. J Magn Reson Imaging 23:473–476

299. Rashid W, Davies G, Chard D, et al (2006) Increasing cord atrophy in early relapsing-remitting multiple sclerosis: a 3 year study. J Neurol Neurosurg Psychiatry 77:51–55

300. Cawley N, Tur C, Prados F, et al (2018) Spinal cord atrophy as a primary outcome measure in phase II trials of progressive multiple sclerosis. Mult Scler 24:932–941 . doi: 10.1177/1352458517709954

301. Schlaeger R, Papinutto N, Panara V, et al (2014) Spinal cord gray matter atrophy correlates with multiple sclerosis disability. Ann Neurol 76:568–580 . doi: 10.1002/ana.24241

302. Chien C, Scheel M, Schmitz-Hübsch T, et al (2018) Spinal cord lesions and atrophy in NMOSD with AQP4-IgG and MOG-IgG associated autoimmunity. Mult Scler 1352458518815596

303. Liu Y, Wang J, Daams M, et al (2015) Differential patterns of spinal cord and brain atrophy in NMO and MS. Neurology 84:1465–1472 . doi: 10.1212/WNL.0000000000001441

304. Schneider R, Bellenberg B, Kleiter I, et al (2017) Cervical cord and ventricle affection in neuromyelitis optica. Acta Neurol Scand 135:324–331 . doi: 10.1111/ane.12601

305. Kearney H, Rocca M, Valsasina P, et al (2014) Magnetic resonance imaging correlates of physical disability in relapse onset multiple sclerosis of long disease duration. Mult Scler 20:72–80 . doi: 10.1177/1352458513492245

306. Cohen AB, Neema M, Arora A, et al (2012) The relationships among MRI-defined spinal cord involvment, brain involvment and disability in multiple sclerosis. J Neuroimaging 22:122–128 . doi: 10.1111/j.1552-6569.2011.00589.x.The

307. Aymerich FX, Auger C, Alonso J, et al (2018) Cervical Cord Atrophy and Long-Term Disease Progression in Patients with Primary-Progressive Multiple Sclerosis. AJNR Am J Neuroradiol 39:399–404 . doi: 10.3174/ajnr.A5495

308. Lukas C, Knol DDL, Sombekke MMH, et al (2015) Cervical spinal cord volume loss is related to clinical disability progression in multiple sclerosis. J Neurol Neurosurg Psychiatry 86:410–418 .

doi: 10.1136/jnnp-2014-308021

309. Brownlee WJ, Altmann DR, Alves Da Mota P, et al (2017) Association of asymptomatic spinal cord lesions and atrophy with disability 5 years after a clinically isolated syndrome. Mult Scler 23:665–674 . doi: 10.1177/1352458516663034

310. Lukas C, Sombekke M, Bellenberg B, et al (2013) Relevance of spinal cord abnormalities to clinical disability in multiple sclerosis: MR imaging findings in a large cohort of patients. Radiology 269:542–552

311. Tsagkas C, Magon S, Gaetano L, et al (2019) Preferential spinal cord volume loss in primary progressive multiple sclerosis. Mult Scler 25:947–957 . doi: 10.1177/1352458518775006

312. Bonati U, Fisniku LK, Altmann DR, et al (2011) Cervical cord and brain grey matter atrophy independently associate with long-term MS disability. J Neurol Neurosurg Psychiatry 82:471–472 . doi: 10.1136/jnnp.2010.205021

313. Singhal T, Tauhid S, Hurwitz S, et al (2017) The Effect of Glatiramer Acetate on Spinal Cord Volume in Relapsing-Remitting Multiple Sclerosis. J Neuroimaging 27:33–36 . doi: 10.1111/jon.12378

314. Dupuy SL, Khalid F, Healy BC, et al (2016) The effect of intramuscular interferon beta-1a on spinal cord volume in relapsing-remitting multiple sclerosis. BMC Med Imaging 16:56 . doi: 10.1186/s12880-016-0158-4

315. Papinutto N, Bakshi R, Bischof A, et al (2018) Gradient nonlinearity effects on upper cervical spinal cord area measurement from 3D T1-weighted brain MRI acquisitions. Magn Reson Med 79:1595–1601 . doi: 10.1002/mrm.26776

316. Valsasina P, Rocca MA, Horsfield MA, et al (2015) A longitudinal MRI study of cervical cord atrophy in multiple sclerosis. J Neurol 262:1622–1628 . doi: 10.1007/s00415-015-7754-z

317. Alcaide-Leon P, Cybulsky K, Sankar S, et al (2018) Quantitative spinal cord MRI in radiologically

isolated syndrome. Neurol Neuroimmunol NeuroInflammation 5:1–9 . doi: 10.1212/NXI.0000000000000436

318. Toosy AT, Kou N, Altmann D, et al (2014) Voxel-based cervical spinal cord mapping of diffusion abnormalities in MS-related myelitis. Neurology 83:1321–1325 . doi: 10.1212/WNL.0000000000000857

319. Martin AR, Aleksanderek I, Cohen-Adad J, et al (2016) Translating state-of-the-art spinal cord MRI techniques to clinical use: A systematic review of clinical studies utilizing DTI, MT, MWF, MRS, and fMRI. NeuroImage Clin 10:192–238 . doi: 10.1016/j.nicl.2015.11.019

320. Cohen Y, Anaby D, Morozov D (2017) Diffusion MRI of the spinal cord : from structural studies to pathology. NMR Biomed 30:e3592 . doi: 10.1002/nbm.3592

321. Stroman P, Wheeler-Kingshott C, Bacon M, et al (2014) The current state-of-the-art of spinal cord imaging: methods. Neuroimage 84:1070–1081 . doi: 10.1016/j.neuroimage.2013.04.124

322. Wheeler-Kingshott C, Hickman S, Parker G, et al (2002) Investigating cervical spinal cord structure using axial diffusion tensor imaging. Neuroimage 16:93–102

323. Klawiter E, Schmidt R, Trinkaus K, et al (2011) Radial diffusivity predicts demyelination in ex vivo multiple sclerosis spinal cords. Neuroimage 55:1454–1460 . doi: 10.1016/j.neuroimage.2011.01.007.Radial

324. Samson RS, Lévy S, Schneider T, et al (2016) ZOOM or Non-ZOOM ? Assessing Spinal Cord Diffusion Tensor Imaging Protocols for Multi- Centre Studies. PLoS One 11:e0155557 . doi: 10.1371/journal.pone.0155557

325. Farrell JAD, Smith SA, Gordon-lipkin EM, et al (2008) High b-Value q-Space Diffusion-Weighted MRI of the Human Cervical Spinal Cord In Vivo : Feasibility and Application to Multiple Sclerosis. Magn Reson Med 59:1079–1089 . doi: 10.1002/mrm.21563

326. Abdel-Aziz K, Schneider T, Solanky BS, et al (2015) Evidence for early neurodegeneration in

the cervical cord of patients with primary progressive multiple sclerosis. Brain 138:1568–82 . doi: 10.1093/brain/awv086

327. Cortese R, Prados F, Moccia M, et al (2017) Evidence for progressive neurodegeneration in the cervical cord of patients with early primary progressive MS during 3-year follow-up. Mult Scler 23:262

328. Zhang H, Schneider T, Wheeler-Kingshott C, Alexander D (2012) NODDI: practical in vivo neurite orientation dispersion and density imaging of the human. Neuroimage 16:1000–1016

329. Grussu F, Schneider T, Tur C, et al (2017) Neurite dispersion: a new marker of multiple sclerosis spinal cord pathology? Ann Clin Transl Neurol 4:663–679 . doi: 10.1002/acn3.445

330. Grussu F, Schneider T, Zhang H, et al (2015) Neurite orientation dispersion and density imaging of the healthy cervical spinal cord in vivo. Neuroimage 111:590–601 . doi: 10.1016/j.neuroimage.2015.01.045

331. Tona F, Cawley N, Grussu F, et al (2016) Neurite orientation dispersion and density imaging (NODDI) of the spinal cord in relapsing remitting multiple sclerosis. ECTRIMS 2016 - London. Mult Scler 22:513–514

332. By S, Xu J, Box BA, et al (2017) Application and evaluation of NODDI in the cervical spinal cord of multiple sclerosis patients. NeuroImage Clin 15:333–342 . doi: 10.1016/j.nicl.2017.05.010

333. By S, Xu J, Box BA, et al (2018) Multi-compartmental diffusion characterization of the human cervical spinal cord in vivo using the spherical mean technique. NMR Biomed 31:e3894 . doi: 10.1002/nbm.3894

334. Panagiotaki E, Schneider T, Siow B, et al (2012) Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. Neuroimage 59:2241–2254 . doi: 10.1016/j.neuroimage.2011.09.081

335. Oh J, Saidha S, Chen M, et al (2013) Spinal cord quantitative MRI discriminates between

disability levels in multiple sclerosis. Neurology 80:540–547

336. Combès B, Kerbrat A, Ferré JC, et al (2018) Focal and diffuse cervical spinal cord damage in patients with early relapsing–remitting MS: A multicentre magnetisation transfer ratio study. Mult Scler 1352458518781999 . doi: 10.1177/1352458518781999

337. Liu Z, Pardini M, Yaldizli O, et al (2015) Magnetization transfer ratio measures in normal-appearing white matter show periventricular gradient abnormalities in multiple sclerosis. Brain 138:1239–1246 . doi: 10.1093/brain/awv065

338. Kearney H, Yiannakas MC, Samson RS, et al (2014) Investigation of magnetization transfer ratio-derived pial and subpial abnormalities in the multiple sclerosis spinal cord. Brain 137:2456–2468 . doi: 10.1093/brain/awu171

339. Laule C, Kozlowski P, Leung E, et al (2008) Myelin water imaging of multiple sclerosis at 7 T : Correlations with histopathology. Neuroimage 40:1575–1580 . doi: 10.1016/j.neuroimage.2007.12.008

340. Laule C, Yung A, Pavolva V, et al (2016) High-resolution myelin water imaging in post-mortem multiple sclerosis spinal cord : A case report. Mult Scler 22:1485–1489 . doi: 10.1177/1352458515624559

341. Minty EP, Bjarnason TA, Laule C, Mackay AL (2009) Myelin Water Measurement in the Spinal Cord. Magn Reson Med 61:883–892 . doi: 10.1002/mrm.21936

342. Liu H, MacMillian EL, Jutzeler CR, et al (2017) Assessing structure and function of myelin in cervical spondylotic myelopathy Evidence of demyelination. Neurology 89:602–610

343. Laule C, Vavasour IM, Zhao Y, et al (2010) Two-year study of cervical cord volume and myelin water in primary progressive multiple sclerosis. Mult Scler 16:670–677 . doi: 10.1177/1352458510365586

344. Kolind S, Seddigh A, Combes A, et al (2015) Brain and cord myelin water imaging: A

progressive multiple sclerosis biomarker. NeuroImage Clin 9:574–580 . doi: 10.1016/j.nicl.2015.10.002

345. Combes AJE, Matthews L, Lee JS, et al (2017) Cervical cord myelin water imaging shows degenerative changes over one year in multiple sclerosis but not neuromyelitis optica spectrum disorder. NeuroImage Clin 16:17–22 . doi: 10.1016/j.nicl.2017.06.019
346. Lema A, Bishop C, Malik O, et al (2017) A Comparison of Magnetization Transfer Methods to Assess Brain and Cervical Cord Microstructure in Multiple Sclerosis. J Neuroimaging 27:221–226 . doi: 10.1111/jon.12377
347. Smith AK, Dortch RD, Dethrage LM, Smith SA (2014) Rapid , high-resolution quantitative magnetization transfer MRI of the human spinal cord. Neuroimage 95:106–116 . doi: 10.1016/j.neuroimage.2014.03.005
348. Wyss PO, Hock A, Kollias S (2017) The Application of Human Spinal Cord Magnetic Resonance Spectroscopy to Clinical Studies : A Review. Semin Ultrasound, CT, MRI 38:153–162 . doi: 10.1053/j.sult.2016.07.005
349. Gómez-Ansón B, MacManus D, Parker G, et al (2000) In vivo 1H-magnetic resonance spectroscopy of the spinal cord in humans. Neuroradiology 42:515–517
350. Hock A, Wilm B, Zandomeneghi G, et al (2016) Neurochemical profile of the human cervical spinal cord determined by MRS. NMR Biomed 29:1464–1476 . doi: 10.1002/nbm.3589
351. Bellenberg B, Busch M, Trampe N, et al (2013) 1H-Magnetic Resonance Spectroscopy in diffuse and focal cervical cord lesions in Multiple Sclerosis. Eur Radiol 23:3379–3392 . doi: 10.1007/s00330-013-2942-7
352. Hock A, Henning A, Boesiger P, Kollias SS (2013) H-MR Spectroscopy in the Human Spinal Cord. AJNR Am J Neuroradiol 34:1682–1689
353. Petracca M, Vancea RO, Fleysher L, et al (2016) Brain intra- and extracellular sodium

concentration in multiple sclerosis: A 7 T MRI study. Brain 139:795–806 . doi: 10.1093/brain/awv386

354. Paling D, Solanky BS, Riemer F, et al (2013) Sodium accumulation is associated with disability and a progressive course in multiple sclerosis. Brain 136:2305–2317 . doi: 10.1093/brain/awt149

355. Conrad BN, Barry RL, Rogers BP, et al (2018) Multiple sclerosis lesions affect intrinsic functional connectivity of the spinal cord. Brain 141:1650–1664 . doi: 10.1093/brain/awy083

356. Solanky BS, Riemer F, Golay X, Wheeler-Kingshott CAM (2013) Sodium quantification in the spinal cord at 3T. Magn Reson Med 69:1201–1208 . doi: 10.1002/mrm.24696

357. Solanky BS, Prados F, Yiannakas MC, et al (2018) Associations between tissue sodium concentration, age and cross-sectional area in the healthy spinal cord. ISMRM

358. Wheeler-Kingshott C, Stroman P, Schwab J, et al (2014) The current state-of-the-art of spinal cord imaging: applications. Neuroimage 84:1082–1093 . doi: 10.1016/j.neuroimage.2013.07.014

359. Kadam P, Bhalerao S (2010) Sample size calculation. Int J Ayurveda Res 1:55–57 . doi: 10.4103/0974

360. Van Den Elskamp IJ, Boden B, Dattola V, et al (2010) Cerebral atrophy as outcome measure in short-term phase 2 clinical trials in multiple sclerosis. Neuroradiology 52:875–881 . doi: 10.1007/s00234-009-0645-1

361. Molyneux PD, Miller DH, Filippi M, et al (2000) The use of magnetic resonance imaging in multiple sclerosis treatment trials: power calculations for annual lesion load measurement. J Neurol 247:34–40 . doi: 10.1007/s004150050007

362. Sormani MP, Bruzzi P, Miller DH, et al (1999) Modelling MRI enhancing lesion counts in multiple sclerosis using a negative binomial model: Implications for clinical trials. J Neurol Sci

163:74–80 . doi: 10.1016/S0022-510X(99)00015-5

363. Sormani MP, Rovaris M, Bagnato F, et al (2001) Sample size estimations for MRI-monitored trials of MS comparing new vs standard treatments. Neurology 57:1883–1886

364. van den Elskamp IJ, Knol DL, Vrenken H, et al (2010) Lesional magnetization transfer ratio: a feasible outcome for remyelinating treatment trials in multiple sclerosis. Mult Scler 16:660–669 . doi: 10.1177/1352458510364630

365. Sormani MP, Calabrese M, Signori A, et al (2011) Modeling the distribution of new MRI cortical lesions in multiple sclerosis longitudinal studies. PLoS One 6:e26712 . doi: 10.1371/journal.pone.0026712

366. Chataway J, Nicholas R, Todd S, et al (2011) A novel adaptive design strategy increases the efficiency of clinical trials in secondary progressive multiple sclerosis. Mult Scler 17:81–88 . doi: 10.1177/1352458510382129

367. Smith SM, Zhang Y, Jenkinson M, et al (2002) Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage 17:479–89 . doi: 10.1006/nimg.2002.1040

368. Freeborough P, Fox N (1997) The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. IEEE Trans Med Imaging 16:623–629 . doi: 10.1109/42.640753

369. Leung K, Clarkson M, Bartlett J, et al (2010) Robust atrophy rate measurement in Alzheimer's disease using multi-site serial MRI: tissue-specific intensity normalization and parameter selection. Neuroimage 50:516–523 . doi: 10.1016/j.neuroimage.2009.12.059

370. Prados F, Cardoso MJ, Leung KK, et al (2015) Measuring brain atrophy with a generalized formulation of the boundary shift integral. Neurobiol Aging 36:S81–S90 . doi: 10.1016/j.neurobiolaging.2014.04.035

371. Radue E-W, Sprenger T, Gaetano L, et al (2015) Teriflunomide slows brain volume loss in relapsing MS: a SIENA analysis of the TEMSO MRI datase. ECTRIMS

372. Fazekas F, Sørensen P, Filippi M, et al (2005) MRI results from the European Study on Intravenous Immunoglobulin in Secondary Progressive Multiple Sclerosis (ESIMS). Mult Scler 11:433–440 . doi: 10.1191/1352458505ms1196oa

373. Prados F, Cardoso M, Yiannakas M, et al (2016) Fully automated grey and white matter spinal cord segmentation. Sci Rep 6:36151

374. Prados F, Yiannakas M, Cardoso M, et al (2016) Computing spinal cord atrophy using the Boundary Shift Integral: a more powerful outcome measure for clinical trials? ECTRIMS

375. Prados F, Yiannakas M, Cardoso M, et al (2016) Atrophy computation in the spinal cord using the Boundary Shift Integral. ISMRM

376. De Leener B, Kadoury S, Cohen-Adad J (2014) Robust, accurate and fast automatic segmentation of the spinal cord. Neuroimage 98:528–536 . doi: 10.1016/j.neuroimage.2014.04.051

377. Jones DK, Basser PJ (2004) “Squashing peanuts and smashing pumpkins”: How noise distorts diffusion-weighted MR data. Magn Reson Med 52:979–993 . doi: 10.1002/mrm.20283

378. Tristán-Vega A, García-Pérez V, Aja-Fernández S, Westin C (2012) Efficient and robust nonlocal means denoising of MR data based on salient features matching. Comput Methods Programs Biomed 105:131–144 . doi: 10.1016/j.cmpb.2011.07.014

379. Tustison N, Avants B, Cook P, et al (2010) N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 29:1310–1320 . doi: 10.1109/TMI.2010.2046908

380. De Leener B, Mangeat G, Dupont S, et al (2017) Toplogically preserving straightening of spinal cord MRI. J Magn Reson Imaging 46:1209–1219 . doi: 10.1002/jmri.25622

381. Smith S, De Stefano N, Jenkinson M, Matthews P (2001) Normalized accurate measurement

of longitudinal brain change. J Comput Assist Tomogr 25:466–475

382. Reuter M, Schmansky N, Rosas H, Fischl B (2012) Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage 61:1402–1418 . doi: 10.1016/j.neuroimage.2012.02.084

383. Leung K, Ridgway G, Ourselin S, et al (2012) Consistent multi-time-point brain atrophy estimation from the boundary shift integral. Neuroimage 59:3995–4005 . doi: 10.1016/j.neuroimage.2011.10.068

384. Modat M, Cash DM, Daga P, et al (2014) Global image registration using a symmetric block-matching approach. J Med Imaging 1:024003 . doi: 10.1117/1.JMI.1.2.024003

385. Lewis E, Fox N (2004) Correction of differential intensity inhomogeneity in longitudinal MR images. Neuroimage 23:75–83 . doi: 10.1016/j.neuroimage.2004.04.030

386. Lukas C, Sombekke M, Bellenberg B, et al (2013) Relevance of spinal cord abnormalities to clinical disability in multiple sclerosis: MR imaging findings in a large cohort of patients. Radiology 269:542–552 . doi: 10.1148/radiol.13122566

387. Ziegler G, Grabher P, Thompson A, et al (2018) Progressive neurodegeneration following spinal cord injury. Neurology 90:e1257–e1266 . doi: 10.1212/WNL.0000000000005258

388. De Leener B, Cohen-Adad J, Kadoury S (2015) Automatic Segmentation of the Spinal Cord and Spinal Canal Coupled with Vertebral Labeling. IEEE Trans Med Imaging 34:1705–1718 . doi: 10.1109/TMI.2015.2437192

389. Prados F, Moccia M, Johnson A, et al (2020) Generalised boundary shift integral for longitudinal assessment of spinal cord atrophy. Neuroimage 209:116489 . doi: https://doi.org/10.1016/j.neuroimage.2019.116489

390. Rocca MA, Sormani MP, Rovaris M, et al (2017) Long-term disability progression in primary progressive multiple sclerosis: A 15-year study. Brain 140:2814–2819 . doi:

10.1093/brain/awx250

391. Cohen-Adad J, Mareyam A, Keil B, et al (2011) 32-channel RF coil optimized for brain and cervical spinal cord at 3T. Magn Reson Med 66:1198–1208 . doi: 10.1002/mrm.22906

392. Dietrich O, Raya JG, Reeder SB, et al (2007) Measurement of signal-to-noise ratios in MR images: Influence of multichannel coils, parallel imaging, and reconstruction filters. J Magn Reson Imaging 26:375–385 . doi: 10.1002/jmri.20969

393. Geissler A, Gartus A, Foki T, et al (2007) Contrast-to-noise ratio (CNR) as a quality parameter in fMRI. J Magn Reson Imaging 25:1263–1270 . doi: 10.1002/jmri.20935

394. Giovannoni G, Tomic D, Bright JR, Havrdová E (2017) “No evident disease activity”: The use of combined assessments in the management of patients with multiple sclerosis. Mult Scler 23:1179–1187 . doi: 10.1177/1352458517703193

395. Uher T, Vaneckova M, Sobisek L, et al (2017) Combining clinical and magnetic resonance imaging markers enhances prediction of 12-year disability in multiple sclerosis. Mult Scler 23:51–61 . doi: 10.1177/1352458516642314

396. Prados F, Jorge M, Kanber B, et al (2016) A multi-time-point modality-agnostic patch-based method for lesion fi lling in multiple sclerosis. Neuroimage 139:376–384 . doi: 10.1016/j.neuroimage.2016.06.053

397. Prados F, Jorge Cardoso M, Yiannakas M, et al (2017) A patch-based method for lesion in-painting in the spinal cord. ISMRM

398. Amiri H, de Sitter A, Bendfeldt K, et al (2018) Urgent challenges in quantification and interpretation of brain grey matter atrophy in individual MS patients using MRI. NeuroImage Clin 19:466–475 . doi: 10.1016/j.nicl.2018.04.023

399. Battaglini M, Jenkinson M, De Stefano N (2012) Evaluating and reducing the impact of white matter lesions on brain volume measurements. Hum Brain Mapp 33:2062–2071 . doi:

10.1002/hbm.21344

400. Popescu V, Ran NCG, Barkhof F, et al (2014) Accurate GM atrophy quantification in MS using lesion-filling with co-registered 2D lesion masks. NeuroImage Clin 4:366–373 . doi: 10.1016/j.nicl.2014.01.004

401. Massire A, Taso M, Besson P, et al (2016) NeuroImage High-resolution multi-parametric quantitative magnetic resonance imaging of the human cervical spinal cord at 7T. Neuroimage 143:58–69 . doi: 10.1016/j.neuroimage.2016.08.055

402. Massire A, Rasoanandrianina H, Taso M, et al (2018) Feasibility of Single-Shot Multi-Level Multi-Angle Diffusion Tensor Imaging of the Human Cervical Spinal Cord at 7T. Magn Reson Med 80:947–957 . doi: 10.1002/mrm.27087

403. Cohen-Adad J, Zhao W, Keil B, et al (2013) 7-T MRI of the spinal cord can detect lateral corticospinal tract abnormality in amyotrophic lateral sclerosis. Muscle Nerve 47:760–762 . doi: 10.1002/mus.23720

404. Wang Y, Wang Y, Tan S, Lu Z (2016) Spinal cord atrophy in neuromyelitis optica spectrum disorders. Mult Scler Relat Disord 8:9–10 . doi: 10.1016/j.msard.2016.04.007

405. Azodi S, Nair G, Enose-Akahata Y, et al (2017) Imaging spinal cord atrophy in progressive myelopathies: HTLV-I-associated neurological disease (HAM/TSP) and multiple sclerosis (MS). Ann Neurol 82:719–728 . doi: 10.1002/ana.25072

406. El Mendili MM, Querin G, Bede P, Pradat PF (2019) Spinal cord imaging in amyotrophic lateral sclerosis: Historical concepts—novel techniques. Front Neurol 10:1–11 . doi: 10.3389/fneur.2019.00350

407. Paquin ME, Mendili MME, Gros C, et al (2018) Spinal cord gray matter atrophy in amyotrophic lateral sclerosis. Am J Neuroradiol 39:184–192 . doi: 10.3174/ajnr.A5427

408. Olney NT, Bischof A, Rosen H, et al (2018) Measurement of spinal cord atrophy using phase

sensitive inversion recovery (PSIR) imaging in motor neuron disease. PLoS One 13:1–15 . doi: 10.1371/journal.pone.0208255

409. Denecke CK, Aljović A, Bareyre FM (2019) Combining molecular intervention with in vivo imaging to untangle mechanisms of axon pathology and outgrowth following spinal cord injury. Exp Neurol 318:1–11 . doi: 10.1016/j.expneurol.2019.04.003

410. Chevis CF, Da Silva CB, D'Abreu A, et al (2013) Spinal cord atrophy correlates with disability in Friedreich's ataxia. Cerebellum 12:43–47 . doi: 10.1007/s12311-012-0390-6

411. Behmanesh B, Gessler F, Quick-Weller J, et al (2017) Regional Spinal Cord Atrophy Is Associated with Poor Outcome After Surgery on Intramedullary Spinal Cord Ependymoma: A New Aspect of Delayed Neurological Deterioration. World Neurosurg 100:250–255 . doi: 10.1016/j.wneu.2017.01.026

412. Lorenzi RM, Palesi F, Castellazzi G, et al (2020) Unsuspected Involvement of Spinal Cord in Alzheimer Disease. Front Cell Neurosci 14:1–10 . doi: 10.3389/fncel.2020.00006

413. Wilhelms W, Bellenberg B, Köster O, et al (2017) Progressive spinal cord atrophy in manifest and premanifest Huntington's disease. J Neurol Neurosurg Psychiatry 88:614–616 . doi: 10.1136/jnnp-2016-315000

414. Antonescu F, Adam M, Popa C, Tuţă S (2018) A review of cervical spine MRI in ALS patients. J Med Life 11:123–127

415. Gray E (2019) Introducing the efficient clinical trials platform. https://www.mssociety.org.uk/research/latest-research/latest-research-news-and-blogs/introducing-the-efficient-clinical-trials-platform