Quantitative MRI and machine learning for the diagnosis and prognosis of Multiple Sclerosis

Antonio Ricciardi

A dissertation submitted to the
University College London for the degree of

Doctor of Philosophy

July 2021

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

Signed,

Abstract

Multiple sclerosis (MS) is an immune-mediated, inflammatory, neurological disease affecting myelin in the central nervous system, whose driving mechanisms are not yet fully understood. Conventional magnetic resonance imaging (MRI) is largely used in the MS diagnostic process, but because of its lack of specificity, it cannot reliably detect microscopic damage. Quantitative MRI provides instead feature maps that can be exploited to improve prognosis and treatment monitoring, at the cost of prolonged acquisition times and specialised MR-protocols.

In this study, two converging approaches were followed to investigate how to best use the available MRI data for the diagnosis and prognosis of MS. On one hand, qualitative data commonly used in clinical research for lesion and anatomical purposes were shown to carry quantitative information that could be used to conduct myelin and relaxometry analyses on cohorts devoid of dedicated quantitative acquisitions. In this study arm, named bottom-up, qualitative information was up-converted to quantitative surrogate: traditional model-fitting and deep-learning frameworks were proposed and tested on MS patients to extract relaxometry and indirect-myelin quantitative data from qualitative scans. On the other hand, when using multi-modal MRI data to classify MS patients with different clinical status, different MR-features contribute to specific classification tasks. The top-down study arm consisted in using machine learning to reduce the multi-modal dataset dimensionality only to those MR-features that are more likely to be biophysically meaningful with respect to each MS phenotype pathophysiology.

Results show that there is much more potential to qualitative data than lesion and tissue segmentation, and that specific MRI modalities might be better suited for investigating individual MS phenotypes. Efficient multi-modal acquisitions informed by biophysical findings, whilst being able to extract quantitative information from qualitative data, would provide huge statistical power through the use of large, historical datasets, as well as constitute a significant step forward in the direction of sustainable research.

Acknowledgements

Thank you to Prof Claudia Gandini Wheeler-Kingshott, for her expertise and constant professional and human support, whether from her office around the corner or through a Teams call around the world.

Thank you to Dr Francesco Grussu and Dr Baris Kanber, for the hands-on help provided, not only in understanding how things work, but also in *making them actually work*.

Thank you to Dr Ferran Prados and Dr Baris Kanber again, making data processing on XNAT possible.

Thank you to Prof Danny Alexander and Prof Olga Ciccarelli for supervising my work throughout the years and providing useful insights on the computer science and clinical sides of MRI research.

Thank you to Tessa and Leafy for having me as their guest, despite my poor social skills, and *the plague* too.

This work has been supported by the EPSRC-funded UCL Centre for Doctoral Training in Medical Imaging (EP/L016478/1); the H2020-EU.3.1 CDS-QUAMRI (634541) grant; the UK Multiple Sclerosis Society (892/08) and the Department of Health's National Institute for Health Research Biomedical Research Centres (BRC R&D03/10/RAG0449) (supporting the NMR Unit);

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Several neurological disorders lack objective criteria for patients stratification with regard to subtypes at early stages of the disease, leading to inaccurate prognoses and making it impossible to have reliable personalised treatment plans. This applies, specifically, also to multiple sclerosis (MS), an inflammatory demyelinating disease affecting the myelinated axons in the central nervous system, damaging the myelin and the axons to varying degrees. The factors at the base of MS pathogenesis are only partially understood, and the mechanisms driving its complex and unpredictable evolution are still unclear [1].

1.1 Problem statement

MS diagnosis relies on lesion-based evidence explaining clinical symptoms, as specified in the 2017 McDonald criteria [2]. Lesion load is usually assessed through T_2-weighted magnetic resonance imaging (MRI), that reveals pathological tissue as hyperintense regions. *Qualitative* scans have represented the workhorse of MS clinical research for decades, and large historical qualitative datasets have been built over the years. However, conventional qualitative MRI lacks specificity, as different histopathological substrates of tissue damage might produce the same patterns of the MR-signal and could not be told apart. Furthermore, whilst conventional MRI readouts are sensitive to certain macroscopic aspects of MS, they do not detect early, widespread microscopic damage, unlike quantitative approaches. For example, in addition to the well-characterised inflammatory white matter lesions, studies using magnetisation transfer imaging have shown that the normal appearing white matter of the majority of MS patients has significant abnormalities otherwise invisible to traditional T_2-weighted MRI [3].

On the other hand, having access to a very rich collection of features might be detrimental to the classical human-based recognition process of finding disease biomarkers, as they might be embedded into a combination of features which makes the information they carry difficult to be visually identified and exploited. Therefore, advanced statistical machine learning and pattern recognition techniques have been applied to a wide variety of neurodegenerative disorders to support the diagnostic process, such as in traumatic brain injury [4] and Alzheimer's disease [5, 6].

With the rising number of specialised quantitative MR-sequences to investigate MS pathophysiology, only few of which eventually reaching clinical fruition, and the growing interest towards multi-modal, big-data approaches, it has become imperative to assess which modalities justify increased acquisition times and costs for their added values to clinical phenotyping and patient management. At the same time, given the ample amount of qualitative scans available, being able to extract quantitative information from routinely acquired images, and thus taking full advantage of the MR-modalities already at hand, would provide great statistical power through the employment of large historical datasets for quantitative analyses, as well as representing a key step forward in the direction of *sustainable* and *efficient* research.

1.2 Aims

This problem was tackled through two converging pathways:

- a *bottom-up* arm, which aimed to *enhance*, or *up-convert* the qualitative content of routine scans to quantitative information sensitive to MS — this approach was further divided into two studies, one using a classical model-fitting approach, and one implementing a dedicated advanced deep-learning algorithm;

- a *top-down* arm, which aimed to *reduce* a multi-dimensional dataset *down* to those MRI modalities that most likely correlate with MS pathophysiology, using machine learning feature selection and classification analysis.

The dual-approach has been instrumental in the definition of a work plan that could contribute in a systematic way towards the stated problem, with the specific objectives and overall contributions from each arm being summarised as follows.

1.2.1 Bottom-up

In the *bottom-up* study, traditional model fitting and deep learning were employed in order to extract quantitative information from MR-images otherwise used only for lesion segmentation and anatomical purposes. Simple modalities acquired routinely for clinical assessment are often dismissed from quantitative analyses because they are traditionally labelled as qualitative: it is thus key to assess whether computational models allow qualitative scans to be exploited for indices sensitive to myelin, as well as other quantitative metrics. This can be rephrased under the *bottom-up hypothesis* that, in the absence of dedicated quantitative scans, standard qualitative images can be used to infer summary relaxometry and myelin-sensitive indices that well correlate with their respective ground truth.

The myelin relaxation, or *MyRelax*, framework[1], was used to extract proton density (PD), macromolecular tissue volume (MTV), T_2 and T_1 maps from qualitative PD-, T_2- and T_1-weighted scans, in this thesis often referred to, collectively, as *qualitative images* or *qualitative scans*. A U-Net was then used to extract magnetisation transfer ratio (MTR) from the same qualitative scans by means of deep learning. In this context, the prefix *QuaSI-*, standing for *qualitative scans for indirect-*, has been introduced to better distinguish quantitative maps produced from qualitative scans, from ground truth, e.g. QuaSI-PD produced via MyRelax as opposed to PD acquired through dedicated sequences.

With this in mind, the *bottom-up* study can be decomposed and summarised into three objectives:

1. **MyRelax validation**: to assess the accuracy and reproducibility of QuaSI-PD, -T_2, -T_1 maps obtained from the qualitative scans using the MyRelax framework, by comparing them with the quantitative PD, T_2, T_1 maps obtained using gold standard quantitative MRI sequences.

2. **MyRelax MS application**: to evaluate the applicability of the MyRelax framework to MS, with QuaSI-MTV maps produced using MyRelax being compared to MTR, to test how much information attributable to myelin is shared by the two modalities. T_1-/T_2-weighted ratio images (T_1w/T_2w) were also compared to the MTR maps for the same reason.

3. **U-Net MS application**: to implement a deep learning network to extract MTR information directly from the qualitative scans — QuaSI-MTR — bypassing traditional model fitting.

[1] Courtesy of Dr Francesco Grussu.

1.2.2 Top-down

The *top-down* study consisted in using machine learning to reduce the dimensionality of the multi-modal MRI dataset only to those feature that are more likely to be *biophysically meaningful* with respect to characterising MS progression, informing future acquisitions and investigation.This was applied to a multi-modal MRI dataset including anatomical, relaxometry, diffusion and sodium quantitative measurements for healthy controls and MS patients with different MS subtypes.

The development of a decision system based on *targeted* multi-modal quantitative MRI, advanced feature extraction and multi-parametric classification would support the implementation of clinical applications, aiding personalised clinical management so that patients could receive the treatment that best suits their own phenotype. This would improve the accuracy of prognoses, especially at early stages where macroscopic alterations like atrophy are not as predominant, and microstructural and functional information might be most meaningful. This would also allow to pinpoint which MRI modalities might contribute more to the diagnosis and patient follow-up, resulting in good candidates for clinical optimisation, and which ones are not as likely to benefit from the added acquisition time and cost.

1.2.3 Contributions

With respect to the *bottom-up* and *top-down* objectives, three key contributions can be delineated:

- **MyRelax: myelin and relaxation imaging**, reporting the contributions from the *MyRelax validation* and *MyRelax MS application bottom-up* objectives [7];
- **Deep learning MTR from qualitative images**, following the *U-Net MS application bottom-up* objective [8];
- **Biophysically meaningful features for classification of MS phenotypes**, associated to the *top-down* study [9].

1.3 Outline

In the *Background* part, a description of myelin and its physiological role in the central nervous system is given, together with a brief overview of MS as a demyelinating disease, its symptoms, diagnosis and phenotypes. A general introduction of MRI fundamentals,

the current state-of-the-art, main contributions and limitations of myelin imaging and quantitative MRI techniques are then outlined, followed by an introduction to machine learning and a literature review on its applications to modern neuroimaging.

Three parts follow, each related to one of the three key contributions delineated above; in each part, the *Introduction* chapter frames the context of the study, the *Methods* chapter describes the MRI protocols and image analysis techniques used, with the outcomes being reported and interpreted in the *Results* and *Discussions* chapters, respectively.

A *Conclusions and future works* chapter reflects on the limitations and future works for each key contribution, with closing remarks on the overall contribution of the whole body of work.

Part I

Background

Chapter 2

Myelin: what is it and what is its role?

Myelin is a lipidic substance forming a multi-layered sheath surrounding axons, creating an electrically insulating membrane preventing electric current travelling within the fibre to leave the cell.

In the central nervous system (CNS, which includes brain, spinal cord and optic nerve), myelin is supplied by cells called *intrafascicular oligodendrocytes*, whilst *Schwann* cells insulate the axons of the peripheral nerves. Myelin is distributed along the axons discontinuously: the insulating sheath spreads along most of the fibre's length, with gaps at regular intervals 1–2 µm wide along the axon called *nodes of Ranvier*. These gaps are rich in voltage-gated ion channels, and allow sodium ions to move across the axonal membrane, depolarising it, and initiating the generation of an *action potential*. Myelinated segments are devoid of these channels, constraining ion transfer across membrane only at the nodes of Ranvier. The action potential appears then to *jump* from one node to the next in a propagation process called *saltuatory* conduction. This process allows signal transmission speed to reach around 150 m/s, whereas in unmyelinated axons it cannot exceed 10 m/s [10]. Disruption of the myelin sheath causes the impulses travelling through the affected fibres to be distorted or interrupted, producing a wide variety of symptoms, ranging from physical conditions, such as vision and balance deterioration, fatigue, bladder problems and stiffness and/or spasms, to degradation of memory and cognitive functions.

Both myelin chemical composition and spatial distribution are fundamental for fast and efficient signal propagation within the brain. The loss or damage of such insulating sheath is called *demyelination*, and can result in the disruption of signal transmission and consequent neuronal degeneration.

Chapter 3

Multiple Sclerosis

Multiple sclerosis (MS) is an immune-mediated, inflammatory, neurological disease of the CNS, characterised by an abnormal response of the body's immune system affecting myelin and damaging neurons to various degrees. Because the exact antigen the immune cells are instructed to attack is still unknown, MS is considered an *immune mediated* disease rather than *autoimmune*, although this continues to be the subject of debate in the scientific community [11, 12]. A wide array of anticancer drugs has been repurposed for the treatment of MS inflammatory symptoms due to their immunosuppressive and immunomodulatory role. These drugs regulate the responses of the central nervous system immune system by inhibiting the activation and proliferation of T-cells, B-cells, lowering antibody production and deactivating macrophages attacking myelin [13].

3.1 Pathophysiology

MS is first characterised by inflammation and acute demyelination, then followed by gliosis, which leads to the formation of scarring tissue, or lesions, also called *sclerosis*. The consequent axonal damage might result in axonal transection, leading to axonal loss via Wallerian degeneration, that is the process of anterograde degeneration, spreading forward along the distal part of the axon, following axonal injury. Neurodegeneration is defined as the progressive loss of neuronal structure and/or function, eventually culminating with the death of neurons; it has been reported since the first stages of the disease, leading to a higher rate of atrophy compared to controls, and is the main cause of accumulation of cognitive and physical disability. Whilst Wallerian degeneration has been shown to contribute significantly to axonal loss within lesions, other mechanisms

of axonal degeneration may also come into play at different stages of the disease [14].

3.2 Epidemiology

Studies performed at a global level have led researchers to believe that the disease is triggered in a genetically predisposed individual by one or more environmental factors, whose specific nature is at the moment still unknown. It appears that women are affected more often than men, with a ratio of about 3:1, which suggests the involvement of hormones in the development of the disease. In addition, a combination of geographic and ethnic factors is believed to have an impact, with a higher incidence of MS being observed at higher latitude compared to regions near the equator [14, 15].

3.3 Phenotypes

The process of inflammation, demyelination and neurodegeneration starts often at a sub-clinical level. In the majority of MS patients, the disease begins to manifests with episodes of reversible neurological deficits lasting at least 24 hours, a condition called *clinically isolated* syndrome (CIS). CIS can be either a *monofocal* or *multifocal* episode, where the former is characterised by a single neurologic symptom, which is often *optic neuritis*[1], whilst the latter is defined by multiple simultaneous symptoms (for example optic neuritis accompanied by sensorimotor conditions in the distal part of the body) caused by lesions or inflammation *foci* affecting more than one area. People who experience a CIS may or may *not* manifest further symptoms during their life, therefore, whilst the episode is suggestive of MS, it does not meet alone the criteria for a diagnosis of *clinically defined* MS. However, it has been shown that when CIS is accompanied by asymptomatic brain lesions similar to those seen in MS, the risk of conversion to clinically defined MS by 10 years ranges between 60% and 80%, whereas, in the absence of lesions, the likelihood of developing MS in the same timespan is as low as 20% [16].

In case of conversion to clinically defined MS, or progressive disease course from the onset, patients may be categorised into three MS phenotypes [17], whose progression in time is summarised in Figure 3.1:

- Relapsing-remitting MS (RRMS) is the most common disease course, affecting

[1] Inflammation or lesion of the optic nerve or pathways that control eye movements and visual coordination which may result in blurring or desaturation of vision, reduction of the peripheral vision or blindness in one eye. A dark spot, or *scotoma*, may occur in the centre of the visual field.

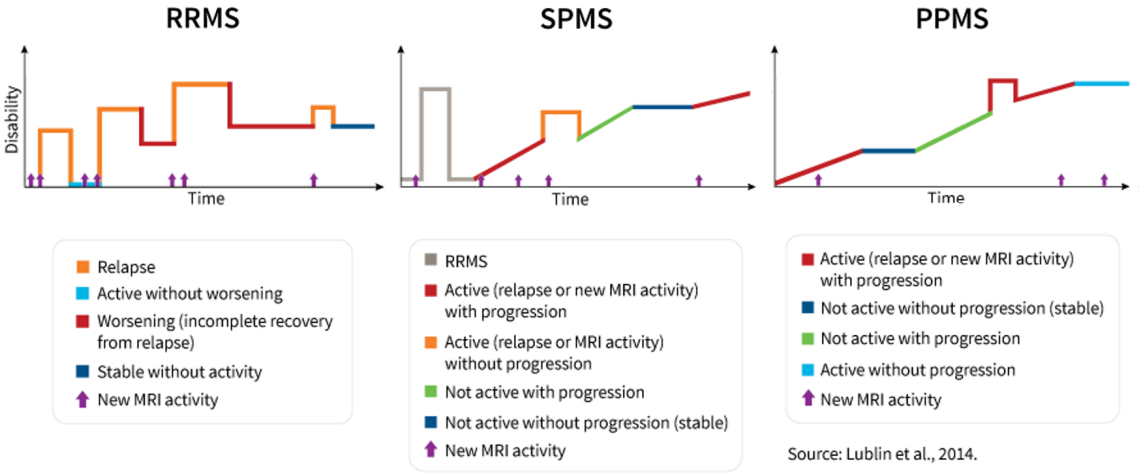


Figure 3.1: MS phenotypes classification and progression. *RRMS*: relapsing remitting MS; *SPMS*: secondary progressive MS; *PPMS*: primary progressive MS. Figures from https://www.nationalmssociety.org/What-is-MS/Types-of-MS in June 2021.

approximately 85% of MS patients. It is identified by clearly defined episodes of new neurologic symptoms or exacerbations of pre-existing ones, called *relapses*. Such attacks are then followed by periods of partial or complete recovery, or *remissions*, during which all symptoms may either disappear or persist and become permanent. Either way, remissions are characterised by the apparent absence of progression of the disease.

- Secondary progressive MS (SPMS) follows an initial relapsing-remitting course, with a progressive worsening of neurologic function and consequent accumulation of disability over time. Less than 20% of the patients who are diagnosed with RRMS will eventually transition to SPMS [18]. Patients may or may not continue to experience relapses caused by inflammation: the disease gradually shifts from the inflammatory process typical of RRMS to a more steadily progressive phase characterised by axonal damage or loss.

- Primary progressive MS (PPMS) is characterised by worsening of neurologic function and accrual of disability from the onset of symptoms, with rare relapses. Unlike SPMS, PPMS is the first and only phase of the disease for approximately 15% of people with MS, and involves much less inflammation (and more neuronal loss) compared to RRMS. As a result, PPMS patients usually present fewer brain lesions than RRMS patients, and the lesions tend to contain fewer inflammatory cells. PPMS patients usually present also more lesions in the spinal cord than in the brain. Together, such differences make PPMS more debilitating and more

difficult to diagnose and treat than MS relapsing forms.

3.4 Diagnosis

When diagnosing CIS patients, that is in absence of reoccurring relapses involving neurological functions, the diagnostic process follows the guidelines defined by the McDonald criteria, based on the evidence of three conditions: the insurgence of at least two different lesions in two independent regions of the CNS (*dissemination in space* criterion, DIS) occurred at dates at least one month apart from each other (*dissemination in time* criterion, DIT), and the presence of chronic inflammation (*inflammatory* criterion). Inflammation can be determined by performing a cerebrospinal fluid (CSF) oligoclonal band screen, which can be used in substitution for DIT [1]. After ruling out any other possible diagnosis, one for MS can be done.

The McDonald criteria include specific guidelines for using MRI and other analyses to aid the physician in the diagnostic process [19]. As of 2021, the latest update to the McDonald Criteria is the 2017 revision, which builds upon the 2010 revision by providing additional avenues for obtaining supporting evidence of lesion dissemination (both DIS and DIT). As per the updated criteria, both asymptomatic and now symptomatic MRI lesions can be considered in determining DIS or DIT (not including lesions in the optic nerve in a patient presenting with optic neuritis). Furthermore, lesions in cortical grey matter have also been included for DIS assessment, now demonstrated by evidence of one or more lesions suggestive of MS in at least two of four CNS areas: periventricular, cortical or juxtacortical, infratentorial sites, and the spinal cord [2].

Chapter 4

Magnetic Resonance Imaging

Because of the conjoint criteria that need to be met, magnetic resonance imaging (MRI) plays a fundamental role in the MS diagnostic process. MRI is a non-invasive diagnostic tool that exploits the magnetic properties of nuclear spins to reconstruct the anatomical structure of the examined tissue, whilst also investigating its biological and physiological properties. In this chapter, the fundamentals of MRI have been presented, mostly with reference to Brown et al. (2014) *Magnetic resonance imaging: physical principles and sequence design* [20]. Whilst not but a glance at the vast MRI landscape, this section aims to give the reader the basic notions necessary to understand the context of the reported work.

4.1 Fundamentals of nuclear spin–magnetic field interaction

Nuclear spin represents an intrinsic property of particles that manifests when they interact with a magnetic field. Nuclear magnetic resonance (NMR), and equivalently MRI, exploit the interaction of unpaired spin-1/2 particles with magnetic fields imposed by the diagnostic apparatus. The nuclear species most typically targeted in medical applications is water hydrogen nuclei ${}^{1}_{1}\mathrm{H}$, that is *protons*, due to the abundance of water, and thus ${}^{1}_{1}\mathrm{H}$ nuclear spins, within biological tissues. Specialised techniques might however focus on different nuclear species, such as ${}^{23}_{11}\mathrm{Na}$ with spin 3/2, in case of sodium MRI[1], which instead is used to investigate neuronal cell function and integrity [21]. The nuclear magnetic moment $\vec{\mu}$ is associated to the intrinsic spin angular momentum $\vec{S}$

[1] ${}^{123}_{11}\mathrm{Na}$ abundance in the human body is about 1000 times lower than ${}^{1}_{1}\mathrm{H}$: 80 mM and 88 M respectively.

according to the relation

$$\vec{\mu} = \gamma \vec{S} \tag{4.1}$$

where γ is defined as *gyromagnetic ratio* which, in the case of ${}^1_1\mathrm{H}$ protons, is equal to $\gamma = 267.513 \cdot 10^6\,\mathrm{rad/sT}$. The sum of the magnetic moments over a volume V affected by a constant external magnetic field defines the local magnetic moment per unit of volume, or *magnetisation*:

$$\vec{M} = \frac{1}{V} \sum_i \vec{\mu}_i \tag{4.2}$$

4.1.1 Zeeman effect and thermal equilibrium

The application of a static magnetic field $\vec{B}_0 = B_0 \hat{z}$ induces the partial alignment of the magnetic moments parallel to the field. This is an energetically favourable state, given the classical formulation of the magnetic potential energy $U = -\vec{\mu} \cdot \vec{B}_0$ is minimised for $\vec{\mu} \parallel \vec{B}_0$. According to the quantum formulation, this can be explained by the *Zeeman effect*, i.e. the discrete splitting of energy levels induced by the application of a magnetic field (also referred to as *Zeeman splitting*). The energy levels (or *eigenvalues*) are given by

$$E_{m_s} = -\vec{\mu} \cdot \vec{B}_0 = -\gamma \vec{S} \cdot \vec{B}_0 = -\gamma S_z B_0 = -\gamma \hbar m_s B_0 \tag{4.3}$$

where $S_z = \hbar m_s$ is the z-component of the spin vector, $m_s = [-s, -s+1, ..., s-1, s]$ is the *magnetic spin quantum number* describing the possible states associated to a particle with *spin quantum number* s, and $\hbar = h/2\pi$, with h being Plank's constant. For ${}^1_1\mathrm{H}$, $s = 1/2$ and $m_s = \pm 1/2$, leading to the two spin states usually referred as *spin-up* $(+1/2)$ and *spin-down* $(-1/2)$, with associated energy eigenvalues $E_+ = -\gamma\hbar B_0/2$ and $E_- = \gamma\hbar B_0/2$ respectively. The energy gap between levels is therefore

$$\Delta E = E_- - E_+ = \gamma \hbar B_0 \tag{4.4}$$

At equilibrium, this produces a net *longitudinal component of the magnetisation* M_z aligned along the $\hat{z}$ axis with magnitude M_0. The value of M_0 depends on the temperature of the sample: in regimes close to the human body temperature ($\sim$310 K), this relation can be approximated by *Curie's law*:

$$M_0 = \frac{\mu^2 \rho}{3 k_B T} B_0 \tag{4.5}$$

where ρ represents the *density of spins* within the sample, also referred to as *proton density* in conventional $^1_1$H MRI, k_B is the Boltzmann's constant, and T the temperature of the sample (expressed in kelvin units). For $^1_1$H, $\mu^2 = \gamma^2 S^2 = \gamma \hbar^2 s(s+1)$, with $s = 1/2$.

4.1.2 Precession

The application of $\vec{B}_0$ also establishes a motion of precession of the magnetic moments around the z axis that can be described in the classic formalism[2] by

$$\frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} = \gamma \vec{\mu} \times \vec{B}_0 \tag{4.6}$$

which can be solved along the x, y, z components as

$$\begin{cases} \mu_x(t) = \mu_x(0)\cos(\omega_0 t) + \mu_y(0)\sin(\omega_0 t) \\ \mu_y(t) = \mu_y(0)\cos(\omega_0 t) - \mu_x(0)\sin(\omega_0 t) \\ \mu_z(t) = \mu_z(0),\ \forall t \end{cases} \tag{4.7}$$

where the angular velocity vector is $\vec{\omega}_0 = -\omega_0 \hat{z}$, indicating a left-handed rotation with respect to the z-axis, and the precessional frequency, referred to as *Larmor frequency*, is defined as

$$\omega_0 = \gamma B_0 \tag{4.8}$$

The precession causes therefore each magnetic moment to accumulate a phase ϕ_0 over time according to the relation

$$\begin{aligned} \frac{\mathrm{d}\phi_0}{\mathrm{d}t}\hat{z} = \vec{\omega}_0 = -\omega_0 \hat{z} \\ \phi_0 = -\omega_0 t \end{aligned} \tag{4.9}$$

4.2 RF pulse and resonance condition

The equilibrium state can be perturbed by the application of a different magnetic field $\vec{B}_1$ orthogonal to $\vec{B}_0$. $\vec{B}_1$ is produced by a radiofrequency (RF) *coil* placed around the

[2] A similar solution can be obtained in terms of expectation values $\langle\mu_x\rangle, \langle\mu_y\rangle, \langle\mu_z\rangle$ through the quantum formulation.

sampled material whilst in transmit phase and over only a short period of time τ, or *pulse*, and it is therefore also referred to as RF pulse.

$\vec{B}_1$ rotates in the xy-plane at frequency ω_{RF}: when $\omega_{\text{RF}} = \omega_0$, $\vec{B}_1$ rotates at the Larmor frequency and this condition is called *on-resonance* (hence the *resonance* in M*R*I). The RF pulse induces a second precessional motion of the magnetic moments with frequency $\omega_1 = \gamma B_1$ which, similarly to what expressed by equation (4.9), causes a rotation of the average magnetisation, around the axis of application of $\vec{B}_1$, by an angle

$$\theta = -\omega_1 \tau = -\gamma B_1 \tau \tag{4.10}$$

called *flip-angle*[3]. This causes the magnetic moments to be tipped away from the $\hat{z}$ axis towards the transverse plane, whilst precessing at the Larmor frequency in a state of *phase coherence*, as represented in Figure 4.1. This reduces the magnitude of the longitudinal magnetisation, and the phase coherence causes the emergence of a net *transverse magnetisation* $\vec{M}_\perp = M_x\hat{x} + M_y\hat{y}$ lying in the xy-plane, which can be equivalently expressed in complex notation as

$$M_+ = M_x + iM_y = M_\perp e^{-i(\omega_0 t - \phi_0)} \tag{4.11}$$

where $M_\perp$ indicates the magnitude of the transverse magnetisation vector, and its x- and y-components can be expressed as

$$M_x = \text{Re}(M_+); \quad M_y = \text{Im}(M_+) \tag{4.12}$$

After the RF pulse, the magnitudes of the longitudinal and transverse components of the magnetisation therefore become

$$M_z(0^+) = M_z(0^-)\cos(\theta); \quad M_\perp(0^+) = M_z(0^-)\sin(\theta) \tag{4.13}$$

where the RF pulse is assumed to be applied instantaneously at $t = 0$; $M_z(0^-)$, $M_z(0^+)$, and $M_\perp(0^+)$ are, in order, short hand notations for M_z *immediately before*, and M_z and $M_\perp$ *immediately after* the RF pulse. If the system is at thermal equilibrium when the RF pulse is applied, then $M_z(0^-) = M_0$ and equation (4.13) become

$$M_z(0^+) = M_0\cos(\theta); \quad M_\perp(0^+) = M_0\sin(\theta) \tag{4.14}$$

[3] For $^1_1$H, $\theta = \pi/2$ can be achieved in $\tau = 1\,\text{ms}$ with $B_1 = 5.9\,\mu\text{T}$

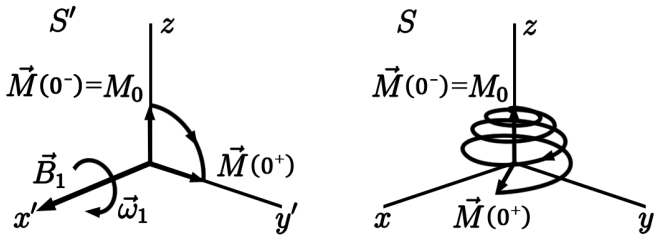


Figure 4.1: The application of the on-resonant B_1 pulse represented in the on-resonant reference frame $S'[x', y', z]$, also rotating at the Larmor frequency, and in the laboratory reference frame $S[x, y, z]$, where precession can be observed. In this example, $M_z(0^-) = M_0$ and $\theta = \pi/2$.

The optimal flip-angle value depends on the purpose of the MR-sequence and its parameters, as it will be explained in section 4.7. For example, for signal readout, a $\theta = \pi/2$ *excitation* RF pulse is often applied to maximise the transverse magnetisation, and consequently the signal amplitude too (see section 4.8), although a smaller θ may be more efficient to ensure shorter acquisition times.

4.3 Longitudinal relaxation

After an RF pulse, which suppose is applied at $t = 0$, the system relaxes to the more energetically favourable state through dipole-dipole energy exchange between the magnetic moments and the surrounding environment, historically referred to as *lattice*. The longitudinal component of the magnetisation M_z returns to its equilibrium value M_0, whilst the transverse component decays. This process is characterised by the *longitudinal*, or *spin-lattice*, *relaxation time* T_1 according to a simple exponential model:

$$\frac{\mathrm{d}M_z}{\mathrm{d}t} = \frac{1}{T_1}(M_0 - M_z) \tag{4.15}$$

with solution

$$\begin{aligned} M_z(t) &= M_0 - (M_0 - M_z(0^+))e^{-t/T_1} \\ &= M_0 - (M_0 - M_z(0^-)\cos(\theta))e^{-t/T_1} \end{aligned} \tag{4.16}$$

with $M_z(0^+)$ depending on the magnitude of the longitudinal magnetisation before the RF pulse and the flip-angle θ as in equation (4.13).

If the pulse sequence is repeated, the time between iterations is called *repetition time* T_R. For any n-th repetition, equation (4.16), valid for $n = 0$, can be generalised to

$$M_z(t_n) = M_0 - (M_0 - M_z(0^-)\cos(\theta))e^{-t_n/T_1}, \quad 0 \leq t_n < T_R, \quad n = 0, 1, 2, ... \tag{4.17}$$

with $t_n = t - nT_R$. If $T_R \gg T_1$, the spin-lattice relaxation can be assumed to have completely relaxed to thermal equilibrium between repetitions, and every n-th iteration will begin with $M_z(t_n = 0^-) = M_z(t = nT_R^-) = M_0, \ \forall n$.

Studies have shown that T_1 varies as a function of B_0: the best fitting of $^1_1$H brain MRI data in white matter, grey matter and blood has resulted in the empirical function:

$$T_1 = C(\gamma B_0)^\beta \tag{4.18}$$

with scaling factor C varying for the different tissues[4], and $\beta \simeq 1/3$ for all of them [22]. No change in T_1 with increased field strength is observed however in the CSF. This phenomenon is related to the efficiency of the spin-lattice relaxation, which is determined by the degree of frequency coupling between the spin precession and the magnetic noise in the lattice due to molecular dynamics [23]. It can be classically described in terms of *molecular tumbling*: if occurring at a rate close to Larmor frequency, it provides the spins an efficient pathway for thermal relaxation, leading to a shorter T_1 compared to spins in a molecular environment where the tumbling rate is considerably higher or lower than Larmor frequency. Free moving water molecules (such as those constituting the CSF) and small molecules are characterised by a wide spectrum of tumbling rates, which results in most water molecules not being resonating at the Larmor frequency: this causes T_1 relaxation in water to be inefficient, hence the long T_1. On the other hand, large molecules with very low mobility present a tight hydration layer whose water molecules exhibit very restricted, low tumbling rates, concentrated mainly below the Larmor frequency, which also results in inefficient spin-lattice relaxation and long T_1. In the middle, water molecules bound to moderately sized macromolecules (e.g. fat, middle sized proteins) are only moderately restricted, presenting a relatively larger spectral component matching Larmor frequency: this leads to an overall more efficient spin-lattice relaxation and shorter T_1. Increasing B_0 results in a higher ω_0, which may reduce the spectral overlap between moderately sized proteins tumbling rates and Larmor frequency, making the spin-lattice relaxation less efficient, increasing T_1 as a function of B_0 [24].

Contrast agents, such as *gadolinium* compounds, can be used to catalyse the longitudinal relaxation process via proton-electron interactions, reducing T_1. Contrast agents have not been used in this project, therefore an in-depth description of how they

[4] $C_{\text{WM}} = 7.1 \cdot 10^{-4}$, $C_{\text{GM}} = 1.2 \cdot 10^{-3}$, $C_{\text{blood}} = 3.4 \cdot 10^{-3}$. WM: white matter, GM: grey matter.

work is left elsewhere; more information can be found in Caravan et al. (1999) which, in addition of providing a clear and detailed review of gadolinium chelates structure, function and applications, is also a source of quite interesting quotes:

> *While it is odd enough to place patients in large superconducting magnets and noisily pulse water protons in their tissues with radio waves, it is odder still to inject into their veins a gram of this potentially toxic metal ion which swiftly floats among the water molecules, tickling them magnetically.* [25]

4.4 Transverse relaxation

The phase coherence of the transverse component of the magnetisation likewise decays over time, leading to the extinction of the induced signal. This process is characterised by the *transverse relaxation time* T_2: it is mediated by spin-spin energy exchange (akin to T_1-relaxation, also referred to as T_1 *contribution to* T_2), spin-spin *flip-flop* dipolar interactions and fluctuations in the local static field due to different molecular configurations causing spins to precess at different frequencies (referred to as *secular contribution to* T_2 [26]).

The magnitude of the transverse magnetisation also decays according to an exponential model:

$$\frac{\mathrm{d}M_\perp}{\mathrm{d}t} = -\frac{M_\perp}{T_2} \tag{4.19}$$

with solution

$$M_\perp(t) = M_\perp(0^+)e^{-t/T_2} = M_z(0^-)\sin(\theta)e^{-t/T_2} \tag{4.20}$$

Similarly to equation (4.17), this can be also generalised for any n-th repetition and $t_n = t - nT_R$ such that:

$$M_\perp(t_n) = M_z(0^-)\sin(\theta)e^{-t_n/T_2}, \quad 0 \leq t_n < T_R, \quad n = 0, 1, 2, \ldots \tag{4.21}$$

Amorphous tissues composed of relatively small and free moving molecules (such as the CSF) do not support magnetic field inhomogeneities and are therefore characterised by long T_2. As the tissue structure increases in complexity, with bigger molecular size and motion becoming more restricted, spin-spin relaxation becomes more efficient due

to the emergence of local magnetic field domains, reducing T_2 [27]. The presence of paramagnetic agents, namely iron, whether endogenous, such as iron ions within deoxyhemoglobin or ferritin, or exogenous in the form of iron-based contrast agents, also produce local magnetic field inhomogeneities which, coupled with the random motion of water molecules in close proximity of iron deposits (see section 4.17), result in faster T_2 decay [28, 29]. According to dipolar relaxation theory, T_2 relaxation is expected to be field-independent, however studies have shown a shortening of T_2 for very high fields (7 T and above) as well, which is believed to be due to susceptibility and microscopic diffusion effects [30, 31]. This T_2 reduction is sequence-dependent and can be mitigated with ultra-short T_E sequences (see next section).

In addition to the random shifts in the magnetic field due to the microscopic environment and water molecules random motion, *static* inhomogeneities may be also imparted, e.g. from defects in the magnet or magnetic field gradients (see section 4.10). This contributes to transverse magnetisation decay with a relaxation time T_2' that results in a faster dephasing due to an overall shorter transverse relaxation time $T_2^* < T_2$ such that

$$\frac{1}{T_2^*} = \frac{1}{T_2'} + \frac{1}{T_2} \tag{4.22}$$

T_2^* effects have been neglected so far, although one can substitute $T_2 \to T_2^*$ to account for them.

4.5 Spin-echo

In the general case where $T_2 \to T_2^*$, extrinsic T_2' effects can be strong enough to cause rapid transverse magnetisation decay. Because of the static nature of these shifts, *spin-echo* sequences can reverse their effects by inverting the spin phase with a $\theta = \pi$ RF *refocusing* pulse.

By expressing local field inhomogeneities at a point $\vec{r}$ as $\Delta B(\vec{r})$, each magnetic moment $\vec{\mu}(\vec{r})$ at that position experiences a local static magnetic field $B(\vec{r}) = B_0 + \Delta B(\vec{r})$; because of this, each nuclear spin precesses at frequency $\omega(\vec{r}) = \gamma B(\vec{r})$. In the frame of reference resonant at the Larmor frequency, the spatial distribution of precession frequencies causes the magnetic moments to accumulate a phase as described in equation (4.9):

$$\phi(\vec{r}, t) = -\gamma \Delta B(\vec{r}) t \tag{4.23}$$

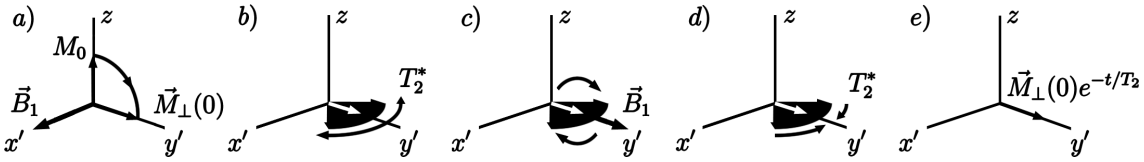


Figure 4.2: Schematics of a typical spin-echo experiment in the $S'[x', y', z]$ resonant reference frame: a) a $\pi/2$ excitation pulse is applied at thermal equilibrium; b) the transverse magnetisation decays due to T_2* relaxation; c) after a time $t = T_E/2$, a π refocusing pulse is applied, inverting the dephasing; d) the transverse magnetisation refocuses, recovering the dephasing due to T_2' relaxation; e) at $t = T_E$, a spin-echo is produced.

If then a π RF pulse is applied at $t = \tau$, the dephasing accumulated just before the pulse $\phi(\vec{r}, \tau^-) = -\gamma \Delta B(\vec{r})\tau$ is flipped to $\phi(\vec{r}, \tau^+) = \gamma \Delta B(\vec{r})\tau$. After a further time τ, at $t = 2\tau$, the phase will thus be

$$\phi(\vec{r}, 2\tau) = \phi(\vec{r}, \tau^+) - \gamma \Delta B(\vec{r})\tau = 0 \tag{4.24}$$

The refocusing induces a peak in the transverse magnetisation magnitude affected only by T_2 decay called *spin-echo*, with magnitude:

$$M_\perp(T_E) = M_\perp(0^+)e^{-T_E/T_2} \tag{4.25}$$

where the time interval $T_E = 2\tau$ is called *echo-time*. An exampled is shown in Figure 4.2.

Spin echo measurements usually employ a $\pi/2$ excitation pulse for signal readout, which allows to express the initial transverse magnetisation, for any given iteration of the sequence after a repetition time T_R, as

$$M_\perp(0^+) = M_z(0^-)\sin(\pi/2) = M_0\Big(1 - e^{-T_R/T_1}\Big) \tag{4.26}$$

where equation (4.16) was used to expand $M_z(0^-)$, and equation (4.25) can be rewritten as

$$M_\perp(T_E, T_R) = M_0\Big(1 - e^{-T_R/T_1}\Big)e^{-T_E/T_2} \tag{4.27}$$

Multiple spin-echoes can be elicited as long as refocusing pulses are applied, each at a T_E interval from the previous one. This results in a train of spin echoes occurring at times multiple of T_E whose peak intensities over time define an envelope function that decays as a function of T_2.

4.6 Bloch equation

By extending equation (4.6) to the entire population of spins in a given volume (i.e. by substituting the magnetisation $\vec{M}$ to the magnetic moment $\vec{\mu}$), and combining it with equations (4.15) and (4.19), it is possible to describe the dynamics of the magnetisation in the presence of an external field $\vec{B}_{\text{ext}}$ through the vector equation:

$$\frac{\mathrm{d}\vec{M}}{\mathrm{d}t} = \gamma\vec{M} \times \vec{B}_{\text{ext}} + \frac{1}{T_1}(M_0 - M_z)\hat{z} - \frac{1}{T_2}\vec{M}_\perp \tag{4.28}$$

referred to as the *Bloch equation*, where $\vec{B}_{\text{ext}} = B_0\hat{z}$ in presence of only the static field, or $\vec{B}_{\text{ext}} = B_0\hat{z} + B_1(\cos(\omega_{\text{RF}}t)\hat{x} - \sin(\omega_{\text{RF}}t)\hat{y})$ when also applying the rotating RF field (with $\omega_{\text{RF}} = \omega_0$ in case of an on-resonance pulse). For the static field case, e.g. after the RF pulse has been applied, the solution for equation (4.28) along the $\hat{z}$ component is the same as equation (4.16), whilst for the transverse component can be expressed in complex notation as

$$\begin{aligned} M_+(t) &= M_\perp(0^+)e^{-i(\omega_0 t - \phi_0)}e^{-t/T_2} \\ &= M_\perp(0^+)(\cos(\omega_0 t - \phi_0) - i\sin(\omega_0 t - \phi_0))e^{-t/T_2} \end{aligned} \tag{4.29}$$

where $M_\perp(0^+)$ is the magnitude of the initial transverse magnetisation, which depends on the flip-angle and the longitudinal magnetisation before the RF pulse (see equation (4.13)).

4.7 Steady-state

After an RF pulse with flip-angle θ, the longitudinal and transverse components of the magnetisation are described by equations (4.13). During the T_R between an RF pulse and the next, the magnetisation relaxes to thermal equilibrium according to the Bloch equation. It will be assumed that transverse magnetisation is completely decayed between repetitions, whether naturally, if $T_R \gg T_2$, or by purposely dephasing it through the application of magnetic field gradients called *spoilers*, which introduce a T_2'-like dephasing effect (see section 4.15 for more details on sequence implementation).

If $\theta = \pi/2$, the system immediately follows a periodic pattern, or *steady state*: using the total time notation $M_z(t) = M_z(t_n + nT_R)$, $\forall n$, this results in the longitudinal magnetisation building-up from $M_z(nT_R^+) = 0$ to a steady state value $M_z((n+1)T_R^-)$

determined by equation (4.17), for every n-th period T_R. If also $T_R \gg T_1$, the system goes periodically from $M_z(nT_R^+) = 0$ to thermal equilibrium, with steady-state value $M_z((n+1)T_R^-) = M_0,\ \forall n$. The steady-state value is then directly transposed into transverse magnetisation by the RF pulse, as defined by equation (4.21).

If however $\theta < \pi/2$ and T_R is smaller or comparable to T_1, the steady-state will be only reached after a transient number of iterations $n \geq N$, specifically when the loss in the longitudinal magnetisation due to the RF pulse θ-tipping is equal to the amount recovered due to spin-lattice relaxation during the T_R period. This is usually the case for fast imaging techniques using small θ and short T_R to speed up the acquisition process. In these cases, it is important to calculate the optimal flip-angle, defined as *Ernst angle* θ_E, resulting in the maximum transverse magnetisation since, as it will be shown in next section, it directly translates into relative maximum signal amplitude.

Equations (4.17) and (4.21) can be re-written in terms of total time $t = t_n + nT_R$ and steady-state values as

$$M_z((n+1)T_R^-) = M_0 - (M_0 - M_z(nT_R^-)\cos(\theta))e^{-T_R/T_1} \tag{4.30}$$

$$M_\perp((n+1)T_R^-) = M_z(nT_R^-)\sin(\theta)e^{-T_R/T_2} \tag{4.31}$$

where, for a naturally spoiled sequence, it is easy to observe from equation (4.31) that $M_\perp((n+1)T_R^-) \to 0$ for $T_R \gg T_2$. If the system is indeed in steady-state, then

$$M_z((n+1)T_R^-) = M_z(nT_R^-) = M_z^{SS}, \quad \forall n \geq N \tag{4.32}$$

and equations (4.30) can be re-written as

$$M_z^{SS} = M_0 - (M_0 - M_z^{SS}\cos(\theta))e^{-T_R/T_1} \tag{4.33}$$

or, equivalently, by solving for M_z^{SS}:

$$M_z^{SS} = M_0 \frac{1 - e^{-T_R/T_1}}{1 - \cos(\theta)e^{-T_R/T_1}} \tag{4.34}$$

The transverse magnetisation, as given by equation (4.21), can now be expressed in terms of M_z^{SS} as a function of θ:

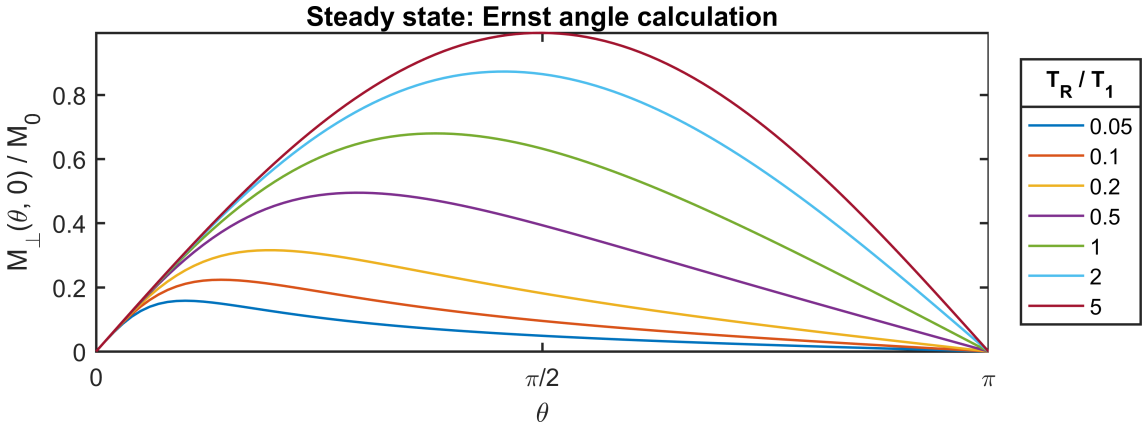


Figure 4.3: Ernst angle for different T_R/T_1. Whilst the absolute maximum of the transverse magnetisation $M_\perp(\theta_E, 0) = M_0$ is obtained for $\theta_E = \pi/2$ and $T_R \gg T_1$, relative steady state magnetisation maxima can be achieved for shorter T_R at lower angles. In this graph, T_2 decay has been ignored by setting $t_n = 0$.

$$\begin{aligned} M_\perp(\theta, t_n) &= M_z^{SS} \sin(\theta) e^{-t_n/T_2} \\ &= M_0 \frac{1 - e^{-T_R/T_1}}{1 - \cos(\theta) e^{-T_R/T_1}} \sin(\theta) e^{-t_n/T_2}, \quad 0 \leq t_n < T_R, \quad \forall n \geq N \end{aligned} \tag{4.35}$$

From equation (4.35) and Figure 4.3, one can easily see that, for $T_R \gg T_1$, the transverse magnetisation is indeed maximised for an Ernst angle $\theta_E = \pi/2$, however, using a $\pi/2$ flip angle with short T_R, results in a sub-optimal steady-state condition. The Ernst angle can be found analytically by calculating the derivative of equation (4.35) with respect to θ, and setting it equal to zero, which results in

$$\theta_E = \arccos\left(e^{-T_R/T_1}\right) \tag{4.36}$$

4.8 Signal detection

The rotation of the transverse magnetisation $\vec{M}_\perp$ precessing in the xy-plane around the direction of $\vec{B}_0$ produces a variation of magnetic flux as described by *Faraday's law* of electromagnetic induction. The same coil used to generate the $\vec{B}_1$ RF pulse is then used to sample the signal induced by the change of magnetic flux in the coil sections that intersect with the xy-plane[5]. The magnetic flux Φ_M produced through the coil

[5]Different coils may also be used for excitation (*transmit* coil) and acquisition (*receive* coil) depending on the specific application and MR-scanner. E.g. in a spinal cord imaging session, a *body* transmit coil, usually integrated in the scanner, may be employed for the excitation, whilst a more targeted receive coil may be used for the signal acquisition.

by the magnetic field $\vec{B}_M$ associated to the magnetisation is given by

$$\Phi_M(t) = \int_S \vec{B}_M(\vec{r}, t) \cdot \mathrm{d}\vec{S} = \int_S (\vec{\nabla} \times \vec{A}_M(\vec{r}, t)) \cdot \mathrm{d}\vec{S} = \oint \mathrm{d}\vec{l} \cdot \vec{A}_M(\vec{r}, t) \quad (4.37)$$

with S being in this context the coil surface and A_M the vector potential associated to the magnetisation. By representing each magnetic moment as a current loop, it is possible to express the vector potential in terms of the effective current density $\vec{J}_M(\vec{r}, t) = \vec{\nabla} \times \vec{M}(\vec{r}, t)$ as

$$\vec{A}_M(\vec{r}, t) = \frac{\mu_0}{4\pi} \int \mathrm{d}^3 r' \frac{\vec{J}_M(\vec{r}\,', t)}{|\vec{r} - \vec{r}\,'|} = \frac{\mu_0}{4\pi} \int \mathrm{d}^3 r' \frac{\vec{\nabla}' \times \vec{M}(\vec{r}\,', t)}{|\vec{r} - \vec{r}\,'|} \quad (4.38)$$

with μ_0 being the magnetic permeability in vacuum. By plugging equation (4.38) in equation (4.37) and manipulating the products between vectors, it can be shown that the flux can be re-written as

$$\Phi_M(t) = \int \mathrm{d}^3 r' \vec{M}(\vec{r}\,', t) \cdot \left[\vec{\nabla}' \times \left(\frac{\mu_0}{4\pi} \oint \frac{\mathrm{d}\vec{l}}{|\vec{r} - \vec{r}\,'|} \right) \right] \quad (4.39)$$

One can notice that the term within round brackets is equivalent to the expression of the vector potential that the receiving coil itself *would* produce if traversed by a unit of current I, evaluated at position $\vec{r}\,'$. Because the curl of a vector potential is a magnetic field $\vec{B}$, it is therefore possible to define the term within square brackets as

$$\vec{B}^{\mathrm{rec}}(\vec{r}\,') = \frac{\vec{B}(\vec{r}\,')}{I} = \vec{\nabla}' \times \left(\frac{\mu_0}{4\pi} \oint \frac{\mathrm{d}\vec{l}}{|\vec{r} - \vec{r}\,'|} \right) \quad (4.40)$$

or *receive* field, i.e. the magnetic field per unit of current that *would* be produced by the coil in any point $\vec{r}\,'$ of non-zero magnetisation. This enables to relate the magnetic flux produced by the magnetic moments through the receive coil to the magnetic flux that *would* be produced by the coil through the magnetic moments: an example of the *principle of reciprocity*.

The signal measured, referred to as *free induction decay* (FID), can thus be expressed as

$$\begin{aligned} \mathrm{FID} &\propto -\frac{\mathrm{d}\Phi_M(t)}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}^3 r \vec{M}(\vec{r}, t) \cdot \vec{B}^{\mathrm{rec}}(\vec{r}) \\ &= -\frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}^3 r [M_x(\vec{r}, t) B_x^{\mathrm{rec}}(\vec{r}) + M_y(\vec{r}, t) B_y^{\mathrm{rec}}(\vec{r}) + M_z(\vec{r}, t) B_z^{\mathrm{rec}}(\vec{r})] \end{aligned} \quad (4.41)$$

The negative time derivative can be taken inside the integral and applied to each

component of the magnetisation independently. The negative time derivative of the longitudinal magnetisation can be expressed through equation (4.16), by making the dependency on position $\vec{r}$ explicit:

$$-\frac{\mathrm{d}}{\mathrm{d}t}M_z(\vec{r},t) = \frac{1}{T_1(\vec{r})}(M_0 - M_z(0))e^{-t/T_1(\vec{r})} \tag{4.42}$$

For the remaining two components, by knowing from equation (4.12) that $M_x = \mathrm{Re}(M_+)$; $M_y = \mathrm{Im}(M_+)$, it is useful to calculate the time derivative of equation (4.29):

$$-\frac{\mathrm{d}}{\mathrm{d}t}M_+(\vec{r},t) = \left(i\omega_0 + \frac{1}{T_2(\vec{r})}\right)M_\perp(\vec{r},0)e^{-i(\omega_0 t - \phi_0(\vec{r}))}e^{-t/T_2(\vec{r})} \tag{4.43}$$

For standard medical applications, $B_0 \sim 1\,\mathrm{T}$ and $T_1, T_2 \sim 10$–$1000\,\mathrm{ms}$, which means $\omega_0 \gg (1/T_1,\ 1/T_2)$ since the Larmor frequency is four orders of magnitude larger than $1/T_1$ and $1/T_2$: when considering the sum of the terms in equation (4.41), it is therefore possible to neglect any term with a multiplicative factor of $1/T_1$ or $1/T_2$ with respect to ω_0, which allows to neglect equation (4.42), and approximate (4.43) to

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}M_+(\vec{r},t) &\simeq i\omega_0 M_\perp(\vec{r},0)e^{-i(\omega_0 t - \phi_0(\vec{r}))}e^{-t/T_2(\vec{r})} \\
&\simeq i\omega_0 M_\perp(\vec{r},0)[\cos(\omega_0 t - \phi_0(\vec{r})) - i\sin(\omega_0 t - \phi_0(\vec{r}))]e^{-t/T_2(\vec{r})} \\
&\simeq \omega_0 M_\perp(\vec{r},0)[\sin(\omega_0 t - \phi_0(\vec{r})) + i\cos(\omega_0 t - \phi_0(\vec{r}))]e^{-t/T_2(\vec{r})}
\end{aligned} \tag{4.44}$$

which leads to

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}M_x(\vec{r},t) = -\mathrm{Re}\left(\frac{\mathrm{d}}{\mathrm{d}t}M_+(\vec{r},t)\right) &\simeq \omega_0 M_\perp(\vec{r},0)\sin(\omega_0 t - \phi_0(\vec{r})))e^{-t/T_2(\vec{r})} \\
-\frac{\mathrm{d}}{\mathrm{d}t}M_y(\vec{r},t) = -\mathrm{Im}\left(\frac{\mathrm{d}}{\mathrm{d}t}M_+(\vec{r},t)\right) &\simeq \omega_0 M_\perp(\vec{r},0)\cos(\omega_0 t - \phi_0(\vec{r})))e^{-t/T_2(\vec{r})}
\end{aligned} \tag{4.45}$$

By substituting equation (4.45) in (4.41) and neglecting the term associated to the longitudinal magnetisation due to the previous considerations, the signal can be expressed as

$$\begin{aligned}
\mathrm{FID} &\propto \omega_0 \int \mathrm{d}^3 r M_\perp(\vec{r},0)e^{-t/T_2(\vec{r})}[B_x^{\mathrm{rec}}(\vec{r})\sin(\omega_0 t - \phi_0(\vec{r})) \\
&\qquad + B_y^{\mathrm{rec}}(\vec{r})\cos(\omega_0 t - \phi_0(\vec{r}))] \\
&\propto \omega_0 \int \mathrm{d}^3 r M_\perp(\vec{r},0)B_\perp^{\mathrm{rec}}(\vec{r})e^{-t/T_2(\vec{r})}\sin(\omega_0 t - \phi_0(\vec{r}) + \theta_B(\vec{r}))
\end{aligned} \tag{4.46}$$

where the generic substitutions $B_x^{\mathrm{rec}} = B_\perp^{\mathrm{rec}}\cos(\theta_B)$, $B_y^{\mathrm{rec}} = B_\perp^{\mathrm{rec}}\sin(\theta_B)$ have been

performed.

The FID is then *demodulated* by decomposing it through an *in-phase* and a *quadrature* channel, which corresponds to multiplying the signal by a $\sin(\omega_0 t)$ and $-\cos(\omega_0 t)$ factor respectively. This process is called *quadrature demodulation* and yields two signal components with a $\pi/2$ phase difference, which can be therefore conveniently represented in complex notation as *real* and *imaginary* channel: s_{Re} and s_{Im}. The demodulated signal expressed in complex notation $s = s_{\mathrm{Re}} + i s_{\mathrm{Im}}$ presents no oscillation due to the precession[6] which is equivalent to the signal as it would be detected from the perspective of an on-resonant rotating reference frame. With reference to equation (4.46) and considering only the oscillating term, the two signal components s_{Re} and s_{Im} can be expressed as

$$\begin{aligned} s_{\mathrm{Re}}(t) &\propto \sin(\omega_0 t)\sin(\omega_0 t + \phi(\vec{r})) \\ &\propto \frac{\cos(\phi(\vec{r})) - \cos(2\omega_0 t + \phi(\vec{r}))}{2} \xrightarrow[\text{pass}]{\text{low}} \frac{\cos(\phi(\vec{r}))}{2} \propto \mathrm{Re}(e^{-i\phi(\vec{r})}) \\ s_{\mathrm{Im}}(t) &\propto -\cos(\omega_0 t)\sin(\omega_0 t + \phi(\vec{r})) \\ &\propto -\frac{\sin(\phi(\vec{r})) + \sin(2\omega_0 t + \phi(\vec{r}))}{2} \xrightarrow[\text{pass}]{\text{low}} -\frac{\sin(\phi(\vec{r}))}{2} \propto \mathrm{Im}(e^{-i\phi(\vec{r})}) \end{aligned} \tag{4.47}$$

where $\phi(\vec{r}) = \theta_B(\vec{r}) - \phi_0(\vec{r})$ is the total accumulated phase, and a *low-pass* filter allows to eliminate the high frequency component, keeping only the static envelope of the signal, assuming perfect demodulation. The demodulated signal can then be expressed as

$$s(t) \propto \omega_0 \int \mathrm{d}^3 r M_\perp(\vec{r}, 0) B_\perp^{\mathrm{rec}}(\vec{r}) e^{-t/T_2(\vec{r})} e^{i(\phi_0(\vec{r}) - \theta_B(\vec{r}))} \tag{4.48}$$

4.9 Signal weighting

Equation (4.48) is particularly interesting because makes the relationship between FID signal, the MR-system and parameters, and sample properties discussed so far, explicit. This can be better observed upon supposing the RF coil to be uniform enough that the initial magnetisation phase ϕ_0, the receive field phase θ_B, and the transverse component of the receive field $B_\perp$ can be considered independent from position[7]. For a standard spin-echo acquisition ($\pi/2$ excitation pulse), the demodulated signal for an echo-time

[6] A residual, low-frequency oscillation in time might still be present due to imperfect demodulation. This may cause signal loss over time which resembles a T_2' effect.

[7] In practice, this assumption is often not met, in which case a *receive bias field* can be defined. It appears as a low-frequency, smooth artifact that makes the image appear more or less bright in the areas where the receive coil is more or less sensitive. This is often corrected for in post-processing [32].

T_E and repetition time T_R can be expressed as

$$\begin{aligned} s(T_E, T_R) &= \omega_0 \Lambda B_\perp \int \mathrm{d}^3 r M_\perp(\vec{r}, 0) e^{-T_E/T_2(\vec{r})} \\ &= \omega_0 \Lambda B_\perp \int \mathrm{d}^3 r M_0(\vec{r}) \Big(1 - e^{-T_R/T_1(\vec{r})}\Big) e^{-T_E/T_2(\vec{r})} \\ &= \gamma \Lambda B_\perp \frac{\mu^2}{3 k_B T} B_0^2 \int \mathrm{d}^3 r \rho_0(\vec{r}) \Big(1 - e^{-T_R/T_1(\vec{r})}\Big) e^{-T_E/T_2(\vec{r})} \end{aligned} \tag{4.49}$$

where equations (4.5), (4.8) and (4.26) have been used, and the proportionality constant Λ has been introduced, which includes the constant phase terms, as well as gain factors from the detection system. The direct proportionality between the signal and the magnitude of the static field B_0^2 explains the need for higher MR-fields to achieve higher signal-to-noise ratios. By introducing the *effective spin density* $\rho(\vec{r})$:

$$\rho(\vec{r}) = \gamma \Lambda B_\perp \frac{\mu^2}{3 k_B T} B_0^2 \rho_0(\vec{r}) \tag{4.50}$$

it is possible to express equation (4.49) as

$$s(T_E, T_R) = \int \mathrm{d}^3 r \rho(\vec{r}) \Big(1 - e^{-T_R/T_1(\vec{r})}\Big) e^{-T_E/T_2(\vec{r})} \tag{4.51}$$

This relation is especially important because it highlights how the modulation of the MR-system parameters T_E and T_R enables the induction of *weighted* signal. Suppose to acquire MR-signals from two samples with different ρ, T_1 and T_2. Short T_R and T_E (compared to T_1 and T_2 respectively) cause the signals to be T_1-weighted: the short T_E causes T_2 decay to be negligible and comparable in the two cases, meaning that the difference between signals is mainly driven by the difference between initial longitudinal magnetisations which, due to the short T_R, have not reached equilibrium yet, and is therefore dictated by the different T_1's (longer T_1 means slower recovery and thus lower signal). On the other hand, long T_R and T_E cause the signals to be T_2-weighted, since the long T_R ensures the longitudinal magnetisation is at equilibrium before subsequent excitation, making the signal T_1-independent, whilst the long T_E causes T_2-decay to be more prominent and affect the two signals (longer T_2 means slower decay and thus higher signal). Finally, long T_R and short T_E mute dependency on both T_1 and T_2, causing the signals to be ρ-weighted and differ based on the actual difference in spin densities of the two samples. For standard ${}^1_1\mathrm{H}$ MRI, the signal would be referred to as *proton density*–, or PD-weighted, although the same concept applies when acquiring ρ-weighted signal from other atomic species, such as in ${}^{23}_{11}\mathrm{Na}$ MRI. Examples of T_1-, T_2-

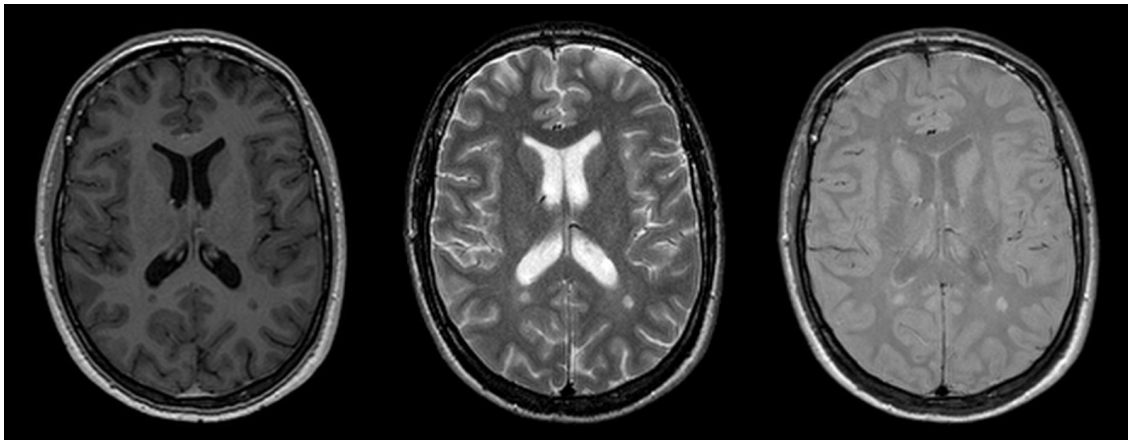


Figure 4.4: From left to right: examples of T_1-, T_2- and PD-weighted images.

and PD-weighted images are shown in Figure 4.4. Alternatively, for fast acquisitions using very short $T_R \ll T_1$, it is still possible to achieve PD-weighting despite the short repetition time by employing a flip-angle much smaller than the Ernst angle $\theta \ll \theta_E$ for the chosen T_R. This can be observed by substituting equation (4.35) in equation (4.48), that is by re-writing the signal for any given flip-angle θ as

$$\begin{aligned} s(\theta) &= \int \mathrm{d}^3 r \rho(\vec{r}) \frac{1 - e^{-T_R/T_1(\vec{r})}}{1 - \cos(\theta) e^{-T_R/T_1(\vec{r})}} \sin(\theta) e^{-T_E/T_2(\vec{r})} \\ &\simeq \int \mathrm{d}^3 r \rho(\vec{r}) \frac{1 - e^{-T_R/T_1(\vec{r})}}{1 - \left(1 - \frac{\theta^2}{2}\right) e^{-T_R/T_1(\vec{r})}} \theta e^{-T_E/T_2(\vec{r})}, \qquad \theta \ll \theta_E \\ &\simeq \int \mathrm{d}^3 r \frac{\rho(\vec{r})\theta}{1 + \frac{T_1}{2T_R}\theta^2} e^{-T_E/T_2(\vec{r})}, \qquad \theta \ll \theta_E, \quad T_R \ll T_1 \end{aligned} \tag{4.52}$$

where the following approximations have been used for $x \ll 1$: $\sin(x) \approx x$, $\cos(x) \approx 1 - x^2/2$, $e^{-x} \approx 1 - x$. Given equation (4.36), θ_E can be approximated as

$$\theta_E \simeq \arccos\left(1 - \frac{T_R}{T_1}\right), \quad T_R \ll T_1 \tag{4.53}$$

from which it is possible to show[8] that $\theta_E \simeq \sqrt{2T_R/T_1}$. Equation (4.52) thus becomes

$$s(\theta) \simeq \int \mathrm{d}^3 r \frac{\rho(\vec{r})\theta}{1 + \left(\frac{\theta}{\theta_E}\right)^2} e^{-T_E/T_2(\vec{r})}, \quad \theta \ll \theta_E, \quad T_R \ll T_1 \tag{4.54}$$

losing all dependency on T_1, showing that, by employing a small θ and short T_E, it is possible to obtain PD-weighted signal even from a short T_R sequence, albeit at lower

[8] $\cos(x) \approx 1 - x^2/2 \quad \Rightarrow \quad \cos(\sqrt{2y}) \approx 1 - y \quad \Rightarrow \quad \arccos(1 - y) \approx \sqrt{2y}$

signal intensity.

4.10 Fourier imaging

The expression for the signal reported in equation (4.48) is given by the sum of the spin contributions within the sample. Due to the molecular environments they are in, identical spins can experience different local magnetic fields, hence precessing at different frequencies. The difference in frequencies caused by the microscopical properties of the spin neighbourhood is called *chemical shift*, and can be exploited in NMR measurements to probe the molecular composition of the sample. The signal can be decomposed in its constituent frequencies, or *spectral components*, through a Fourier-transform, thus resulting in the signal *spectrum*:

$$\hat{s}(\omega) = \mathscr{F}[s(t)](\omega) = \int \mathrm{d}t\ s(t) e^{i\omega t} \tag{4.55}$$

In addition to acquiring the signal spectrum, the objective of MRI is also to reconstruct the 3D spatial distribution of the spin populations that have contributed to its generation. The reconstructed 3D image is constituted by volumetric pixel units called *voxels*, whose dimensions depend on the MR-sequence parameters used for the acquisition and define the image resolution. To produce such images, three spatially constant magnetic field *gradients* are applied as part of the MR-protocol. The direction of the magnetic fields is parallel to $\vec{B}_0 = B_0\hat{z}$ and their magnitude varies linearly along the x-, y- and z-axes, respectively. The gradient vector

$$\vec{G}(t) = G_x(t)\hat{x} + G_y(t)\hat{y} + G_z(t)\hat{z} \tag{4.56}$$

may be defined. Each gradient component i ideally follows a *boxcar* profile, that is $G_i(t) = G_i$ when the gradient is on, and $G_i(t) = 0$ otherwise, with $i = x, y, z$, although a *trapezoidal* time-profile is also used as a more realistic approximation. The total magnetic field along the z-axis can then be written as

$$B_z(\vec{r}, t) = B_0 + \vec{G}(t) \cdot \vec{r} \tag{4.57}$$

The gradients introduce a spatial dependency to the precessional frequency:

$$\omega(\vec{r}, t) = \gamma B_z = \omega_0 + \gamma \vec{G}(t) \cdot \vec{r} = \omega_0 + \omega_G(\vec{r}, t) \tag{4.58}$$

as well as the phase ϕ_G accumulated by the magnetic moments in the on-resonant reference frame:

$$\begin{aligned} \frac{\mathrm{d}\phi_G}{\mathrm{d}t}\hat{z} = \vec{\omega}_G = -\omega_G\hat{z} \\ \phi_G(\vec{r},t) = -\int_0^t \omega_G(\vec{r},t')\mathrm{d}t' = -\gamma\vec{r}\cdot\int_0^t \vec{G}(t')\mathrm{d}t' \end{aligned} \tag{4.59}$$

With reference to equation (4.48), and following same steps as for equations (4.49)–(4.51), under the assumption of RF coil uniformity (which is acceptable at the voxel-scale), the expression for the demodulated signal in the presence of the magnetic field gradients is given by

$$s(t) = \int \mathrm{d}^3 r\rho(\vec{r})e^{i\phi_G(\vec{r},t)} \tag{4.60}$$

where the relaxation effects have been omitted, or the associated factor equivalently incorporated into the effective spin density ρ.

A key step for MRI is to introduce the concept of k-space, where $\vec{k} = (k_x, k_y, k_z)$ is a vector defined as

$$\vec{k} = \frac{\gamma}{2\pi}\int_0^t \vec{G}(t')\mathrm{d}t' = \gamma\!\!\!^{-}\int_0^t \vec{G}(t')\mathrm{d}t' \tag{4.61}$$

where $\gamma\!\!\!^{-} = \gamma/2\pi$ and the time dependency is considered implicit in $\vec{k} = \vec{k}(t)$. Equation (4.59) can then be re-written as

$$\phi_G(\vec{r},t) = -2\pi\vec{k}\cdot\vec{r} \tag{4.62}$$

By re-writing equation (4.60) as a function of $\vec{k}$, one can express it in terms of the 3D Fourier transform of the effective spin density ρ:

$$s(\vec{k}) = \int \mathrm{d}^3 r\rho(\vec{r})e^{-i2\pi\vec{k}\cdot\vec{r}} = \mathscr{F}[\rho(\vec{r})](\vec{k}) \tag{4.63}$$

with $\vec{k}$ acting as the signal constituent *spatial frequency*. By applying an inverse Fourier transform to the left and right hand sides of equation (4.63), it is therefore possible to reconstruct the spatial distribution of the effective spin density from the signal acquired over time:

$$\rho(\vec{r}) = \mathscr{F}^{-1}[s(\vec{k})](\vec{r}) \tag{4.64}$$

4.11 Slice selection

Each spatial gradient introduces a dependency on one spatial direction, hence the need for three orthogonal magnetic field gradients. In 2D MR-protocols, the gradient along the z-axis is used to select the z-coordinate of the xy-plane to be imaged. Whilst being applied, the *slice select* gradient G_z^{SS} causes the spins located in the slice with coordinate z to precess at a frequency

$$\omega(z) = \omega_0 + \gamma G_z^{\text{SS}} z \tag{4.65}$$

If an RF pulse rotating at a frequency $\omega_{\text{RF}} = \omega(\bar{z})$ is then applied simultaneously with the gradient, only the spins in the slice at $z = \bar{z}$ will be excited due to the *spatially selective* on-resonant condition. In this section, it is convenient to express precession frequencies in Hz units ν, rather than rad/s, such that $\omega = 2\pi\nu$.

So far the RF pulse has been assumed to be resonant with only one specific frequency at a time. However, given equation (4.65), this corresponds to an infinitely thin slice, which is obviously not practical. In reality, the RF pulse is designed to target a range of frequencies $\Delta\omega = 2\pi\Delta\nu$, which in turn determines the thickness of the excited slice:

$$\Delta z = \frac{\Delta\omega}{\gamma G_z^{\text{SS}}} = \frac{\Delta\nu}{\cancel{\gamma} G_z^{\text{SS}}} \tag{4.66}$$

Slice thickness depends therefore on the *bandwidth* of the RF pulse $\text{BW}_{\text{RF}} = \Delta\nu$, that is the range of frequencies in Hz units resonant with the pulse, and is given analytically by the Fourier transform of the temporal profile of the RF pulse $B_1(t)$:

$$\text{BW}_{\text{RF}}(\nu) = \mathscr{F}[B_1(t)](\nu) \tag{4.67}$$

In the limit case where $\omega_{\text{RF}} = \omega$ (where the spatial dependency is implicit in $\omega = \omega(z)$), the temporal profile of the RF pulse can be expressed using complex notation as

$$B_1(t) \propto e^{i\omega_{\text{RF}} t} = e^{i2\pi\nu_{\text{RF}} t} \tag{4.68}$$

which in the frequency domain, from equation (4.67), corresponds to a bandwidth proportional to a Dirac delta function $\text{BW}_{\text{RF}}(\nu) = \delta(\nu - \nu_{\text{RF}})$. In reality, to excite a range of frequencies $\Delta\nu$ around the central frequency ν_{RF}, the magnitude of the RF pulse is modulated over time according to a $\text{sinc}(\pi\Delta\nu t)$ function. This can be derived by considering a sum of RF pulses as expressed in equation (4.68), each resonating at

a frequency $\nu \in N$, with $N = [\nu_{RF} - \Delta\nu/2, \nu_{RF} + \Delta\nu/2]$ or, equivalently:

$$\begin{aligned} B_1(t) &\propto \int_{\nu \in N} \mathrm{d}\nu e^{i2\pi\nu t} \\ &\propto \frac{e^{i2\pi\nu t}}{2it}\Big|_{\nu_{RF}-\frac{\Delta\nu}{2}}^{\nu_{RF}+\frac{\Delta\nu}{2}} \\ &\propto \frac{e^{i\pi\Delta\nu t} - e^{-i\pi\Delta\nu t}}{2i\pi t} e^{i2\pi\nu_{\mathrm{RF}} t} \\ &\propto \frac{\sin(\pi\Delta\nu t)}{\pi t} e^{i2\pi\nu_{\mathrm{RF}} t} \\ &\propto \mathrm{sinc}(\pi\Delta\nu t) e^{i2\pi\nu_{\mathrm{RF}} t} \end{aligned} \tag{4.69}$$

In the frequency domain, this is equivalent to a rectangular function, or *boxcar* $\Pi((\nu - \nu_{\mathrm{RF}})/\Delta\nu)$ centred on $\nu = \nu_{\mathrm{RF}}$ and of bandwidth $\Delta\nu$. This can be shown in terms of Fourier transform by observing the first line in equation (4.69) where, knowing that $\Pi(\nu) = 1$ for $\nu \in N$, and $\Pi(\nu) = 0$ otherwise, it is possible to notice that:

$$\begin{aligned} B_1(t) &\propto \int_{\nu \in N} \mathrm{d}\nu e^{i2\pi\nu t} \\ &\propto \int_{-\infty}^{\infty} \mathrm{d}\nu \Pi\Big(\frac{\nu - \nu_{\mathrm{RF}}}{\Delta\nu}\Big) e^{i2\pi\nu t} \\ &\propto \mathscr{F}^{-1}\Big[\Pi\Big(\frac{\nu - \nu_{\mathrm{RF}}}{\Delta\nu}\Big)\Big](t) \end{aligned} \tag{4.70}$$

Figure 4.5 shows an example of B_1 profile and corresponding slice selection. Analytically, B_1 is a continuously defined function over time which goes to zero only at $|t| = \infty$, which is clearly not practical. As a result, the boxcar excitation frequency profile, and the sharp slice thickness are also ideal limits. In practice, the RF pulse is time-truncated to include only a certain number of zero-crossings, with this number being key in the RF pulse design, as it determines how close the actual excitation profile and slice boundaries will be to the ideal ones.

4.12 Rephasing and gradient echo

The application of the slice select gradient also introduces an additional dephasing term similar to the T_2' effects discussed in section 4.4. Given the RF pulse is applied simultaneously to the slice select gradient, as soon as the magnetic moments are tipped away from their equilibrium position and a transverse magnetisation starts to build up, the different precessional frequencies across the slice thickness given by equation (4.65)

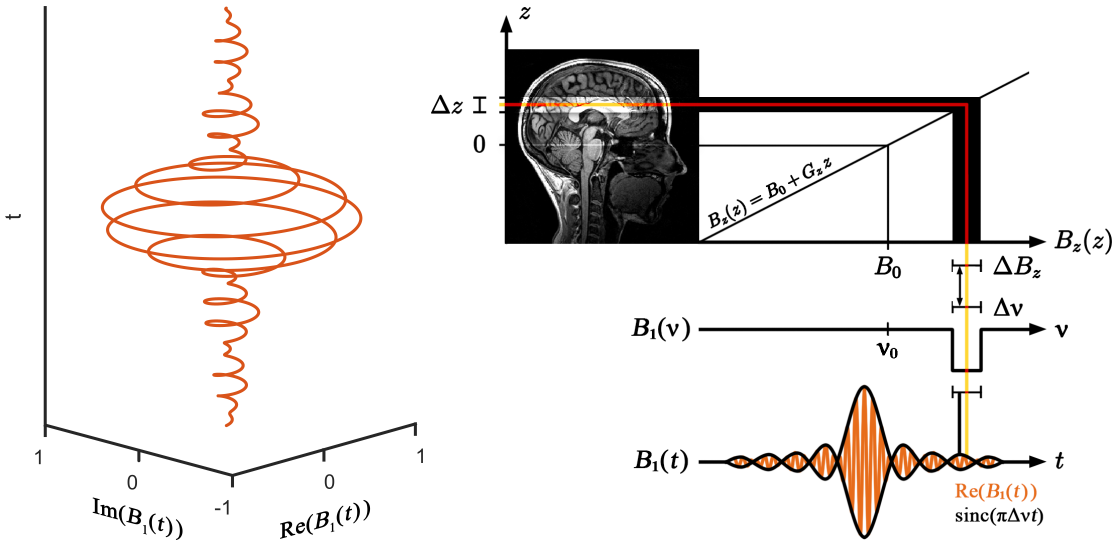


Figure 4.5: The graph on the left shows an example of the $B_1(t)$ profile in complex form. The scheme on the right represents the slice selection process: the frequency of the sinc envelope defines the bandwidth of the pulse, and thus the thickness of the imaged slice, whilst the frequency of the complex phase determines its centre.

cause a dephasing of the spins proportional to the area under the gradient profile as given by equation (4.59). For a static gradient, this is equivalent to

$$\phi_{G_z}(z,t) = -\gamma G_z^{SS} z t \tag{4.71}$$

with consequent signal loss as described by equation (4.60). For this reason, a second *rephasing* gradient lobe with same amplitude but opposite sign is usually applied, after the RF pulse is terminated, to offset the dephasing due to the first gradient. The rephasing induces a new peak in the transverse magnetisation called *gradient echo*, with T_E echo time. Every time the transverse magnetisation, and the signal with it, decays due to the application of a magnetic field gradient, the dephasing can be recovered by adding a rephasing gradient along the same direction, thus producing a gradient echo. Since the only dephasing recovered is the one imposed by the gradient in the first place, the amplitude of the gradient echo will be affected by T_2^* decay; otherwise, the general principle is equivalent to what described in section 4.5.

The duration of the rephasing lobe (and, equivalently, the gradient echo time T_E) depends on the flip-angle of the RF-pulse: for small flip-angles, the spin tipping can be approximated to take place instantaneously at the mid-point of the RF pulse ($t = 0$), with the dephasing being caused only by the second half of the slice select gradient. Assuming pulse and slice select gradient to be applied simultaneously over a τ_z time interval $-\tau_z/2 <$

$t < \tau_z/2$, the dephasing is therefore assumed to occur over $0 < t < \tau_z/2$, such that $\phi_{G_z}(z, \tau_z/2) = -\gamma G_z^{\text{SS}} z \tau_z/2$. By applying a rephasing lobe with amplitude $G_z^{\text{RP}} = -G_z^{\text{SS}}$ over $\tau_z/2 < t < \tau_z$, after the RF pulse has terminated, the total dephasing is recovered:

$$\phi_{G_z}(z, \tau_z) = -\gamma G_z^{\text{SS}} z \frac{\tau_z}{2} - \gamma G_z^{\text{RP}} z \frac{\tau_z}{2} = 0 \tag{4.72}$$

For small flip-angles, the rephasing lobe duration is thus 50% the slice select, and $T_E = \tau_z$; larger flip-angles instead induce a wider dephasing, requiring a longer rephasing: for $\pi/2$ flip-angle, the ideal rephasing time can be shown to be 50.6% of the slice select gradient duration. For slice select and rephasing gradients with different amplitudes, the *areas* under the gradient profiles need to be matched, rather than just the gradients duration.

4.13 Phase encoding, frequency encoding and k-space coverage

Once the rephasing is complete, the slice is ready to be imaged. The signal produced by the spins in the slice is given by equation (4.63), with $k_z = 0$ due to G_z being off, and the integration over z being limited to the thickness of the slice Δz:

$$s(k_x, k_y) = \int \mathrm{d}x \int \mathrm{d}y \int_{z_0 - \frac{\Delta z}{2}}^{z_0 + \frac{\Delta z}{2}} \mathrm{d}z \rho(x, y, z) e^{-i2\pi(k_x x + k_y y)} \tag{4.73}$$

In order to reconstruct the spatial distribution of the effective spin density in the slice, it is therefore necessary to sample the signal at different k_x and k_y throughout the k-space. The centre of the k-space $(k_x, k_y) = (0, 0)$ is defined by the absence of gradients and its neighbourhood contains *low-frequency* spatial information, which constitutes the bulk of the image, as it determines the overall image contrast and shapes; *higher-frequency* spatial information, that is sharper edges and details, can be accessed by sampling (k_x, k_y) points farther away from the centre of the k-space. Different (k_x, k_y) points can be sampled by applying *frequency encoding* G_x^{FE} and *phase encoding* G_y^{PE} gradients, respectively.

The phase encoding gradient G_y^{PE} is applied over a time τ_y, inducing a spatial dependency between the phase accumulated by the magnetic moments and their y-coordinate, similarly to what is expressed by equation (4.71). This is equivalent to a shift in the k-space over the phase encoding direction equal to

$$\delta k_y = \not{\gamma} G_y^{\text{PE}} \tau_y \tag{4.74}$$

The frequency encoding gradient G_x^{FE} is applied in two lobes with opposite sign to elicit a gradient echo, as described in 4.12. For a fully sampled k-space, the negative (dephasing) lobe has $1/2$ the area of the positive (rephasing) lobe, and is applied simultaneously, or soon after the phase encoding. In the simple case of both lobes having the same amplitude, the dephasing lobe has a duration $\tau_x/2$, whilst the rephasing lobe follows over a time interval τ_x, eliciting a gradient echo in the mid-point. The signal induced by the gradient echo is then sampled during the rephasing lobe at intervals $\Delta\tau_x$, with the frequency encoding gradient effectively introducing a spatial dependency between the spins local precessional frequencies and their x-coordinate. The dephasing and rephasing lobes correspond to shifts in the frequency encoding direction δk_x^- and δk_x^+ respectively, given by

$$\begin{aligned} \delta k_x^- &= -\not{\gamma} G_x^{\text{FE}} \frac{\tau_x}{2} = -K_x \\ \delta k_x^+ &= \not{\gamma} G_x^{\text{FE}} \tau_x = 2K_x \end{aligned} \tag{4.75}$$

Each $(\delta k_x, \delta k_y)$ pair delineates a 2D shift in the k-space from one point to another $(k_x, k_y) \to (k_x + \delta k_x, k_y + \delta k_y)$: the path connecting these points is called *trajectory*. Starting at the centre of the k-space, the $(\delta k_x^-, \delta k_y)$ shift leads to $(0, 0) \to (-K_x, \delta k_y)$; the rephasing gradient moves then the trajectory by δk_x^+ along the frequency encoding direction, that is $(-K_x, \delta k_y) \to (K_x, \delta k_y)$, covering the entire k-line $k_x \in [-K_x, K_x]$ and phase encoding coordinate $k_y = \delta k_y$. By changing the k_y-coordinate, that is by sampling the signal over different k-lines, it is possible to span the trajectory across the entire k-space. In a standard 2D gradient echo sequence, schematised in Figure 4.6, this is done by repeating the sequence described above, with the time between each RF pulse and the next one being the repetition time T_R. At each new iteration, the value of G_y^{PE} is changed by a multiple of ΔG_y^{PE}, or *step*, through the range $[-G_{y,\text{max}}^{\text{PE}}, G_{y,\text{max}}^{\text{PE}}]$, where ${G_{y,\text{max}}}^{\text{PE}}$ is the maximum amplitude of the phase encoding gradient. This defines a step in the k_y direction $\Delta k_y = \not{\gamma} \Delta G_y^{\text{PE}} \tau_y$, which allows to span over the phase encoding range $k_y \in [-K_y, K_y]$ at discrete Δk_y intervals, with $K_y = \not{\gamma} G_{y,\text{max}}^{\text{PE}} \tau_y$.

Depending on the stepping order, the phase encoding range can be covered according to different *ordering schemes*. In the *sequential ordering* scheme, the k_y-range is covered linearly from $-K_y$ to K_y stepping by Δk_y at each iteration. Alternatively, a *centric reordering* scheme can be used, with the k-space trajectory starting from $k_y = 0$ and then moving back and forth with increasing discrete shifts $0 \to \Delta k_y \to -\Delta k_y \to 2\Delta k_y \to$

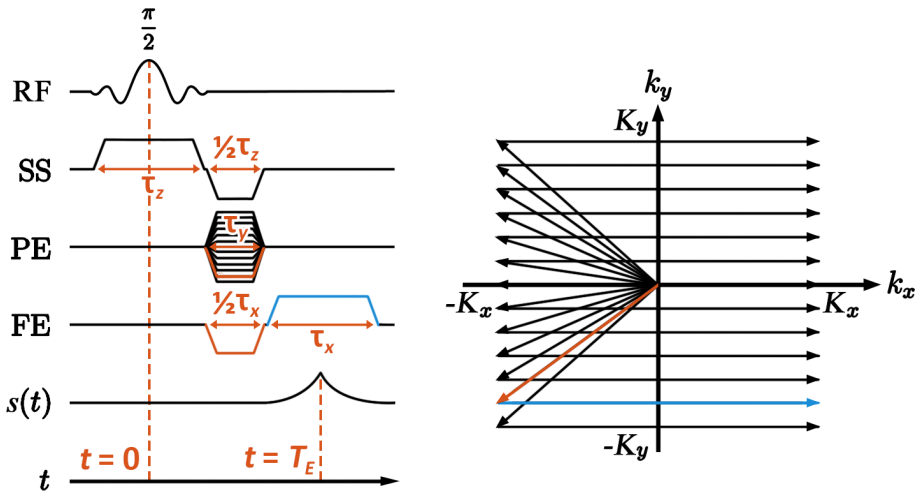


Figure 4.6: On the left, standard 2D gradient echo sequence: notice the rephasing lobe of the slice-selection gradient, and the stepping of the phase-encoding gradient. On the right, the corresponding k-space coverage, with each new k-line being acquired at every new iteration of the sequence.

$-2\Delta k_y \rightarrow \ldots \rightarrow K_y \rightarrow -K_y$. A variation of the centric reordering, the *reverse-centric reordering* scheme is equivalent to the former, but in the opposite direction, that is from the periphery of the k-space to the centre. Depending on the MR-sequence, the ordering scheme might have an effect in terms of how the sequence parameters, such as the T_E, affect the signal, and should therefore be taken into consideration during data processing.

In three spatial dimensions, i.e. a 3D acquisition, the gradient encoding is the same, with the addition of a phase encoding gradient G_z^{PE} over the z direction. The slice-select gradient G_z^{SS} is applied to excite a thicker slice, called *slab*, divided into *partitions*. Each partition is reconstructed through the application of a phase encoding gradient G_z^{PE}, whose amplitude is stepped at each iteration to span through all the partitions.

4.14 Field of view, Nyquist sampling criterion and resolution

The MR-signal is collected over a *matrix* of uniformly distributed points in the k-space, whose granularity is determined by the sampling time interval $\Delta\tau_x$ and the stepping of the phase encoding gradient ΔG_y^{PE}. The matrix step dimensions $(\Delta k_x, \Delta k_y)$ over the frequency and phase encoding directions are defined as

$$\begin{aligned}\Delta k_x &= \gamma G_x^{\mathrm{FE}} \Delta\tau_x \\ \Delta k_y &= \gamma \Delta G_y^{\mathrm{PE}} \tau_y\end{aligned} \tag{4.76}$$

with the number of steps being referred to as *matrix size*. Considering only the discrete sampling over the frequency encoding direction, with a matrix size of $2n$, the *measured* signal s_m can be expressed as the product of the signal s (equation (4.63)) and a *combing function* given by the sum of $2n$ Dirac delta functions separated by Δk_x intervals, with a scaling factor Δk_x:

$$\begin{aligned}s_m(k_x) &= s(k_x) \cdot \left[\Delta k_x \sum_{p=-n}^{n-1} \delta(k_x - p\Delta k_x)\right] \\ &= \Delta k_x \sum_{p=-n}^{n-1} s(p\Delta k_x)\delta(k_x - p\Delta k_x)\end{aligned} \tag{4.77}$$

The *reconstructed* spin density $\hat{\rho}$ is given by the inverse-Fourier transform of equation (4.77):

$$\begin{aligned}\hat{\rho}(x) &= \int \mathrm{d}k_x s_m(k_x) e^{i2\pi k_x x} \\ &= \Delta k_x \int \mathrm{d}k_x \sum_{p=-n}^{n-1} s(p\Delta k_x)\delta(k_x - p\Delta k_x) e^{i2\pi k_x x} \\ &= \Delta k_x \sum_{p=-n}^{n-1} s(p\Delta k_x) \int \mathrm{d}k_x \delta(k_x - p\Delta k_x) e^{i2\pi k_x x} \\ &= \Delta k_x \sum_{p=-n}^{n-1} s(p\Delta k_x) e^{i2\pi p\Delta k_x x}\end{aligned} \tag{4.78}$$

From this, it is already possible to see that the reconstructed spin density is periodic, i.e. it is translationally invariant when shifted by a period $1/\Delta k_x$:

$$\begin{aligned}\hat{\rho}\Big(x + \frac{1}{\Delta k_x}\Big) &= \Delta k_x \sum_{p=-n}^{n-1} s(p\Delta k_x) e^{i2\pi p\Delta k_x (x + 1/\Delta x)} \\ &= \Delta k_x \sum_{p=-n}^{n-1} s(p\Delta k_x) e^{i2\pi p\Delta k_x x} e^{i2\pi p} \\ &= \hat{\rho}(x)\end{aligned} \tag{4.79}$$

where $e^{i2\pi p} = 1$, due to $p \in [-n, n-1]$ being integer. The reconstructed density function is therefore constituted of $2n$ identical copies of $\hat{\rho}$ repeating over the x dimension

with uniform spacing $L_x = 1/\Delta k_x$. In order to avoid overlapping of neighbouring copies of the reconstructed density, known as *ghosting* artifact, the dimension of the imaged sample A_x must be smaller than L_x, such that $\Delta k_x < 1/A_x$. This condition is referred to as the *Nyquist sampling criterion*, and is used to calculate the upper limit for the sampling time interval $\Delta\tau_x$ above which ghosting appears.

The same process can be applied to the phase encoding direction, leading to a spatial interval $L_y = 1/\Delta k_y$. In this case, the Nyquist sampling criterion is used to calculate the upper limit for the phase encoding stepping ΔG_y^{PE}. The image *field of view* (FOV) can thus be defined as

$$\text{FOV} = [L_x, L_y] = \left[\frac{1}{\Delta k_x}, \frac{1}{\Delta k_y}\right] \tag{4.80}$$

The truncation of the measured signal in equation (4.77) to a finite sum over $2n$ points is the result of the finite dimensions of the sampled k-space. By recalling that signal sampled at farther points of the k-space contains higher spatial frequency information, it derives that this approximation limits the amount of detail that can be reconstructed from the signal, which in turn imposes a lower limit to the smallest resolvable distance between objects, or *resolution*, causing the emergence of a *blurring* artifact. As a result, the spin density also needs to be discretised over a matrix in the physical space, thus the division of MR-images into voxels, whose dimensions is determined by the image resolution.

For a 2D protocol, the voxel z dimension is equal to the slice thickness Δz, whilst the in-plane resolution $(\Delta x, \Delta y)$ is determined by the physical matrix size. It can be shown that the matrix size for the physical space is the same as the k-space matrix size, that is the number of voxels within the image FOV over the x and y directions is equal to the number of samples over the frequency and phase encoding directions respectively. Therefore, for a matrix size $(2n_x, 2n_y)$, the in-plane resolution is given by

$$\Delta x = \frac{L_x}{2n_x} = \frac{1}{2n_x \Delta k_x};\ \Delta y = \frac{L_y}{2n_y} = \frac{1}{2n_y \Delta k_y} \tag{4.81}$$

where it can be clearly seen that in the limit of an infinite matrix size, the resolvable distance would be infinitely small, and the resolution infinitely high. It is thus possible to define the spin density within a voxel $\hat{\rho}_{\text{MRI}}$, as displayed in an MR-image, as the reconstructed spin density at that physical position, multiplied by the volume of the voxel:

$$\hat{\rho}_{\text{MRI}}(\vec{r}) = \hat{\rho}(\vec{r})\Delta x \Delta y \Delta z \tag{4.82}$$

4.15 MR-sequences

An MR-sequence is a programmed set of RF pulses and magnetic field gradients defined by specific parameters, such as flip-angle, T_E and T_R, resulting in MR-images highlighting specific properties of the sample. To follow, the description of some standard MR-sequences, with a focus on those used in for this project.

4.15.1 Spin echo

The physical process at the base of a spin echo has been discussed in section 4.5. In a standard 2D spin echo sequence, shown in Figure 4.7, a $\pi/2$ excitation pulse is applied at $t = 0$ simultaneously with a slice select gradient G_z^{SS}, followed by a rephasing lobe G_z^{RP} as described in sections 4.11 and 4.12. A phase encoding gradient G_y^{PE} with stepped amplitude every T_R, and a dephasing lobe along the frequency encoding direction G_x^{FE} are then applied as described in section 4.13. The dephasing gradient is however applied with a positive amplitude, instead of a negative one. At $t = T_E/2$, the refocusing π RF pulse is applied together with a slice select gradient. In this case, no rephasing lobe G_z^{RP} is necessary as the dephasing accumulated during the first half of the π RF pulse due to the gradient is automatically recovered during the second half due to the π pulse symmetry. Since the role of the refocusing pulse is to invert the spins phase, this also inverts the phase accumulated due to the dephasing G_x^{FE} gradient, effectively making it equivalent to a dephasing lobe with negative amplitude (which is why it was initially set as positive). At $t = T_E$, the spin echo occurs and it is sampled simultaneously with the application of the rephasing lobe of the frequency encoding gradient. The rephasing gradients (G_z^{RP} and G_x^{FE}) offset the dephasing accumulated due to the initially applied gradients, whilst the π RF pulse recovers the dephasing due to T_2' effects over the T_E. The signal amplitude at the peak is therefore only affected by T_2 decay and, for given T_E and T_R sequence parameters, can be determined according to equation (4.51).

4.15.2 Multi-echo spin echo

As described in section 4.5, given a spin echo sequence, it is possible to keep eliciting spin echoes after the first one by applying multiple π refocusing pulses after the first, each accompanied by a slice select gradient. The k-space is covered independently for each new spin-echo. Since every echo results in an independent image, the phase encoding ordering is not important as long as the entire k-space is covered. Once the sequence

is complete and the signal Fourier-transformed, each echo will result in an MR-image (or a *volume* of the same image) with different T_2-weighting due to the increasing T_E. The volumes can then be fitted voxel-wise to extrapolate a map of quantitative T_2.

In the simplest case, two refocusing pulses are applied, which result in two spin echoes, the first at a *short* T_E, and the second at a *long* T_E. This is often referred to as *dual echo* spin echo. Coupled with a *long* T_R, one can notice from equation (4.51) that this allows to generate PD- and T_2-weighted images respectively within the same sequence.

4.15.3 Fast/turbo spin echo

Instead of generating a separate image for each echo, multiple echoes can be employed to speed up the acquisition process, as shown in Figure 4.7: by sampling each echo over a different line of the k-space, multiple k-lines can be covered during the same T_R. This sequence is an application of RARE (*Rapid Acquisition with Relaxation Enhancement*), but is also often referred to as *fast spin echo* or *turbo spin echo* (TSE) depending on the MRI scanner manufacturer (since Philips data has been used in this project, it will be referred to as TSE). The number of echoes, and thus k-lines, acquired in the same T_R is called *echo train length* or *turbo factor*. Dual echo TSE has been for decades the workhorse of clinical MRI, only recently starting to be replaced by its 3D counterpart, as it allows the generation of two images with clinically relevant weighting, and relatively high resolution, using a single fast sequence.

TSE follows the same gradient scheme described for the multi-echo spin echo, with the difference that instead of a single phase encoding gradient being applied at the beginning of every iteration, phase encoding lobes with equal amplitude but opposite sign are applied before and after each spin echo, stepping the value of the amplitude at each new echo. This allows to move the k-space trajectory to a new line during each new echo, and reset it to $k_y = 0$ between echoes. Unlike multi-echo spin echo, the multiple spin echoes are combined to generate a single image, and the phase encoding ordering scheme must be taken into consideration when modelling the signal. In particular, if the refocusing pulses are applied every T_E, each echo will be generated at a different *effective* echo time: $T_E^1 = T_E$, $T_E^2 = 2T_E$, $T_E^3 = 3T_E$ and so on, with T_E^n referring to the echo time of the n-th spin echo. As a result, the final MR-image will not be described by a single T_E value, and the T_E associated with the sequence will be representative only of the first spin echo. For this reason, an *effective echo time* T_E^{eff} is defined, corresponding to the time between the excitation pulse and the occurrence of the signal peak of the spin echo sampled along

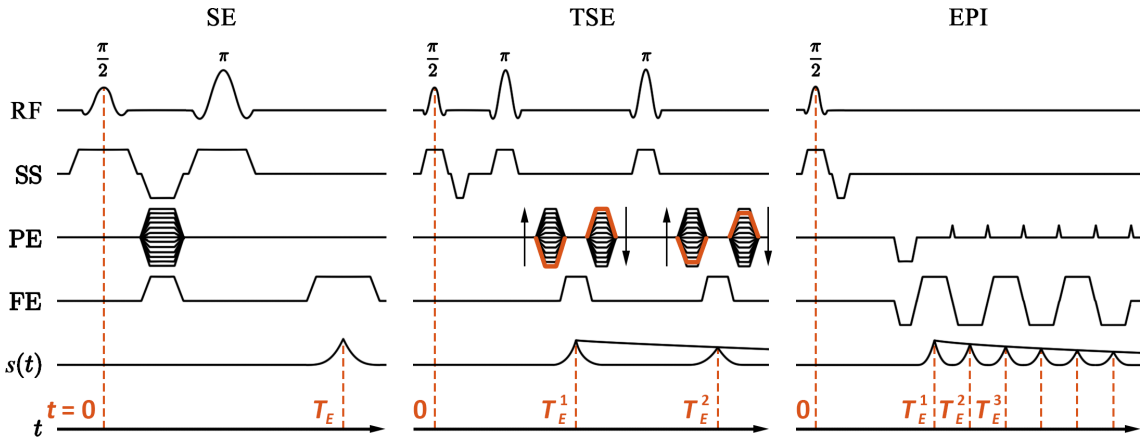


Figure 4.7: Diagrams for three standard 2D MR-sequences: spin-echo (SE), turbo spin echo (TSE) and echo planar imaging (EPI). TSE and EPI enable to acquire multiple k-lines in the same T_R: the paired phase encoding gradients in a TSE allow to sample a new k-line for each echo, whilst returning the k-space trajectory to $k_y = 0$ between echoes; EPI instead uses short phase encoding gradients, or *blips*, to switch to a new k-line, following a snake-like pattern. The envelope of TSE peaks decays according to T_2, whilst for EPI it follows T_2^* decay. In both cases, the resulting image is not defined by a single T_E and an *effective* echo time must be defined.

the central k-space line, that is $k_y = 0$. Since the peak of the echo is also sampled in the middle of the k-line, that is $k_x = 0$, this allows to associate the T_E^{eff} to the signal sampled at the centre of the k-space $s(k_x = 0, k_y = 0)$. By recalling that the central area of the k-space is the one containing the bulk of the image information in terms of overall contrast and shapes, it becomes clear why T_E^{eff} is used as the effective echo time for the entire image. In case of centric reordering scheme, $T_E^{\text{eff}} = T_E$ since the first spin echo is also sampled at $k_y = 0$, but this may not be true for other phase encoding schemes.

4.15.4 Gradient echo

The standard 2D gradient echo sequence (GRE, or *fast field echo*, FFE for Philips manufacturer) has been already described in section 4.13. It can be greatly sped-up through the addition of *spoiler* gradients along all directions: these are applied at the beginning of every iteration specifically to dephase, or *spoil*, the remnant transverse magnetisation between repetitions. As described in section 4.7, this sequence, referred to as spoiled *fast low-angle shot* (FLASH), allows to employ much shorter T_R's and smaller flip angles, resulting in an overall shorter acquisition time. FLASH sequences can be both 2D and 3D, and can achieve PD-, T_2- and T_1-weighting simply by varying T_R, T_E and flip angle.

4.15.5 Echo planar imaging

Similarly to TSE for spin echo, *echo planar imaging* readout (EPI) uses multiple gradient echoes to sample multiple k-lines within the same T_R, as shown in Figure 4.7. This is done by applying a rapid train of G_x^{FE} gradients in the frequency encoding direction, all with same amplitude and alternate sign. Each gradient rephases the magnetisation during the first half, eliciting a gradient echo at the mid-point, and dephases it during the second half, preparing it for the next echo. Positive lobes move the k-space trajectory left-to-right, whilst negative lobes move it right-to-left. Between each echo and the next, a short phase encoding gradient pulse, or *blip*, is applied, moving the trajectory one line up, covering the entire k-space in a snake-like path. As for the gradient echo, the envelope of the peaks decays in time as a function of T_2^*. Similarly to TSE, an effective echo time T_E^{eff}, corresponding to the echo time relative to the gradient echo sampled at the centre of the k-space, is defined. EPI is also prone to geometric distortions and artifacts due to B_0 inhomogeneities and eddy currents, although correction methods exist that mitigate their effects [33].

4.15.6 Inversion preparation

It is possible to precede any acquisition sequence by a preparatory module consisting of a π RF pulse applied at time $t = 0$, which inverts the longitudinal magnetisation so that $M_z(0^+) = -M_0$. This is done for different purposes, such as to obtain heavily T_1-weighted images or to suppress the signal from a certain tissue. In the case of an *inversion recovery* sequence, the magnetisation is left to recover via T_1 relaxation as described in section 4.3 for a time T_I, or *inversion time*, according to equation (4.16):

$$M_z(\vec{r}, T_I^-) = M_0\left(1 - 2e^{-T_I/T_1(\vec{r})}\right) \tag{4.83}$$

The acquisition sequence, e.g. a spin echo, is then applied at $t = T_I$, with the transverse magnetisation for an echo being

$$M_\perp(\vec{r}, T_E + T_I) = M_0\left|1 - 2e^{-T_I/T_1(\vec{r})}\right|e^{-T_E/T_2(\vec{r})} \tag{4.84}$$

with the absolute value due to the fact that, once the longitudinal magnetisation is tilted into the transverse plane by the $\pi/2$ excitation pulse, only its magnitude is considered, as its sign is translated into a phase offset. This causes the signal also to be modulated and locally reduced by a function of T_1. By repeating the same sequence and acquiring different images at a different T_I, it is possible to reconstruct the T_1 relaxation curve

in each voxel, which can be fitted to extrapolate a spatial map of quantitative T_1.

A π preparatory pulse may be also applied prior of a 3D-FLASH sequence to produce fast, isotropic, heavily T_1-weighted images. Being a 3D sequence, multiple k_z-lines are acquired within the same T_R by stepping the G_z^{PE} gradient, with the effective inversion time T_I^{eff} being defined as the time between the inversion pulse and the acquisition of the central $k_z = 0$ line. The sequence, including the preparatory module, is then repeated stepping the G_y^{PE} gradient. This sequence, referred to as *magnetisation prepared rapid gradient echo* (MP-RAGE, or 3D T_1-*turbo field echo*, 3D T_1-TFE for Philips manufacturer), is a staple for anatomical MRI due to the sharp contrast and resolution.

Alternatively, a single T_I can be used, specifically tailored so that $M_z(t = T_I) = 0$ for a given T_1, that is $T_I = \ln(2) \cdot T_1$. In brain MRI, this is often used to suppress the signal produced by fluids, due to their characteristically long T_1, and the sequence is referred to as *fluid attenuated inversion recovery*, or FLAIR. In the case of CSF, average $T_1 \sim 4000\,\text{ms}$, therefore $T_I \sim 2700\,\text{ms}$. A FLAIR is also characterised by relatively long T_E and T_R to achieve a T_2-weighted contrast, and is largely used in clinical applications for the identification of brain lesions. The *short tau inversion recovery*, or STIR, is a similar signal suppression technique, using a short T_I to instead attenuate the signal produced by fat. Because of fat short T_1 varying considerably as a function of B_0 ($T_1 = 288\,\text{ms}$ at $1.5\,\text{T}$, $T_1 = 371\,\text{ms}$ at $3\,\text{T}$ [34]), magnetic field strength needs to be taken into account when calculating STIR T_I.

4.16 Myelin imaging

Due to the abundance of water in biological tissues, imaging of water hydrogen nuclei ${}^1_1\text{H}$ constitutes the typical objective of MRI. Water in the body can be found *free* or *bound* to other molecules. These different water states, or *pools*, determine different magnetic and imaging properties of the water protons. Unlike protons in the free water pool, water protons bound to lipids or other macromolecules are tightly packed, restricting their motion and resulting in T_2-values in the order of magnitude of microseconds. When imaged via conventional MRI techniques whilst employing a $T_E > 1\,\text{ms}$, bound pool protons are effectively invisible since, due to their extremely short T_2, the MR-signal they produce decays before it can be acquired. A few specialised techniques with *ultra-short* T_E [35] have been developed in order to detect the signal coming from bound protons, whilst other specialised methods have been developed to perform myelin mapping in an indirect way. Nonetheless, it is important to remember that each method has its own

assumptions and limitations, hereby described with reference to Heath et al. (2017) *Advances in noninvasive myelin imaging* review [27], and that, as a recent meta-analysis performed by Mancini et al. (2020) reports:

> *A holy grail of myelin imaging does not exist, at least as long as we consider histology to be the ground truth. Given that we have to pick or poison...* [36]

4.16.1 Magnetisation transfer

Despite being invisible to conventional MRI, water protons bound to macromolecules exhibit a much wider range of resonance frequencies, and this property can be exploited to indirectly image them. In *magnetisation transfer* (MT) imaging, an off-resonance RF pulse is applied to saturate the macromolecular protons, whilst leaving free water protons unexcited. By exploiting the T_2 properties of the bound macromolecular pool, cross-relaxation with the free water pool via spin-spin interactions causes the free water protons to partially saturate and, if subsequently imaged, to manifest a reduced MR-signal depending on the local macromolecular content. This effect of exchanging energy between water pools is known as magnetisation transfer. Under the assumption that macromolecular content is a good representative for myelin, MT can be exploited as an indirect measure for myelin content.

The simplest application of this effect is MT ratio (MTR) imaging, based on the ratio between an image with MT-weighting and one without (that is, a *reference* image). By adding an independently acquired T_1 map, and maps of B_0 and B_1 fields for inhomogeneity correction to the MTR setup, macromolecular proton fraction (MPF) can be estimated, representing the relative amount of protons in the macromolecular pool. Quantitative MT employs multiple MT pulses at different off-resonances in order to characterise the bound pool properties via multi-parametric modelling. To robustly fit all the parameters in these models, many different measurements are required, leading to long acquisition times. Additional assumptions can be included to simplify the models and the MR-protocol, although they may not necessarily be valid across subjects of differing age or disease status.

4.16.2 Macromolecular tissue volume

Under the same assumption of myelin abundance in the CNS macromolecular pool, macromolecular tissue volume (MTV) can also be used as an indirect metric for myelin content. MTV expresses the fraction of tissue volume in each voxel occupied by the

water in the bound pool. It is defined as a function of proton density (see section 4.9), as MTV $= 1 -$ PD: if PD represents the relative concentration of free water protons, then $1 -$ PD indicates the relative density of the macromolecular pool. PD-mapping is conceptually one of the most basic MRI measures, and several estimation techniques have been described in literature, however the resulting quantitative PD maps are usually corrupted by inhomogeneities in the coil receiver field. In order to account for this bias, different *bias field correction* methods have been developed: i) some combine data from multichannel coils into a single channel, whilst others keep data from the multiple channels separate during the estimation; ii) different regularisation assumptions can be adopted to overcome the ill-posed nature of the problem; iii) the techniques differ as to whether relying on a single global brain analysis or a set of local calculations that are integrated in a final step. In this work, a specific form of local regularisation has been used, which has been shown to lead to high PD-estimation accuracy [37].

4.16.3 Myelin water imaging

Myelin water imaging uses multi-echo spin echo with short T_E's to image water protons trapped within myelin layers. The signal produced by these protons contributes to 10–15% of the overall signal in a white matter voxel, referred to as *myelin water fraction* (MWF). Since T_2 relaxation time is related to the local environment with which the imaged protons interact, more restricted or dense tissue microstructure will give rise to shorter T_2, due to the higher rate of dipole-dipole interactions. Within a single white matter voxel, different compartments will present different ranges of T_2, with prolonged T_2 as the tissue architecture becomes *less* geometrically restricted: myelin water ($T_2 \sim$ 10–50 ms) $\rightarrow$ extra-cellular water ($T_2 \sim$ 60–90 ms) $\rightarrow$ free water ($T_2 \geq 120$ ms). In a multi-echo spin echo sequence, signal in a thin slice is measured multiple times as it decays away due to T_2 relaxation effects: the observed signal is then fitted to a model of multiple nonexchanging compartments with unique T_2 values. MWF can then be quantified by measuring the relative signal contribution of components with $T_2 < 50$ ms to the overall MR-signal. Extending this technique to multiple slices is however problematic due to induced MT effects in neighbouring slices, which lead to confounding measurements.

Fast 3D volumetric imaging techniques are available, such as *multicomponent driven equilibrium single pulse observation of T_1 and T_2* (mcDESPOT), which however come with limitations based on the complexity of the underlying models. Additionally, in spite of the high correlation with myelination, the multi-exponential model does not take into account exchange between compartments, and is only valid under the assumption

that the duration of the measurements over the signal decay takes place at a timescale much shorter than the rate of exchange. Given, however, the duration of a typical measurement is within 64–128 ms, this assumption is likely unmet, which has been demonstrated to lead to underestimation of MWF.

4.16.4 Susceptibility imaging

Susceptibility mapping exploits myelin susceptibility contrast to measure its distribution in the CNS. Magnetic susceptibility is defined as the degree to which tissue is magnetised by an external magnetic field and describes how the magnetic environment is perturbed by the presence of the magnetised material, as tissue and materials with differing susceptibility values influence local MR signal magnitude and phase. The phase of MR signal encodes the precession of proton spins due to the local resonance frequency offset: spins near diamagnetic regions (low relative susceptibility value) will precess at a slightly lower frequency than the resonance frequency of the main magnetic field, whilst spins near paramagnetic regions (high relative susceptibility value) will precess at a slightly higher frequency, leading to different phase accumulation. Susceptibility field inhomogeneities contribute to T_2^* decay and can thus be probed by using gradient echo sequences. *Susceptibility weighted imaging* (SWI) can be implemented by combining and filtering magnitude and phase information together to enhance the contrast of susceptibility differences in tissues. These differences are driven by a number of factors, including the microstructural organisation and chemical composition of tissues: for example, iron (paramagnetic) and myelin (diamagnetic) content in the CNS are known to be major contributors to susceptibility contrast.

Unlike SWI, *quantitative susceptibility mapping* (QSM) seeks to quantify the susceptibility shift in tissues, rather than simply enhancing contrast due to local susceptibility differences. In spite of the strong correlations between QSM and white matter myelin, it has been shown that the assumption of an isotropic, scalar susceptibility value within white matter is not valid and values measured using QSM are influenced by the orientation of the white matter fibres with respect to the main magnetic field. Tensorial approaches must then be adopted (*susceptibility tensor imaging*, STI), which however significantly increase acquisition times and costs. At the same time, it is known that iron deposition occurs in MS and Alzheimer's disease, and may represent a confounding factor for myelin quantification via QSM or STI.

4.17 Diffusion MRI

MRI enables to investigate different aspects of both the anatomical and functional structure of the brain. For example, measuring the diffusion of water molecules along axons due to thermal agitation allows to reconstruct the microstructural properties and neuronal architecture that supports the brain connectivity and anatomical integrity. Diffusion MRI is an extremely vast branch of MRI, characterised by a whole spectrum of techniques and objectives: in this section, the very basic notions of diffusion MRI will be explained, with a focus on the elements that had a specific role in this project.

4.17.1 Brownian motion

In 1827, the botanic Robert Brown observed through a microscope the motion of pollen particles dispersed in water, but was unable to determine the underlying mechanism:

> *These motions were such as to satisfy me, after frequently repeated observation, that they arose neither from currents in the fluid, nor from its gradual evaporation, but belonged to the particle itself.* Philosophical Magazine (1828)

This phenomenon was then described macroscopically by Fick twenty years later through his two *laws of diffusion*, and then microscopically by Einstein in 1905. Nonetheless, this random motion of particles in liquids or gas is referred to as *Brownian* motion.

In the context of NMR, water molecules diffusing through external field inhomogeneities cause irreversible magnetisation dephasing that cannot be recovered via spin-echo techniques due to its intrinsic randomness. The trajectory delineated by a single spin as a result of diffusion can be expressed in the 1D case along the x-axis as a succession of small steps $\lambda\epsilon_i$, with length λ and random direction, forward or backward, given by $\epsilon_i = \pm 1$, occurring every τ_d seconds. At each step i, the spin encounters a variation in the local magnetic field $(\mathrm{d}B_z/\mathrm{d}x)\lambda\epsilon_i \simeq G_x\lambda\epsilon_i$, where the magnetic field variation can be approximated by a constant gradient G_x for small enough step size λ. The magnetic field perceived by the spin after a number j of steps, and a time $t = j\tau_d$ is given by the initial magnetic field $B_z(t = 0)$ plus a field inhomogeneity term ΔB_z given by the sum of the local field variations along the trajectory:

$$B_z(j\tau_d) = B_z(0) + \Delta B_z(j\tau_d) = B_z(0) + G_x\lambda \sum_{i=1}^{j} \epsilon_i \tag{4.85}$$

The amount of relative dephasing ϕ_k accumulated by the k-th spin along the diffusion trajectory after N steps is then given by

$$\phi_k = -\sum_{j=1}^{N} \gamma \tau_d \Delta B_z(j\tau_d) = -\gamma \tau_d G_x \lambda \sum_{j=1}^{N} \sum_{i=1}^{j} \epsilon_i \tag{4.86}$$

When considering the entirety of the spin population, the *central limit theorem* can be applied, and the spin dephasing can be described by a Gaussian probability distribution:

$$P(\phi) = \frac{e^{-\phi^2/(2\langle \phi_k^2 \rangle)}}{\sqrt{2\pi \langle \phi_k^2 \rangle}} \tag{4.87}$$

with mean $\langle \phi_k \rangle = 0$ and variance for $N \gg 1$:

$$\langle \phi_k^2 \rangle = \frac{\gamma^2 \tau_d^2 G_x^2 \lambda^2 N^3}{3} = \frac{\gamma^2 G_x^2 \lambda^2 t^3}{3\tau_d} \tag{4.88}$$

where $t = N\tau_d$ allows to implicitly take into account diffusion time dependency[9].With reference to equation (4.29), this adds one more dephasing term to the complex transverse magnetisation, given by the weighted sum of all the diffusing components:

$$\begin{aligned} M_+^{\text{diff}} = \int \mathrm{d}\phi P(\phi(t)) e^{i\phi} &= \frac{1}{\sqrt{2\pi \langle \phi_k^2 \rangle}} \int \mathrm{d}\phi e^{i\phi - \phi^2/(2\langle \phi_k^2 \rangle)} \\ &= \frac{e^{-\langle \phi_k^2 \rangle/2}}{\sqrt{2\pi \langle \phi_k^2 \rangle}} \int \mathrm{d}\phi e^{-(\phi - i\langle \phi_k^2 \rangle)^2/(2\langle \phi_k^2 \rangle)} \\ &= e^{-\langle \phi_k^2 \rangle/2} \int \mathrm{d}\phi P(\phi(t) - i\langle \phi_k^2 \rangle) \\ &= e^{-\gamma^2 G_x^2 \lambda^2 t^3/(6\tau_d)} \equiv e^{-bD} \end{aligned} \tag{4.89}$$

where the integral of the probability distribution is, by definition, equal to unity, and the variance has been expanded as in equation (4.88), with $D = \lambda^2/(2\tau_d)$ being the

[9] Each step assumes random direction $\epsilon_i = \pm 1$, thus $\langle \epsilon_i \rangle = 0$, and $\langle \phi_k \rangle = 0$. For the variance:

$$\langle \phi_k^2 \rangle \propto \langle \Big[\sum_{j=1}^{N} \sum_{i=1}^{j} \epsilon_i \Big]^2 \rangle = \langle [\epsilon_1 + (\epsilon_1 + \epsilon_2) + (\epsilon_1 + \epsilon_2 + \epsilon_3) + ...]^2 \rangle = \langle [N\epsilon_1 + (N-1)\epsilon_2 + ... + \epsilon_N]^2 \rangle$$

When expanding the square and calculating the average, the cross products vanish since each step is independent from the others: $\langle \epsilon_i \epsilon_j \rangle = \langle \epsilon_i \rangle \langle \epsilon_j \rangle = 0$, $i \neq j$. This leaves only the sum of squares, with $\langle \epsilon_i^2 \rangle = 1$:

$$\langle \phi_k^2 \rangle \propto N^2 \langle \epsilon_1^2 \rangle + (N-1)^2 \langle \epsilon_2^2 \rangle + ... + \langle \epsilon_N^2 \rangle = \sum_{p=1}^{N} p^2 = \frac{N(N+1)(2N+1)}{6} \to \frac{N^3}{3}, \text{ for} N \gg 1$$

diffusion coefficient, and $b = \gamma^2 G^2 t^3/3$ the *b-value*. By incorporating the diffusion decay term, the complex transverse magnetisation in the resonant frame of reference can thus be expressed as

$$M_+(t) = M_+(0^+)e^{-t/T_2} M_+^{\text{diff}} = M_+(0^+)e^{-t/T_2} e^{-bD} \tag{4.90}$$

4.17.2 Diffusion Weighted Imaging

Diffusion weighted imaging (DWI) exploits the brownian diffusion process to reconstruct the microscopical architecture of the imaged sample. DWI measures are usually conducted using a bipolar pulsed-gradient spin echo (PGSE) sequence with EPI readout, where a *diffusion encoding gradient* $\vec{G}^{\text{diff}} = G^{\text{diff}}\hat{r} = G_x^{\text{diff}}\hat{x} + G_y^{\text{diff}}\hat{y} + G_z^{\text{diff}}\hat{z}$ with duration δ is applied before and after the π refocusing pulse, with a time interval Δ between the two applications. For a spin in position $\vec{r}_1$, the first gradient induces a dephasing

$$\phi_1 = -\gamma \vec{r}_1 \cdot \vec{G}^{\text{diff}} \delta \tag{4.91}$$

with $\vec{r}_1$ supposed to be constant during the duration δ of the gradient. The π refocusing pulse inverts the accumulated phase, such that $\phi_1 \to -\phi_1$. During the time Δ between the two gradient lobes, the spin diffuses within the sample, eventually reaching a position $\vec{r}_2$ when the second gradient is applied. This induces a rephasing

$$\phi_2 = -\gamma \vec{r}_2 \cdot \vec{G}^{\text{diff}} \delta \tag{4.92}$$

such that the total dephasing results in

$$\phi = \phi_2 - \phi_1 = -\gamma(\vec{r}_2 - \vec{r}_1) \cdot \vec{G}^{\text{diff}} \delta \tag{4.93}$$

Effectively, the diffusion gradients act as extrinsic magnetic field inhomogeneity, and an equation similar to (4.89) for the transverse magnetisation diffusion decay can be derived. The exact solution, with the gradient pulse duration δ not being negligible compared to the pulse interval Δ, has been first proposed by Le Bihan et al. (1985) [38] to be:

$$M_+^{\text{diff}} = e^{-\gamma^2 (G^{\text{diff}})^2 \delta^2 (\Delta - \delta/3) D} = e^{-bD} \tag{4.94}$$

with $b = \gamma^2 (G^{\text{diff}})^2 \delta^2 (\Delta - \delta/3)$, which takes into account the PGSE sequence parameters, being the original definition of Le Bihan's (hence the name) *b*-value.

Diffusion in tissues is, however, not *free*, but either *hindered* or *restricted* depending on the microscopic environment. Extra-cellular water is considered hindered by the interactions of water molecules with different obstacles, such as macromolecules, fibres and membranes. The diffusion coefficient D measured in each voxel appears therefore lower, and is defined as *apparent diffusion coefficient* (ADC). Intra-cellular water is considered restricted by being confined within the enclosed cellular environment. Unlike hindered diffusion, the time evolution of the displacement of restricted water molecules is not Gaussian and depends on the size and shape of the enclosing compartment; the associated ADC decreases with the diffusion time Δ, and thus the b-value as well, rather than being a constant property of the tissue, due to water molecules eventually reaching and *bouncing-off* the cell membrane [39].

For stationary spins ($\vec{r}_2 = \vec{r}_1$), or spins that diffused orthogonally to the direction of the gradient ($(\vec{r}_2 - \vec{r}_1) \perp \vec{G}^{\text{diff}}$), the dephasing induced by the first lobe is completely recovered by the second one, and there is no diffusion encoding decay in the transverse magnetisation. Otherwise, spins will experience a net dephasing as a function of the diffusion trajectory relative to the direction of the gradient, with maximum magnetisation diffusion encoding decay for $(\vec{r}_2 - \vec{r}_1) \parallel \vec{G}^{\text{diff}}$. By applying $\vec{G}^{\text{diff}}$ along different directions, the signal produced by water moving along each direction decreases proportionally to the water diffusivity along that direction, as the recovery of the magnetisation coherence does not match the initial spoil. In case of water moving *isotropically*, diffusion weighted signal is diminished uniformly along all encoded directions: this is the case of CSF, tissues rich in cell bodies, and lesions; on the other hand, water trapped within a highly oriented structure, such as a neuron axon, will be able to diffuse mainly along its principal direction, that is *anisotropically*, and its signal will decrease accordingly only when the diffusion encoding is applied parallel to that direction.

4.17.3 Diffusion Tensor Imaging

To take into account the directionality of the diffusion process, the water displacement in each voxel and b-encoding can be expressed in terms of symmetric matrices [40] called respectively *diffusion tensor* $\mathbf{D}$ and *b-tensor* $\mathbf{b}$:

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_{xx} & b_{xy} & b_{xz} \\ b_{yx} & b_{yy} & b_{yz} \\ b_{zx} & b_{zy} & b_{zz} \end{bmatrix} \tag{4.95}$$

with the ij ($i, j = x, y, z$) terms coupling diffusion displacements along i-j directions, and $ij = ji$ due to symmetry. In particular, for a standard PGSE sequence with *perfect* boxcar diffusion gradients, the b_{ij} terms can be expressed as

$$b_{ij} = \gamma^2 G_i^{\text{diff}} G_j^{\text{diff}} \left[\delta^2 \left(\Delta - \frac{\delta}{3}\right)\right] \tag{4.96}$$

but become more complex when taking into account transient times or non-constant gradients. It is worth recalling that the diffusion tensor and its products, e.g. eigenvalues, are all voxel-wise quantities, even though the dependency on voxel position $\vec{r}$ is left understood (e.g. $\mathbf{D} = \mathbf{D}(\vec{r})$) for better readability.

The tensors in equation (4.95) result in a diffusion decay term:

$$M_+^{\text{diff}} = e^{-\mathrm{b}\odot\mathrm{D}} = e^{-(b_{xx}D_{xx}+b_{yy}D_{yy}+b_{zz}D_{zz}+2b_{xy}D_{xy}+2b_{yz}D_{yz}+2b_{zx}D_{zx})} \tag{4.97}$$

where $\odot$ denotes the Hadamard (i.e. element-wise) product and the matrix symmetry has been taken into account when expanding it. From this it follows that at least six non-collinear diffusion encoded measures must be acquired to reconstruct the diffusion tensor in each voxel, plus one with no diffusion encoding, referred to as a b_0-image. This technique is thus called *diffusion tensor imaging* (DTI), and it is the simplest model able to represent the directionality of diffusion in tissues.

The diffusion tensor defines a so called *diffusion ellipsoid* [41]: a 3D representation of the distance covered by water molecules, in each voxel, over a certain diffusion time τ_D. The eigenvalues of $\mathbf{D}$ — λ_1, λ_2, λ_3 — can be intended as the local apparent diffusion coefficients along the three orthogonal directions determined by the corresponding eigenvectors, which define the orientation of the ellipsoid's three axes of symmetry. The ellipsoid can thus be expressed analytically as

$$\frac{x_\lambda^2}{2\lambda_1\tau_D} + \frac{y_\lambda^2}{2\lambda_2\tau_D} + \frac{z_\lambda^2}{2\lambda_3\tau_D} = 1 \tag{4.98}$$

where x_λ, y_λ, z_λ are the coordinates in the reference frame generated by the eigenvectors. The diffusion tensor and its eigenvalues in each voxel can be used to compute quantitative diffusion maps based on the following diffusion metrics.

Mean diffusivity (MD) measures the average ADC along all directions, and is therefore invariant with respect to the ellipsoid orientation. It is used to highlight lesions in which tissue microstructure has been disrupted, resulting in more isotropic diffusion, and

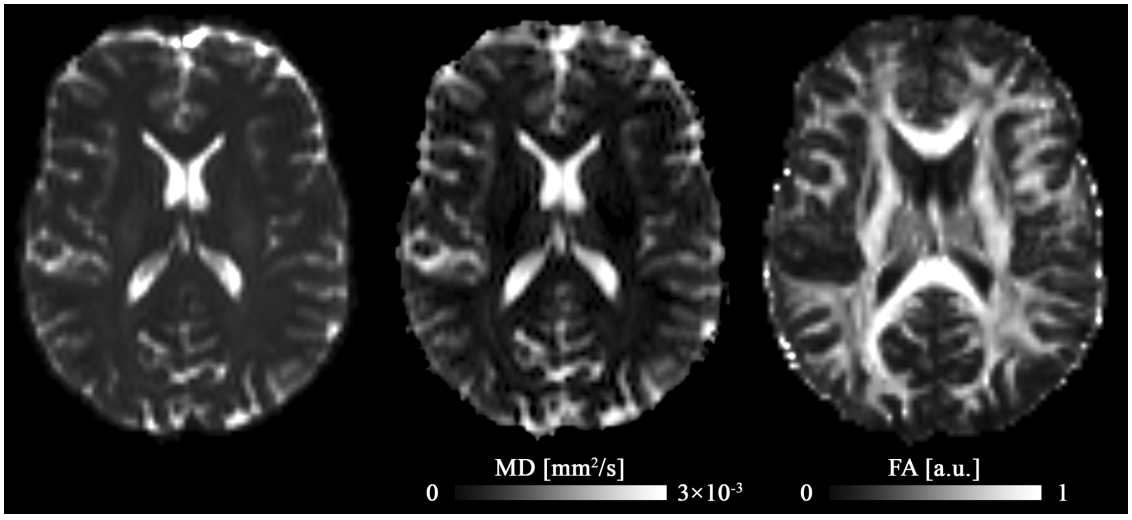


Figure 4.8: From left to right: examples of a b_0-image, MD and FA maps.

higher MD. It is defined as

$$\mathrm{MD} = \frac{\mathrm{Tr}(\mathbf{D})}{3} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} = \langle \lambda \rangle \tag{4.99}$$

Fractional anisotropy (FA) is an index of the degree of orientation coherence within the tissue, being higher in voxels characterised by highly oriented structures, i.e. high anisotropy, such as white matter tracts, and lower in voxels with no or disrupted microstructure, such as CSF and lesions. It is defined as

$$\mathrm{FA} = \frac{\sqrt{3[(\lambda_1 - \langle \lambda \rangle)^2 + (\lambda_2 - \langle \lambda \rangle)^2 + (\lambda_3 - \langle \lambda \rangle)^2]}}{\sqrt{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \tag{4.100}$$

Examples of a b_0-image, MD and FA maps are shown in Figure 4.8.

4.17.4 Advanced techniques

DTI metrics are inherently non-specific to tissue microstructure, and the DT-model relies on the assumption of Gaussian displacement of the water molecules over time, which fits free and hindered diffusion, but it is not met in the case of restricted diffusion. Furthermore, because the geometric translation of the diffusion tensor is an ellipsoid, DTI can only indicate the main diffusion direction in each voxel, which is acceptable in regions where axons are coherently aligned along a single direction, but performs poorly in case of *crossing fibres*. To overcome these limitations, more advanced techniques have been developed over time.

High angular resolution diffusion imaging (HARDI) [42] acquisitions employ high, possibly multiple *b*-values (or *shells*) and a large number of diffusion encoding directions (with a minimum of 45 [43]) to compute fibre orientation density functions (ODF), showed in Figure 4.9. Spherical deconvolution [44] and q-ball imaging [45] are examples of the first methods to compute ODFs, however others have been proposed since then with successful results. Multi-compartment models have been introduced to take into account the microstructural properties of brain tissues in terms of microscopic environments, and overcome some of the limitations of DTI. Behrens et al. (2003) [46] were among the first to propose an alternative to DTI in the form of a *ball-and-stick* model, distinguishing between an intra-cellular compartment, modelled as a zero-radius cylinder (*stick*), and an extra-cellular compartment defined by isotropic diffusion (*ball*). The *neurite*[10] *orientation dispersion and density imaging* (NODDI) technique, proposed by Zhang et al. (2012) [47] greatly contributed towards clinically viable multi-compartment tissue modelling, distinguishing between intra-cellular, extra-cellular, and CSF compartments. In the NODDI model, the intra-cellular compartment refers to the highly restricted intra-neurite diffusion environment, modelled as *sticks* with different degrees of orientation dispersion; the extra-cellular compartment refers to the space around the neurites, occupied by glial cells and cell bodies (in grey matter), with water diffusion in this environment being hindered and described by an anisotropic Gaussian model; finally, the CSF compartment free water diffusion is modelled as an isotropic Gaussian model.

In this work, the more recent *spherical mean technique* (SMT) has been employed, described in detail as follows.

4.17.5 Spherical Mean Technique

SMT is a recent multi-compartment method proposed by Kaden et al. (2016) [48, 49] that aims to quantify the parameters regulating the per-axon diffusion process by relying on the insight that, for a fixed configuration of the diffusion encoding gradients, that is for a certain *b*-value, the spherical mean (SM) of the diffusion-weighted signal over the gradient directions does not depend on the underlying neurite orientation distribution. Once the SM of the signal over the diffusion encoding direction has been computed for each *b*-value, the parameters of a multi-compartment microscopic diffusion model (MCMicro) can be estimated by least-squares fitting the SM-version of the model to the SM-signal. The MCMicro model separates between an intra-neurite and an

[10] Collective term for dendrites and axons.

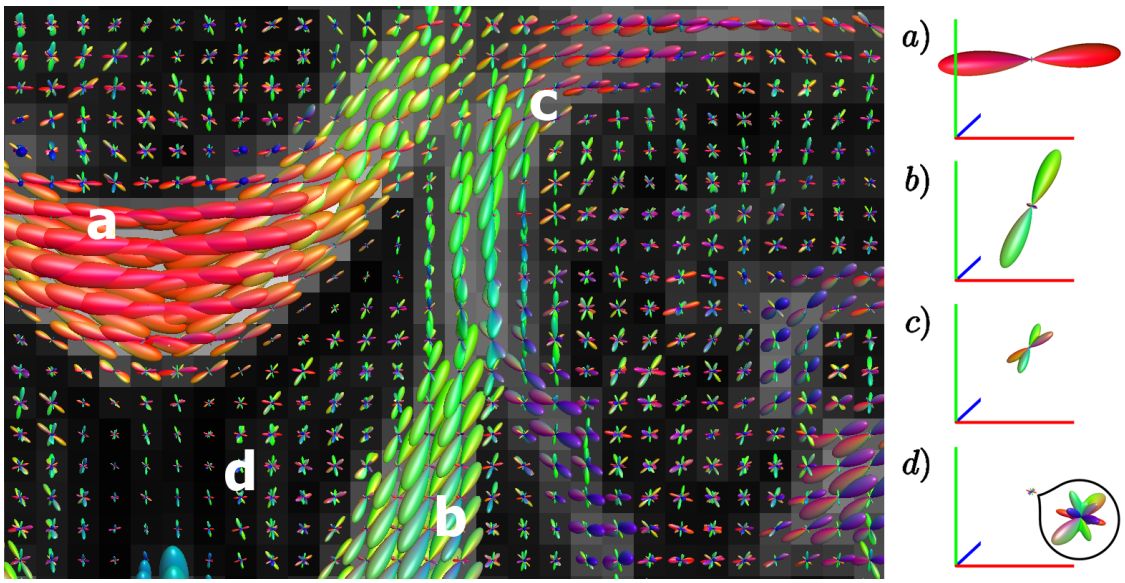


Figure 4.9: ODF overlayed onto an FA map. The image shows different tissue architectures in the periventricular region: a) corpus callosum — highly oriented fibres in the left-right direction; b) anterior thalamic radiation — highly oriented fibres in the anterior-posterior direction; c) crossing fibres, given by the commixture of left-right and anterior-posterior oriented fibres populations; d) CSF, characterised by diffusion isotropy.

extra-neurite component, where the former consists of dendrites and axons, and the latter of neuronal bodies, glial cells and extra-cellular space. According to this model, and using Kaden's notation, the DW-signal $h_b(g,\omega)$ for a given b-value and normalised diffusion gradient direction $g \in S^2$, with S^2 being the unit sphere, generated by a microscopic environment oriented along a direction $\omega \in S^2$, can be modelled as

$$h_b(g,\omega) = \nu_{\text{int}} h_b^{\text{int}}(g,\omega) + (1-\nu_{\text{int}}) h_b^{\text{ext}}(g,\omega) \tag{4.101}$$

where h_b^{int} and h_b^{ext} indicate the signal from the intra- and extra-neurite water pools respectively, whilst ν_{int} denotes the intra-neurite volume fraction. Assuming that the diffusion process within neurites can be modelled as within a zero-radius cylinder (i.e. a *stick*), the transverse microscopic diffusion coefficient can be neglected and the microscopic signal generated from the intra-neurite compartment becomes

$$h_b^{\text{int}}(g,\omega) = e^{-b\langle g,\omega\rangle^2 \lambda} \tag{4.102}$$

where $0 < \lambda < \lambda_{\text{free}}$ is the intrinsic diffusivity along the neurite axis, and $\langle g,\omega\rangle \in [-1,1]$ indicates the spherical distance between ω and g. The upper bound λ_{free} denotes the diffusion coefficient in free-water, which is about $3.05 \cdot 10^{-3}\,\text{mm}^2/\text{s}$ at body temperature

(37 °C). The extra-neurite signal can be expressed instead as

$$h_b^{\text{ext}}(g,\omega) = e^{-b\langle g,\omega\rangle^2\lambda} e^{-b(1-\langle g,\omega\rangle^2)\lambda_\perp^{\text{ext}}} \tag{4.103}$$

where the two factors on the right side of the equation model the diffusion process in the neighbouring areas parallel and perpendicular to the neurites, respectively. $\lambda_\perp^{\text{ext}}$ is the transverse extra-neurite microscopic diffusion coefficient, and is expressed as $(1-\nu_{\text{int}})\lambda$ according to the first-order *tortuosity approximation*.

Once the microscopic diffusion coefficients and compartments volume fractions have been recovered, and the diffusion process modelled in each voxel, ODFs can be estimated by means of spherical deconvolution. A summary statistics that can be extracted from the neurite orientation distributions is the relative *entropy* $H(p)$ of the ODF in each voxel, defined as the Kullback-Leibler divergence of $p(\omega) = \text{ODF}$ with respect to a uniform distribution $q(\omega) = 1/(4\pi), \forall\omega \in S^2$. Specifically:

$$H(p) = \int_{S^2} p(\omega) \log\left(\frac{p(\omega)}{q(\omega)}\right) \mathrm{d}\omega \tag{4.104}$$

$H(p) \approx 0$ if $p \approx q$, that is if the ODF in a certain voxel approaches the uniform distribution, such as in areas of isotropic diffusion like the CSF. On the other hand, in regions of high orientation coherence and diffusion anisotropy, such as the corpus callosum, $H(p) \gg 1$, as the ODF is locally very different from the uniform distribution.

4.18 MRI in MS

Thanks to its noninvasiveness, versatility and sensitivity to different contrasts, MRI plays a key role in the diagnosis and follow-up of MS patients. In this section, with reference to the MRI principles and techniques described previously, some of the main applications and limitations of MRI in the context of MS have been explored.

4.18.1 The clinico-radiological paradox

T_2-weighted images constitute an essential tool in the MS diagnostic process, as they can be acquired in less than 5 minutes, they are easy to interpret and enable clinicians to quickly and precisely distinguish scarred regions from normal appearing tissue. However, the number and volume of lesions explain only a small fraction of the diversity of clinical disability in MS, and this mismatch has been defined as the *clinico-radiological*

paradox [50, 51]. Furthermore, as described below, studies have shown that normal appearing tissues in MS patients present abnormalities otherwise invisible to conventional qualitative MRI [3]. Variable degrees of normal appearing white matter alteration have been in fact shown to precede new lesion formation, are detected in all MS phenotypes (albeit with different strengths), correlate with the level of physical disability and cognitive impairment, and are only modestly correlated with the total amount of macroscopic lesions. It appears thus evident that, whilst conventional MRI is an important qualitative diagnostic tool in MS, it lacks specificity in pinpointing widespread microscopic damage, and that proper quantification of myelin content and tissue derangement in the CNS, in addition to measures of lesion dissemination, is key for the diagnosis and prognosis of MS.

4.18.2 Relaxometry

Relaxometry, that is the measurement of relaxation times from MR images, represents a fundamental quantitative tool for the characterisation of MS. Increase in T_2 may be caused by inflammation, demyelination and axonal loss, whilst reduced T_2 may be observed as a result of iron deposition (see section 4.4). Both histological and MRI studies have shown the involvement of iron deposition in brain in neurologic diseases, although it has a role in normal ageing as well, and it still unclear whether it is the cause of the neuronal degeneration or a mere marker. Regions of reduced T_2 due to, presumably, ferruginated neurons have been observed in MS patients in grey matter, particularly in deep grey matter structures including globus pallidus, putamen, caudate and red nuclei, substantia nigra and thalamus [52]. T_1 has also been shown to have a role in the characterisation of MS, with quantitative T_1 studies showing that alterations in T_1 can highlight damaged areas otherwise invisible on T_2-weighted scans: increase of T_1 in normal appearing white matter and grey matter has been observed in MS patients, with SPMS patients experiencing greater alterations compared to RRMS ones. Furthermore, T_1 prolongation in grey matter appears to significantly correlate with physical disability, whilst increases in white matter do with brain atrophy. However, although T_1 is indeed sensitive to microscopic alterations, a change in measured T_1 lacks the specificity necessary to be used alone as a marker for the underlying pathology [53].

A different approach, which stirs away from quantitative MRI, consists of using qualitative T_1-weighted and T_2-weighted images to enhance myelin contrast in the brain. Several studies have suggested that myelin content of cortical areas covaries with both T_1-weighted and T_2-weighted intensities, but in opposite directions. In marmorsets, strong positive correlation has been observed between T_1-weighted intensities and

histologically measured myelin content (regions with high concentration of myelin appear hyperintense), whilst low T_2-weighted intensities are observed in regions rich in iron, strongly co-localised with myelin in the CNS [54]. As a result, myelin contrast can be enhanced by calculating the ratio between the two maps. Fast scanning times and conceptual simplicity make this technique well-suited for clinical investigations, although T_1-/T_2-weighted ratio alterations observed in pathological cases may be not only due to demyelination, but also inflammation, oedema, iron accumulation or atrophy. Furthermore, it is worth noting that T_1-/T_2-weighted ratio is still a qualitative measure, and can therefore be potentially characterised by intensity scale inconsistencies across datasets despite the application of intensity calibration techniques, especially in case of scans acquired with very different pulse sequences and in presence of disease [55].

4.18.3 Myelin

Demyelination in the CNS is one of the core symptoms associated to MS, thus quantification of myelin content (see section 4.16) in clinics is instrumental for an accurate understanding of MS pathophysiology and prediction of disability.

In MT-imaging, MTR has been shown to be highly sensitive to demyelination, with reduced values being observed in enhancing lesions, followed by a rapid recovery in acute ones, suggesting a clear correlation between demyelination and remyelination, and MTR transient changes. Widespread reduced MTR has also been observed in normal appearing tissues of MS patients, as well as prior to lesion formation [56]. MTR has been shown to strongly associate with myelin content and residual axons in post-mortem histological studies [57], as well as to robustly correlate with physical disability [58]. Despite being also influenced by inflammation and oedema [59] and not a *direct* measure of myelin content, and thus caution should be employed when interpreting results, MTR has represented an undeniably popular tool for *indirect* myelin mapping in clinical research over the years. Similarly to MTR, MPF has been found to be significantly lower in normal appearing tissues and lesions of SPMS patients compared to RRMS, which in turn exhibit a lower MPF in normal appearing white matter and grey matter compared to healthy controls. MPF has also been shown to correlate with clinical disability, and overall outperform both MTR and quantitative T_1 in the detection of tissue alterations [60].

Relatively new in the landscape of myelin imaging, $\mathrm{MTV} = 1 - \mathrm{PD}$ has recently started to be employed in MS clinical research as an indirect measure for myelin content. Reduced MTV values have in fact been observed in MS compared to healthy controls, both in

lesions and normal appearing tissue, working synergistically with diffusion imaging and T_1 mapping [61], although with shorter acquisition times compared to MTR and from routinely available MR-scans, making this modality highly appealing in clinical research.

Through the quantification of short-T_2 component, MWF has also been reported as a robust indicator of demyelination in MS: higher free water content and, in turn, reduced MWF have been observed in normal appearing white matter of MS patients compared to healthy control [62]. Although still potentially co-dependent on inflammation and oedema, strong association with myelin content has however been shown in histopathological studies [63], with reduced MWF in specific white matter functional systems significantly correlating with clinical disability [64].

QSM studies have shown that reduction in myelin content (less diamagnetic) results in a dramatic increase in white matter susceptibility, making susceptibility imaging a potentially powerful tool for the quantification of demyelination in normal appearing tissues and the anatomical reconstruction of demyelinating plaques. Like other MR-modalities, QSM cannot however be taken as a direct index of myelin content, as similarly to demyelination, iron deposition (more paramagnetic) causes increases in QSM local values, and whilst increased iron deposits have been reported in MS concomitantly with myelin loss, it has been also shown to correlate with ageing in healthy controls [65].

4.18.4 Microstructure

With microstructural alterations being at the core of neurodegenerative diseases, DWI offers a key tool in probing MS pathophysiology in terms of axonal degeneration (see section 4.17). Several diffusion-related techniques have in fact been employed in the context of MS, with DTI being a popular choice in MS clinical research for its simplicity and clinical viability: increased MD and decreased FA have been observed in MS-lesions compared to healthy tissue, indicative of the microstructure disruption due to the demyelination process, with similar behaviour being observed in normal appearing white matter of MS patients as well [52]. Due to DTI inability to differentiate crossing-fibres and correctly model tissue microstructural compartments, more advanced models have been progressively employed instead.

NODDI has been used extensively, providing a clinically feasible method for mapping neurite orientation dispersion and density. NODDI has provided evidence, across multiple studies, of reduced neurite density in lesions and in both brain and spinal cord normal appearing white matter, as well as in spinal cord normal appearing grey matter, in

MS patients compared with healthy controls. Contradictory results have been however observed on the basis of ODI measurements, with high inter-study variability, making the interpretation of NODDI results challenging [66].

More recently, SMT-derived parameters have been shown to be sensitive to different degrees of brain tissue damage [67], working synergistically with DTI metrics in the characterisation of MS lesions [68], with reduced intra-neurite volume fraction also being observed in spinal cord normal appearing white matter of MS patients compared with healthy controls [66]. The recent development of a multiband acquisition technique has allowed to halve total scan time, reducing it to $\sim$10 min, making SMT a potentially useful tool in MS clinical research.

4.18.5 Physiology

Alterations in neuronal physiology are ultimately the root of clinical disability. Different MRI techniques have been developed in order to investigate and quantify the state of cellular physiology and function in the context of MS.

Functional MRI (fMRI) assesses brain activation by measuring changes in *blood-oxygen level dependent* (BOLD) signal, under the assumption that an increase in neuronal activity corresponds to a proportional local haemodynamic response: this relationship is called *neurovascular coupling*. The whole fMRI paradigm relies on the assumption that such mechanism is intact. Pathology, including MS, might alter the neurovascular cascade, for example by causing vasodilation and increasing cerebral blood flow, which can then influence BOLD signal. In MS, *perfusion imaging* studies have shown a cerebral blood flow decrease in normal appearing white matter in patients compared to controls, together with an increase in lesion load, suggesting that the microcirculation may be indeed influenced by inflammation [52]. In such cases, a correction step might be needed for a correct interpretation of fMRI the results [69]. That being said, fMRI has revealed the existence of adaptive processes limiting cognitive impairment, acting through the recruitment of supplementary brain areas. Although the specific mechanisms are still unknown, they are thought to respond to the damaged neuronal circuits, compensating for tissue degeneration and preserving cognitive performance within certain limits depending on the disease progression [70].

Given the fundamental role sodium ions provide in neuronal physiology, sodium imaging offers a direct marker for neurons functional integrity. Quantitative sodium MRI allows to measure tissue sodium concentration (TSC) in brain by calibrating the acquired spin

density–weighted signal using the one produced by phantoms with known concentrations of $^{23}_{11}$Na sodium ions (see section 4.9). Compared to other MRI techniques, it is still relatively little used in MS clinical research, despite the renown physiological importance of sodium in the brain for the propagation of action potentials. This is mainly due to the limitations intrinsic to the measuring process: the very low concentration of sodium ions in the brain compared to water (80 mM and 88 M respectively) and the short $T_2 \sim 0.5$–5 ms associated to 60% of the signal contribute to an inherently low signal to noise ratio. In the last few years, interest in quantitative sodium MRI has been revived by the shortening of acquisition time through ultra-short T_E sequences and the increased resolution given by better scanners, and applications to the MS field did not fail to emerge. Significant differences in TSC have been in fact reported in both white matter and grey matter regions between healthy controls and MS patients [71, 72]. Advances in sodium imaging have also offered a unique chance to probe neuronal activation by accessing signal changes directly linked to sodium ions flux across the cell membrane, rather than indirectly via BOLD signal. A recent work by Gandini Wheeler-Kingshott et al. (2018) has shown that quantitative *functional* sodium imaging (fNaI) at 3T is potentially sensitive to sodium concentration changes during finger tapping in regions of the CNS linked to motor control, suggesting that fNaI is sensitive to distributed functional alterations and may constitute a powerful tool for the investigation of motor disability accrual during MS progression [73].

Outside of the imaging sphere, *MR-spectroscopy* has been used to measure the concentration of metabolites in the brain, reporting changes that appeared to correlate with the degree of activity of MS lesions [74].

Chapter 5

Machine learning

All the presented techniques are but a small fraction of the spectrum of available MRI methods, with new ones being constantly developed, each coming with its own intrinsic limitations and assumptions to be met, and many requiring specialised pulse sequences, often too long or expensive to be included in standard clinical protocols. In this landscape of high-dimensional and complex data, *machine learning* represents a powerful set of statistical tools for patient classification and the identification of those modalities, or *features* of the acquired data, that best correlate with disease phenotypes: with quantitative MRI offering a window into biological properties of tissues, this would enable to get a better understanding of the disease itself. Machine learning general purpose and its potential to learn from the data without the need for strong, prior assumptions has enabled the permeation of artificial intelligence throughout all fields of medical imaging [75].

Whilst machine learning refers to the study and implementation of algorithms *learning-from-data* as a whole, a subset of machine learning called *deep learning*, characterised in particular by the use of *neural networks*, has emerged to extract useful information directly from the image local contrast and topology. In the past few years, deep learning has proven to be particularly suitable for image processing, revolutionising the field of neurosciences, as it has been successfully used to automatise processes performed until recently manually or semi-automatically, namely lesion segmentation, or to greatly speed up post-processing pipelines, that would take hours to complete, to mere seconds.

In this chapter, elements of machine learning relevant to this work are described, mostly with reference to Hastie et al. (2009) *The Elements of Statistical Learning* [76].

5.1 Algorithm categorisation

Machine learning algorithms can be distinguished into different categories depending on different factors, e.g. the data provided during the learning stage, objective, and architecture. To follow, two main categorisations are presented.

5.1.1 Supervised vs unsupervised

To train a *supervised* algorithm, both input and output data are provided. The output data type varies greatly depending on the learning task: it may be a categorical *label* representative of the subject group (e.g. healthy control vs patient), a binary label associated to an image voxel (e.g. normal tissue vs lesion), a continuous value (e.g. probability map), or a combination of them. In any case, the objective of a supervised algorithm is to learn the function mapping the input to the provided output, which acts as ground truth, so that a *prediction* can be performed when applied to new input data. In the case of MS, the ground truth may be provided by an expert, such as hand-drawn lesion masks, or for example using clinical observations to assess the expanded disability status scale (EDSS) to determine the patient's MS subtype. The output may also be the result of a traditional model fitting performed on the input data before training, in which case the objective of the machine learning algorithm is to learn the fitted model and return the output for new data in a fraction of the time. Examples of supervised algorithms include *support vector machines*, *random forests* and various applications of neural networks, described more in detail below, as they were used throughout this project.

Unsupervised algorithms are trained instead to find patterns in the input data without providing a ground truth output during training, drawing inferences using similarity metrics defined for the specific algorithms. A common example of unsupervised learning is *clustering* analysis, which aims to divide the data into K clusters of elements closely matched to each other. Clustering strategies include K-means, hierarchical clustering, and mixed Gaussian model. A recently developed unsupervised algorithm, *Subtype and Stage Inference* (SuStaIn), uses a mixture of linear z-score models, with the z-scores calculated with respect to healthy controls, to categorise patients' disease progression into data-driven phenotypes, based on the patterns of biomarker *temporal evolution* [77]. SuStaIn provides a way to model phenotypic and temporal variation at once, whilst traditional clustering methods can only focus on one at a time. SuStaIn has been applied to Alzheimer's disease data and, very recently, to MS as well [78].

5.1.2 Classification vs regression

Algorithms that are trained to categorise the input into a certain group or class are called *classifiers*. Support vector machines and random forests are also examples of supervised classifiers, often used for patient classification. Neural networks may also fall in this category, e.g. *convolutional* neural networks, described more in detail below, employed for lesion segmentation, where the classification occurs at a voxel-wise level, and the classification task consists in assigning any given voxel to either the lesion or healthy tissue group.

If the output is instead a continuous value, the prediction task is called *regression*. Machine learning model fitting, as well as techniques devised to improve image quality, e.g. denoising, artifact correction, often using deep neural networks, are examples of regression.

5.2 Performance scores

In the case of classifiers, three quantities are often adopted to describe performances: *accuracy* is defined as the fraction of true positives and true negatives correctly classified; *sensitivity* is the rate of true positives correctly classified, whilst *specificity* corresponds to the true negative rate. All these scores assume values within $[0, 1]$. Accuracy alone works well for *balanced* classification problems, that is tasks where all classes have the same number of instances, with accuracy of 1 indicating perfect classification, and 0.5 corresponding to *random chance*. However, in *imbalanced* problems, accuracy will be biased towards the majority class, as the random chance value will rise to the relative size of the majority class over the total number of instances, and is therefore necessary to report sensitivity and specificity as well, which makes interpretation of the results more cumbersome.

That being said, in binary classification — with classes being here defined as *positive–negative* — the prediction is often based on a continuous score X, often scaled to represent the *probability* of an instance belonging to a certain group, which is then compared to a *threshold* T to determine the *hard* classification output: if $X > T$, the instance is classified as positive, and negative otherwise. Varying the value of T will therefore change the prediction for any given instance, and thus alter the performance scores. By plotting the *true positive rate* (TPR = sensitivity) against the *false positive rate* (FPR $= 1 -$ specificity) at various threshold settings in a 2D $[0, 1] \times [0, 1]$ space, it is possible to define a *receiver operating characteristic* (ROC) curve that describes the classifier performance at different degrees of confidence. Different classifiers will

produce different curves, with the *diagonal* indicating random chance, and curves closer to the (FPR, TPR) $= (0, 1)$ top-left corner indicating better performances. Two points are however shared by all ROC curves:

(FPR, TPR) $= (0, 0)$, for $T = 1$: all instances are classified as negative.

(FPR, TPR) $= (1, 1)$, for $T = 0$: all instances are classified as positive.

Apart from these two points, a *perfect* classifier (or perfectly separated data) will produce probability scores $X = 0$ for *all* negative-classified instances, and $X = 1$ for *all* positive-classified ones: therefore (FPR, TPR) $= (0, 1)\ \forall T \neq \{0, 1\}$. A *good* classifier (or highly separated data) will produce probability scores clustered around $X \sim 0$ for negative instances, and $X \sim 1$ for positive ones: in most cases, varying T will not change the hard classification, and most instances will therefore be correctly classified; however, the hard classification for some instances classified with lower confidence (e.g. $X = 0.4$ for a negative instance, or $X = 0.6$ for a positive one) will change with T, producing a ROC curve overall below the (FPR, TPR) $= (0, 1)\ \forall T \neq \{0, 1\}$ *perfect* one. A *random* classifier (or heavily overlapping data) will produce an array of probability scores for all instances distributed within $X = [0, 1]$, with no discernible clustering: hard classification will therefore change strongly with the value of T, with (FPR, TPR) delineating the $(0, 0) \rightarrow (1, 1)$ ROC diagonal.

The *area under the curve* (AUC) is a summary ROC performance score that encapsulates the information of accuracy, sensitivity and specificity within a single quantity. It can be used in place of them, particularly in case of imbalanced binary classification tasks, with AUC $= 1$ for a *perfect* classifier and AUC $= 0.5$ for random chance.

For regression, several norm and/or similarity functions quantifying the difference between the predicted and expected outputs (e.g. L_1 or L_2 norm), already calculated internally to the algorithm to drive the learning process, may be used.

5.3 Support Vector Machine

Support vector machine (SVM) is a popular architecture designed to solve binary classification tasks as a geometric problem consisting of separating two classes distributed in the feature space $\mathbb{R}^p$ using a hyperplane, defined by

$$\{\boldsymbol{x} \in \mathbb{R}^p : f(\boldsymbol{x}) = \boldsymbol{x} \cdot \boldsymbol{\beta} + \beta_0 = 0\} \tag{5.1}$$

where $\{\boldsymbol{\beta}, \beta_0\}$ are the hyperplane parameters, and $\text{sign}(f(\boldsymbol{x})) \in \{-1, 1\}$ returns the binary classification for a data point $\boldsymbol{x}$ in the feature space. Given a set of linearly separable data points composed of N pairs $\{(\boldsymbol{x}_i, y_i), i = 1, ...N\}$, with feature vector $\boldsymbol{x}_i \in \mathbb{R}^p$ and label $y_i \in \{-1, 1\}$, the closest ones to the separating hyperplane are called *support vectors* (whose set is hereby defined as S). The function $f(\boldsymbol{x})$ can then be scaled so that

$$\begin{aligned} &\forall i : \boldsymbol{x}_i \in S, f(\boldsymbol{x}_i) = \boldsymbol{x}_i \cdot \boldsymbol{\beta} + \beta_0 = 1, \quad \text{if } y_i = 1 \\ &\forall i : \boldsymbol{x}_i \in S, f(\boldsymbol{x}_i) = \boldsymbol{x}_i \cdot \boldsymbol{\beta} + \beta_0 = -1, \quad \text{if } y_i = -1 \end{aligned} \tag{5.2}$$

without altering (5.1). As a result, data points further away will be subject to

$$\begin{aligned} &\forall i : \boldsymbol{x}_i \notin S, f(\boldsymbol{x}_i) = \boldsymbol{x}_i \cdot \boldsymbol{\beta} + \beta_0 > 1, \quad \text{if } y_i = 1 \\ &\forall i : \boldsymbol{x}_i \notin S, f(\boldsymbol{x}_i) = \boldsymbol{x}_i \cdot \boldsymbol{\beta} + \beta_0 < -1, \quad \text{if } y_i = -1 \end{aligned} \tag{5.3}$$

which can be summarised as

$$f(\boldsymbol{x}_i) = y_i(\boldsymbol{x}_i \cdot \boldsymbol{\beta} + \beta_0) \geq 1, \forall i = 1, ..., N \tag{5.4}$$

In the feature space, equation (5.2) defines two supporting hyperplanes, each distant $1/\|\boldsymbol{\beta}\|$ from the hyperplane in (5.1), and $2/\|\boldsymbol{\beta}\|$ from each other. The goal of a SVM algorithm is to maximise this distance, or *margin*, using convex optimisation or, equivalently, to minimise $\|\boldsymbol{\beta}\|$ under the constraint defined in equation (5.4), which ensures that all data points are lying on or outside of the supporting hyperplanes, and not within the margin.

In case the two classes overlap in the feature space, whether due to noise or imperfect labelling, the data points are no longer linearly separable. One option is to still maximise the margin between the supporting hyperplanes (or, equivalently, minimise $\|\boldsymbol{\beta}\|$), whilst allowing some data points to be on the *wrong* side of the margin. This can be done by introducing *slack variables* $\xi_i \geq 0, i = 1, ..., N$ for each data point, which represent the amount by which the prediction expressed by $f(\boldsymbol{x}_i)$ is on the wrong side of its margin, and modify the constraint in equation (5.4) as

$$f(\boldsymbol{x}_i) = y_i(\boldsymbol{x}_i \cdot \boldsymbol{\beta} + \beta_0) \geq 1 - \xi_i, \forall i = 1, ..., N \tag{5.5}$$

By bounding the sum of the slack variables to a constant K, it is possible to limit the total amount by which predictions fall on the wrong side of the margin. This leads to the definition of a generic SVM for a non separable dataset:

$$\min_{\boldsymbol{\beta},\beta_0,\xi_i} \|\boldsymbol{\beta}\|, \text{subject to } \begin{cases} f(\boldsymbol{x}_i) = y_i(\boldsymbol{x}_i \cdot \boldsymbol{\beta} + \beta_0) \geq 1 - \xi_i, \forall i \\ \xi_i \geq 0, \forall i;\ \sum_{i=1}^{N} \xi_i \leq K \end{cases} \tag{5.6}$$

which can be re-expressed in the computationally equivalent form:

$$\begin{gathered} \min_{\boldsymbol{\beta},\beta_0,\xi_i} \frac{1}{2}\|\boldsymbol{\beta}\|^2 + C\sum_{i=1}^{N} \xi_i \\ \text{subject to } f(\boldsymbol{x}_i) = y_i(\boldsymbol{x}_i \cdot \boldsymbol{\beta} + \beta_0) \geq 1 - \xi_i, \forall i;\ \xi_i \geq 0, \forall i \end{gathered} \tag{5.7}$$

where the *cost*, or *regularisation* parameter C has the same role of the constant K in equation (5.6), as it regulates the trade-off between the maximised margin and the degree of misclassification. Equation (5.7) can be solved using Lagrange multipliers, which leads to the solution for $\boldsymbol{\beta}$ in the form

$$\boldsymbol{\beta} = \sum_{i=1}^{N} \alpha_i y_i \boldsymbol{x}_i \tag{5.8}$$

with α_i being nonzero only for those data points for which equality in the constraint (5.5) is met, that is the support vectors. As a result, only support vectors contribute to the characterisation of the solution, hence their name.

In more complex scenarios, data points are non-linearly separable in the feature space because the relationship between features is inherently non-linear in the first place. This relationship can be however linearised by enlarging the feature space using M-dimensional basis expansions $\{h_m(\boldsymbol{x})\},\ m = 1, ..., M$, such as polynomials, so that data points mapped into the higher-dimensional space become linearly separable. The hyperplane function expressed in equation (5.1) gets transposed in terms of a non-linear function

$$f(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x}) \cdot \boldsymbol{\beta} + \beta_0 \tag{5.9}$$

where $\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), ..., h_M(\boldsymbol{x}))$ is the feature vector expanded into the M-dimensional basis. From equation (5.8), with $\boldsymbol{h}(\boldsymbol{x})$ taking the place of $\boldsymbol{x_i}$, $f(\boldsymbol{x})$ can be written as

$$\begin{aligned} f(\boldsymbol{x}) &= \sum_{i=1}^{N} \alpha_i y_i \boldsymbol{h}(\boldsymbol{x}) \cdot \boldsymbol{h}(\boldsymbol{x_i}) + \beta_0 \\ &= \sum_{i=1}^{N} \alpha_i y_i K(\boldsymbol{x}, \boldsymbol{x}_i) + \beta_0 \end{aligned} \tag{5.10}$$

with $K(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{h}(\boldsymbol{x}) \cdot \boldsymbol{h}(\boldsymbol{x}')$ or, more generally, $K(\boldsymbol{x}, \boldsymbol{x}') = \langle \boldsymbol{h}(\boldsymbol{x}), \boldsymbol{h}(\boldsymbol{x}') \rangle$ being defined as *kernel*. Through this so called *kernel trick*, the optimisation problem does not depend explicitly on $\boldsymbol{h}(\boldsymbol{x})$, and only the knowledge of the kernel function is required. Popular kernels include dth-degree polynomials $K(\boldsymbol{x}, \boldsymbol{x}') = (1 - \langle \boldsymbol{x}, \boldsymbol{x}_i \rangle)^d$ and radial basis function (RBF) with Gaussian kernel $K(\boldsymbol{x}, \boldsymbol{x}') = e^{-\gamma \|x - x_i\|^2}, \gamma > 0$.

5.4 Random forest

Random forest (RF) is a modification of *bagging*, or *bootstrap aggregation*: an ensemble method that uses bootstrap samples of the dataset to train a collection of classifiers, and then returns the average (for regression) or the majority vote (for classification) prediction, thereby reducing the overall variance. This can be observed by considering B independent and identically distributed (i.i.d.) random variables with variance σ^2: their average has variance $\bar{\sigma}^2 = \sigma^2 / B$, which tends to zero as B increases. Decision trees represent ideal classifiers for bagging, since they can easily characterise complex structures within the data, as long as they are deep enough. However, such low bias also means that they are prone to overfit the training data, hence performing poorly on unknown data points. Averaging (or taking the mode of the prediction in the classification case) reduces not only variance of the prediction, but also the risk of overfitting, increasing the degree of generalisability of the classifier.

However, bagging trees are not independent, but only identically distributed (i.d.) with a positive pairwise correlation coefficient ρ. The variance of the average is therefore

$$\bar{\sigma}^2 = \rho \sigma^2 + \frac{1 - \rho}{B} \sigma^2 \tag{5.11}$$

which tends to the first term as B increases. In other words, the non-independency of the bagged trees sets a lower bound to the variance reduction.

RFs tries to compensate for that by reducing the correlation between trees, hence decreasing ρ. This is done by using a randomly selected subset of features as candidates for each splitting, instead of using all of them. In this way, each bagged tree is grown from an independent set of features. For classification, the amount of features randomly selected at each split is set in general to the square root of the total. At each node, the best variable with associated split-point, among those selected, is the one that best suits the splitting criterion. In classification tasks, it is usually either *Gini index*:

$$\sum_{k=1}^{K} p_{mk}(1 - p_{mk}) \tag{5.12}$$

or *cross-entropy*:

$$-\sum_{k=1}^{K} p_{mk} \log(p_{mk}) \tag{5.13}$$

They both measure nodes *impurity*, that is the amount p of observations of two or more classes $k = 1, 2, ...K$ within the same node m. For example, a node containing observations belonging only to one class will have an impurity measure of 0, that is the minimum, whilst a node with observations belonging to two classes in the same proportion will have a maximum impurity of 0.5. Previous studies have shown that the choice between Gini index and cross-entropy impurity measures has little to no effect on the performance of the final classifier [79].

In addition to its robustness to overfitting, RF are particularly useful also for their ability to automatically perform feature ranking, that is to assign to each feature a score corresponding to their relative contribution to the classification. Such *variable importance* score is defined by the improvement in the split-criterion attributed to the feature, accumulated over all the trees in the forest. Variable importance is normalised so that the sum of the importances across all features is equal to 1.

5.5 Neural Networks

Neural networks encompass a wide range of machine learning techniques employing one or multiple processing layers, also called *hidden* layers, whose output is passed further down the network and is not directly observed, hence the name. Each layer is composed of *neurons*, that is mathematical functions that take the output from neurons in the previous layers as input, and integrate them through a, usually, nonlinear[1] *activation* function. The output is then fed to the neurons in the next layer, following a hierarchical structure that resembles the functional organisation of the primary visual cortex.

Perhaps due to this biological correlate and the uncanny feeling spread by media about artificial intelligence, deep neural networks have taken over the machine learning scene,

[1] The non-linearity of activation functions is what effectively allows to build *deep* neural networks. Without nonlinear activation functions, the entire network would behave as a *single-layer perceptron* — a neural network whose input nodes are fed directly to the outputs — regardless of the number of hidden layers, as the combination of multiple linear functions is still a linear function.

both in terms of technical applications and popular perception. It is therefore useful to remember, in this writer's opinion, what has been noted by Hastie et al. (2009):

> *There has been a great deal of hype surrounding neural networks, making them seem magical and mysterious. As we make clear in this section, they are just nonlinear statistical models*. [76]

5.5.1 Single hidden layer neural network

An example of a basic single hidden layer neural network is shown in Figure 5.1. Assuming an input vector with L components $X = [X_1, ..., X_L]$ and target with K components $Y = [Y_1, ..., Y_K]$[2], the hidden feature vector $Z = [Z_1, ..., Z_M]$ and predicted output $\hat{Y} = [\hat{Y}_1, ..., \hat{Y}_K]$ components can be modelled as:

$$\begin{aligned} Z_m &= \sigma(\alpha_{0,m} + \alpha_m^T X), & m &= 1, ..., M \\ \hat{Y}_k &= g_k(\beta_{0,k} + \beta_k^T Z) = g_k(T_k), & k &= 1, ..., K \end{aligned} \tag{5.14}$$

with $\{\alpha_m, \beta_k\}$ the unknown network parameters, or *weights* (the $\{\alpha_{0,m}, \beta_{0,k}\}$ intercepts are referred to as *bias*), σ the activation function and g_k the output function. The activation function most used currently is the *rectified linear unit* (ReLU), defined as:

$$\sigma(x) = \text{ReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{else} \end{cases} \tag{5.15}$$

or, equivalently, $\text{ReLU}(x) = \max(0, x)$, due to its computational efficiency compared to a *sigmoid* $\sigma(x) = 1/(1 + e^{-x})$, which was the standard beforehand. Alternatively, if setting the input equal to 0 for negative input values results in poor learning performances, a *leaky* ReLU might be implemented instead as $\max(\alpha x, x)$, with α being a small scaling factor, e.g. $\alpha = 0.01$. Potential advantages of leaky ReLU over standard ReLU depend on the specific deep learning application, as no significant differences in training performance between the two have been reported in general [80].

As output function, the *identity* $g_k(T_k) = T_k$ is usually chosen for regression; for K-class probabilistic classification, the *softmax* function is often selected:

$$g_k(T_k) = \frac{e^{T_k}}{\sum_{k=1}^{K} e^{T_k}} \tag{5.16}$$

[2] E.g.: classification task with subject records composed of L clinical measurements classified into one of K groups; regression task with input and target MR-images composed of L and K voxels respectively.

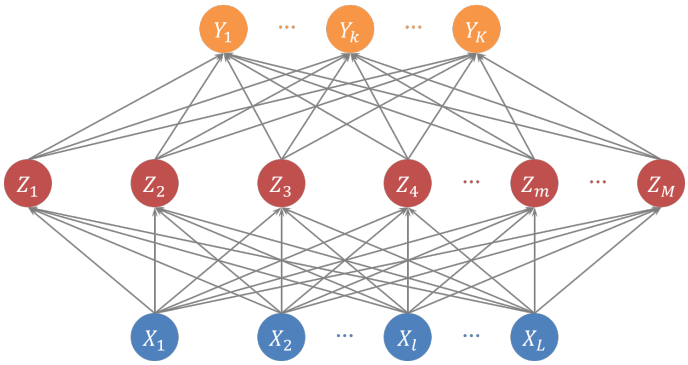


Figure 5.1: Single hidden layer neural network

from which a *hard* classification can be obtained by calculating $\text{argmax}_k(g_k(T_k))$.

5.5.2 Back-propagation

The aggregate set of unknown weights θ, consisting of

$$\begin{aligned} \{\alpha_{0,m}, \alpha_m = [\alpha_{1,m}, ..., \alpha_{P,m}]; m = 1, ..., M\}, \quad & M(P+1) \text{ weights} \\ \{\beta_{0,k}, \beta_k = [\beta_{1,k}, ..., \beta_{M,k}]; k = 1, ..., K\}, \quad & K(M+1) \text{ weights} \end{aligned} \tag{5.17}$$

can be determined by fitting the model to the training data. To do so, an objective, or *loss*, function $f(\theta)$ is defined to quantify the distance between the network predicted output and the target data, given the current set of parameters. Whilst the loss function can assume any form that best suits the specific task, *squared error*[3] over training data $(X^i, Y^i),\ i = 1, ..., N$ can be used as a general example:

$$f(\theta) = \sum_{i=1}^{N} f^i(\theta) = \sum_{i=1}^{N} \sum_{k=1}^{K} (Y_k^i - \hat{Y}_k^i)^2 \tag{5.18}$$

Loss function minimisation (which in this case is reduced to a standard *least-squares* problem) can be achieved via *gradient descent*, which in this context follows the so-called *back-propagation* algorithm. The derivatives of each term $f^i(\theta)$ with respect to the weights are:

[3]Denoting the squared error as SE, other standard loss functions include *mean squared error* $\text{MSE} = \text{SE}/N$, *mean absolute error* $\text{MAE} = \sum_{i=1}^{N} |Y^i - \hat{Y}^i|$ and *root mean squared error* $\text{RMSE} = \sqrt{\text{MSE}}$. MAE and RMSE are also sometimes referred to, improperly, as L_1 and L_2 loss, respectively.

$$\begin{aligned}\frac{\partial f^i}{\partial \beta_{m,k}} &= -2(Y_k^i - \hat{Y}_k^i) g_k'(\beta_k^T Z^i) Z_m^i = \delta_k^i Z_m^i \\ \frac{\partial f^i}{\partial \alpha_{l,m}} &= -\sum_{k=1}^{K} 2(Y_k^i - \hat{Y}_k^i) g_k'(\beta_k^T Z^i) \beta_{m,k} \sigma'(\alpha_m^T X^i) X_l^i = s_m^i X_l^i\end{aligned} \tag{5.19}$$

where the defined quantities δ_k^i, s_m^i are the *errors* at the level of the output and hidden layer, respectively. From their definition, the errors satisfy the relation:

$$s_m^i = \sigma'(\alpha_m^T X^i) \sum_{k=1}^{K} \beta^{m,k} \delta_k^i \tag{5.20}$$

called *back-propagation equation*. In the general case of a neural network with $p = 1, ..., P$ layers, equation (5.20) is still valid, as it applies to any given $(p-1) \to p$ pair of connected layers. The back-propagation equation enables therefore to recursively express the error at level $(p-1)$ as a function of the error at the p-th level and thus, once the error at the level of the P-th layer is calculated, to efficiently[4] back-propagate the errors throughout all layers via simple, 1-to-1 *local* updates: this is the heart of back-propagation.

5.5.3 Network fitting

Network fitting to the training data follows an iterative process that terminates on convergence. Given the derivatives of the loss function, weights at the $(r+1)$-th iteration can be updated via gradient descent as

$$\begin{aligned}\beta_{m,k}^{(r+1)} &= \beta_{m,k}^{(r)} - \gamma^{(r)} \sum_{i=1}^{N} \frac{\partial f^i}{\partial \beta_{m,k}^{(r)}} \\ \alpha_{l,m}^{(r+1)} &= \alpha_{l,m}^{(r)} - \gamma^{(r)} \sum_{i=1}^{N} \frac{\partial f^i}{\partial \alpha_{l,m}^{(r)}}\end{aligned} \tag{5.21}$$

where the update coefficient at the r-th iteration $\gamma^{(r)}$ is referred to as *learning rate*. More sophisticated update rules can be implemented to further optimise the gradient descent, such as gradient descent with *momentum* or *adaptive learning*.

The weight updates in equations (5.21) take into consideration the entire training set, which is referred to as *batch gradient descent*. Whilst conceptually straightforward

[4] Notice how the use of ReLU activation reduces the derivative of the activation function in equation (5.20) to $\sigma'(x) = 0$ if $x < 0$, and $\sigma'(x) = 1$ otherwise.

and computationally stable, batch gradient descent is also computationally expensive in terms of memory usage. Alternatively, updates can be performed on *mini-batch* of N elements, called *mini-batch gradient descent*, or one training element at a time (remove the summation symbols in equations (5.21)), until all elements in the training set are picked: since elements are usually chosen at random, this method is called *stochastic gradient descent*. The set of training iterations required to go through the entire training set defines one *training epoch*. The learning rate for batch gradient descent is usually a constant, whilst it is set to decrease — learning rate *annealing* — for mini-batch and stochastic gradient descent (e.g. $\gamma^{(r)} \propto 1/r$), although more advanced optimisation processes can do it internally.

Upon initialising the network weights[5], the back-propagation method follows a *two-pass* algorithm:

1. **forward pass**: use equation (5.14) to calculate $\hat{Y}_k^i$ given the current weights;
2. **backward pass**: calculate the errors δ_k^i at the output layer and back-propagate them via equation (5.20) to obtain the errors at all hidden layers; use the errors to compute the gradients in equation (5.19), and use them to update the weights via equation (5.21).

5.6 Convolutional neural network

Convolutional neural networks (CNN) are a class of neural networks that have revolutionised the field of *machine vision*, that is the automatic recognition of elements and patterns in images, through the use of stacked convolutional layers. The output of each convolutional layer is a tensor $\mathbf{g}$ defined element-wise as the dot product between the input tensor $\mathbf{f}$ and a sliding convolutional kernel, or *filter*, $\mathbf{k}$ centred on that element. For a 3×3 2D-filter kernel, this would be:

$$g_{ij} = (\mathbf{k} * \mathbf{f})_{ij} = \sum_{l=-1}^{1} \sum_{m=-1}^{1} k_{l,m} f_{(i+l),(j+m)}, \quad \text{with} \quad \mathbf{k} = \begin{bmatrix} k_{-1,-1} & k_{0,-1} & k_{1,-1} \\ k_{-1,0} & k_{0,0} & k_{1,0} \\ k_{-1,1} & k_{0,1} & k_{1,1} \end{bmatrix} \tag{5.22}$$

[5]Weights can be initialised to random values, zeroes, or informed by previous learning. The latter case is called *pre-training*, where a network trained on a similar, previous task is used as starting point for the current task, referred to as *fine-tuning*. Generally speaking, using pre-trained networks for weight initialisation for relatively different tasks is a form of *transfer-learning*.

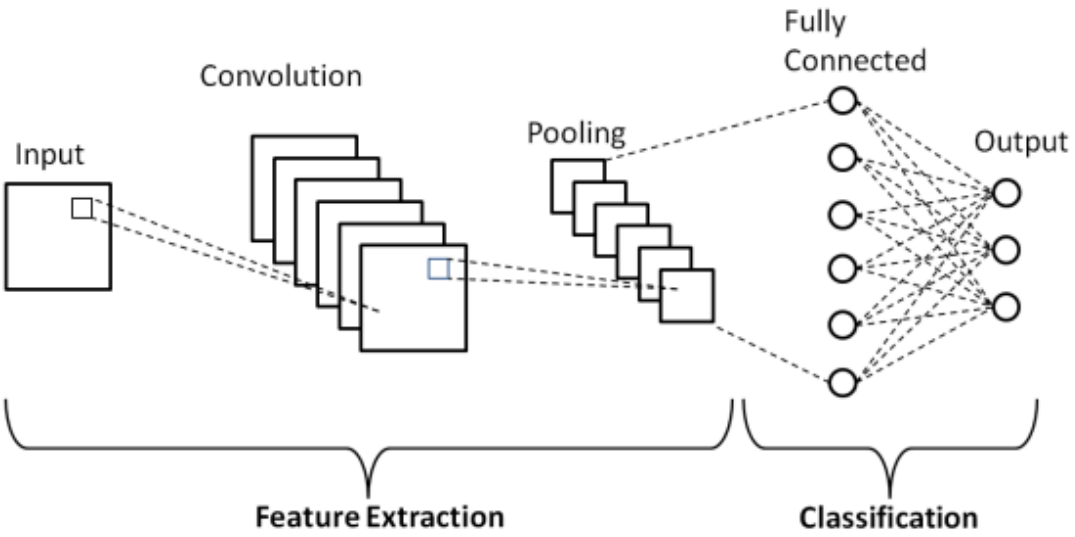


Figure 5.2: Example of basic CNN. In the convolutional layer, a sliding filter kernel is convolved with the input tensor, producing a number of output channels equal to the number of filters applied. The output of the convolutional layer is then down-sampled through pooling. Additional convolutional/pooling layers can be added for deeper networks. The output of the *feature extraction* block is then fed to a fully connected layer to perform the learning task, e.g. classification. Figure from Phung & Rhee (2019) [81].

Through convolution ($*$), the input image gets converted to a *feature map* defined by the specific filter employed in the convolution. Multiple filters, each mapping a different feature of the data, can be used within the same convolutional layer, resulting in a number of *output channels* equal to the number of filter kernels applied. To learn progressively more global, high-level features, the output is then reduced in size, or *down-sampled*, throughout the network architecture by setting a kernel *stride* — the step-size the kernel undergoes whilst sliding over the input image — greater than 1, or using *pooling* layers to merge the information of neighbouring tensor elements into one, usually through average- or max-pooling. The result of each layer is then fed to next layer, which enables to process the data through levels of increasing abstraction. An example of a basic 2D-CNN is shown in Figure 5.2.

The filter kernel weights are learned directly from the data through back-propagation, upon opportunely adapting the updating rules to the CNN model. Each filter kernel is therefore learnt to match a more or less abstract feature of the training data which is particularly relevant for the specific task. In the context of image recognition, trained kernels may reflect more or less complex topological structures such as lines, curves, all the way up to abstract, archetypal patterns.

5.7 U-Net

U-Net is an application of CNN for automatised biomedical image segmentation, first proposed by Ronneberg et al. (2015) [82], and counting almost 20 thousand citations at the time of writing. U-net uses a CNN encoding pathway to learn the high-level features of the image, followed by a decoding pathway to bring the output back to image space. *Shortcut connections* between the encoding and the decoding pathways allow to recover the spatial information lost due to the down-sampling. For a more detailed description of the U-Net architecture, see section 11.6.

5.8 Machine learning in MS

Machine learning algorithms have been proven a useful tool in the study of MS. Some of the results obtained through the use of different methods and architectures are hereby presented.

5.8.1 Support vector machine

The first study about machine learning on MS was published in 2012 by Bendfeldt et al. (2012) [83]: a SVM was trained over a dataset of grey matter segmentation features to distinguish MS patients at different disease stages (task I: *early* versus *late*), with different white matter lesion loads (task II: *low* versus *high*) and MS types (task III: *benign* versus *non-benign*). An accuracy of 85%, sensitivity of 82.3% and specificity of 88.2% was reached on task I; 83%, 85% and 80% respectively on task II; 77%, 76.9% and 76.9% respectively on task III.

Similar studies with increasing features- and classification-complexity followed. In a work published by Wottschel et al. (2015) [84], SVMs were applied to lesion-based features, including lesion count, load, centrality, size profile and average proton density and T_2 intensity in the lesions to predict the conversion (or non conversion) of CIS patients to clinically definite MS within 1 and 3 years from the first episode. The classification at 1 year reached an accuracy of 71.4%, with a sensitivity of 77% and a specificity of 66%, whilst performances at three years were respectively 68%, 60% and 76%. The most relevant features for classification at 1 year were the type of presentation, that is whether the CIS episode affected the spinal cord, the optic nerve or other anatomical regions of the CNS, gender and lesion load; on the other hand, lesion

count, PD average intensity in the lesions, lesions centrality, EDSS and age showed the highest predictive power at 3 year. In particular, the distance of the lesions from the vertical axis of the brain appeared to be an important predictor for conversion, with a shorter distance being associated to more probable future attacks.

5.8.2 Random forest

A work by Eshaghi et al. (2016) [85], used the RF ensemble classification algorithm to distinguish between MS patients and those experiencing neuromyelitis optica (NMO), a condition causing inflammation and demyelination of the optic nerve and spinal cord, whose clinical and MRI characteristics are very similar to MS.

The dataset was built around grey matter-based metrics obtained from MRI acquisitions as part of a clinical protocol. They included thickness, volume and surface area of cortical regions of interest (ROIs), and volume of basal ganglia, for a total of 157 features. Furthermore, since the data were acquired in two different centres, an additional variable *centre* was added to check whether the results were independent on the origin of the data and, therefore, reproducible on different sites. The aim of the classifier was to distinguish between MS versus NMO patients (task I), MS versus healthy controls (task II) and NMO versus healthy controls (task III). For task I, an accuracy of 74%, sensitivity of 77% and specificity of 72% were reached. For task II, performances were 92%, 94% and 90% respectively, whilst for task III they were 88%, 89% and 88% respectively. The volume of deep grey matter structures and the thickness of the insular cortex, which appeared to be reduced in MS patients compared to NMO, showed the greatest prediction power in task I. For task II and III, the volumes of the parahippocampal gyri, the middle frontal gyrus (for task II) and the superior temporal gyrus (for task III) showed the highest importance in the classification task. As expected, the variable *centre* showed the lowest importance among all the features for all tasks.

5.8.3 Convolutional neural network

First proposed by Krizhevsky et al. (2012) [86] for ImageNet[6] classification, it counts, at the time of writing, more than 72 thousand citations. Applications in biomedical imaging were delayed by a few years, due to the scarcity of similar large, publicly available datasets at the time. A recent example in MS is the one proposed by Zhang

[6] A freely accessible visual database containing, to date, almost fifteen million labelled images [87]

et al. (2018), using a 10-layers deep CNN for patients vs healthy control classification, achieving an accuracy, sensitivity and specificity of 98% on all three accounts [88].

5.8.4 U-Net

In MS, Brosch et al. (2016) have published several lesion segmentation works between 2014 and 2016, the most recent of which used a 3D U-net which showed comparable performances with the best state-of-the-art methods, and outperformed freely available popular segmentation methods on a large MS clinical trial dataset [89].

5.8.5 Deep learning model fitting

On the regression front, a variety of deep learning model fitting algorithms have been proposed. This can be virtually applied to any model fitting problem, with the trained network substituting the traditional model fitting on new data. Whilst this process does not add any new information, as the predicted output could be produced through traditional model fitting anyway, and with the network performing *at best* just as well, it is a great way to cut down on post-processing times *if* enough data variability is provided during training.

A recent example that can be applied to MS is T_2-fitting for the purposes of myelin water imaging using neural networks, as done by Lee et al. (2020) [90]. Another deep learning regression application is what has been defined by Alexander et al. (2017) as *image quality transfer* (IQT), whose aim is:

> *To transfer the rich information available from one-off experimental medical imaging devices to the abundant but lower-quality data from routine acquisitions* [91].

Whilst specifically applied to diffusion data, the IQT framework and basic concept can be applied to any pair of images that are not explicitly related by an underlying model, and is not limited only to a *super-resolution* task.

5.8.6 Other applications

Applied to a big dataset of MS patients (6322 for training, 3068 for testing), SuStaIn has identified three MS subtypes based on different patters of tissue alteration over time, defined by Eshaghi et al. (2021) as cortex-led, normal-appearing white matter-led,

and lesion-led [78]. Patients in the lesion-led category have shown the highest risk of disability progression and relapse rate, as well as positive response to treatment in selected trials. These results may be of particular interest not only for prediction of disability accumulation, but also to define different MS categories that better reflect disease evolution.

In addition to MRI-based metrics, clinical scales and patients-reported outcomes have been used as features in the classification task as well. A work by Fiorini et al. (2015) [92] used questionnaires and anthropometric measures to train a linear classifier to distinguish RRMS from progressive and benign MS forms. After a phase of feature selection, the classifier reached an accuracy of 78%, over a baseline of 63% for the same task. Despite this last results are not exceptional, the inexpensiveness, non-invasiveness and ease of acquiring such features make them worth to be considered as part of a more extended dataset.

Part II

MyRelax:
myelin and relaxation imaging

Chapter 6

Introduction

The myelin relaxation — *MyRelax* — framework[1] allows to extract quantitative relaxometry measurements, as well as indirect myelin content indicators, from routine qualitative images commonly used for lesion delineation and anatomical purposes through a traditional model fitting approach. In this study, the MyRelax framework was used to extract quantitative proton density (PD), T_2 and T_1 maps, together with macromolecular tissue volume (MTV) maps from PD-, T_2- and T_1-weighted qualitative scans, which have represented for decades a staple in clinical MRI research. Being able to extract quantitative information from qualitative data would provide great statistical power through the implementation of quantitative studies on readily available retrospective historical datasets, as well as represent a key step towards the future of sustainable research.

The *QuaSI-* prefix (*qualitative scans for indirect-*) has been introduced to better distinguish quantitative data produced indirectly from qualitative scans (e.g. QuaSI-PD produced via MyRelax), from their counterparts acquired through dedicated MRI sequences.

For *MyRelax validation*, MyRelax was applied to a *prospective* cohort of healthy controls, and the resulting QuaSI-PD, -T_2, -T_1 maps were compared with the respective maps obtained by fitting data acquired through quantitative MRI sequences.

For *MyRelax MS application*, QuaSI-MTV maps and T_1-/T_2-weighted ratio maps, useful indicators for brain myelin content, were compared with magnetisation transfer ratio (MTR) maps on a *retrospective* dataset of both healthy controls and MS patients, to assess whether qualitative scans can be used to produce surrogates for myelin imaging.

[1] Courtesy of Dr Francesco Grussu

Chapter 7

Methods

7.1 Cohort

For this study, two datasets were employed, both including qualitative PD-, T_2- and T_1-weighted images acquired using identical scanner parameters.

7.1.1 Prospective cohort: MyRelax validation

The first dataset, of only healthy controls (HC), was composed of qualitative data acquired with conventional sequences, and quantitative data acquired using gold standard sequences for PD and relaxometry measurement. It was used for the *MyRelax validation* objective, and consisted of 3 HC (2 men, age: 27±2 years old), with one repetition, acquired *prospectively*, i.e. specifically for this application. This dataset will be referred to simply as *MyRelax validation cohort*.

7.1.2 Retrospective cohort: MyRelax MS application

The second, richer, dataset named *GML02 cohort* included both HC and MS patients. It was already employed for several other studies — e.g. Pardini et al. (2016) [93] — and is therefore composed of *retrospective* data. The MRI modalities used in this study were qualitative PD-, T_2- and T_1-weighted scans, and MTR acquired via a standard sequence. The cohort consisted of 34 HC (16 men, age: 38±11 years old) and 109 MS patients (34 men, age: 48±11 years old) with different MS phenotype and disease duration. The cohort included 2-years follow-up scans which were treated as independent acquisitions to increase the dataset statistical power.

7.2 MRI protocol

All MRI data were acquired on a 3T Philips Achieva scanner with a 32-channels head-coil. This study was approved by the local ethical committee.

7.2.1 MyRelax validation protocol

For *MyRelax validation*, the qualitative scans included:

1. Qualitative scans:
   (a) **PD/T2**. Dual-echo 2D PD-/T_2-weighted turbo spin-echo (TSE).
   (b) **T1**. 2D T_1-weighted spin-echo (SE).
2. Anatomical scan: **3DT1**. 3D sagittal T_1-weighted MP-RAGE: magnetisation-prepared rapid gradient echo (GRE).

The following gold standard specialised sequences were then acquired:

1. Quantitative PD ground truth:
   (a) **vFA-GRE**. 3D spoiled GRE at two different flip angles.
   (b) **ME-GRE**. Multi-echo 3D GRE.
   (c) **B1**. Dual-TR flip angle map.
2. Quantitative T_2 ground truth: **ME-SE**. Multi-echo 2D SE.
3. Quantitative T_1 ground truth:
   (a) **IR-EPI**. 2D echo-planar imaging (EPI) inversion recovery with non spatially selective adiabatic inversion pulse, where slice shuffling mechanism and single shot EPI readout were used to speed up the acquisition as in Clare et al. (2001) [94].
   (b) **blip-up/down**. 2D EPI images acquired with both phase-encoding blips *up* and *down* to correct maps with EPI readout for susceptibility induced distortions.

Details on the *MyRelax validation* MRI protocol are reported in Table 7.1.

Table 7.1: *MyRelax validation* MRI protocol.

	scan time	Res [mm]	FOV [mm]	slices	slice orientation	sequence	TE [ms]	TR [ms]	TI [ms]	flip angle[°]
Qualitative scans										
PD/T2	4:02	RL = 1 AP = 1 FH = 3	RL = 240 AP = 240 FH = 150	50	transverse	TSE	19/85	3500		90
T1	5:43	RL = 1 AP = 1 FH = 3	RL = 240 AP = 240 FH = 150	50	transverse	SE	10	625		90
Anatomical scan										
3DT1	6:32	RL = 1 AP = 1 FH = 1	RL = 180 AP = 256 FH = 256	180	sagittal	GRE	3.1	6.9	823	8
Ground-truth										
vFA-GRE	∼12:12	RL = 1 AP = 1 FH = 1	RL = 170 AP = 240 FH = 256	170	sagittal	GRE	2.4	30		4/25
ME-GRE	6:06	RL = 1 AP = 1 FH = 1	RL = 170 AP = 240 FH = 256	170	sagittal	GRE	2.4−14.4 Δ = 2.4 6 TEs	30		25
B1	2:56	RL = 6 AP = 4 FH = 4	RL = 170 AP = 240 FH = 256	29	sagittal	GRE	2.3	30/150		60
ME-SE	19:25	RL = 1 AP = 1 FH = 3	RL = 180 AP = 240 FH = 150	50	transverse	SE	14−112 Δ = 14 8 TEs	3500		90
IR-EPI	∼12:00	RL = 2.5 AP = 2.5 FH = 2.5	RL = 192 AP = 222 FH = 125	50	transverse	EPI	34	12000	40−4000 Δ = 440 10 TIs	90
blip-up/down	1:36	RL = 2.5 AP = 2.5 FH = 2.5	RL = 220 AP = 220 FH = 125	50	transverse	EPI	42	8000		90

7.2.2 MyRelax MS application protocol

For *MyRelax MS application*, the qualitative scans consisted of the same PD/T2 and T1 reported above, which were used to produce QuaSI-MTV maps and T_1w/T_2w. A 3DT1 was also acquired for lesion-filling and tissue segmentation.

The magnetisation transfer protocol — **MT** — consisted of two dual-echo 3D-GRE, with and without a MT saturation pulse, as described in Pardini et al. (2016):

> *High-resolution magnetisation transfer imaging using a 3D slab-selective FFE sequence with two echoes:* $1 \times 1 \times 1\,\text{mm}^3$*, repetition time* $= 6.4\,\text{ms}$*, echo time* $= 2.7/4.3\,\text{ms}$*,* $\alpha = 9°$ *with and without sinc Gaussian-shaped magnetisation transfer pulses of nominal* $\alpha = 360°$*, offset frequency* $1\,\text{kHz}$*,*

duration 16 ms. *A turbo field echo (TFE) readout was used, with an echo train length of four, TFE shot interval* 32.5 ms, *giving a total time between successive magnetisation transfer pulses of* 50 ms, *and scan time of 25 min.* [93]

Details on the GML02 MRI protocol are reported in Table 7.2.

Table 7.2: GML02 MRI protocol.

	scan time	Res [mm]	FOV [mm]	slices	slice orientation	sequence	TE [ms]	TR [ms]	TI [ms]	flip angle[°]
Qualitative scans										
PD/T2	4:02	RL = 1 AP = 1 FH = 3	RL = 240 AP = 240 FH = 150	50	transverse	TSE	19/85	3500		90
T1	5:43	RL = 1 AP = 1 FH = 3	RL = 240 AP = 240 FH = 150	50	transverse	SE	10	625		90
Anatomical Scan										
3DT1	6:32	RL = 1 AP = 1 FH = 1	RL = 256 AP = 256 FH = 180	180	sagittal	GRE	3.1	6.9	823	8
Ground-truth										
MT	∼25:00	RL = 1 AP = 1 FH = 1	RL = 256 AP = 256 FH = 180	180	sagittal	GRE	2.7/4.3	6.4	823	9

7.3 MyRelax validation

7.3.1 Preprocessing

For the MyRelax cohort, brain segmentation was performed on the anatomical 3DT1 using UCL software *geodesic information flow* (GIF) [95]. Registrations were performed using the NiftyReg software package [96]. Data acquired with EPI readout was corrected for susceptibility-induced distortions applying the FSL `topup` [97, 98] tool to the blip-up/down pair of images. The FSL toolbox was extensively used throughout all studies for data visualisation and general image manipulation. Preprocessing was performed locally using MATLAB 2017b–2019b [99].

7.3.2 Image analysis: MyRelax framework

The MyRelax framework enables to calculate QuaSI-PD, -MTV and quantitative QuaSI-T_2 and -T_1 maps voxel-wise from a set of three qualitative images with different

Table 7.3: Average PD, T_2, T_1 in brain

	PD[p.u.] [101]	T2[ms]	T1[ms]
GM	0.78-0.82	100	810
WM	0.70	90	680
CSF	1	3000	3000

MR-contrast (three TEs, two TRs) through a simple three-points fitting approach. The fitting is performed *analytically*, rather than through traditional least-square optimisation, making it computationally inexpensive and fast to run. The framework was initially implemented in Python 2.7, and subsequently adapted to Python 3 [100].

Bloch equations

QuaSI-PD, -T_2, -T_1 maps were computed from the PD/T2 TSE and T1 SE qualitative scans by solving the associated Bloch equations in each voxel:

$$\begin{aligned} S_{\text{PD}} &= S_0 e^{-\text{TE}_1/T_2}(1 - e^{-\text{TR}'_1/T_1}) \quad (a) \\ S_{T_2} &= S_0 e^{-\text{TE}_2/T_2}(1 - e^{-\text{TR}'_1/T_1}) \quad (b) \\ S_{T_1} &= S_0 e^{-\text{TE}_3/T_2}(1 - e^{-\text{TR}_2/T_1}) \quad (c) \end{aligned} \tag{7.1}$$

where S_{PD}, S_{T_2} and S_{T_1} represent PD-, T_2- and T_1-weighted MRI signal in any given voxel, whilst S_0, T_2 and T_1 are the unknown *apparent* PD, and quantitative transverse and longitudinal relaxation times respectively, i.e. QuaSI-T_2 and QuaSI-T_1. The *apparent* PD differs from a proper PD map because corrupted by the receiver *bias field* produced by the spatial inhomogeneities in the receiver coil sensitivity profile (see section 4.9). The framework also provides a means to correct for bias field, allowing to obtain quantitative PD maps as well, i.e. QuaSI-PD. Typical values of PD, T_2 and T_1 in the grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) at 3T are reported in Table 7.3.

TE_i is the effective echo time[1] associated to the i-th MRI contrast: $\text{TE}_1 = 19$ ms, $\text{TE}_2 = 85$ ms, $\text{TE}_3 = 10$ ms. $\text{TR}'_1 = \text{TR}_1 - \text{TF} \times \text{ES}$ corresponds to the effective PD/T2 repetition time accommodated for the TSE acquisition, as defined by Rydberg et al. (1995) [102], with $\text{TF} = 10$ being the TSE turbo factor, $\text{ES} = 9.4$ ms the echo spacing, and $\text{TR}_1 = 3500$ ms. The repetition time associated to the T_1-weighted spin echo, $\text{TR}_2 = 625$ ms, did not require any adjustment for the lack of a turbo factor.

[1] In the case of TSE, the effective TE indicates the echo time corresponding to the central line of the k-space (see section 4.15).

QuaSI-T_2

With reference to equation (7.1), T_2 is analytically computed from equations (a) and (b), which present the same TR:

$$T_2 = \frac{\mathrm{TE}_1 - \mathrm{TE}_2}{\log(S_{T_2}/S_{\mathrm{PD}})} \tag{7.2}$$

QuaSI-T_1

Once T_2 is known, T_1 can be calculated by numerically finding the convergence value of the iterative function

$$T_1^{n+1} = -\frac{\mathrm{TR}_2}{\log(1 - m(1 - e^{-\mathrm{TR}_1/T_1^n}))} \tag{7.3}$$

obtained by rearranging the ratio between equations (c) and (a), where n is the iteration index, and

$$m = \frac{S_{T_1}}{S_{\mathrm{PD}}} e^{-(\mathrm{TE}_3 - \mathrm{TE}_1)/T_2} \tag{7.4}$$

QuaSI-PD

S_0 is then calculated by substituting equations (7.2) and (7.3) into any one of equations (7.1) (a)–(c). In order to correct for the bias field, the apparent PD S_0 can be expressed as a function of the true PD and the receiver coil sensitivity profile (RP):

$$S_0 = k\mathrm{RP} \cdot \mathrm{PD} \tag{7.5}$$

where k is a spatially invariant scaling constant. Receiver bias field correction was performed as described by Volz et al. (2012) [103], using the T_1 map calculated in the previous step to take advantage of the linear correlation, frequently reported in literature, between the inverse of T_1 and PD:

$$\frac{1}{\mathrm{PD}} \approx A + \frac{B}{T_1} \tag{7.6}$$

This relationship holds only in *normal appearing* WM (NAWM) and GM, whilst it is not generally true in CSF, and neither it has been proven to hold in MS lesions. For this reason, both CSF and lesions were excluded from the following steps.

The parameters A and B are calculated within a recursive algorithm composed by four steps:

1. At each iteration i, a *pseudo* PD map (pPD_i) is computed voxel-wise in NAWM and GM using (7.6):

$$\frac{1}{\text{pPD}_i} = A_i + \frac{B_i}{T_1} \tag{7.7}$$

   where A and B have been initialised to $A_0 = 0.916$ and $B_0 = 436\,\text{ms}$, as described in Volz's paper. Forgoing the scaling constant in (7.5), an approximation of the RF field is thus calculated as

$$\text{RP}'_i = \frac{S_0}{\text{pPD}_i} \tag{7.8}$$

2. Since RP'_i values are calculated in NAWM and GM only, 3D polynomial fitting is used to smooth the map over the whole brain, and obtain a new guess for the RP map at each iteration, denoted as RP_i.

3. A candidate for the true PD map is then calculated using equation (7.5) as

$$\text{PD}_i = \frac{S_0}{\text{RP}_i} \tag{7.9}$$

   Since k in equation (7.5) is unknown, PD_i is rescaled to the median CSF intensity within the ventricles, so that the median value of PD_i in the ventricles is 1.

4. Finally, A_i and B_i are linearly fitted from the values of PD_i and T_1 using (7.6).

The algorithm is iterated until A_i and B_i values converge, which usually happens within about 5 to 7 iterations. The final RP and QuaSI-PD maps are the ones generated in the last iteration.

7.3.3 Image analysis: ground truth

QuaSI-PD, -T_2 and -T_1 maps calculated using the MyRelax framework were correlated to the corresponding ground truth, calculated as follows.

Quantitative T_2

T_2 maps were calculated by exponentially fitting the ME-SE data in each voxel:

$$S_{\mathrm{ME}} = S_0^{T_1} e^{-\mathrm{TE}/T_2} \tag{7.10}$$

with respect to T_2, where $S_0^{T_1} = S_0(1 - e^{-\mathrm{TR}/T_1})$ is the T_1-weighted apparent PD, which was not used. The first echo was excluded from the fit in order to reduce the effects of stimulated echoes.

Although skipping the first echo has been shown not to be the best adjustment for exponential T_2-fitting [104], and more sophisticated techniques should be employed instead, such as *extended phase graphs* (EPG), ME-T_2 mapping is not the main focus of the study, and exponential fitting has shown to be sufficient for assessing the degree of reproducibility and correlation with QuaSI-T_2 maps. Furthermore, when testing for the effect of EPG for T_2 fitting on a single subject, no particularly meaningful differences were observed with the exponential fitting (differences were around 5% within tissue). Therefore, in the assumption that the errors caused by the simpler fitting method are reproduced similarly across subjects, exponential fitting was chosen to produce the ground truth T_2 map.

Quantitative T_1

Ground truth T_1 maps were calculated from distortion-corrected IR-EPI images by fitting the model

$$S_{\mathrm{IR}} = S_0^{T_2} |1 - 2e^{-\mathrm{TI}/T_1}| \tag{7.11}$$

with respect to T_1, where TI is the inversion time, whilst $S_0^{T_2} = S_0 e^{-\mathrm{TE}/T_2}$ is the T_2-weighted apparent PD, which was left unused.

Quantitative PD

PD was extracted from the vFA-GRE images by solving the Bloch equations in each voxel

$$\begin{aligned} S_{\mathrm{FA}_1} &= S_0 e^{-\mathrm{TE}_1/T_2^*}(1 - e^{-\mathrm{TR}/T_1}) \frac{\sin(\mathrm{FA}_1)}{1 - \cos(\mathrm{FA}_1) e^{-\mathrm{TR}/T_1}} \\ S_{\mathrm{FA}_2} &= S_0 e^{-\mathrm{TE}_1/T_2^*}(1 - e^{-\mathrm{TR}/T_1}) \frac{\sin(\mathrm{FA}_2)}{1 - \cos(\mathrm{FA}_2) e^{-\mathrm{TR}/T_1}} \end{aligned} \tag{7.12}$$

with respect to T_1:

$$T_1 = \mathrm{TR}/\log\left(\frac{\cos(\mathrm{FA}_2) - m\cos(\mathrm{FA}_1)}{1 - m}\right) \tag{7.13}$$

where FA_1 and FA_2 are the two *effective* flip angles, given by the product between the nominal flip angles (4° and 25°) and the value of the B_1 map in each voxel, and

$$m = \frac{S_{\mathrm{FA}_1}}{S_{\mathrm{FA}_2}} \frac{\sin(\mathrm{FA}_2)}{\sin(\mathrm{FA}_1)} \tag{7.14}$$

T_2^* can be calculated from the ME-GRE using exponential fitting and, once T_1 and T_2^* have been calculated, S_0 can be computed in each voxel using either equation in (7.12). After correcting for the receiver bias field as described previously, the ground truth PD map can be extracted.

7.3.4 Statistical analysis

MyRelax and ground truth maps were rigidly registered onto 3DT1 space to use the same high-resolution segmentation maps for all modalities. In order to correlate QuaSI-PD, -T_2 and -T_1 maps obtained with the MyRelax method to the corresponding ground truth, the histogram of each map in WM, pons, cortical and deep GM was interpolated with a nonparametric kernel-smoothing distribution, and the peak value computed. Linear fitting was run and Pearson correlation coefficient calculated.

The following MATLAB functions were employed:

- `histfit`, for nonparametric kernel-smoothing;
- `fitlm`, for linear fitting;
- `corrcoef`, for Pearson correlation coefficient calculation.

7.4 MyRelax MS application

7.4.1 Preprocessing

For the GML02 cohort, lesions were identified and semi-automatically segmented by an experienced clinician[2] on the PD-weighted images prior to segmentation, using JIMv6.0 [105], and then linearly registered to anatomical 3DT1 space (see Pardini et al. (2016) [93]). 3DT1 images were then lesion-filled [106]. Brain tissue segmentation was performed on the 3DT1 using the GIF package. Cross-modality non-linear registration [107] to MNI [108] space at $2 \times 2 \times 2\,\mathrm{mm}^3$ resolution was performed for voxel-wise statistical analysis. Lesion filling, brain segmentation and registrations were run on the medical image data management tool XNAT [109], following standardised pipelines for MS data post-processing.

[2] Courtesy of Dr Declan T. Chard.

7.4.2 Image analysis

QuaSI-MTV and T_1-/T_2-weighted ratio (T_1w/T_2w) were calculated from the qualitative scans and compared with MTR. Examples are shown in Figure 7.1.

QuaSI-MTV

QuaSI-PD maps were produced from the PD-, T_2- and T_1-weighted images using MyRelax. QuaSI-MTV was then calculated as $1 - \text{PD}$.

T_1-/T_2-weighted ratio

T_1w/T_2w maps were calculated on each subject by simply performing the ratio between the two maps voxel-wise.

MTR

MTR was used as the reference — indirect — metric for myelin content. For each MT-scan, i.e. with and without the MT saturation pulse, the two echoes were averaged to increase signal-to-noise ratio (SNR), producing a pair of images MT^{on} and MT^{off}, respectively [93]. MTR maps were then calculated voxel-wise as

$$\text{MTR} = \frac{\text{MT}^{\text{off}} - \text{MT}^{\text{on}}}{\text{MT}^{\text{off}}} \tag{7.15}$$

A $1 \times 1 \times 3\,\text{mm}^3$ mean-filter, matching the qualitative scans resolution, was applied to further increase axial SNR, whilst preserving original resolution.

7.4.3 Statistical analysis

The degree of voxel-wise matching information between modalities was assessed by registering QuaSI-MTV, T_1w/T_2w and MTR maps of all subject to MNI space at $2\times2\times2\,\text{mm}^3$ resolution. The similarity of patterns of alterations between HC and MS patients in normal appearing tissue for each modality were assessed by performing voxel-wise t-test between the two groups, excluding lesions and CSF. Pearson correlation coefficient between MTR and, respectively, MTV and T_1w/T_2w across subjects was also calculated, again excluding voxels within lesions and CSF. In both cases, Benjamini/Yekutieli *false discovery rate* (FDR) correction [110] was implemented for multiple comparisons.

The following Python functions were used:

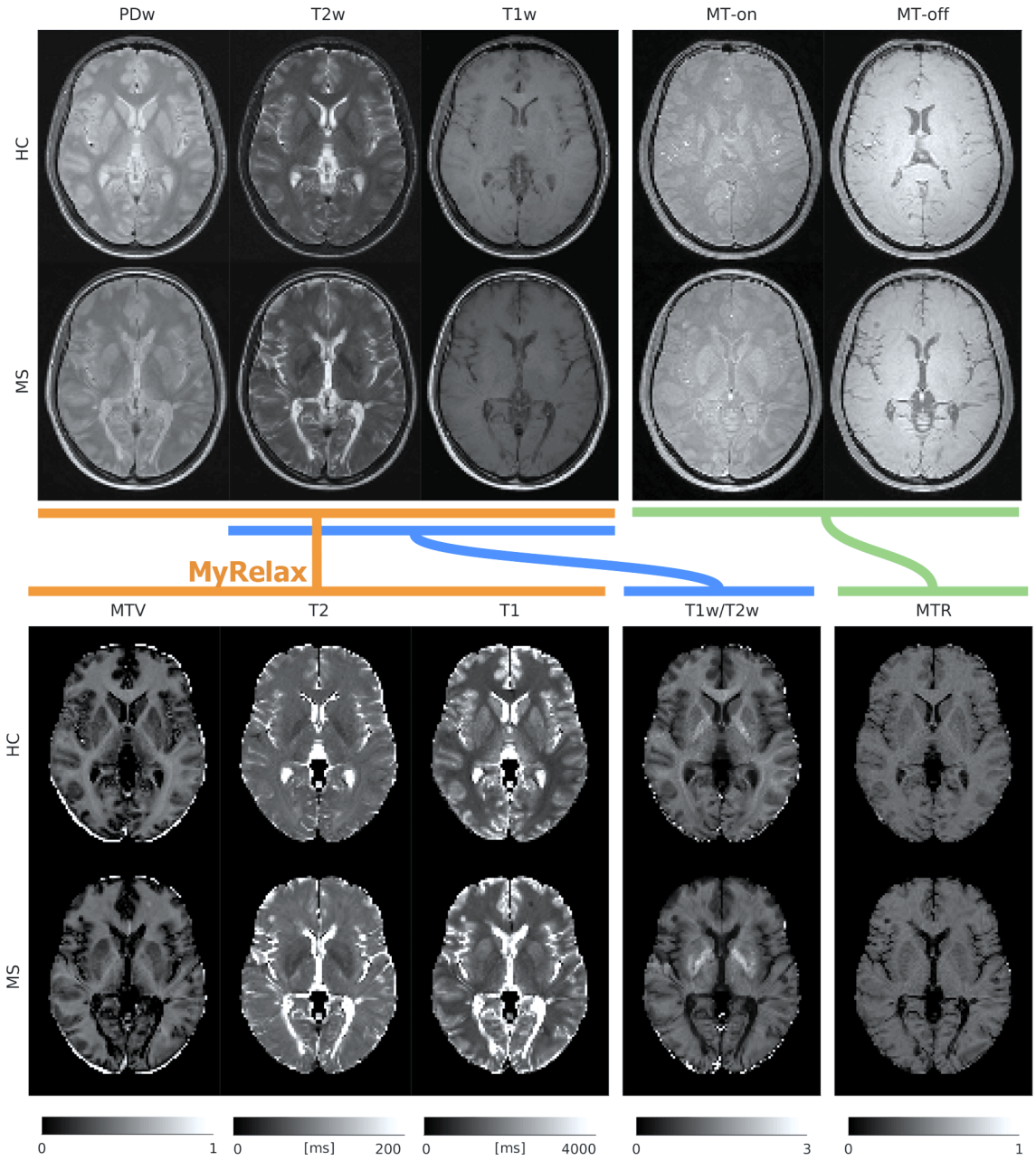


Figure 7.1: Overview of the MRI modalities investigated for the *MyRelax MS application* objective, with examples for one HC and one MS patient. MTV, T_2 and T_1 maps were extracted from PD-/T_2-weighted turbo spin-echo and T_1-weighted spin-echo qualitative scans. T_1w/T_2w maps were calculated as the voxel-wise ratio between the qualitative T_1- and T_2-weighted images. MTR maps were produced from the 3D gradient-echo images with and without MT-weighting.

- `stats.ttest_ind`, from the `scipy` package, for the t-test analysis;
- `stats.pearsonr`, from the `scipy` package, for the Pearson correlation coefficient calculation;
- `stats.multitest.multipletests`, from the `statsmodels` package, for multiple comparisons correction.

Chapter 8

Results

In this chapter, results for the *MyRelax validation* and *MyRelax MS application bottom-up* objectives are reported. The first section details the validation of the MyRelax framework through the correlation of QuaSI- maps with their respective ground truth. The second section reports the comparison between QuaSI-MTV, T_1w/T_2w and MTR behaviour with respect to MS.

8.1 MyRelax validation

QuaSI-PD, -T_2 and -T_1 maps obtained through the MyRelax framework from the qualitative PD-, T_2-, T_1-weighted scans and ground truth maps for a single subject are shown in Figure 8.1. Correlation plots between MyRelax and ground truth regional values are shown in Figure 8.2. Linear regression coefficients β_0, β_1, and Pearson correlation coefficients r are reported in Table 8.1. Despite MyRelax regional values not distributing identically as the ground truth ones, results were reproducible across subjects and strong correlation ($r \geq 0.94$) was observed. This suggests that qualitative images are affected by weights and/or biases, perhaps scanner-dependent, that should be incorporated into the Bloch equations to better fit the data. Given that the mismatch between MyRelax and ground truth maps seems to be reproducible across subjects, the mismatch could in fact be considered systematic, i.e. fixed for a given scanner and acquisition protocol, meaning that it could alternatively be regressed out via *calibration*. As a proof of concept, calibration was performed on each subject, using the remaining three as calibration cohort to calculate the calibration function, e.g. using the linear regression coefficients calculated on subjects 1, 1-rescan, and 2 to correct subject 3 maps.

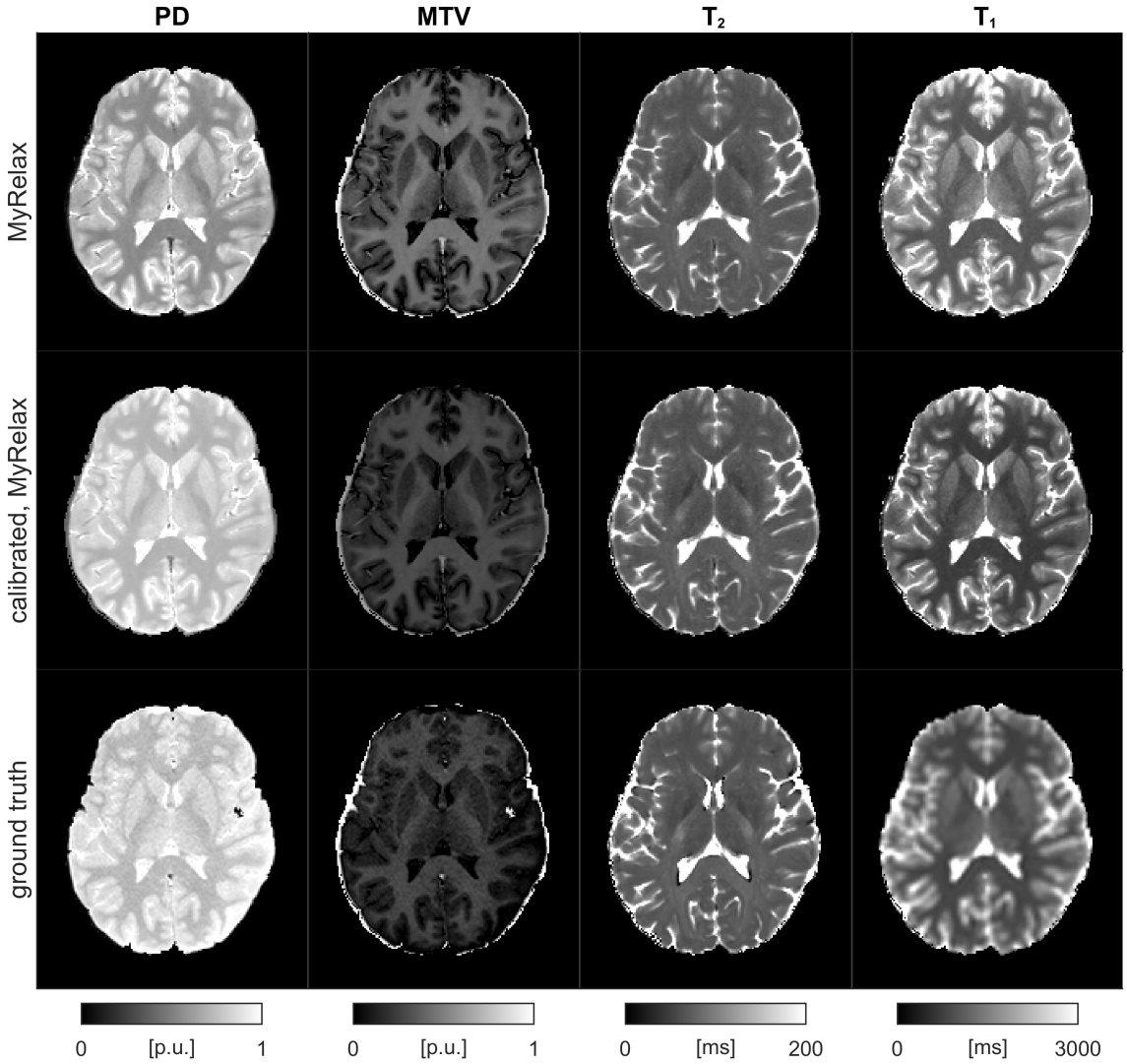
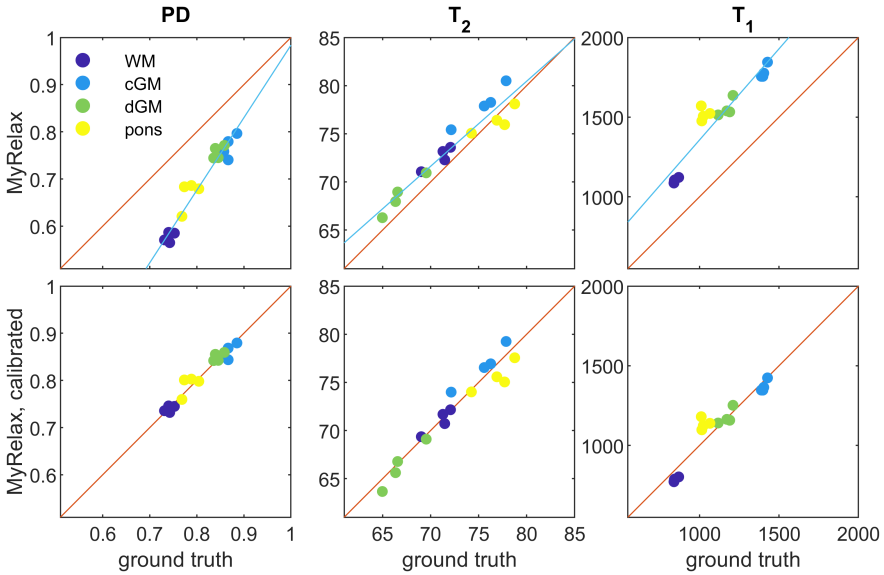


Figure 8.1: *MyRelax validation*. Top row: examples of QuaSI-PD, -T_2, and -T_1 maps for subject 3 of the *MyRelax validation* cohort. Middle row: same maps after calibration, using subjects 1, 1-rescan and 2 as calibration cohort. Bottom row: ground truth maps for the same subject. P.u.: percentage units within $[0, 1]$.

The post-calibration maps for subject 3 are also shown in Figure 8.1, with the calibrated regional values being reported in Figure 8.2 as well. The calibrated QuaSI-maps showed indeed much closer similarity to the ground truth, with the calibrated QuaSI-regional values distributing almost identically to their respective reference values.

8.2 MyRelax MS application

Tissue contrast in white matter and cortical grey matter appeared to be preserved across modalities; the sub-cortical region appeared instead visibly hyper-intense in T_1w/T_2w,



	PD
β_0	-0.52 ± 0.09
β_1	1.50 ± 0.11
r	0.97
	T_2
β_0	10.25 ± 5.29
β_1	0.88 ± 0.07
r	0.96
	T_1
β_0	223.79 ± 122.76
β_1	1.14 ± 0.11
r	0.94

Figure 8.2 & Table 8.1: *MyRelax validation*. Top row: MyRelax QuaSI-PD, -T_2 and -T_1 maps compared with the corresponding ground truth images (each dot of a given colour indicates one subject's regional value). Bottom row: calibration proof of concept. Table: regression ($\beta_{0,1}$) and Pearson correlation (r) coefficients. *cGM*: cortical GM, *dGM*: deep GM; the red line indicates $y = x$, the blue line indicates the linear fitting $y = \beta_0 + \beta_1 x$.

but not in QuaSI-MTV and MTR maps. Examples of QuaSI-MTV, T_1w/T_2w, and ground truth MTR for the same MS subject are shown in Figure 8.3.

Voxel-wise t-statistic maps between HC and MS are shown in Figure 8.4. Statistical significance was assessed upon performing Benjamini/Yekutieli FDR correction for multiple comparisons. Statistically significant alterations were observed in the principal NAWM bundles — corpus callosum, optic radiations and corticospinal tracts — with HC exhibiting overall significantly higher values than MS (i.e. positive t-statistic) for all three modalities. Voxels with negative t-statistic values, showing the opposite trend, were also observed in deep GM and internal capsule for T_1w/T_2w, but not — or not as predominantly — in QuaSI-MTV or MTR.

This difference in local trends between HC and MS populations was further investigated by inspecting the original t-statistic maps, prior to FDR correction and thresholding. Clusters of negative correlation values were observed in deep GM, mainly localised within thalamus, putamen and caudate nucleus, for both QuaSI-MTV and MTR, whilst extending to the internal capsule WM tract for T_1w/T_2w. Whilst not satisfying the requirements for statistical significance, this result suggests that MTR behaves more similarly to QuaSI-MTV in terms of patterns of regional alterations, than to T_1w/T_2w. This also seems to indicate that T_1w/T_2w might be co-dependent on one or more

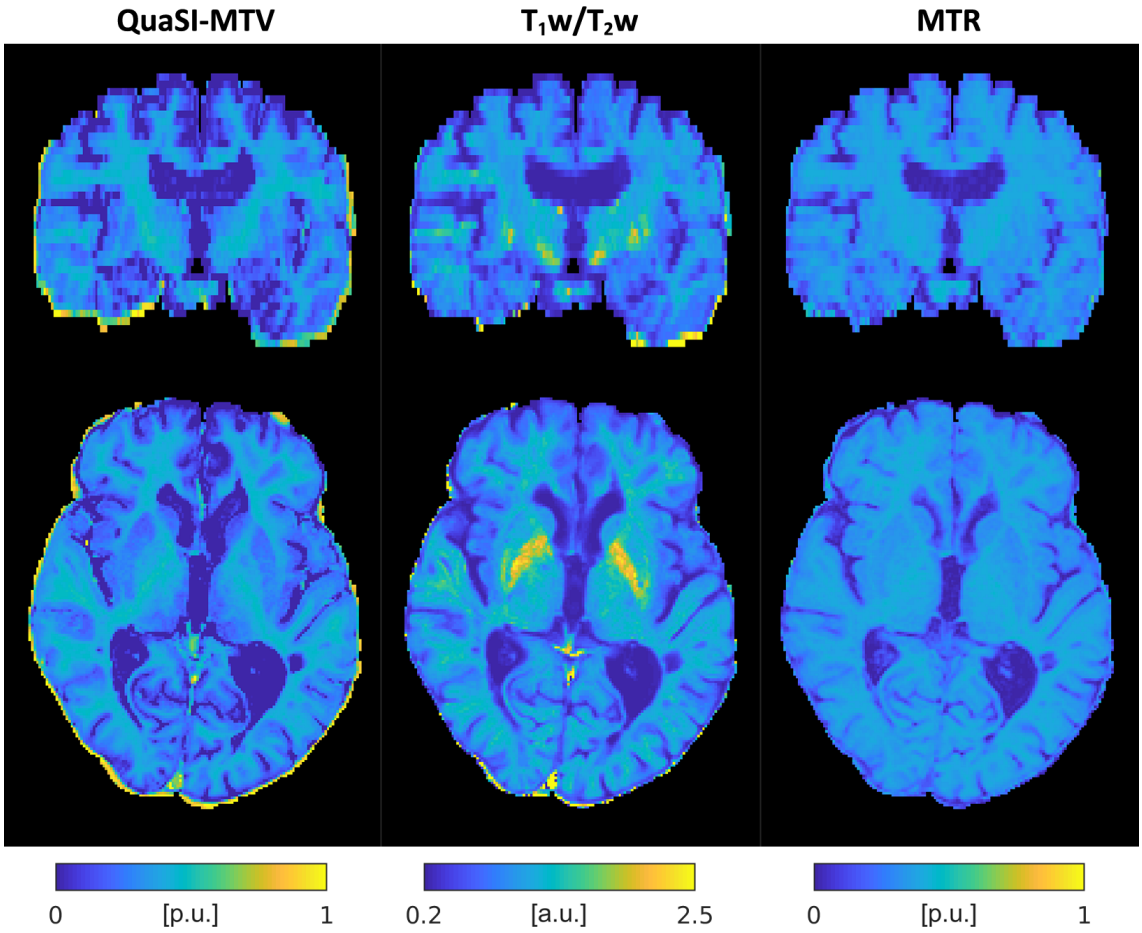
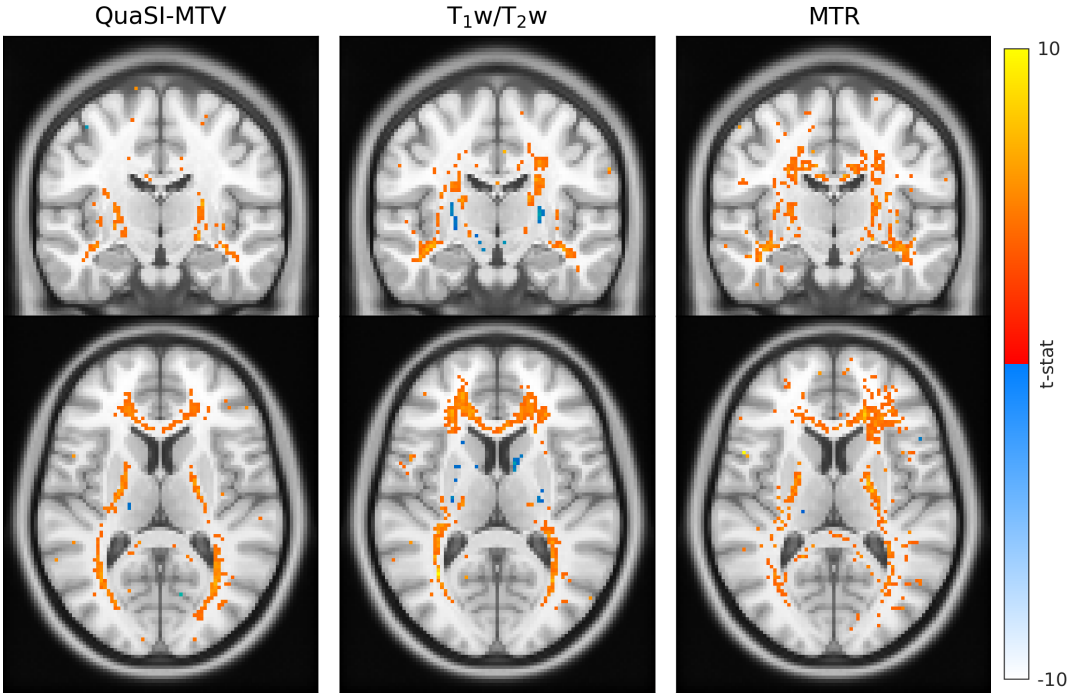


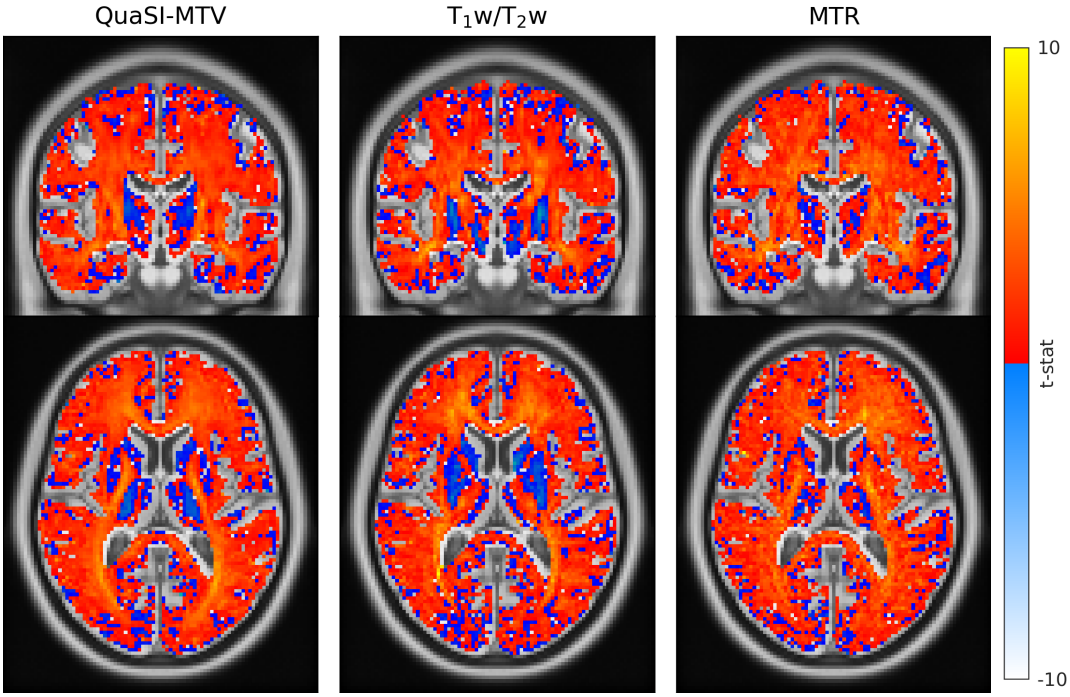
Figure 8.3: *MyRelax MS application*. Examples of QuaSI-MTV, T_1w/T_2w, and MTR for the same patient. Relative hyper-intensity can be observed in the sub-cortical region T_1w/T_2w. P.u.: percentage units within $[0, 1]$; a.u.: arbitrary units within $[0, \infty)$.

factors that otherwise do not affect QuaSI-MTV and MTR.

Voxel-wise Pearson correlation maps for MTR versus QuaSI-MTV, and MTR versus T_1w/T_2w were also calculated, applying FDR correction as above for statistical significance, as shown in Figure 8.5. Significant moderate positive correlation between MTR and both QuaSI-MTV and T_1w/T_2w was observed in NAWM ($r = 0.41 \pm 0.10$ and $r = 0.43 \pm 0.11$ respectively). Significant moderate negative correlation ($r = -0.36 \pm 0.08$) was also observed in the internal capsule WM tract when correlating MTR to T_1w/T_2w, but not to QuaSI-MTV. This result further points towards greater similarity in behaviours between MTR and QuaSI-MTV, than T_1w/T_2w.



(a) T-statistic maps after FDR correction.



(b) Original t-statistic maps.

Figure 8.4: *MyRelax MS application*. T-test between HC and MS groups for QuaSI-MTV, T_1w/T_2w, and MTR maps. a) Statistically significant positive t-statistic values (HC>MS) were observed for all three modalities in NAWM; statistically significant negative t-statistic values (HC<MS) were observed in the claustrum WM tract for T_1w/T_2w, but not QuaSI-MTV or MTR. b) Original t-statistic maps showed similar patterns of alteration for QuaSI-MTV and MTR, but not for T_1w/T_2w.

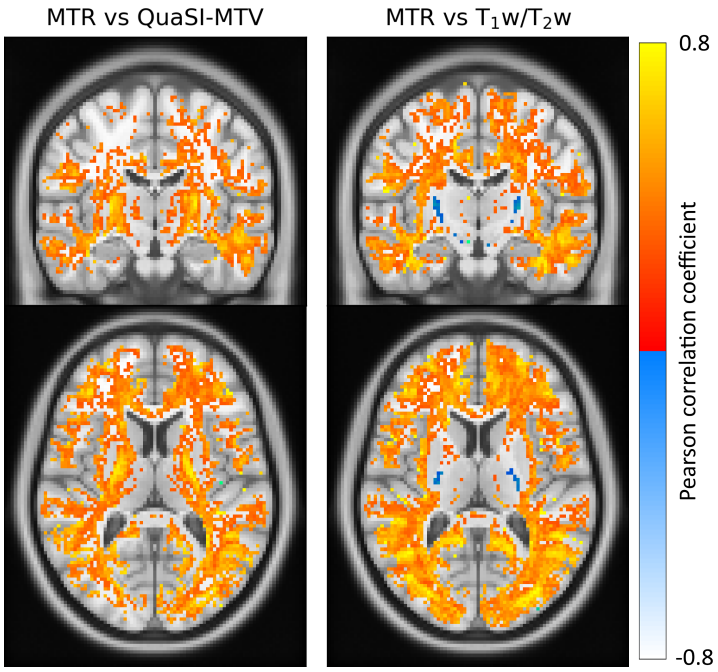


Figure 8.5: *MyRelax MS application*. Pearson correlation coefficient maps in normal appearing tissue after FDR correction.

Chapter 9

Discussions

9.1 MyRelax validation

Through the *MyRelax validation* objective, it has been shown that it is feasible to extract high-level quantitative information from qualitative data in a *bottom-up* way, via traditional model fitting. Previous studies have shown the feasibility of this approach, with qualitative images being used to extract quantitative information, limiting however its applicability to T_1 and T_2 maps, and in particular using purposely acquired qualitative data, that do not match the standard acquisition set-up [111, 112]. In this study, this approach was extended to PD and MTV maps, providing a means to access myelin-sensitive information, demonstrating that it works even with standard TSE/SE images.

9.1.1 Limitations

In addition to the small sample size, as it was already observed in section 8.1, using conventional qualitative scans instead of specialised ones comes at the — predictable — cost of quantitative maps that do not match the ground truth identically. It is likely that the standard acquisitions are indeed affected by additional effects not fully described by the MyRelax framework and specific to the MR-sequence employed, so different results may be observed for qualitative scans acquired with different readouts, on different MR-scanners.

Whilst a more complex system of equations tailored over the specifics of each MR-protocol could be devised to obtain more accurate results, it is worth noticing that the mismatch between QuaSI- maps and the respective ground truth images appears to be

reproducible across the subjects, and could therefore be regressed out via *calibration*.

9.1.2 Calibration

The simplified calibration performed as proof of concept on one subject at a time, using the remaining three as calibration cohort, produced very promising results, which is particularly significant when considering the abundance of this kind of qualitative data in pre-existing large historical datasets, currently used only for lesion segmentation and anatomical analyses. Upon retrieving, or prospectively acquiring, if necessary, ground truth data on a small calibration cohort, calibrated QuaSI-PD, -T_2 and -T_1 maps could thus be obtained from any qualitative data acquired on the same scanner with the same MR-sequence as the calibration cohort.

For definitive proof, a proper calibration study including both HC and MS patients data acquired on different scanners, should be performed. That being said, calibration is not strictly necessary when using MyRelax maps for group comparisons, since relative differences between groups would be preserved regardless of linear scaling, as it was done for the *MyRelax MS application* objective.

9.2 MyRelax MS application

The quantitative potential of retrospective qualitative datasets was further investigated through the *MyRelax MS application* objective. The qualitative scans were used to extract QuaSI-MTV maps via MyRelax, as well as T_1w/T_2w maps. Both sets were compared to ground truth MTR to assess their behaviour with respect to MS pathology.

9.2.1 MTR vs QuaSI-MTV vs T_1w/T_2w

From the patterns of alteration observed when comparing HC to MS patients and the correlation analysis, MTR showed closer matching information with QuaSI-MTV than T_1w/T_2w, in particular in deep GM and internal capsule regions, where notably hyper-intensity and negative correlation with MTR were observed. These results are compatible with high concentration of iron deposition observed in MS in deep GM [52], which causes reduced T_2, and in turn increased T_1w/T_2w. The lack of a similar behaviour for QuaSI-MTV and MTR suggests a lower co-dependency on similar confounding effects: it is thus sensible to conclude that MyRelax does enable to produce QuaSI-MTV maps

that, even when *not* calibrated, provide added value with respect to specificity to MS compared to T_1w/T_2w maps, despite being both derived from the same qualitative data.

9.2.2 Limitations

In addition to the limitations already disclosed for the MyRelax framework in the previous section, it is unfortunately impossible at this stage to conclude whether QuaSI-MTV is a good, specific indicator for demyelination or not, as MTR itself, which was used as ground truth, is not uniquely specific to myelin, but it is also influenced by other factors such as inflammation and neurodegeneration. Nonetheless, MTR is still largely used in clinical research and as outcome measure in clinical trials, as an indirect metric sensitive to myelin, and surrogates for when MTR is unavailable or missing might be of potential use.

Summary

- MyRelax can be used to generate reproducible QuaSI-PD, -T_2 and T_1 maps from qualitative data that well correlate with the respective ground truth maps.
- QuaSI-MTV was shown to behave similarly to MTR with respect to MS.
- Calibration could be implemented for more accurate results.
- Results suggest myelin content information could be extracted directly from qualitative data as well.
- Whilst dedicated quantitative scans are to be preferred, MyRelax might still provide a way to perform quantitative studies when these are unavailable or missing.

Part III

Deep learning MTR from qualitative images

Chapter 10

Introduction

Through the *MyRelax validation* and *MyRelax MS application bottom-up* objectives, the **MyRelax: myelin and relaxation imaging** contribution has shown that QuaSI-MTV maps produced from qualitative images behave similarly to MTR when testing for differences between healthy controls and MS patients. The *bottom-up hypothesis* of myelin-content information being present already within the qualitative data, and therefore accessible via the proper mathematical methods despite the lack of an explicit model, was thus further explored with a data-driven *deep learning* approach.

As part of the *U-Net MS application* objective, deep learning was employed to infer such a model through the use of a 2D U-Net, defined and trained using PyTorch [113]. Mapping the qualitative PD-, T_2- and T_1-weighted scans to reference MTR maps enabled generating QuaSI-MTR maps, learning the relationship between qualitative and myelin imaging directly from the data.

Chapter 11

Methods

11.1 Cohort

A subset of the GML02 cohort (see section 7.2.1) of 48 subjects (20 men, age: 43±12 years old) was used, composed of 16 HC, 9 RRMS, 17 SPMS and 10 PPMS patients, with no follow-up scans. Of these, 24 subjects were used for training, 12 for validation and 12 for testing. The training set was composed of 8 HC, 3 RRMS, 9 SPMS and 4 PPMS patients. Validation and test set were both composed of 3 HC, 3 RRMS, 3 SPMS and 3 PPMS patients each. All sets were age and gender matched. This study was approved by the local ethical committee.

11.2 Preprocessing

The preprocessing described for the *MyRelax validation* objective (see section 7.4.1) applies for the *U-Net MS application* as well. In addition, the 1×1×3 mean-filtered MTR maps (see section 7.4.2) were linearly registered onto the qualitative scans space and used as the network target data. This enabled to focus the learning process solely on QuaSI-MTR regression, being the central point of this objective, rather than a deep-learning regression/registration hybrid model.

11.3 Training

The training was run on an Nvidia Quadro P2000 5GB GPU, which limited the batch size to 12 before saturation. The set of batches required to exhaust all training data

defined a *training epoch*. The input data per training iteration consisted therefore of a batch of 12 sets of images, each set being composed of 3 volumes, or *channels*, corresponding to the PD-, T_2- and T_1-weighted scans for the same brain axial slice, with size 240×240. The *target* data consisted of 12 single-channel images, each corresponding to the respective MTR slice.

In tensor terms, the input and target data consisted each of a tensor with size 12×3×240×240 and 12×1×240×240, respectively. In this context, the term *slice* will refer to the i-th input/target tensor pair with size {input: 1×3×240×240; target: 1×1×240×240}, with $i = 1, ..., 12$. With this understood, channels will also be left implied, referring to the slice size simply as 240×240.

11.3.1 Data selection

Slices containing less than 10% non-zero voxels were excluded to reduce spurious learning, for an average of about 40 useful slices per subject. Each slice was cropped down from its 240×240 original size to 224×224 to fit the network. Since the excluded voxels were empty on all instances, the cropping did not lead to any loss of information. For each batch, slices were picked at random, without repetition, across all subjects, until all the slices of all the training set had been used.

11.3.2 Data normalisation

Being composed of qualitative data, with voxel values not bound within a certain interval, the input data required normalisation. First, for each input slice X, outliers were excluded using the *interquartile range* (IQR) method. Then, X was divided by the updated maximum, such that $X \in [0, 1]$. Denoting the i-th percentile as P_i, this can be expressed in pseudo-algorithm form as:

$$\begin{aligned} Q_1 &= P_{25}(X) \\ Q_3 &= P_{75}(X) \\ \text{IQR} &= Q_3 - Q_1 \\ X[X < (Q_1 - 1.5 \cdot \text{IQR})] &= Q_1 - 1.5 \cdot \text{IQR} \\ X[X > (Q_3 + 1.5 \cdot \text{IQR})] &= Q_3 + 1.5 \cdot \text{IQR} \\ X &= X/\max(X) \end{aligned} \tag{11.1}$$

Notice X is a $3\times224\times224$ sized tensor, so the normalisation is conducted on all three channels jointly to ensure the PD-, T_2- and T_1-weighted images are scaled by the same quantity.

Since MTR is, by definition, defined within $[0, 1]$, no normalisation was required for the target slices.

11.3.3 Data augmentation

Slices were rotated in plane, as they were loaded, by an angle randomly sampled from a normal distribution with centre 0° and standard deviation 10° to artificially increase training data variability. The rotation was performed using the `ndimage.rotate` from the `scipy` package.

11.3.4 Loss function

The loss function $f(y', y)$ quantifying the distance between the prediction y' and the target y was defined as the sum of three objective functions:

$$f(y', y) = \text{RMSE}(y', y) + \text{RMSE}(\mathbf{G}(y'), \mathbf{G}(y)) + \text{DSSIM}(y', y) \tag{11.2}$$

- $\text{RMSE}(y', y)$ is the root mean square error between the prediction and the target[1]. By minimising this term, the difference in the overall contrast between the two images is also minimised.
- $\text{RMSE}(\mathbf{G}(y'), \mathbf{G}(y))$ is the root mean square error between prediction and target *edge* maps. The minimisation of this term aims to preserve the edge sharpness of the predicted image and minimise blurriness.

  Edge detection was performed using the *Sobel* operator

  $$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2} \tag{11.3}$$

  where $\mathbf{G}_{x,y}$ represent approximations of gradient operators along the x, y directions, respectively. The operation consists of a 3×3 kernel convolution with the input tensor $\mathbf{A}$ such that:

[1] NaN errors were observed during training, which resulted in the script to crash. It was discovered this to be caused by the `torch.sqrt` (PyTorch tensor square root) function producing NaN during gradient back-propagation when 0 was presented as the function argument. The substitution $\sqrt{x} \rightarrow \sqrt{x + \epsilon}$, with $\epsilon = 10^{-15}$, was thus performed on all instances of `torch.sqrt`.

$$\mathbf{G}_x(\mathbf{A}) = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * \mathbf{A}; \quad \mathbf{G}_y(\mathbf{A}) = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A} \tag{11.4}$$

The $\mathbf{G}_{x,y}$ filtering was computed via the `filter.SpatialGradient` function, imported from the `kornia` package [114].

- $\text{DSSIM}(y', y) = [1 - \text{SSIM}(y', y)]/2$, is the *structural dissimilarity index*, derived from SSIM, the *structural similarity index*. SSIM is a perception-based metric assessing the degree of matching information between two images x and y, depending on three measurements:

$$\begin{aligned} \text{luminance:} \quad & l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \\ \text{contrast:} \quad & c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \\ \text{structure:} \quad & s(x, y) = \frac{\sigma_{xy}\mu_y + c_3}{\sigma_x\sigma_y + c_3} \end{aligned} \tag{11.5}$$

  with:

  - $\mu_{x,y}$ the average of x, y respectively;
  - $\sigma_{x,y}$ the variance of x, y respectively;
  - σ_{xy} the covariance of x and y;
  - $c_{1,2} = (k_{1,2}L)^2$, $c_3 = c_2/2$ variable stabilising the division, where:
    - $k_1 = 0.01$ and $k_2 = 0.03$ by default;
    - L the value range of the input images, which in the case of data normalised within $[0, 1]$ is $L = 1$.

  Luminance, contrast and structure are then weighted together, such that:

$$\text{SSIM}(x, y) = l(x, y)^\alpha c(x, y)^\beta s(x, y)^\gamma \tag{11.6}$$

  As in most applications, the default values of $\alpha = \beta = \gamma = 1$ have been used, which result in the SSIM reduced form:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{11.7}$$

Algorithm 1: *Adam*, our proposed algorithm for stochastic optimization. See section 2 for details, and for a slightly more efficient (but less clear) order of computation. g_t^2 indicates the elementwise square $g_t \odot g_t$. Good default settings for the tested machine learning problems are $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 10^{-8}$. All operations on vectors are element-wise. With β_1^t and β_2^t we denote β_1 and β_2 to the power t.

Require: α: Stepsize
Require: $\beta_1, \beta_2 \in [0, 1)$: Exponential decay rates for the moment estimates
Require: $f(\theta)$: Stochastic objective function with parameters θ
Require: θ_0: Initial parameter vector
  $m_0 \leftarrow 0$ (Initialize 1st moment vector)
  $v_0 \leftarrow 0$ (Initialize 2nd moment vector)
  $t \leftarrow 0$ (Initialize timestep)
  while θ_t not converged **do**
    $t \leftarrow t + 1$
    $g_t \leftarrow \nabla_\theta f_t(\theta_{t-1})$ (Get gradients w.r.t. stochastic objective at timestep t)
    $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (Update biased first moment estimate)
    $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ (Update biased second raw moment estimate)
    $\widehat{m}_t \leftarrow m_t / (1 - \beta_1^t)$ (Compute bias-corrected first moment estimate)
    $\widehat{v}_t \leftarrow v_t / (1 - \beta_2^t)$ (Compute bias-corrected second raw moment estimate)
    $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \widehat{m}_t / (\sqrt{\widehat{v}_t} + \epsilon)$ (Update parameters)
  end while
  return θ_t (Resulting parameters)

Figure 11.1: Adam pseudo-code [116]. The term ϵ is used to avoid divisions by zero.

SSIM was computed using the `ssim` function from the `pytorch-msssim` package [115].

11.3.5 Optimisation

Adaptive moment estimation — or Adam [116] — was chosen as optimisation algorithm. Adam is an extension of *stochastic gradient descent* that automatically adapts *learning rates* for each network weight (see section 5.5) from estimates of exponential moving averages of the gradient (*1-st moment*) and the squared gradient (*2-nd raw moment*, or *uncentered variance*). With its almost 75 thousand citations at the time of writing, Adam is *de facto* the current default optimiser for most deep learning tasks.

1-st moment

With reference to Adam pseudo-code in Figure 11.1, the exponential average of the gradient at the iteration t over the past iterations with weight $\beta_1 \in [0, 1)$

$$\begin{aligned} m_t &= (1 - \beta_1)(g_t + \beta_1 g_{t-1} + \beta_1^2 g_{t-2} + ... + \beta_1^{t-1} g_1) \\ &= (1 - \beta_1) \sum_{t'=0}^{t-1} (\beta_1^{t'} g_{t-t'}) \end{aligned} \tag{11.8}$$

acts as an heuristic measure of momentum, averaging out sudden changes in the objective function slope sign along any given direction, effectively dampening oscillations in the parameter space along that direction, and promoting descent along the direction of monotone gradient. The exponential weighting ensures that recent gradient updates are favoured over past ones to quickly adapt to changes in the gradient overall behaviour[2].

2-nd raw moment

The exponential average of the square of the gradient ν_t, with weight $\beta_2 \in [0, 1)$, inherited from the *RMSprop* optimisation method, is conceptually opposite, but works in concert with the 1-st moment estimate: due to the square, oscillating gradient updates will not cancel out, but rather stack additively. This acts as a metric for the overall magnitude of the gradient, with steep slopes contributing towards a high 2-nd raw moment estimate, regardless of the sign. By dividing the learning rate α by $\sqrt{\nu_t}$, the average squared gradient acts as a *brake* on the effective update step-size in parameter space $\Delta_t = \alpha \cdot m_t/\sqrt{\nu_t}$, reducing it when in proximity of steep gradients, whilst its magnitude stays bounded, in most scenarios, by the learning rate setting, i.e. $|\Delta_t| \leq \alpha$.

Bias-correction

Since the moments are initialised to vector of 0's, their estimations will be biased towards zeroes in the initial steps of the optimisation. This can be compensated by computing bias-corrected estimates $\hat{m}_t$ and $\hat{\nu}_t$ after each update, as shown in Figure 11.1.

Implementation

Optimisation was implemented via the `torch.optim.Adam` function, using default $\beta_1 = 0.9$, $\beta_2 = 0.999$ hyperparameters. Although not strictly necessary, given how Adam internally adapts the effective learning rate for each weight, the upper-bound learning rate was still set to gently anneal with the number of epochs n at a 0.99 rate: $\alpha_t = 0.001 \cdot 0.99^t$, initialised at $\alpha_0 = 0.001$ as recommended.

The AMSGrad variant was chosen (by setting `amsgrad=True`), which has been shown to improve convergence of Adam optimisers by using the maximum of past squared

[2] The usual metaphor is that of a ball rolling down a steep ravine with a gently sloping river valley at the bottom. In standard stochastic gradient descent, the ball will oscillate back and forth along the two river banks, making little progress along the direction of the river, since descent is greater along steep gradient slopes. By using the exponential average of the gradient instead, oscillations will soon cancel out after a few iterations, progressing along the river flow.

gradients instead of a moving exponential average [117].

11.4 Model selection and testing

At the end of each epoch, the trained model was applied to the entire validation set and the average *validation loss* was recorded[3] for the purposes of selecting the best model, i.e. the model with the lowest validation loss over all epochs. The training/validation was iterated until validation loss stabilised, which occurred within 100 epochs.

The best model was then applied to the test set. The difference map between the predicted QuaSI-MTR and ground truth MTR was calculated.

11.5 ResNet encoding

With reference to section 5.7, a *residual neural network* (ResNet) [118] has been used for the encoding arm of the U-Net employed in this work, specifically `models.resnet18` blocks imported from the `torchvision` package. Like U-Net, ResNet is also a class of *convolutional neural network* (CNN, see section 5.6) that makes use of shortcut connections to connect otherwise non-consecutive layers. The *basic block* of the network, as shown in Figure 11.2, is defined by a CNN arm — the *residual* — and an *identity* arm, added together. *Down-sampling* can be implemented by setting a kernel stride of 2 or more, which determines the scaling factor; a 1×1 convolutional layer with the same stride needs to be added to the shortcut connection to preserve tensor size coherence with the residual pathway.

The training consists therefore in optimising the residuals rather than the entire underlying mapping. This has been shown to mitigate the *degradation* problem that accompanies deeper architectures in the form of *vanishing gradients*, by providing, through the skip connection, an alternative and simple pathway for the gradient to propagate. The advantage of this architecture is evident when considering the extreme example of the network trained to fit an identity mapping since, as pointed out by He et al. (2016) in the ResNet original paper:

> *It would be easier to push the residual to zero than to fit an identity mapping by a stack of nonlinear layers.* [118]

[3] The average *training loss* over all training batches was also recorded every epoch, although only to monitor the network learning performances and it was not used to inform model selection in any way.

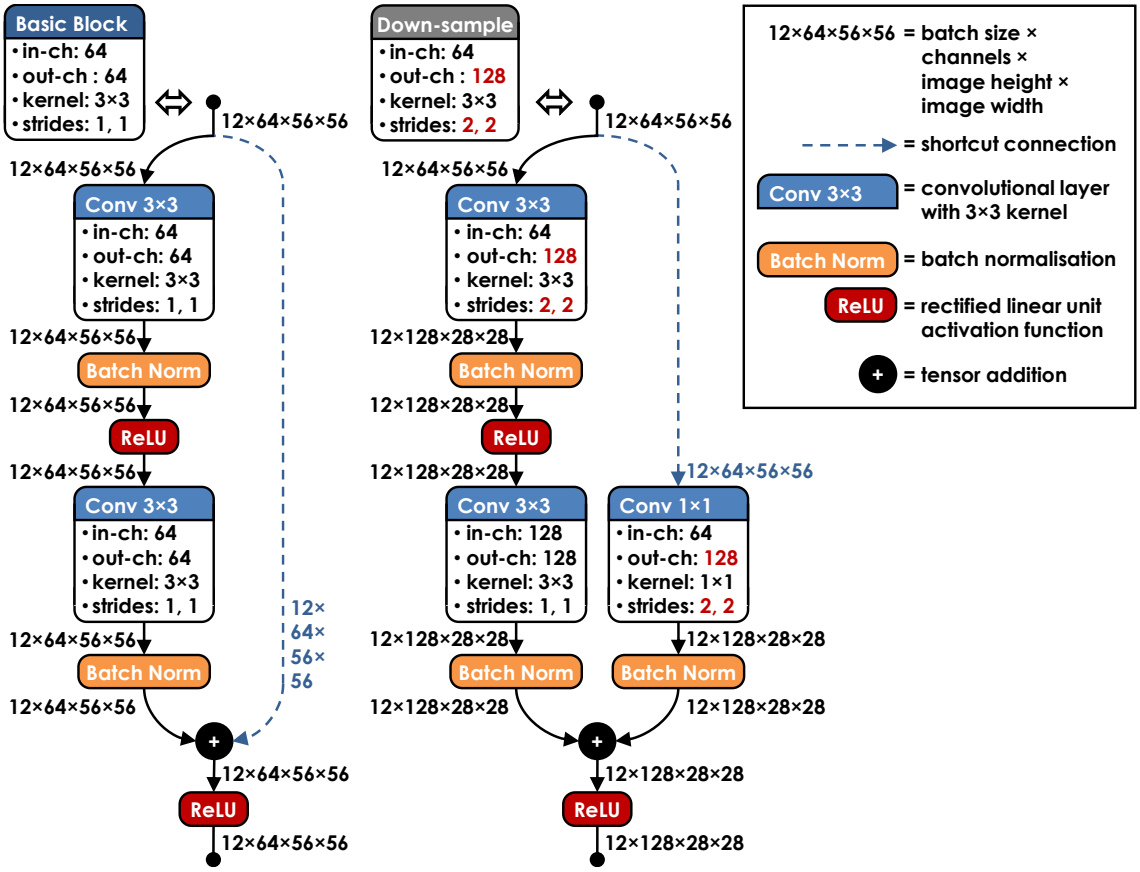


Figure 11.2: ResNet building blocks. The *basic block* is composed of a CNN residual pathway and a shortcut connection added together. *Down-sampling* by a factor of 2 is implemented by using a kernel stride of 2. Further feature encoding is implemented as part of the down-sampling by increasing the number of output channels. See section 11.6 for details on *batch normalisation* and *ReLU activation*.

11.6 Network architecture

The network architecture is shown in Figure 11.3, with its building blocks being described as follows.

1. **12**×**3**×**224**×**224**. Tensor size at each stage has been reported as batch size × number of channels × image height × image width.

   As the data is processed through the encoding, or contracting, pathway of the network, the batch size is preserved, the number of channels increases as a result of feature encoding, and the image gets *down-sampled*.

   In the decoding, or expanding, arm of the network, the data gets *up-sampled* to native size, whilst the number of features is condensed to ultimately one output

channel, corresponding to the QuaSI-MTR map.

2. **Conv 3×3**. 2D convolutional layer with 3×3 kernel size; other kernel sizes used in the network are 1×1 and 7×7. Implicit zero-padding of 1 and 3 was set for the 3×3 and 7×7 convolution layers, respectively. Padding and (1, 1) kernel stride ensure input tensor size is preserved at the layer output. On the other hand, higher kernel stride values produce down-sampled output, as the kernel *jumps* over the input tensor, skipping elements. Down-sampling by a factor of 2 was implemented as part of some convolution layers via (2, 2) kernel stride (see Down-sample block).

   2D convolution was chosen for being computationally less expensive, whilst also reflecting the PD/T2 and T1 scans 2D acquisition method, with each axial slice being acquired independently from the others.

3. **ReLU**. Rectified linear unit (ReLU) activation function (see section 5.5).

4. **Batch Norm**. Batch normalisation is used to re-centre and re-scale tensors between layers, by setting the mean to 0 and standard deviation to 1. This operation has been shown to improve numerical stability and network efficiency, and is a standard step in most deep-learning applications.

5. **Max Pool**. 2D max-pooling layer. Pooling is a form of non-linear down-sampling where the input data is divided into partitions, and the values within each partition are aggregated into a single value. The partitioning can be implemented using a sliding kernel with stride equal to the kernel size for non-overlapping partitions, or less for partial overlapping. As for the convolutional layer, the kernel stride determines the down-sampling factor. For max-pooling, the maximum over each partition is selected.

6. **Basic Block**. ResNet basic block as shown in Figure 11.2.

7. **Down-sample**. ResNet down-sample block as shown in Figure 11.2.

8. **Up-sample**. 2D up-sampling layer, using bi-linear interpolation to up-sample the input layer by a factor of 2.

9. **Concat**. Tensor concatenation along the channel dimension. It was used to merge the information carried by shortcut connections into the network expanding pathway.

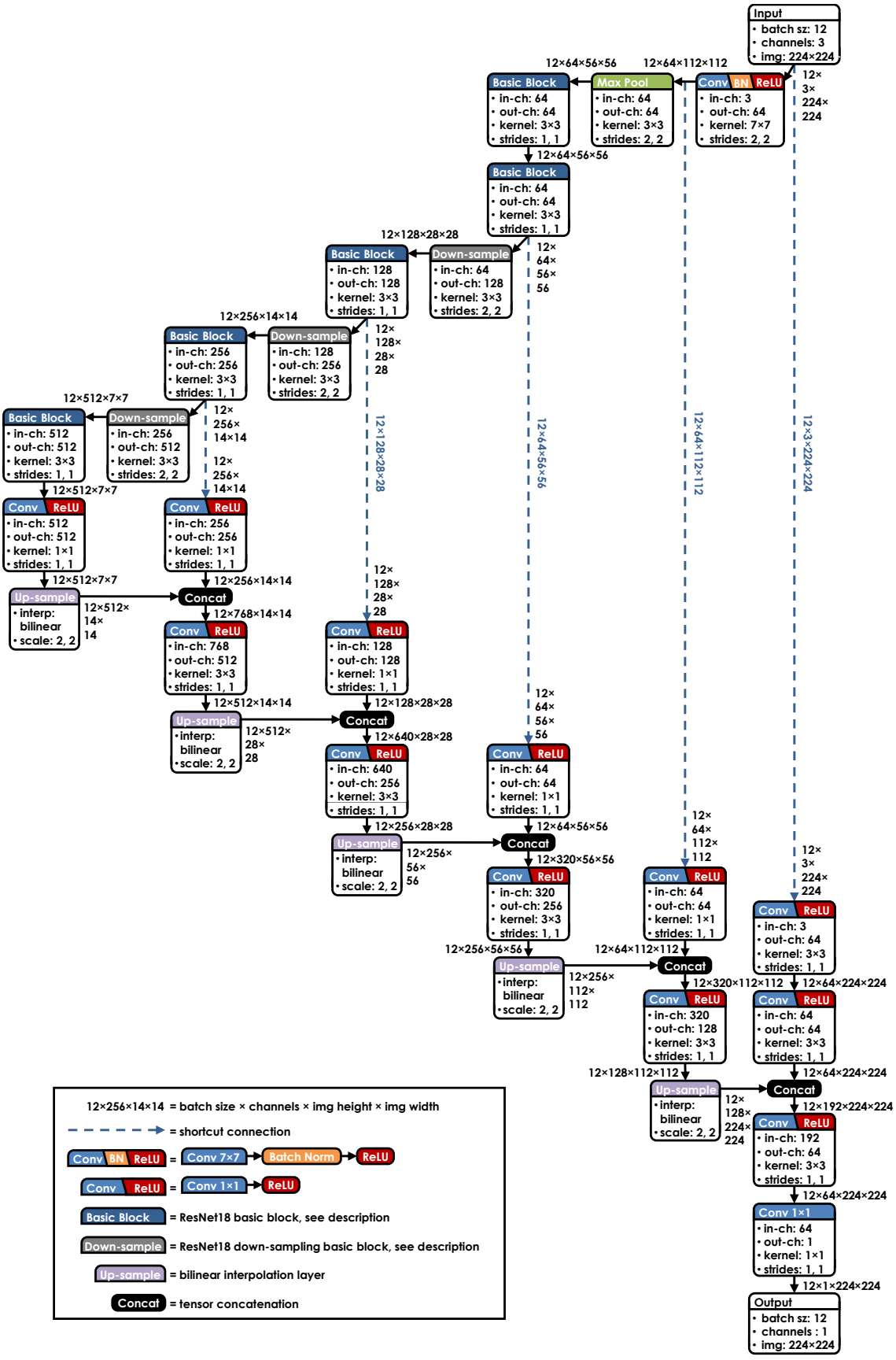


Figure 11.3: U-Net architecture with ResNet18 encoding.

Chapter 12

Results

Examples of synthetic QuaSI-MTR produced through U-Net from qualitative PD-, T_2-, and T_1-weighted images, together with ground truth MTR and residual maps for four subjects belonging to the test-set, are shown in Figure 12.1. QuaSI-MTR maps are qualitatively similar to ground truth MTR, with contrast between tissues being visually comparable. Positive and negative residuals appeared homogeneously distributed across the brain, with no clear emerging patterns suggestive of systematic misrepresented brain structures.

Mean QuaSI-MTR in NAWM, cortical GM, deep GM, and lesions correlated well with the respective MTR regional values, for all subjects in the test-set, as shown in Figure 12.2. Mean regional errors were contained within $\pm 5\%$ of the ground truth MTR, and did not show any particular trend associated to tissue or patient group. Higher spread (one standard deviation within MTR$\pm 10\%$) was however observed for low MTR values, which incidentally correspond to cortical GM areas and some lesions.

12.1 Residuals

Possible causes for increased residual variance at low MTR values include:

- **Low signal-to-noise ratio**. Higher variance in the regression error for low MTR values can be explained by the lower signal-to-noise ratio, which makes low-intensity regions inherently more noisy, and thus more difficult to *learn*.
- **Misregistration**. An important component contributing to errors in cortical GM or, in general, sharp high-contrast borders is misregistration, with the interface

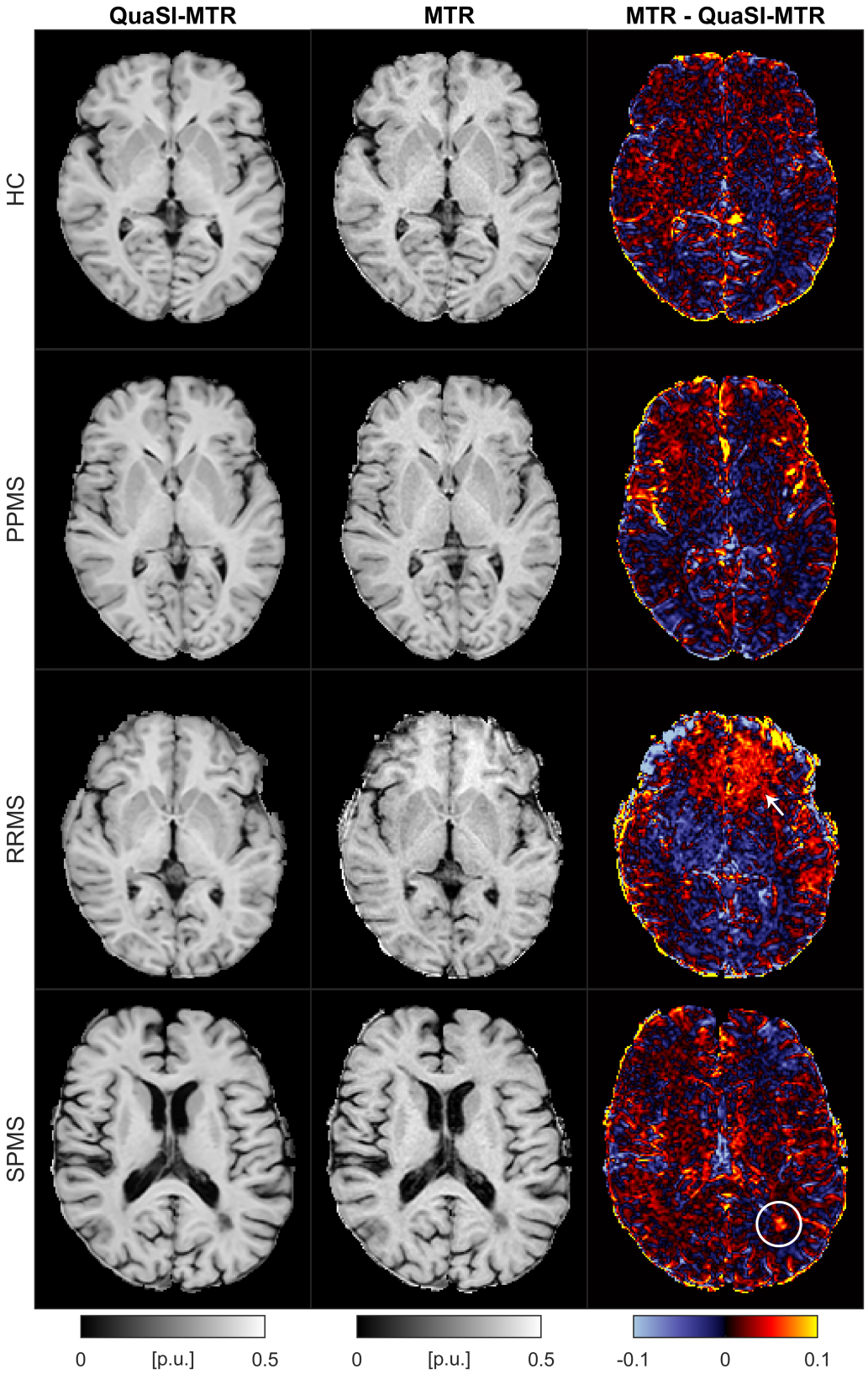


Figure 12.1: Examples of synthetic QuaSI-MTR, ground truth MTR, and error map for four different test-subjects. Highlighted: lesion (circle); possible bias field artifact affecting ground truth MTR (arrow). P.u.: percentage units within $[0, 1]$.

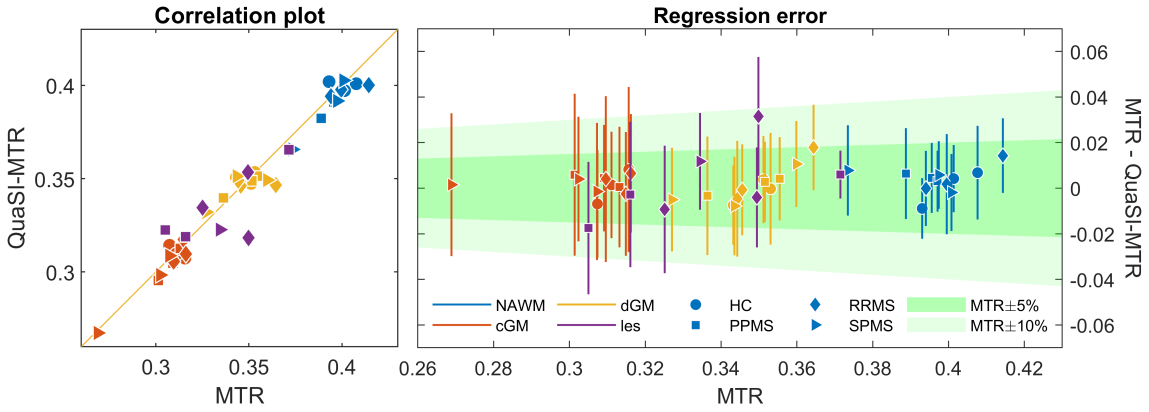


Figure 12.2: Correlation plot shows agreement between QuaSI-MTR and ground truth MTR regional values in the test-set. Residuals distributed symmetrically around 0, with higher spread for low-MTR values, corresponding to cortical GM and some lesions. Vertical lines indicate ±1 standard deviation; cGM: cortical GM, dGM: deep GM, les: lesions.

between brain tissue and CSF representing a prominent example of sharp high-contrast borders. CSF interface is abundant in the GM gyri and sulci, which can be challenging to correctly align when registering brain images from different modalities. Misregistrations can affect training by producing blurry images, as well as causing positive/negative residuals to emerge in the error map due to the misalignment between regressed and reference images. An example of this can be observed in Figure 12.3, where the ridges of positive and negative residuals have been caused by the imperfect alignment between the reference MTR map and the PD/T2 and T1 images, and thus the resulting QuaSI-MTR, rather than by an inherent regression error.

- **Under-representation**. Another source of error variance for cortical GM and some lesions, may be the limited representation of these tissues within the dataset, not only in terms of volume, but also *data variability*. This is especially incisive in the case of lesions, as they not only affect MS patients alone and in small regions, but they are also characterised by different pathophysiology depending on the specific MS phenotype. On the other hand, whilst not scarce with respect to the total intracranial volume, brain cortex is highly heterogeneous due to partial volume effects with both CSF and WM.

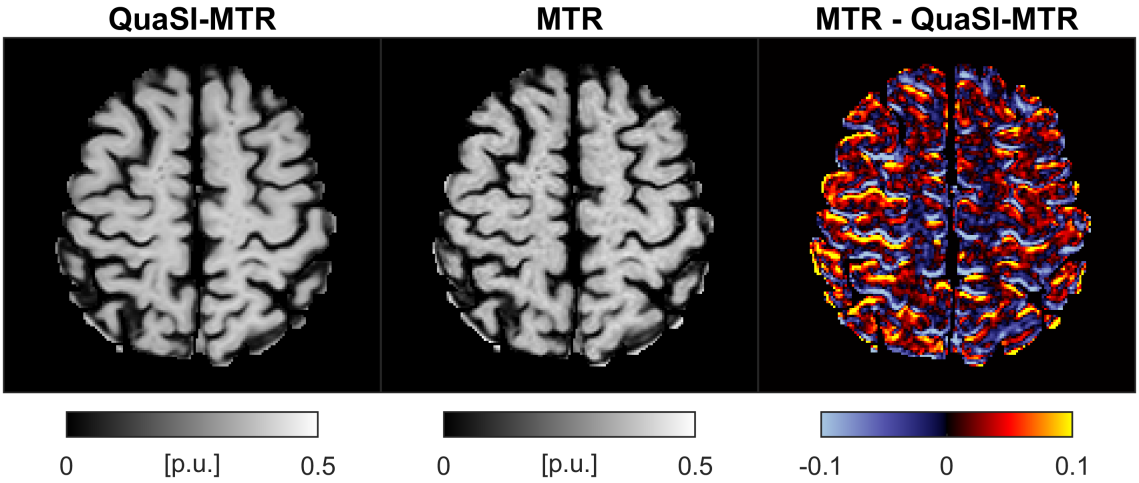


Figure 12.3: Effects of misregistration between QuaSI-MTR and reference MTR maps on the residual maps. P.u.: percentage units within [0, 1].

12.2 Field inhomogeneity effects

Finally, with reference to Figure 12.1, error maps in some subjects also exhibited low-frequency patterns suggestive of B_0 field inhomogeneity (arrow) near air-tissue interfaces (e.g. sinus) [119], which however did not occur for all subjects. The diverse incidence of this effect in the test-set can be observed in Figure 12.4, where reference MTR maps for two test-set RRMS subjects are shown: subject 1 does not present any visible inhomogeneity effects, whilst they are strong for subjects 2 (same subject shown in Figure 12.1).

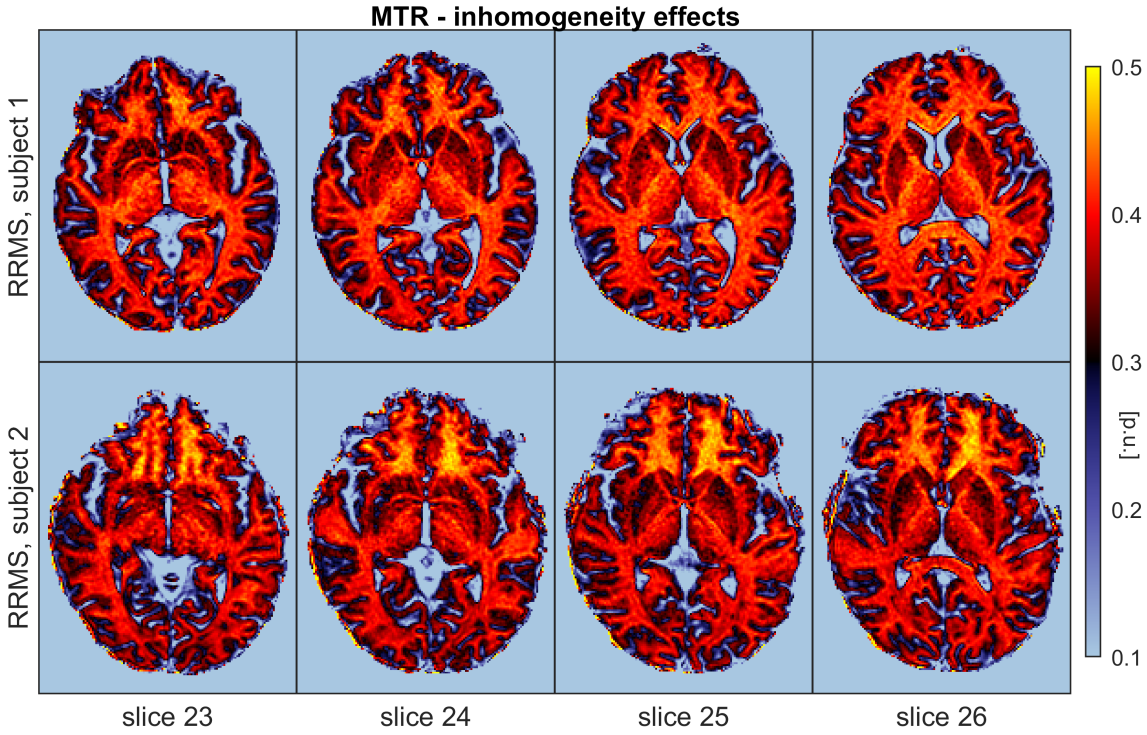


Figure 12.4: Inhomogeneities in ground truth MTR maps, manifesting as hyperintense regions in proximity of air-tissue interfaces (e.g. sinus). Both subjects 1 and 2 belong to the RRMS group, although inhomogeneities are observable only in subject 2 (subject 2 being the one shown in Figure 12.1). QuaSI-MTR maps do not seem to present the same artifact. P.u.: percentage units within $[0, 1]$.

Chapter 13

Discussions

The *U-Net MS application* objective aimed to offer surrogates for MTR using deep learning. Unlike deep learning model fitting applications, popular in medical imaging for using neural networks to replicate and speed-up model fitting, positive results were not granted. In deep learning model fitting, the network target output y is calculated from the input X *a priori* through a pre-existing model: $y = f(X)$; the aim of the training process is therefore to map the fitting function f in terms of neuron weights rather than least-squares, and thus the network *will* produce, eventually, results virtually identical to the traditional least-squares fitting ones. In this case there was no *a priori* certainty that a relationship between qualitative images and MTR could be found, and it was rather hypothesised based on *MyRelax validation* and *MyRelax MS application* objective results. This hypothesis was corroborated as, despite the reduced training-set (24 subjects for training, 12 for validation, and 12 for testing), QuaSI-MTR produced through deep learning from qualitative images did exhibit strong similarity with the ground truth MTR.

13.1 Limitations

This approach faced particular challenges that do not otherwise affect deep learning model fitting: sub-voxel misregistrations between target MTR and input qualitative images, the inherent noise affecting MTR, coupled with the small-sample size, contributed to the slight blurring of the QuaSI-MTR and increased error spread in hypointense or sharp-contrast areas, e.g. cortical GM. Additional data would certainly help reducing the error variance due to under-representation in the training-set, such as lesions or cortical GM interfaces with CSF and WM, although will not necessarily eliminate effects

due to random noise and misregistration.

Additional data would also be required to further investigate the bias field–like artifact observed in some reference MTR images, specifically with B_0 and B_1 mapping performed simultaneously to the MT-acquisition. Proper bias field mapping and correction would be necessary to assess in what amount this effect is due biological variability, rather than bias field itself. With respect to the former, more training data would certainly be required to ensure the U-Net is exposed to a sufficient amount of examples of MTR variability, whilst with respect to the latter, that is in the hypothesis that this is mainly a field inhomogeneity effect, supported by the sparsity across the dataset independent of HC or MS status, this method would allow to generate QuaSI-MTR maps free of bias field effects without the need for post-processing corrections.

13.2 Generalisability

Additional multi-centre data would be necessary to assess how well this method fares with qualitative images acquired with different MR-scanners and protocols. Worse results are to be expected if using the same U-Net trained on the current data to regress QuaSI-MTR maps from qualitative scans acquired with very different acquisition parameters and on different machines, and whilst a rich multi-centric training-set encompassing a wide range of TEs and TRs, and scanner manufacturers might be a solution, it might not be the *best* one for the niche this method aims to fill.

This method is not intended to act as a one-for-all solution for surrogate MTR mapping from any set of qualitative data, but to provide each research centre the opportunity to perform MT-analyses by training their own network tailored to the *specific* qualitative data available to them. Different research centres might have large datasets of qualitative data acquired with different protocols, on different MR-scanners, through different scanner upgrades, and a single network performing comparably well for any dataset might be not only impractical, but also superfluous for any given centre having access to data acquired in the same, or similar, conditions. One of the most important limitations of deep learning approaches in general is in fact their *generalisability* to data not explicitly represented in the training-set, but a greatly generalised network is still not going to perform better, on any given dataset, than a network specifically trained for that dataset alone.

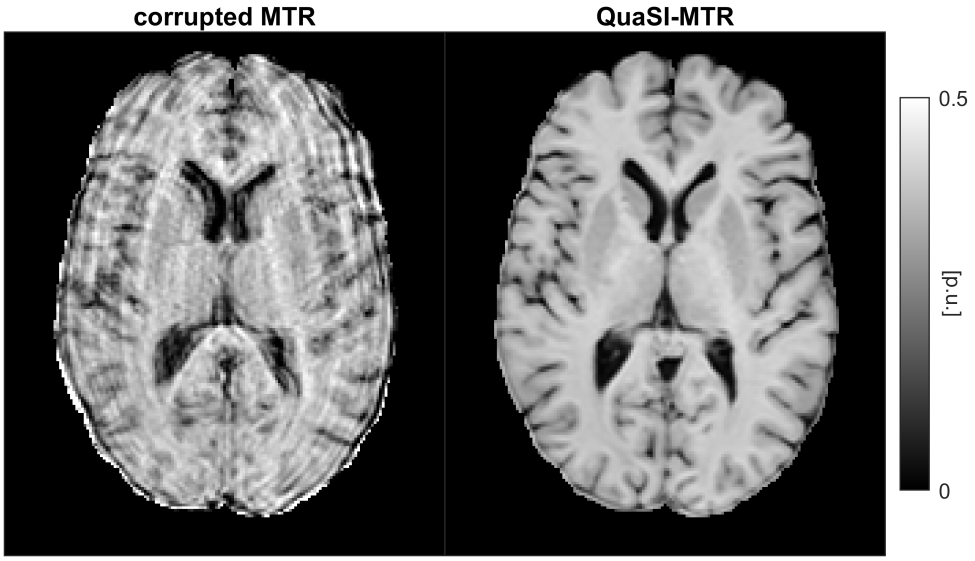


Figure 13.1: QuaSI-MTR use case. In case of corrupt MTR data, QuaSI-MTR can be employed to recover MTR information from qualitative images, rather than discarding the entire subject entry. P.u.: percentage units within $[0, 1]$.

13.3 Fine-tuning

After having identified an archive of qualitative images suitable for this analysis, similarly to the calibration step described for the *MyRelax validation* objective, a small cohort of subjects with both MTR and qualitative data acquired in the same conditions can be gathered from different projects, or prospectively acquired if necessary, and used to train a network optimised to produce QuaSI-MTR maps from those particular qualitative images. A pre-trained network, like the one resulting from this study, can be used as starting point, with the new training-set being used to *fine-tune* the new network to the particular data at hand. So whilst additional data from multiple centres could greatly help in generating a strong and versatile pre-trained network, fine-tuning would enable to tailor the network to the study specifics.

13.4 Missing data

In addition to retrospective, historical datasets of qualitative images, this method could also be applied to new MT-studies, with missing or corrupted MTR maps. Using Figure 13.1 as an example, one can see that instead of discarding the entire subject entry, which might include multi-modal acquisitions and clinical assessment information, and thus took time and resources to gather, QuaSI-MTR could be employed to recover

MTR information from the qualitative scans, using the remaining, correctly acquired data to train (or fine-tune) a network for this purpose.

Summary

- Deep learning can be used to produce QuaSI-MTR that well correlates with ground truth MTR.
- QuaSI-MTR might provide a way to perform MTR studies when dedicated MT-scans are unavailable or missing.

Part IV

Biophysically meaningful features for classification of MS phenotypes

Chapter 14

Introduction

Given the high-dimensional landscape of MRI modalities, being able to identify those that are most likely to provide meaningful information for any given task, and are thus worth clinical optimisation and adoption, is key for an accurate and efficient understanding of MS mechanisms and prognosis of disease progression. Whilst the previous contributions offered possible ways to enrich a dataset by extracting quantitative information from qualitative images, this study aimed, through the *top-down* objective, to decompose an already rich, multi-modal dataset to the constituent MRI features that are most likely to be biophysically meaningful with respect to the MS phenotypes.

A multi-modal dataset of healthy controls (HC), subjects affected by a clinically isolated syndrome (CIS) and clinically defined MS patients with relapsing remitting (RRMS) and secondary progressive (SPMS) MS-phenotypes has been used. The dataset was used to train and test support vector machine (SVM) and random forest (RF) machine learning algorithms in order to explore the correlation between MRI features and MS subtypes.

Chapter 15

Methods

15.1 Cohort

The *top-down* study cohort, called *CIS2014*, consisted of a total of 123 subjects: 29 HC (10 men, age: 35±10 years old), 18 CIS (6 men, age: 47±10 years old), 63 RRMS (15 men, age: 47±8 years old), 13 SPMS (4 men, age: 48±8 years old) patients with same disease duration of 15 years after the first CIS. CIS patients did not manifest any new MS-related symptom over the same 15 years time period.

15.2 MRI protocol

MRI data were acquired on a 3T Philips Achieva scanner. This study was approved by the local ethical committee.

The acquisition protocol included:

1. **PD/T2**. Dual-echo 2D PD-/T_2-weighted turbo spin-echo (TSE).
2. **T1**. 2D T_1-weighted spin-echo (SE).
3. **DWI**. Multi-shell diffusion-weighted EPI.
4. **Na**. Sodium 3D-cone gradient echo (GRE).
5. **3DT1**. 3D sagittal T_1-weighted MP-RAGE — magnetisation-prepared rapid GRE.

Details about the sequences are summarised in Table 15.1. Details about the DWI

shells are reported in Table 15.2.

Two 4% agar phantoms with sodium concentration of 40 mM and 80 mM were placed near the subject's head during the sodium acquisition for calibration purposes. The 3DT1 scan was used for brain tissue segmentation purposes. Anatomical and DW-images were acquired using a 32 channel head coil, whilst sodium imaging was performed using a single channel transmit-receive volume head coil (Rapid Biomedical, Rimpar, Germany).

Table 15.1: CIS2014 MRI protocol details.

	scan time	Res [mm]	FOV [mm]	slices	slice orientation	sequence	TE [ms]	TR [ms]	TI [ms]	flip angle[°]
PD/T2	04:02	RL = 1 AP = 1 FH = 3	RL = 240 AP = 250 FH = 150	50	coronal	TSE	19/85	3500		90
T1	05:43	RL = 1 AP = 1 FH = 3	RL = 240 AP = 240 FH = 150	50	coronal	SE	10	625		90
3DT1	06:32	RL = 1 AP = 1 FH = 1	RL = 256 AP = 256 FH = 180	180	sagittal	GRE	3.1	6.9	823	8
DWI	16:34*	RL = 2.3 AP = 2.3 FH = 2.5	RL = 220 AP = 220 FH = 150	60	coronal	EPI	82	13846*		90
Na	∼40:00	RL = 3 AP = 3 FH = 3	RL = 240 AP = 240 FH = 240	80	coronal	3D cone -GRE	0.22	120		90

* Nominal, actual time depending on heart rate; TR = 12 beats.

Table 15.2: DWI shells.

	b-value [s/mm^2]	Directions
DWI	300	8
	711	15
	2000	30

15.3 Image analysis

15.3.1 Preprocessing

The preprocessing was performed on the medical image data management tool XNAT [109]. The preprocessing included lesion delineation and filling, registration and brain segmentation as described in section 7.4.1.

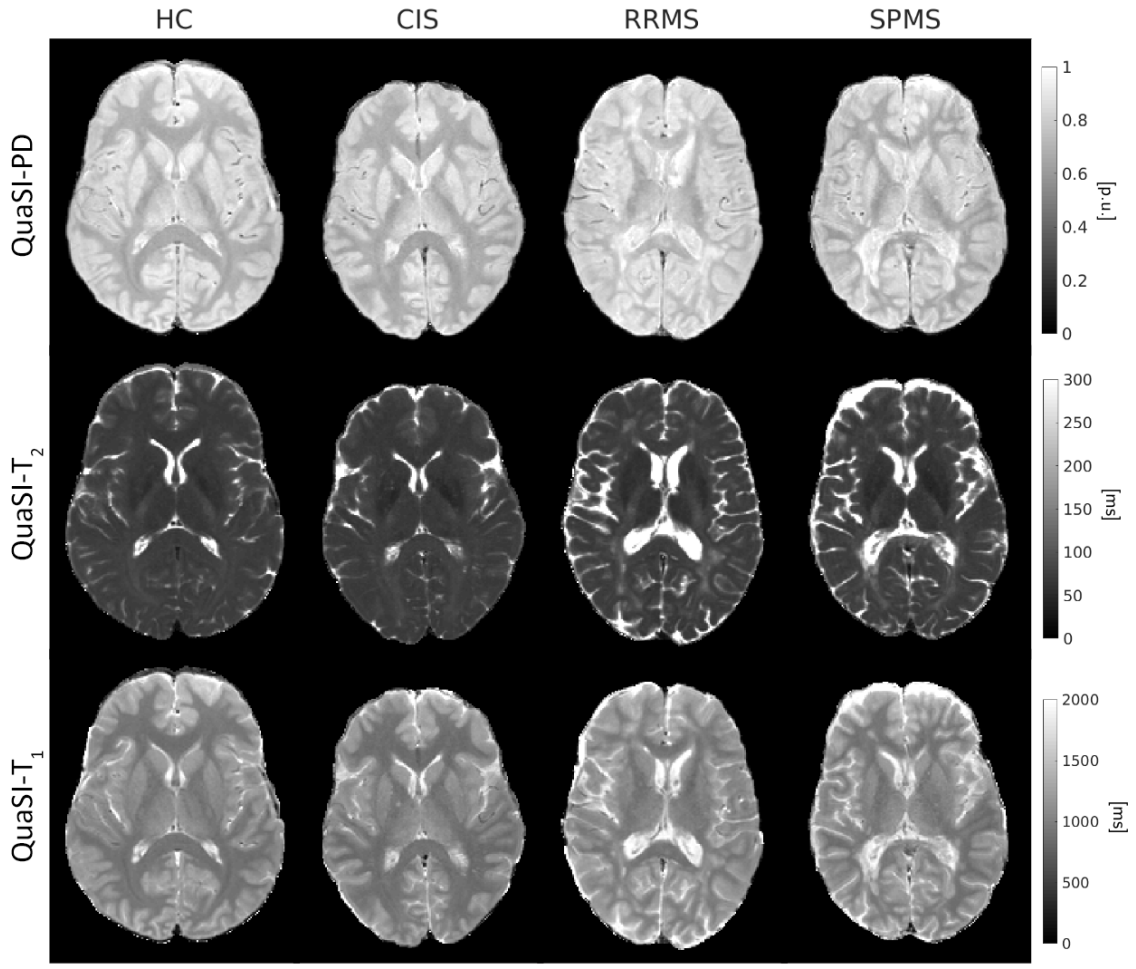


Figure 15.1: Examples of QuaSI-PD, -T_2 and -T_1 maps for HC, CIS, RRMS and SPMS subjects. Periventricular lesions in RRMS and SPMS patients are clearly recognisable as hyperintense regions in all modalities. P.u.: percentage units within $[0, 1]$.

15.3.2 Relaxometry: MyRelax

Relaxometry maps were computed using the MyRelax framework as described in section 7.3.2. Examples of quantitative QuaSI-PD, -T_2 and -T_1 maps for HC, CIS, RRMS, SPMS subjects are shown in Figure 15.1.

15.3.3 Diffusion imaging: spherical mean technique

DWI analysis was performed using *spherical mean technique* (SMT), as described in section 4.17.5. The SMT toolbox [120] was used to compute quantitative maps of intra-neurite volume fraction ν_{in} (*intra*), intrinsic diffusivity λ (*diff*) and orientation entropy $H(q)$ (*entropy*). Example maps are shown in figure 15.2 for HC, CIS, RRMS and SPMS subjects.

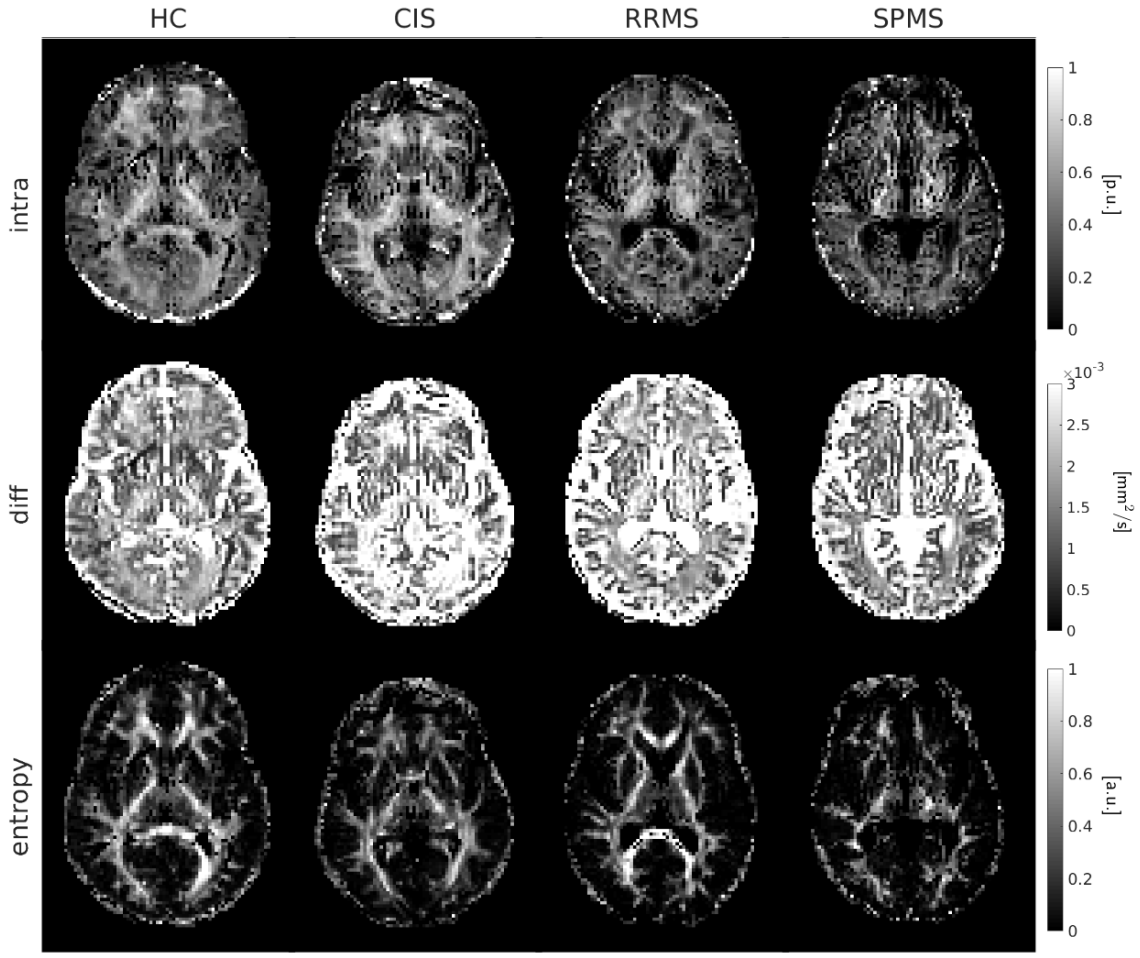


Figure 15.2: Examples of intra-neurite volume fraction (*intra*), intrinsic diffusivity (*diff*) and neurite orientation entropy (*entropy*) maps for HC, CIS, RRMS and SPMS subjects. Periventricular lesions are visible in RRMS and SPMS patients as regions of hypointense intra-neurite volume fraction, which is suggestive of disruption of the tissue microstructure. No lesions were reported for the CIS subjects. P.u.: percentage units within $[0, 1]$; a.u.: arbitrary units within $[0, \infty)$.

15.3.4 Sodium imaging

Total sodium concentration (TSC) was calculated by calibrating the $^{23}_{11}$Na MR-signal in the brain over the one generated by two phantoms with known TSC placed near the head of the subjects at the time of the scan [71]. The phantoms were segmented automatically as described in Prados et al. (2016) [121].

An example of TSC maps for HC, CIS, RRMS and SPMS subjects is shown in Figure 15.3.

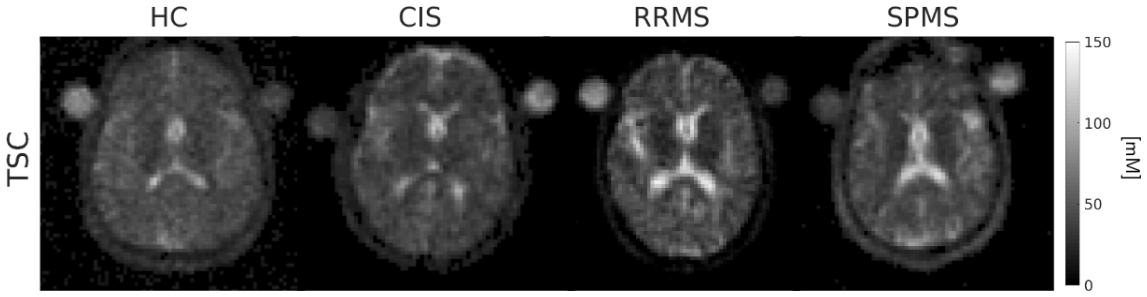


Figure 15.3: Examples of total sodium concentration (TSC) maps for HC, CIS, RRMS and SPMS subjects. No clear alterations can be spotted for MS patients. Notice the presence of the calibration phantoms in the TSC maps.

15.4 Features

The extracted MRI features were regional measurements of

- **relaxometry**: QuaSI-PD, -T_2 and -T_1(the QuaSI- prefix may be omitted in graphs or summaries for better readability);
- **diffusion imaging**:*intra*, *diff*, and *entropy*;
- **sodium imaging**, i.e. TSC;
- **atrophy**, i.e. tissue volume.

Summary statistics for each metric were calculated in white matter (WM), cortical and deep grey matter (GM), for a total of 24 features for each subject.

Summary statistics were calculated by first excluding outliers from the data distribution in each region, following the IQR method described in equation (11.1) in order to reduce artifacts and partial volume effects, and then computing the median value of the resulting distribution. Volumetric features were calculated by counting the number of voxels within each tissue mask, and dividing the result by the *total intra-cranial volume*, accounting for head-size variability.

This process was repeated excluding lesions from the regional distributions: MS patients' lesion masks were dilated with a 3×3×3 uniform kernel (i.e. by one voxel over all dimensions) and subtracted from the tissue segmentation masks. Therefore, in addition to the *baseline* dataset calculated over the whole brain, a *lesion-free* dataset calculated over only normal appearing tissue was also produced. Volumetric features were kept the same for the two datasets.

15.5 Classification

The raw data consisted of a 123×24-sized matrix X containing the summary statistics for the 24 brain ROIs for each of the 123 subjects, and a 123-long array y indicating the subjects' class target *labels* (HC = 0, CIS = 1, RRMS = 2, SPMS = 3). The dataset was used to train and test SVM and RF algorithms, over different binary classification tasks: HC vs MS (that is RRMS and SPMS), CIS vs MS, and all binary permutations of HC, CIS, RRMS and SPMS. This was implemented using Python 3.7.4 [100] and the `scikit-learn` (`sklearn`) package [122].

15.5.1 Data initialisation

X was *standardised* column-wise such that the data distribution in each column had mean of zero and standard deviation of one. Depending on the classification task, subjects not involved in the classification were excluded, and the remaining labels binarised — e.g. in the HC vs {RRMS, SPMS} classification task, the rows of X and elements of y corresponding to CIS subjects were excluded, and {RRMS, SPMS} labels were set to 1: $y = 2, 3 \rightarrow y = 1$. $(X, y)^t$ will refer to the standardised data and target labels filtered for a given classification task t.

15.5.2 Support vector machine

For SVM classification, the `svm.SVC` function from the `sklearn` package was used, selecting *linear* as the kernel of choice. Whilst polynomial and RBF kernels can in fact offer improved fitting performances in general, they also come with increased chance of *overfitting*, which is a concern particularly in datasets with small sample-size. The cost parameter C and the number of features K actually used for the classification were treated as *hyper-parameters* and learned through *cross-validation* (CV) via *grid-search*, with

$$\begin{aligned} C &= 0.1, 1, 10 \\ K &= 1, 2, 3, 5, 7, 10, 12, 15, 18, 24 \end{aligned} \tag{15.1}$$

For each classification task t, a 10-fold CV is implemented to select $1/10$ of the subjects to be used for *testing* and the remaining $9/10$ for *training*. The splitting is *stratified*, ensuring approximately the same class-proportions both in the training and test sets. This is iterated 10 times until the entire dataset is used both for training and testing,

which in turn is repeated 10 times with different splitting orders, for a total of 100 iterations. For each i-th iteration, the splitting follows:

$$(X, y)^t = \begin{matrix} \frac{1}{10} \\ \\ \frac{9}{10} \end{matrix} \begin{cases} (X^{\text{test}}, y^{\text{test}})_i^t \\ (X^{\text{train}}, y^{\text{train}})_i^t \end{cases} \quad i = 1, ..., 100 \tag{15.2}$$

For each i-th iteration, the training set (which for clarity will be referred to as *outer*-training set) is further split into *validation* and *inner*-training sets through 5-fold stratified-CV. This process is again repeated 5 times, for a total of 25 inner-CV j-iterations:

$$(X, y)^t \rightarrow \begin{matrix} \frac{1}{10} \\ \\ \frac{9}{10} \end{matrix} \begin{cases} (X^{\text{test}}, y^{\text{test}})_i^t \\ (X^{\text{out}}_{\text{train}}, y^{\text{out}}_{\text{train}})_i^t \rightarrow \end{cases} \begin{matrix} \frac{1}{5} \\ \\ \frac{4}{5} \end{matrix} \begin{cases} (X^{\text{val}}, y^{\text{val}})_j^t & i = 1, ..., 100 \\ (X^{\text{in}}_{\text{train}}, y^{\text{in}}_{\text{train}})_j^t & j = 1, ..., 25 \end{cases} \tag{15.3}$$

For each j-th iteration, an *analysis of variance* (ANOVA) is run on $(X^{\text{in}}_{\text{train}}, y^{\text{in}}_{\text{train}})_j^t$ to rank the 24 features based on their ability to discriminate between the classes in the task. A SVM_c classifier is then trained on $(X^{\text{in}}_{k\,\text{train}}, y^{\text{in}}_{\text{train}})_j^t$ for every combination of (c, k) with $c, k \in C, K$ as defined in (15.1), where SVM_c indicates a SVM estimator with regularisation parameter set to c, and $X^{\text{in}}_{k\,\text{train}}$ the dataset *reduced* to the first k most discriminating features according to the ANOVA.

The trained model is then applied to the reduced validation set X^{val}_k, producing a *probability* array for the predicted labels $\hat{y}^{\text{val}}$. The classification performance is assessed by comparing the prediction to the target labels y^{val} via *receiver operating characteristic* (ROC) *area under the curve* (AUC)[1], with ROC AUC scores close to 1 indicating good classification performance, and ROC AUC scores around 0.5 corresponding to chance (see section 5.2).

The inner-CV is repeated over $j = 1, ..., 25$, and the average ROC AUC is calculated for each combination of the grid-search, with the best hyper-parameter values $(\bar{c}, \bar{k})$ being the ones associated to the highest mean ROC AUC on the validation sets.

Once the best hyper-parameters have been found, a new ANOVA is then run on the outer-training set, $\bar{k}$-features are selected, and a new $\text{SVM}_{\bar{c}}$ estimator is trained on $(X^{\text{out}}_{\bar{k}\,\text{train}}, y^{\text{out}}_{\text{train}})_i^t$. The trained model is applied to the reduced test set $X^{\text{test}}_{\bar{k}}$, and the

[1] In order to calculate the ROC AUC, *probabilistic* prediction was used.

ROC AUC score calculated between predicted $\hat{y}^{\text{test}}$ and target labels y^{test}. The average ROC AUC across the $i = 1, ..., 100$ iterations indicates the estimator classification performance on the test sets.

In addition to `svm.SVC`, the following `sklearn` functions were employed:

- `model_selection.RepeatedStratifiedKfold`, for the data-splitting;
- `model_selection.GridSearchCV`, for the grid-search CV;
- `feature_selection.SelectKBest`, for the ANOVA feature-ranking;
- `metrics.roc_auc_score`, for the ROC AUC score calculation.

15.5.3 Random Forest

For RF classification, the `ensemble.RandomForestClassifier` function from the `sklearn` package was used, with the number of trees set to 1000. Due to RF robustness against overfitting and internal feature selection, no model selection was required, with the remaining parameters being left to default.

For each classification task, a 10-fold stratified CV with 10 repetitions was implemented as described in (15.2). Unlike the SVM training, inner-CV loop for model selection was not required, otherwise the RF training and testing followed the same pipeline. The RF classification performance was given by the average ROC AUC score on the test set across the $i = 1, ..., 100$ train/test iterations.

Variable importances were averaged across iterations, returning the mean feature ranking for the task; this allowed to identify the features that most contributed to each classification task, and thus are more likely to be biophysically meaningful with respect to MS progression.

The same process was repeated for the lesion-free dataset as well to investigate the effect of lesions (or their absence) on RF classification performances.

15.5.4 Dealing with imbalanced data

The different number of subjects within each group make this dataset *imbalanced*. For this reason, ROC AUC has been used to estimate the classification performance instead of *accuracy*, i.e. the ratio between the number of correctly classified subjects over the total number of subjects classified. Accuracy alone would in fact over-estimate

classification performances that favour indiscriminately the *majority* class. In these circumstances, *true positive rate*, or *sensitivity*, and *false positive rate*, i.e. 1−*specificity*, offer a more meaningful estimation of classification performances, which the ROC AUC conveniently summarise within a single score.

Dataset imbalance has also a chance to affect the training, with the trained model being exposed to more data points belonging to the majority class, which may lead to predictions biased towards that class. This effect can be counteracted with *re-sampling* to artificially balance data during training. In this study, two data re-sampling strategies were explored with both SVM and RF classification, using the `imblearn` package [123]:

- **Random under-sampling**: it allows to randomly sub-sample the majority class such that it matches the size of the minority class. This method effectively reduces the amount of information available to the classifier during training, discarding potentially useful data, which may cause reduction in performances. Under-sampling was implemented using the `under_sampling.RandomUnder Sampler` function.
- **Synthetic minority oversampling technique (SMOTE)**: it allows to generate synthetic data in the minority class such that it matches the size of the majority class. Synthetic instances are generated by interpolating neighbouring data points in the feature space, effectively performing *data augmentation* [124]. SMOTE has being shown to perform better than over-sampling by replication with repetition, however it could lead to poor classification performances in case of data characterised by high within-class variance and between-class similarity, as the newly generated data from the minority class may actually overlap with the majority class in the feature space, a phenomenon known as *over-generalisation*. Variants of the standard SMOTE (e.g. Borderline SMOTE) and/or hybrid re-sampling strategies can be applied in these cases to improve classification [125]. SMOTE was implemented using the `over_sampling.SMOTE` function.

15.5.5 Randomisation

In order to assess the significance of the classification results, the SVM and RF classification pipelines as described in section 15.5.3 were repeated 100 times with randomly permuted labels.

For each iteration $r = 1, ..., 100$ and any given task t, target labels y were randomly permuted, with $\tilde{y}_r$ being the permuted labels, keeping the original data X. A classifier

was then trained and tested on $(X, \tilde{y}_r)^t$ as described previously, and the average test ROC AUC score for each iteration recorded. The distribution of the 100 mean ROC AUC scores was then used as reference to calculate the p-value associated to the classification performances on the original data.

Chapter 16

Results

In this section, the classification ROC AUC scores for SVM and RF algorithms are reported for the different resampling methods (SMOTE, no resampling, and under-sampling). The permutation test outcomes are then shown, followed by the comparison between whole-brain and lesion-free datasets. Finally, the findings related to the MRI feature ranking are presented.

16.1 Classification

Classification results for SVM and random forest RF algorithms on the test-set are shown in Figure 16.1. Mean ROC AUC scores for the best model during validation are also shown for SVM. Median ROC AUC scores for the test-set are reported in Table 16.1; interquartile range (IQR) [25th percentile – 75th percentile] is used as a measure of uncertainty, as opposed to standard deviation, due to the asymmetry of the ROC AUC score distributions.

For both SVM and RF algorithms, and all three resampling strategies, the best classi-

Table 16.1: ROC AUC test-scores.

	SVM			RF		
Tasks	SMOTE	no resampling	under-sampling	SMOTE	no resampling	under-sampling
HC – RR	0.83 [0.72–0.96]	0.86 [0.72–0.95]	0.83 [0.71–0.94]	0.89 [0.75–1.00]	0.90 [0.77–1.00]	0.89 [0.76–1.00]
HC – SP	1.00 [1.00–1.00]	1.00 [1.00–1.00]	1.00 [1.00–1.00]	1.00 [1.00–1.00]	1.00 [1.00–1.00]	1.00 [1.00–1.00]
HC – MS	0.88 [0.81–0.96]	0.90 [0.81–1.00]	0.88 [0.81–0.95]	0.90 [0.81–0.95]	0.92 [0.81–1.00]	0.89 [0.79–0.96]
CIS – RR	0.83 [0.71–0.92]	0.83 [0.70–0.92]	0.86 [0.70–0.93]	0.83 [0.74–1.00]	0.85 [0.71–1.00]	0.83 [0.67–0.95]
CIS – SP	1.00 [1.00–1.00]	1.00 [1.00–1.00]	1.00 [1.00–1.00]	1.00 [1.00–1.00]	1.00 [1.00–1.00]	1.00 [1.00–1.00]
CIS – MS	0.88 [0.75–1.00]	0.88 [0.78–1.00]	0.88 [0.78–1.00]	0.88 [0.75–1.00]	0.88 [0.75–0.97]	0.88 [0.75–1.00]
RR – SP	0.69 [0.50–0.86]	0.83 [0.67–1.00]	0.79 [0.57–1.00]	0.83 [0.67–1.00]	0.83 [0.55–1.00]	0.83 [0.50–1.00]
HC – CIS	0.83 [0.67–1.00]	0.79 [0.67–1.00]	0.83 [0.67–1.00]	0.92 [0.67–1.00]	1.00 [0.67–1.00]	0.83 [0.67–1.00]

RR = RRMS; SP = SPMS; MS = RRMS, SPMS

fication results were obtained when classifying HC or CIS against SPMS patients, with median ROC AUC scores and IQR equal to 1.00 [1.00 – 1.00] for both. Lower and more spread-out ROC AUC scores were recorded when classifying HC or CIS against RRMS patients, with median ROC AUC scores of about 0.84 and IQR $\simeq$ [0.72 – 0.98] for HC vs RRMS using SVM, 0.89 and IQR $\simeq$ [0.76 – 0.98] using RF, and 0.84 [0.71 – 0.95] for CIS vs RRMS with both algorithms. Classification against both MS-subtypes at once fell in between. With regards to MS, both HC and CIS appeared to behave similarly, with CIS performing marginally worse (more spread out IQR), suggesting some shared traits with the MS groups, not observed in HC.

The worse classification performances were observed for SVM and the SMOTE resampling method for the RRMS vs SPMS task, with median ROC AUC of 0.69 [0.50 – 0.86], showing signs of overfitting to the validation-set, likely due to the heaviest class imbalance among the tasks (about 45 RRMS vs 8 SPMS in the inner training-set) and the synthetic over-sampled examples possibly overlapping with the majority class. Random under-sampling with SVM showed higher ROC AUC spread than no resampling, but close median ROC AUC scores: 0.79 [0.57 – 1.00] and 0.83 [0.67 — 1.00] respectively. Better results were observed for RF, with same ROC AUC median values of 0.83, but different IQR depending on the resampling: [0.67 – 1.00] for SMOTE, [0.55 – 1.00] for no resampling, [0.50 – 1.00] for under-sampling. The different results obtained for SVM and RF indicate that the model selection step in the SVM pipeline might exacerbate issues related to imbalanced datasets, particularly in cases of high class imbalance like this classification task.

The classification of HC versus CIS showed interesting results, with median ROC AUC scores of about 0.82 for SVM and 0.92 for RF, and IQR = [0.67 – 1.00] for both. Despite the similar performances when classifying HC and CIS against MS, and CIS subjects not having manifested any MS-related symptoms in the 15 years prior to the acquisition, these results suggest long-term alterations might have accrued in the CIS population following the initial clinically isolated syndrome, discriminating them from HC on an asymptomatic level. A more conservative explanation revolves however on the confounding effect of age differences between the two groups, with HC being, on average, 12 years younger than CIS. These hypotheses and their implications are discussed in detail in section 17.2.

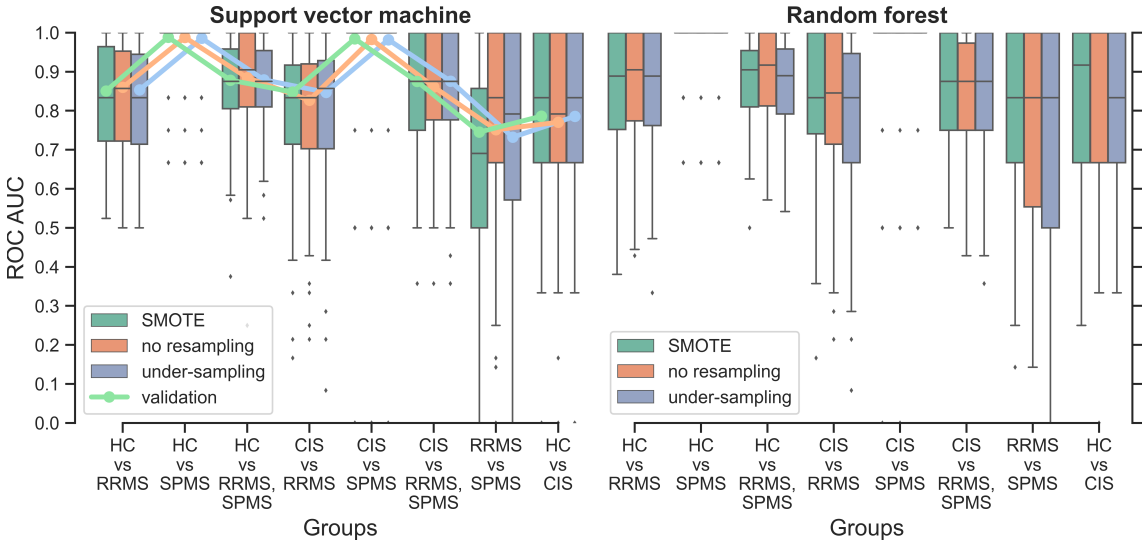


Figure 16.1: Classification results for SVM and RF algorithms, with no data re-sampling, SMOTE and random under-sampling. Best-model mean ROC AUC validation scores are shown for SVM for the three re-sampling methods; no model selection was implemented for RF instead. HC vs SPMS and CIS vs SPMS ROC AUC scores are mostly 1.0.

16.2 Randomisation

Permutation test results for SVM and RF algorithms, and the three resampling strategies, are shown in Figure 16.2. For each permutation iteration, the average ROC AUC score was calculated over the test-set: the resulting distribution was used to assess the p-value associated to the average ROC AUC score for each classification task. All classification results fell below $p < 0.01$, except for the RRMS vs SPMS task, with $p < 0.05$ for SVM and random under-sampling, and not significant results for SVM with SMOTE — as expected given the overfitting.

Additionally, no-resampling appeared to produce *skewed* permutation distributions, i.e. with tails towards high ROC AUC scores, and thus not symmetric around 0.50 as expected for randomly permuted data. Upon further investigation, this issue was again revealed to be caused by the data imbalance, coupled with model selection. This was corroborated on a toy dataset of comparable sample-size as one of the classification tasks exhibiting this behaviour, and exaggerated class imbalance ratio of 1:10 for classes 0 and 1 respectively, with 1 being the majority class. The toy dataset was generated using the `make_classification` function from the `sklearn.datasets` package. Upon training a SVM algorithm with no data resampling, following the same pipeline used above, and applying it to a toy test-set of 10 normally distributed random

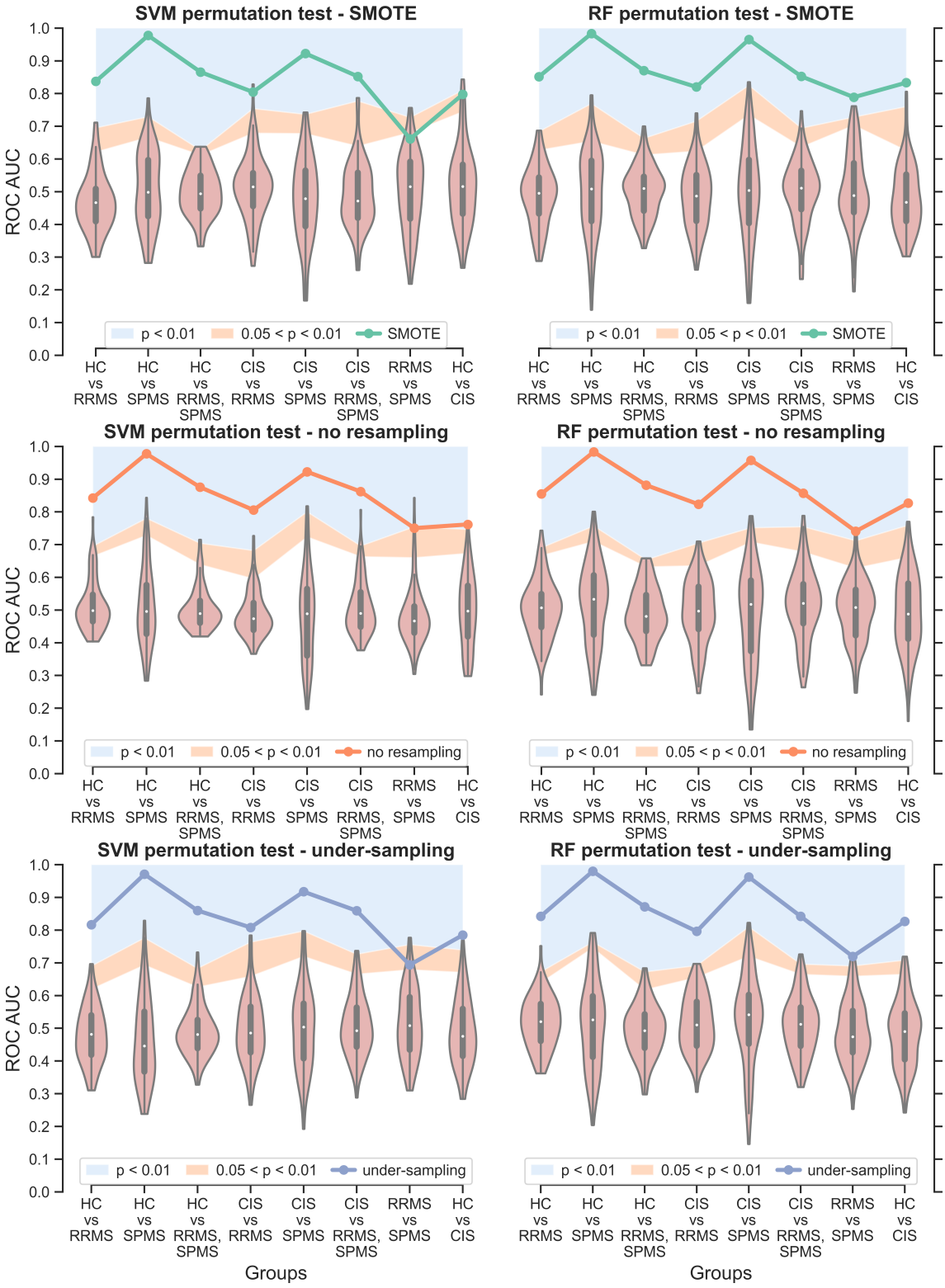


Figure 16.2: Permutation test. Mean ROC AUC scores were compared to the distributions of mean ROC AUC scores over 100 permutations. $p < 0.01$ was observed for most tasks; not significant results were observed using SMOTE (over-fitting). No resampling produced skewed permutation distributions due to the data imbalance. RF did not show differences for the different resampling methods.

data entries, classification probabilities of:

$$P(\hat{y}|1:10) = \{0.74, 0.73, 0.73, 0.73, 0.73, 0.73, 0.73, 0.74, 0.73, 0.74\} \quad (16.1)$$

were obtained. When inverting the class imbalance ratio to 10:1, with 0 being now the majority class, but keeping the rest unchanged, the opposite was observed:

$$P(\hat{y}|10:1) = \{0.27, 0.27, 0.27, 0.27, 0.27, 0.27, 0.27, 0.27, 0.27, 0.27\} \quad (16.2)$$

where $P < 0.50$ determines a prediction for class 0, or class 1 otherwise. In these conditions, the SVM algorithm tends therefore to favour the majority class, which in turn might translate into higher ROC AUC scores if the test-set is also equally imbalanced.

In the case of the dataset used in this study, the imbalance ratio was not as extreme and, in fact, with the exception of SMOTE overfitting in the RRMS vs SPMS task, no other important differences were observed in terms of SVM classification performances between either of the resampling methods (SMOTE or random under-sampling) and no resampling. This phenomenon only emerged when classifying randomly permuted data, thus presumably only when an underlying relationship between classes could not be found, and it was likely exacerbated by the model selection, given no such behaviour was observed for RF. In light of these observations, only RF will be taken into consideration for the following steps.

16.3 Lesions

The results reported so far were produced from a dataset of whole brain regional values, which included lesions when calculating the median value for each tissue distribution. RF classification results for data produced upon excluding lesions from the regional distributions are shown in Figure 16.3. No major differences were observed between the two sets of results.

Whilst the presence, number, location and characterisation of lesions are fundamental features for MS diagnosis, lesions also affect often relatively few voxels and with sparsely heterogeneous patterns across the MS population. These results suggest in fact the effect of lesions might not be as prominent when analysing median regional values, as their information might be lost when computing summary statistics, or overshadowed by more macroscopic alterations (e.g. atrophy, see next section).

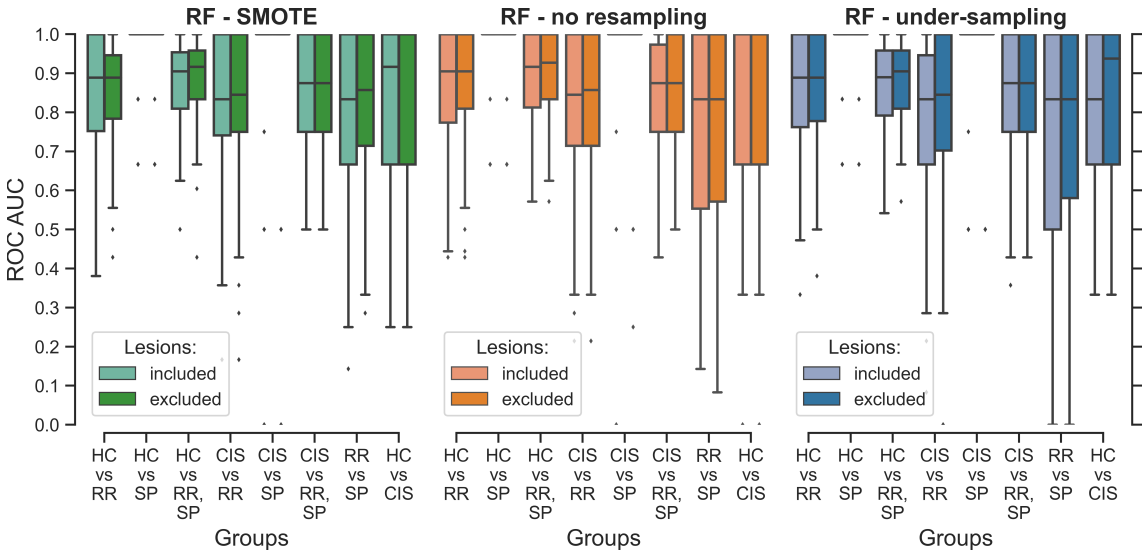


Figure 16.3: Effect of lesions on classification. Similar results were obtained upon excluding lesions prior to calculating regional values, compared to keeping them included. RR = RRMS, SP = SPMS.

16.4 Biophysically meaningful features

For the purpose of investigating possible biophysically meaningful features with respect to MS pathophysiology, RF *variable importances* were observed. Given the similar classification performances between resampling methods and no resampling, in order to reduce the likelihood of spurious results coming from synthetic over-sampling, and information loss associated to under-sampling, no data resampling strategy was implemented. Feature rankings and associated data distributions for each classification task are shown in Figures 16.4 and 16.5. The features contributing to 50% of the classification, i.e. the top ranked features whose combined importances amount to half the total (0.5), have been highlighted.

Feature ranking distributions showed highest skewness for HC or CIS classification against SPMS, which is compatible with the best classification performances across the tasks, indicating fewer features contributing the most to the classification. With respect to the top ranked features contributing to 50% of the classification process, when classifying HC against MS patients, *atrophy* appeared to be the most relevant set of features (↓*vol* in MS), particularly volume loss in deep GM. *Entropy* and *intra-neurite volume fraction*, particularly in WM, also appeared to contribute when classifying against RRMS, suggesting a higher *relative* incidence of microstructural alteration in terms of reduced neurite orientation coherence and integrity with respect to SPMS,

respectively ($\downarrow entropy$, $\downarrow intra$). Similar atrophy contributions were observed when classifying CIS against MS ($\downarrow vol$), although with a higher incidence of microstructural alterations described by reduced intra-neurite volume fraction in WM ($\downarrow intra$), and the emergence at lower ranks of relaxometry-related features, suggestive of inflammation, as well as demyelination, particularly in WM ($\uparrow PD$, $\uparrow T_2$, $\uparrow T_1$).

A different pattern of alteration was observed when classifying MS subtypes against each other. No atrophy contribution emerged from the top-ranked features, which instead included a strong component of diffusion-related microstructural alterations

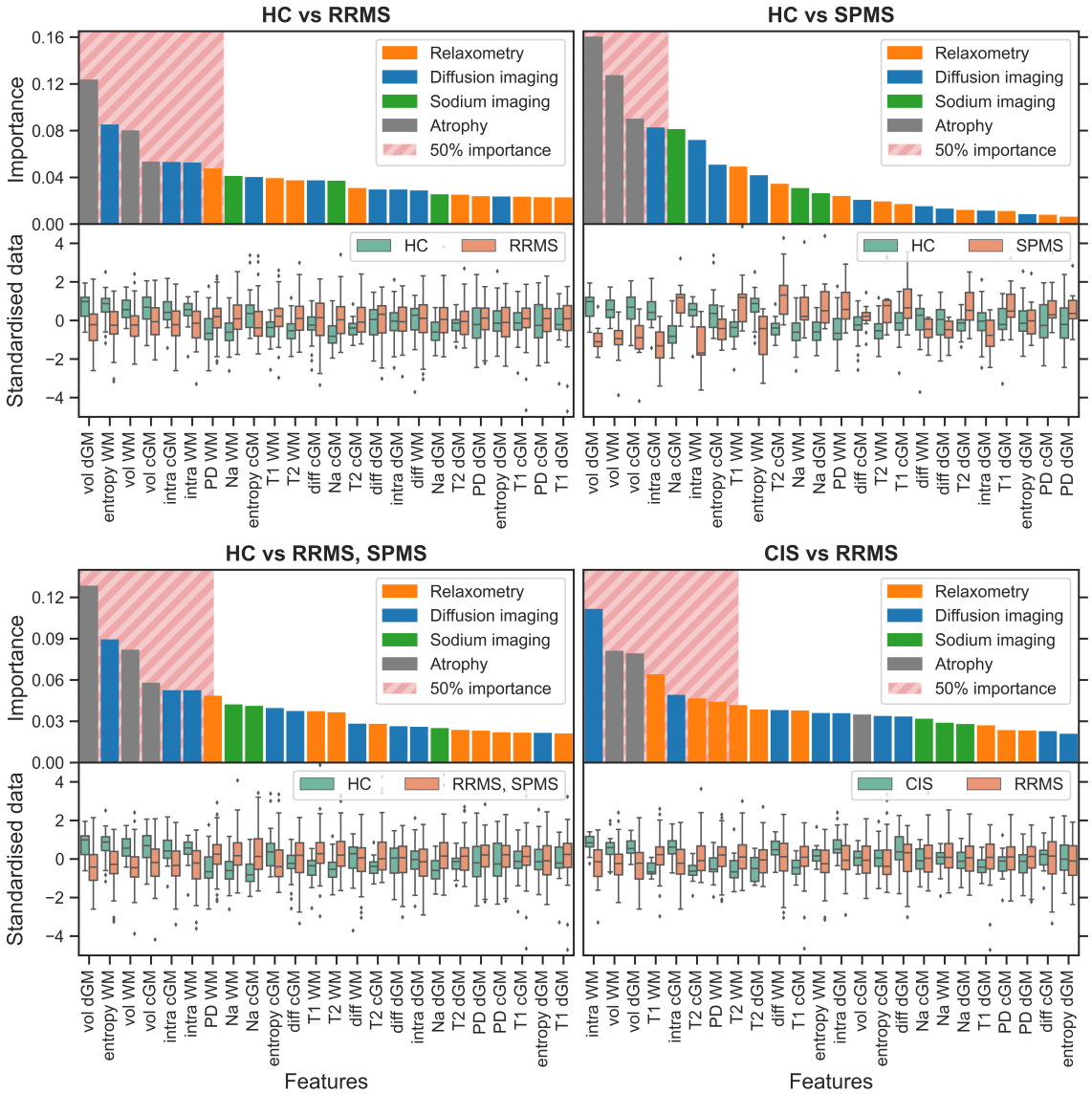


Figure 16.4: Biophysically meaningful features. Atrophy appeared to be the most important feature when classifying HC vs MS patients, with diffusion-related alterations also contributing when classifying against RRMS. Continues to Figure 16.5.

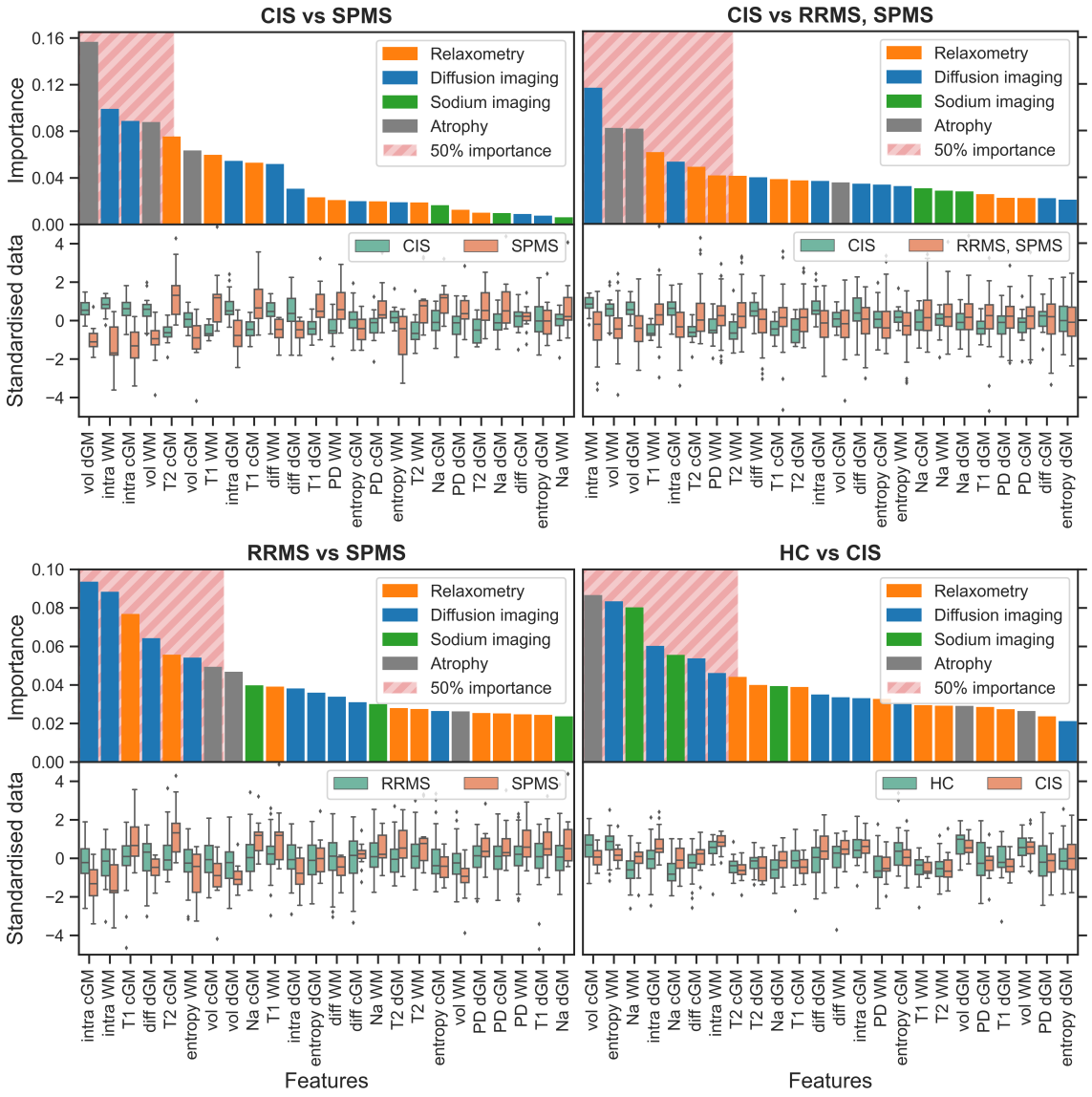


Figure 16.5: Biophysically meaningful features. Continued from Figure 16.4. Atrophy contributed to CIS vs MS classification as well, with stronger incidence of diffusion- and relaxometry-related alterations. Different patterns of alteration were observed in the RRMS vs SPMS classification task, dominated by diffusion and relaxometry alterations, but no atrophy, and HC vs CIS, characterised by cortical atrophy, diffusion and sodium alterations, but no relaxometry. *PD*, T_2, T_1: quantitative QuaSI-PD, -T_2, -T_1; *intra, diff, entropy*: intra-neurite volume fraction, intrinsic diffusivity, neurite orientation dispersion entropy; *Na*: total sodium concentration; *vol*: tissue volume (atrophy); *cGM*: cortical GM; *dGM*: deep GM.

given by lower intra-neurite volume fraction for SPMS patients compared to RRMS, and increased T_1 and T_2 in cortical GM ($\downarrow$*intra*, $\uparrow T_1$, $\uparrow T_2$ in SPMS with respect to RRMS). Given the small sample-size and high class imbalance associated in particular to this classification task, these results are likely to include spurious alterations, for

example the reduced *diffusivity* in deep GM for SPMS compared to RRMS ($\downarrow$*diff*), and thus more data would certainly be necessary to make any inference about the biological correlates of the observed results. However, given the distinct pattern of alterations observed for this classification task, dominated by diffusion and relaxometry metrics in the top-ranking features, one could speculate that whilst RRMS and SPMS differ from HC mostly in terms of structural alterations (atrophy), they appear to differ against each other on a more microstructural and inflammation/demyelination-related level.

Finally, when classifying HC against CIS, yet another distinctively unique pattern of alterations emerged. In addition to atrophy in cortical GM and diffusion-related alterations ($\downarrow$*vol*, $\downarrow$*entropy*, $\uparrow$*diff*, $\uparrow$*intra* in CIS), *total sodium concentration*, both in WM and cortical GM, appeared among the top-ranked features, with increased values for CIS compared to HC ($\uparrow$*Na*). This result is in line with studies showing increased total sodium concentration in cortical GM being associated to cognitive impairment in RRMS patients. The same caveat disclosed beforehand applies to this case as well: the increased *intra* goes against the usual reduction in intra-neurite volume fraction associated to MS, and thus could be spurious, however the fact that it emerged both in deep GM and WM, and that 15 years stable CIS subjects are not indeed MS patients, might be indicative of some other phenomenon worth investigating. The absence of relaxation-related metrics in the top-ranked features is compatible with the lack of pathology in the CIS population, whilst diffusion- and sodium-related alterations might be suggestive, respectively, of long-lasting microstructural and *functional* alterations accrued past the initial clinically isolated syndrome, despite the lack of symptoms. This is also compatible with the different classification results and patterns of alterations observed when discriminating HC against MS, with respect to CIS against MS. The emergence of total sodium concentration among the possible biophysically meaningful features is particularly interesting, since this is the only task where it showed a strong contribution to the classification, and with sodium imaging being mostly a niche modality in the MRI multi-modal landscape, it might be worth further investing.

Chapter 17

Discussions

The *top-down* study helped ranking MRI modalities with respect to different classification tasks, highlighting which ones might be biophysically meaningful in the context of MS characterisation. The dataset included QuaSI-PD, -T_2, and -T_1 maps extracted from qualitative data using the MyRelax framework, intra-neurite volume fraction, intrinsic diffusivity and neurite orientation dispersion entropy diffusion metrics, atrophy, and total sodium concentration in WM, cortical and deep GM. The cohort was composed of HC, MS patients with RRMS and SPMS phenotypes with 15 years disease progression, as well as subjects who did not manifest any new symptom suggestive of MS in the 15 years following their initial clinically isolated syndrome (CIS).

17.1 Limitations

The interpretation of these results is conditional to the small sample-size, and the different models used to fit the multi-modal MRI data, each coming with its own limitations and assumptions. Possible age-confounders due to the average younger HC population compared to CIS and MS patients might also influence the results, as described in detail below.

Overall, these results do not aim to portray a comprehensive picture of the biophysical alterations associated to the MS subtypes investigated, as this goes beyond the scope of machine learning: what follows is intended to discuss where these findings fall within the landscape of multi-modal MRI literature, and how they could help inform further research.

17.2 Atrophy

Atrophy emerged as the most important feature in the classification of HC against MS patients, with atrophy in deep GM scoring consistently higher than cortical GM or WM, for both RRMS and SPMS. When classifying HC against SPMS specifically, atrophy alone contributed to almost 50% of the classification process. This result is in line with previous studies reporting not uniform atrophy within the brain in MS, in particular deep GM showing the highest rate of tissue loss in relapsing-remitting and progressive MS [126]. Deep GM significant involvement in MS neurodegeneration is well known in the scientific community, however a consensus for the incorporation of *global* GM volumetrics into clinical practice has only recently been reached, and the inclusion of deep GM structures (e.g. thalami, basal ganglia) in particular is still debated [127, 128]. Further research is therefore recommended.

Similar result were observed when classifying CIS against MS, with comparable contributions from WM and deep GM, and below-threshold importance in cortical GM against RRMS. This suggests a certain degree of similarity in terms of cortical GM volume loss exists between CIS and RRMS, compared with HC. Cortical GM volume was in fact the most important feature when classifying HC against CIS, with deep GM and WM volume scoring very low in the ranking distribution, which indeed indicates, given the data at hand, cortical atrophy associated to the CIS phenotype.

The most conservative explanation revolves around *age*: it is known from the literature that cortical volume loss is mildly correlated with age [129], and the HC population used for this study had an average age 12 years lower than CIS, RRMS and SPMS, whilst the CIS and MS populations were age-matched. Negative correlation was indeed found within the HC population between both pre-standardised cortical and deep GM volume, and age: slope $\beta_1 = -0.000484$, $p = 0.009$ for cortical GM; $\beta_1 = -0.000066$, $p = 0.006$ for deep GM. Age-adjustment was performed, for each of these features, by subtracting $\beta_1 \times$ age from the data. GM volume distributions before and after adjusting for age are shown in Figure 17.1; t-test was also performed between all group pairs and statistically significant p-values reported as well. The results show that the statistically significant differences between cortical volume of HC vs CIS, and HC vs RRMS can be explained by the age covariance, as the null-hypothesis is no more rejected after correcting for age. This would explain why alterations in cortical GM were observed between HC and both RRMS and CIS, but not between CIS and RRMS, as CIS and RRMS are age-matched. All other comparisons, either for cortical or deep GM volume,

stay however significant before and after age-adjustment, and thus we might deduce that age has no significant effect on them, or changes due to MS are large enough to overcome age-related atrophy. On the other hand, it is not possible to definitely dismiss cortical GM volume loss in CIS and RRMS as an age confounder, as the age-adjustment process comes with its own limitations:

- the correlation between GM volume loss and age might not be actually statistically significant in the first place, if taking into account correction for multiple comparisons, e.g. Bonferroni $p \leq 0.05/n$, with $n = 24$ features, then the threshold for statistical significance would become $p \leq 0.002$;

- the age-adjustment coefficient was calculated on HC, with age range $[20 - 56]$ years, and there is no guarantee it will adequately correct data from CIS or RRMS patients with age range $[33 - 65]$ years, e.g. the correction might over- or under-estimate the effect of atrophy at older age and/or in presence of pathology;

- the observed cortical GM alterations affecting CIS and RRMS populations might not be completely due to age, and slow-rate cortical GM volume loss due to pathology might coexist with the age-related atrophy.

Overall, additional HC age-matched data would be required for further conclusive evidence.

That being said, previous studies reported deep GM volume loss in CIS patients *at presentation*, and in particular thalamic atrophy being associated with *conversion* to clinically defined MS after 2 years [130, 131], which were not however observed in this study. The findings reported in literature do not necessarily conflict with the observed results, and might actually be complementary: it is worth recalling that the CIS subjects analysed in this study present a 15 years-long stable clinical status, and are therefore likely *not* to convert to MS, which explains why no deep GM atrophy has been observed, whilst CIS subjects reporting deep GM atrophy at presentation had likely converted in the same time period.

Finally, no strong atrophy contribution was observed when classifying MS subtypes against each other, with cortical and deep GM volume loss scoring at the 50% cumulative importance threshold. Whilst there are differences in terms of GM volume loss, they do not seem to be as meaningful as other features when discriminating RRMS against SPMS at 15 years follow-up.

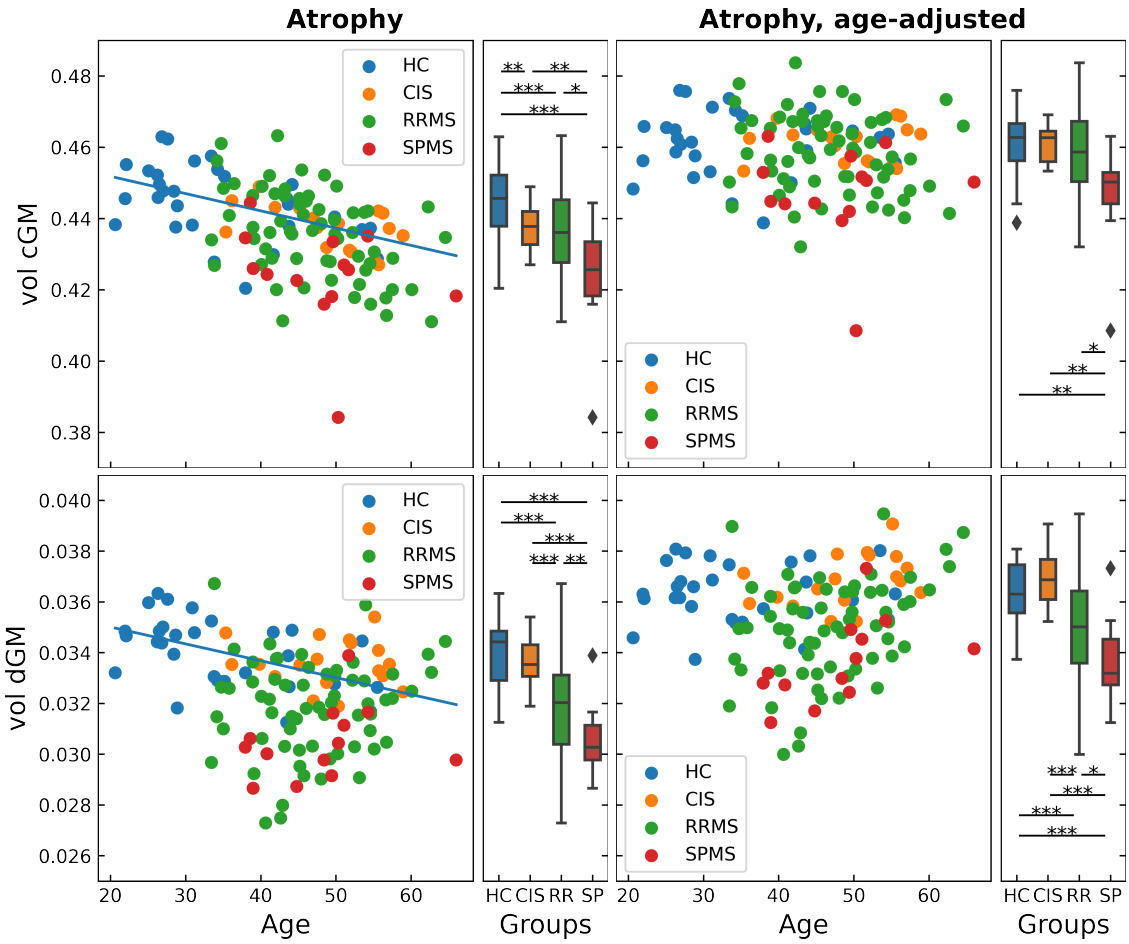


Figure 17.1: Correlation between GM atrophy and age. Top row: cortical GM; bottom row: deep GM. Left column: original data (not standardised), HC used to calculate the correction coefficient β_1 (blue line: $y = \beta_0 + \beta_1 x$); right column: age-adjusted data. Age covariance explains the statistically significant differences in cortical GM for HC vs CIS, and HC vs RRMS. *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$.

17.3 Diffusion-weighted imaging

Diffusion weighted imaging (DWI) metrics emerged in most classification tasks, suggesting a widespread involvement of microstructural alterations in MS pathophysiology.

Against HC, MS patients exhibited higher fibre dispersion (lower neurite orientation dispersion entropy) in WM, and intra-neurite volume fraction loss in WM and cortical GM. Together with atrophy, although with lower relative importance, diffusion metrics contributed to 50% of the classification process. This is in line with previous findings in *diffusion tensor imaging*, reporting reduced fractional anisotropy (FA) in several WM regions [132], compatible with reduced entropy, and increased density of transected neurites, causing intra-neurite volume loss, in both cortical lesions and myelinated

tissue [133] in MS. A very recent study by Johnson et al. (2021) on the same CIS2014 cohort used for this study has shown alterations in SMT intra-neurite volume fraction and entropy in RRMS patients, with respect to HC, to correlate with the associated DTI and NODDI metrics (see section 4.17.4), as well as with physical disability [134].

Similar results were observed in SPMS patients compared to RRMS, indicating a correlation in the severity of microstructural alterations with MS subtype. Slightly reduced intrinsic diffusivity was also observed in SPMS with respect to RRMS, which cannot be easily interpreted: likely a spurious result due to the heavy class imbalance (13 SPMS vs 63 RRMS in total), although there can be room for different explanations, as discussed below.

Reduced intra-neurite volume fraction in WM and cortical GM was also observed when classifying CIS against MS, but not entropy. As with cortical atrophy, this might indicate the presence of microstructural alterations in both the CIS and MS populations but, unlike cortical atrophy, this cannot be similarly explained in terms of age covariance, since no significant correlation between neurite orientation dispersion entropy and age could be found within the HC population ($p > 0.47$). Entropy in WM was indeed the second most relevant feature when classifying HC against CIS, or the first if excluding cortical GM volume as an age confounder, with reduced entropy being observed in CIS subjects. This is also in line with published studies reporting evidence for structural network alterations, namely but not limited to the thalamocortical network, involving reduced FA not explained by WM lesions, in CIS and patients with early MS [135, 136].

Increased intrinsic diffusivity in cortical GM, which could be however explained by age ($\beta_1 = 6 \cdot 10^{-6}$, $p = 0.04$), and alterations in intra-neurite volume fraction were also observed in CIS deep GM and WM with respect to HC, with no statistically significant correlation with age for the latter ($p > 0.15$). Opposite to the reduction observed for intra-neurite volume fraction in MS patients, and to what it would be expected in the context of a neurodegenerative disease, *increased* values were observed in CIS compared to HC. Given the small sample-size, this could also be explained as a spurious result, however given the same behaviour observed both in WM and deep GM, and the widespread contribution of diffusion metrics to the HC vs CIS classification, it is reasonable to explore other options involving microstructural alterations as well.

Apparently counter-intuitive results are in fact not new in the context of DWI, for example *increased* FA has been reported in both cortical lesions and normal appearing cortex, positively correlating with disability in MS [137], as well as in deep GM nuclei of patients

with Huntington disease [138]. This has been explained as selective neurites degeneration with consequent dendritic arborisation loss which, together with increased microglia activation, can cause increased *relative* orientation coherence, and thus higher FA [139].

At the same time, evidence for adaptive network reorganisation mechanisms has been reported, compensating for early tissue damage in CIS subjects at presentation [140]: increased intra-neurite volume fraction in WM and deep GM might therefore be the result of structural connectivity reorganisation to limit the impact of the physiological changes after CIS, e.g. on the thalamocortical network, whose effects are still visible 15 years after the onset in those subjects who did not convert to clinically defined MS. Alternatively, it could be an indicator of *axonal swelling*, which would increase the intra-neurite volume fraction, arising as a byproduct of an adaptive physiological change: this interpretation is particularly fitting when also considering the increased total sodium concentration observed in the same regions, as it is further explored in section 17.5. Either way, histological evidence would be required to further determine the nature of these alterations.

That being said, it is important also to keep in mind the dependency of these findings from the model used to fit the DW-data, and the relative assumptions. The SMT model used in this work assumes the signal to be produced by an intra- and an extra-neurite compartment that are impermeable to each other, and whose relative volume fractions add up to 1. These or similar constraints are common in clinically available multi-compartmental diffusion model, as they help keeping the model complexity and the data required clinically feasible, however they might only partially hold in case of MS pathology.

The model does not account, in fact, for a myelin compartment, as the T_2 associated to the macromolecular water pool would be too short to produce a clinically measurable signal, however the physical presence (or lack of thereof, in MS) of a myelin sheath would still affect the interaction between intra- and extra-cellular volume fractions, as their sum would not add up to 1 anymore, but rather be a function of myelin content. Furthermore, the impermeability approximation is likely sensible in the case of healthy, myelinated neurites, but might in fact not hold for unmyelinated neurites, as they lack the insulating sheath provided by myelin. Due to the predominance of demyelination in MS, these assumptions are likely to affect the fitted parameters, and thus, as the SMT authors have also noted [49], results should be interpreted accordingly.

17.4 Relaxometry

Relaxometry features extracted through the MyRelax framework from qualitative images emerged mainly when classifying CIS against MS patients, and RRMS against SPMS, exhibiting a trend of increasing values correlating with the severity of the clinical status. With respect to CIS, the RRMS population exhibited widespread WM alterations in terms of all relaxometry features, that is increased QuaSI-PD, -T_2, and -T_1 regional values. Increased QuaSI-PD in WM was also observed, at threshold, when classifying RRMS against HC, although with a much lower relative importance than atrophy and DWI-metrics. Increased QuaSI-T_2 compared to CIS was also observed in cortical GM, with a higher relative importance for SPMS. Alterations in cortical GM appeared to be in fact particularly meaningful for the RRMS vs SPMS classification task, where increased QuaSI-T_1 and QuaSI-T_2 were observed in the cortex.

The diffused involvement of WM alterations, particularly increased QuaSI-PD, indicate a strong component of axonal demyelination typical of MS, that becomes especially meaningful in classification tasks where differences in volumetric measurements are not as pronounced. MS subtypes appear to share the same degree of WM alterations, however prolonged relaxation times, particularly QuaSI-T_1, in cortical GM in SPMS compared to RRMS seem to indicate a progression in the tissue degeneration, and a correlation between cortical involvement and clinical disability. These results and relative interpretations are in line with published studies, reporting in MS patients:

- reduced myelin water fraction and MTR in normal appearing WM [53], indicative of demyelination, which in turn causes increased PD;
- diffusively prolonged T_2 not dependent on lesional tissue, also suggestive of demyelination and/or inflammation [141];
- global increase in T_1 with more marked effects observed in cortical GM associated with clinical disability, with worse results in SPMS than RRMS patients [142].

No meaningful contribution of relaxometry features was observed in the HC vs CIS classification task, which is in line with the lack of clinical symptoms in the CIS population. QuaSI-T_2 in cortical and deep GM emerged at- and below-threshold: whilst the distributions for the two populations are largely overlapping, CIS exhibited, on average, slightly *lower* QuaSI-T_2 values than HC, opposite to the T_2 increase observed in pathology, which can be explained by the age difference between the two groups ($\beta_1 = -0.07$, $p = 0.005$ for deep GM, $\beta_1 = -0.09$, $p = 0.0008$ for cortical GM). This result is therefore likely

due to age-related iron deposition, which shortens T_2 relaxation time and has been reported to accumulate in cortical and deep GM with age [143, 144] and not to MS.

Whilst these results are not new, it is worth recalling that these relaxometry maps were not the result of specialised MR-protocols, but were extracted *for free* from otherwise unused qualitative scans. The fact that QuaSI-PD, -T_2, and -T_1 maps on a relatively small dataset managed to replicate clinical outcomes of dedicated quantitative studies shows the utility of the MyRelax framework, and more importantly the often underestimated potential of qualitative data.

17.5 Sodium imaging

Sodium imaging appeared to be particularly meaningful when classifying CIS against HC, with high total sodium concentration observed for CIS subjects in WM and cortical GM, as well as deep GM, although below-threshold. Increased values were also observed, below-threshold, in MS when compared to HC, as well as SPMS to RRMS, although providing a much lower contribution to the classification than other features. No correlation with age was found in the HC population ($p > 0.55$), indicating this result is likely not due to the different average age between groups.

Increased total sodium concentration in MS has been reported in literature, with modest involvement at early stage and growing with disease progression [145]. It has been associated with the over-expression and redistribution of sodium-potassium channels from the Ranvier nodes to newly demyelinated membrane. This is an adaptive response to the disruption of saltuatory conduction caused by demyelination, apt to preserve action potential transmission, limit the onset of neurological deficits, and facilitate recovery. This however also increases the axonal metabolism, as the proliferation of the sodium-potassium *active* pumps comes with higher energy expenditure which, if not satisfied, causes the accumulation of intra-cellular sodium, resulting in increased values of total sodium concentration. In MS, the impaired trophic support from oligodendrocites and mitochondrial dysfunction contribute to energy under-production which, coupled with the increased metabolic need, can lead to axonal degeneration due to metabolic failure secondary to chronic energy deprivation [146].

The increased total sodium concentration observed in MS can thus be explained as the byproduct of sodium channels over-expression which is still present over the 15 years disease progression. This interpretation could also indirectly explain the increased

intra-neurite volume fraction observed in the CIS population, as failure of the sodium-potassium pumps, with consequent intra-cellular accumulation of sodium, might induce axonal swelling through osmosis [147]. Coupled with a reduction in the myelin compartment due to early demyelination, axonal swelling can result in the observed increase of intra-neurite volume fraction. Osmotic swelling might eventually lead to axonal loss, which results instead in reduced intra-neurite volume fraction, both of which have been observed in clinically defined MS [148].

CIS presenting similar alterations to MS, seconded by the absence of sodium features in the CIS vs MS classification, suggests the same neuroprotective mechanisms may be at play in the stable CIS population, but, unlike MS, they do not lead to measurable demyelination, axonal loss, other MS-related symptoms over the same time period. One could speculate that, despite the 15 years-long stable CIS clinical status and the absence of symptoms, these subjects did accrue *silent* long-lasting, if not irreversible, microstructural and functional alterations, which sodium imaging would greatly help further investigate. The ability to adapt to the increased metabolic demand without succumbing to energy failure, or avoiding axonal degeneration by excessive osmotic swelling, might be compensatory or even protective mechanisms, and as such key factors in what determines conversion, or lack thereof, to clinically defined MS.

Summary

- Brain volumetrics offer a reliable indicator of MS pathology that is relatively simple to compute and straightforward to interpret.
  - Better discrimination power could be achieved by separating atrophy in deep and cortical GM.
  - Volumetric measurements might not constitute the most meaningful feature when discriminating among MS subtypes, stable CIS against HC, or non age-matched groups.
- DWI offers a wide breadth of classification markers, as microstructural alterations are at the core of neurodegenerative diseases.
  - DWI also comes with several limitations which should be taken into account:
    * lengthy, specialised acquisition protocols;
    * low-resolution maps, prone to artifacts (Gibbs ringing, Nyquist ghost);

        * dependency on the chosen model-fitting method;
        * challenging interpretation of the results.
    - DWI metrics might not constitute the most meaningful features when discriminating between HC or CIS against MS, as atrophy could be equally or more representative, but with fewer limitations.

- Relaxometry offers a well known indicator for axonal demyelination and inflammation sensitive to MS severity.
    - Relaxometry is particularly meaningful in classification tasks where atrophy is not as important, either due to volume loss being equally present in both groups, or absent.
    - Relaxometry- and indirect myelin-features might not be as informative in terms of microstructural damage as diffusion, however with MyRelax, or equivalent *bottom-up* methods, they could be extracted from qualitative data, providing additional value at no additional cost.
- Sodium imaging offers a specialised indicator for functional and microstructural axonal integrity.
    - Due to the dedicated MR-protocols and low-resolution maps, even lower than DWI, sodium imaging might not represent the most fitting option for classification tasks better defined by volumetrics, relaxometry or DWI.
    - Sodium imaging might constitute a particularly meaningful feature when investigating *silent* (i.e. not associated to atrophy, lesions or other MS-related symptoms) physiological alterations.

Chapter 18

Conclusions and future works

In this study, two converging approaches have been followed to investigate how to best use the available MR-data for understanding the mechanisms of MS.

On one hand, qualitative data commonly used in clinical research as workhorse for lesion count and anatomical purposes have been shown to carry quantitative information that could be used to conduct myelin and relaxometry analyses on cohorts devoid of dedicated quantitative acquisitions. This study arm, named *bottom-up*, as qualitative information was *up-converted* to quantitative surrogate, was conducted on the basis of three objectives:

1. **MyRelax validation**: to assess the accuracy and reproducibility of QuaSI-PD, -T_2, -T_1 maps obtained from the qualitative scans using the MyRelax framework, by comparing them with the quantitative PD, T_2, T_1 maps obtained using gold standard quantitative MRI sequences.
2. **MyRelax MS application**: to evaluate the applicability of the MyRelax framework to MS, with QuaSI-MTV maps produced using MyRelax being compared to MTR, to test how much information attributable to myelin is shared by the two modalities. T_1-/T_2-weighted ratio images (T_1w/T_2w) were also compared to the MTR maps for the same reason.
3. **U-Net MS application**: to implement a deep learning network to extract MTR information directly from the qualitative scans — QuaSI-MTR — bypassing traditional model fitting.

On the other hand, when analysing multi-modal MR-data of healthy controls and subjects with a different 15 years-long stable clinical status, different MR-features appeared to be

meaningful with respect to specific classification tasks. The *top-down* study consisted in using machine learning to reduce the dimensionality of the multi-modal MRI dataset only to those feature that are more likely to be *biophysically meaningful* with respect to characterising MS progression, informing future acquisitions and investigation.

Following these objectives, we have investigated, developed and presented an array of options that, through advanced quantitative MRI and machine learning techniques, build towards a more in-depth characterisation of MS pathophysiology and a more efficient research environment. These findings can be summarised within three main contributions, as detailed below.

18.1 MyRelax: myelin and relaxation imaging

The MyRelax framework provides a way to extract indirect myelin and quantitative relaxometry indices from routinely acquired qualitative scans. The QuaSI-maps so produced have been shown to correlate well with ground truth in a prospective cohort of healthy controls, with QuaSI-MTV behaving similarly to MTR in a retrospective cohort of MS patients as well. Furthermore, by employing MyRelax to produce QuaSI-PD, -T_2 and -T_1 maps used as part of the **Biophysically meaningful features for classification of MS phenotypes** contribution, that would have otherwise lacked these modalities without additional MRI data acquisitions, we have also presented a practical example of how this method can supplement quantitative information when applied to a multi-modal MRI study, at no added cost.

Additional, multi-centric data would be required in future works to estimate the generalisability of this framework, which represents the main limitation of this method. The same results might in fact not be reproducible on data acquired with different MR-protocols and/or on different MR-scanners, in which case a revision to the underlying MyRelax model would be due. On this regard, a calibration step has been proposed as a possible solution to inter-study variability, providing a way to tailor MyRelax maps to the specific acquisition protocols for any given research centre. Future studies may explore the implementation and use of similar frameworks to include and/or produce different contrasts images, e.g. FLAIR, further optimising clinical acquisition protocols.

Overall, through the *MyRelax validation* and *MyRelax MS application* objectives, we have shown that it is indeed possible to exploit, by means of traditional model fitting, routine qualitative data for more than simple lesion delineation and brain tissue segmentation.

Whilst more data is needed for further validation of these results, MyRelax holds potential for providing new avenues for quantitative mapping when dedicated scans are not available, paving the way to quantitative analyses of large, historical MS datasets.

18.2 Deep learning MTR from qualitative images

As a direct follow-up of the **MyRelax: myelin and relaxation imaging** contribution, under the *bottom-up hypothesis* that qualitative scans contain quantitative information, we showed through the *U-Net MS application* objective that deep learning can be indeed used to extract indirect myelin content information directly from qualitative images, with QuaSI-MTR maps well correlating with ground truth MTR in an MS cohort.

Additional data may be required, as well as a multi-centric dataset, to assess the generalisability of the method to different routine scans. Given however that the relatively small dataset employed in this study was still enough to produce acceptable results, it would be possible, for any given research group, to train a specific network tailored to their own qualitative data, instead of relying on one generalised network. Future works would focus on testing the feasibility of *transfer learning* to adapt this, or similar, networks to qualitative data acquired with different MR-protocols and/or MR-scanners via fine-tuning, compared to training a new model using varied, multi-centric data from scratch. Future works may also include testing whether QuaSI-MTR can replicate MTR clinical outcomes, e.g. from concluded clinical trials, on the same MS dataset containing both qualitative and MT-data, as the clinical relevance of this method has yet to be investigated. Finally, whilst specific to MTR in this particular case, this contribution offers a framework for surrogate quantitative MRI mapping that could be applied, virtually, to a wide variety of MRI modalities: future studies may investigate what kind of quantitative information strictly requires dedicated scans to be accessed, which MRI modalities could be otherwise extracted from qualitative images through deep learning with comparable sensitivity to pathology, and everything in between.

Overall, in this study we have shown that deep learning can be employed to transfer information from routine scans to quantitative maps, especially when an explicit physical model mapping the former to the latter is not available, or traditional model fitting approaches are otherwise not feasible. As with MyRelax, being able to access indirect myelin information from qualitative data readily available in the form of QuaSI-MTR maps would be be of great use for sustainable MS research.

18.3 Biophysically meaningful features for classification of MS phenotypes

Machine learning has been shown to be an important tool in medical imaging for its ability to highlight patterns of alterations in the highly-dimensional landscape of MRI modalities. Through the *top-down* objective, we have investigated how different MRI modalities currently available mostly only in the research setting behave with respect to MS pathophysiology, highlighting which ones might be most meaningful in the characterisation of specific MS phenotypes. Different patterns of alterations were in fact observed for different classification tasks, with readily available features — such as volumetric indices — exhibiting strong sensitivity to late-stage pathology, whilst more specialised techniques — such as diffusion and sodium imaging — might be more useful in the study of early or pre-symptomatic stages of MS.

After having highlighted the *interaction* between the different MRI modalities by means of machine learning, it would be interesting to go back to the drawing board, and focus on the *integration* of the different features within hybrid, cross-modality physical models, able to delineate a more comprehensive and accurate picture of the MS-related physiological alterations compared to the sum of the single modalities alone. Future studies may, for example, incorporate MTV/MTR mapping — whether through prospective dedicated scans or via MyRelax/U-Net from retrospective data — within diffusion multi-compartment models to take the myelin/macromolecular compartment into account, in addition to the intra- and extra-cellular environments. Similarly, studies incorporating both sodium and diffusion MRI data within a single model, that is also clinically viable, might provide a much more powerful tool for the characterisation of neuronal physiology and microstructural integrity than diffusion or sodium imaging alone.

Due to MS multifaceted pathophysiology, larger datasets would be required to fully represent MS variability during training, and thus improve classification performances. Given the small cohort, age confounders, and the abundance of features, spurious findings may be present within this study, with more data being required, together with histopathological corroboration, for further confirmation. The dataset used for this study had a very homogeneous cohort in terms of disease duration — all subjects being scanned 15 years after the initial clinical episode — however, due to the lack of longitudinal data, these results are not intended to be used for prediction of MS progression. It can be speculated that the observed alterations are the end point of a 15 years long evolution

path that may still be traced back to the onset, and thus being informative to new patients presenting with CIS based on how similar their MRI status is to the 15 years stable subtypes: an investigation of the evolution over time of the MRI features emerged in this study should follow in future works. Future studies accessing multi-modal MRI datasets with significantly larger sample-size may also explore the emergence of data-driven MS phenotypes (similarly to what has been done by SuStaIn, see section 5.8.6), how they interact with the ones currently defined based on clinical criteria, and how they would affect classification. Overall, this contribution provides numerous avenues for investigation, both in terms of MRI research and computer science, and whilst we have attempted to provide an interpretation for the observed results based on biophysical correlates, these findings are not to be intended as proof of physiological alterations.

Similarly to the **Deep learning MTR from qualitative images** contribution, in this study we provided an example of how machine learning could be of great use in MS clinical research. Feature-ranking through artificial intelligence may in fact aid reducing the wide spectrum of quantitative MR-modalities only to those that are more likely to be meaningful to the task at hand, which might provide further insight for the understanding of MS-related neurodegeneration, as well as direct future studies to more targeted and efficient MR-acquisition protocols.

18.4 Closing statement

MRI is a fundamental tool in the diagnosis and prognosis of MS, with different MRI modalities offering specific insights to MS pathophysiology. In the vibrant landscape of MRI research, new and improved techniques are constantly proposed to progressively broaden the spectrum of specialised options available to clinicians, but only few actually reach fruition in the clinical environment. It is therefore important to investigate which specialised modalities could add biophysically meaningful information to patient characterisation that is worth the process of clinical optimisation, as well as the added scan time and image processing, whilst at the same time taking full advantage of the limited MRI modalities already established in the context of clinical trials. In this work, we have proposed two different approaches — through traditional model fitting and deep learning — to extract quantitative information from well established routine MR-scans, and a machine learning framework to rank advanced MRI features given their involvement in different MS phenotypes classification. Being able to exploit qualitative data for indirect myelin and relaxometry imaging could pave the way to a new wave of quantitative

studies on large historical datasets that might not have been used for much more than lesion counting and tissue segmentation. On the other hand, the MRI feature ranking provided could inform the scientific community with evidence on the use of dedicated MRI modalities for specific tasks that have not yet reached a consensus, e.g. the use of brain sub-cortical volumetry in clinically defined MS, or sodium concentration measurements in the early stages of the disease. Through these contributions, we hope to have made a small step forward in the direction of a more targeted, as well sustainable MRI research, able to provide direct answers to MS patients needs, in the most efficient way.

Bibliography

[1] Goldenberg, M. M. (2012). Multiple sclerosis review. *Pharmacy and Therapeutics*, 37(3), 175.

[2] Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., ... & Cohen, J. A. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology*, 17(2), 162-173.

[3] Zeis, T., Graumann, U., Reynolds, R., & Schaeren-Wiemers, N. (2008). Normal-appearing white matter in multiple sclerosis is in a subtle balance between inflammation and neuroprotection. *Brain*, 131(1), 288-303.

[4] Lui, Y. W., Xue, Y., Kenul, D., Ge, Y., Grossman, R. I., & Wang, Y. (2014). Classification algorithms using multiple MRI features in mild traumatic brain injury. *Neurology*, 83(14), 1235-1240.

[5] Klöppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., ... & Frackowiak, R. S. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain*, 131(3), 681-689.

[6] Lebedev, A. V., Westman, E., Van Westen, G. J. P., Kramberger, M. G., Lundervold, A., Aarsland, D., ... & AddNeuroMed consortium. (2014). Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical*, 6, 115-125.

[7] Ricciardi, A., Grussu, F., Battiston, M., Prados, F., Kanber, B., Samson, R. S., ... & Gandini Wheeler-Kingshott C. A. M., (2019). Myelin-sensitive indices in multiple sclerosis: the unseen qualities of qualitative clinical MRI. In *Scientific Meeting and Exhibition of the International Society for Magnetic Resonance in Medicine — ISMRM 2019*. https://archive.ismrm.org/2019/3312.html

[8] Ricciardi, A., Grussu, F., Prados, F., Kanber, B., Samson, R. S., Alexander, D. C., ... & Gandini Wheeler-Kingshott C. A. M., (2020). QuaSI-MTR (qualitative scans for imaging MTR): deep-learned MTR from routine scans using U-nets. In *Scientific Meeting and Exhibition of the International Society for Magnetic Resonance in Medicine — ISMRM 2019*. https://archive.ismrm.org/2020/1825.html

[9] Ricciardi, A., Grussu, F., Brownlee, W., Kanber, B., Prados, F., Collorone, S., ... & Gandini Wheeler-Kingshott C. A. M., (2018). Biophysically meaningful MRI features for accurate classification of multiple sclerosis phenotypes. In *Scientific Meeting and Exhibition of the International Society for Magnetic Resonance in Medicine — ISMRM 2018*. https://archive.ismrm.org/2018/1981.html

[10] Purves, D., Augustine, G., Fitzpatrick, D., Katz, L., LaMantia, A., McNamara, J., & Williams, S. (2004). Increased conduction velocity as a result of myelination. *Neuroscience*. Neuroscience, 3nd edn. Sinauer Associates, Sunderland, 63-65.

[11] Olsson, T., Barcellos, L. F., & Alfredsson, L. (2017). Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nature Reviews Neurology*, 13(1), 25-36.

[12] von Büdingen, H. C., Bar-Or, A., & Zamvil, S. S. (2011). B cells in multiple sclerosis: connecting the dots. *Current opinion in immunology*, 23(6), 713-720.

[13] Durães, F., Pinto, M., & Sousa, E. (2018). Old drugs as new treatments for neurodegenerative diseases. *Pharmaceuticals*, 11(2), 44.

[14] Dziedzic, T., Metz, I., Dallenga, T., König, F. B., Müller, S., Stadelmann, C., & Brück, W. (2010). Wallerian degeneration: a major component of early axonal pathology in multiple sclerosis. *Brain pathology*, 20(5), 976-985.

[15] National MS Society, *Who gets MS?*, https://www.nationalmssociety.org/What-is-MS/Who-Gets-MS

[16] Marcus, J. F., & Waubant, E. L. (2013). Updates on clinically isolated syndrome and diagnostic criteria for multiple sclerosis. *The Neurohospitalist*, 3(2), 65-80.

[17] National MS Society, *MS subtypes.*, http://www.nationalmssociety.org/What-is-MS/Types-of-MS

[18] University of California, San Francisco MS-EPIC Team:, Cree, B. A., Gourraud, P. A., Oksenberg, J. R., Bevan, C., Crabtree-Hartman, E., ... & Hauser, S. L.

(2016). Long-term evolution of multiple sclerosis disability in the treatment era. *Annals of neurology*, 80(4), 499-510.

[19] Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., ... & Wolinsky, J. S. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of neurology*, 69(2), 292-302.

[20] Brown, R. W., Cheng, Y. C. N., Haacke, E. M., Thompson, M. R., & Venkatesan, R. (2014). *Magnetic resonance imaging: physical principles and sequence design*. John Wiley & Sons.

[21] Christensen, J. D., Barrère, B. J., Boada, F. E., Vevea, J. M., & Thulborn, K. R. (1996). Quantitative tissue sodium concentration mapping of normal rat brain. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 36(1), 83-89.

[22] Rooney, W. D., Johnson, G., Li, X., Cohen, E. R., Kim, S. G., Ugurbil, K., & Springer Jr, C. S. (2007). Magnetic field & tissue dependencies of human brain longitudinal 1H2O relaxation in vivo. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 57(2), 308-318.

[23] Korb, J. P., & Bryant, R. G. (2002). Magnetic field dependence of proton spin-lattice relaxation times. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 48(1), 21-26.

[24] Bushberg, J. T., & Boone, J. M. (2011). *The essential physics of medical imaging*. Lippincott Williams & Wilkins.

[25] Caravan, P., Ellison, J. J., McMurry, T. J., & Lauffer, R. B. (1999). Gadolinium (III) chelates as MRI contrast agents: structure, dynamics, and applications. *Chemical reviews*, 99(9), 2293-2352.

[26] Suzuki, S., Sakai, O., & Jara, H. (2006). Combined volumetric T1, T2 & secular-T2 quantitative MRI of the brain: age-related global changes (preliminary results). *Magnetic resonance imaging*, 24(7), 877-887.

[27] Heath, F., Hurley, S. A., Johansen-Berg, H., & Sampaio-Baptista, C. (2018). Advances in noninvasive myelin imaging. *Developmental Neurobiology*, 78(2), 136-151.

[28] Ordidge, R. J., Gorell, J. M., Deniau, J. C., Knight, R. A., & Helpern, J. A. (1994). Assessment of relative brain iron concentrations using T2-weighted and T2*-weighted MRI at 3 Tesla. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 32(3), 335-341.

[29] Burtea, C., Laurent, S., Vander Elst, L., & Muller, R. N. (2008). Contrast agents: magnetic resonance. *Molecular imaging* I, 135-165.

[30] de Graaf, R. A., Brown, P. B., McIntyre, S., Nixon, T. W., Behar, K. L., & Rothman, D. L. (2006). High magnetic field water & metabolite proton T1 & T2 relaxation in rat brain in vivo. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 56(2), 386-394.

[31] Michaeli, S., Garwood, M., Zhu, X. H., DelaBarre, L., Andersen, P., Adriany, G., ... & Chen, W. (2002). Proton T2 relaxation study of water, N-acetylaspartate, & creatine in human brain using Hahn & Carr-Purcell spin echoes at 4T & 7T. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 47(4), 629-633.

[32] Juntu, J., Sijbers, J., Van Dyck, D., & Gielen, J. (2005). Bias field correction for MRI images. In *Computer Recognition Systems* (pp. 543-551). Springer, Berlin, Heidelberg.

[33] Zeng, H., & Constable, R. T. (2002). Image distortion correction in EPI: comparison of field mapping with point spread function mapping. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 48(1), 137-146.

[34] Gold, G. E., Han, E., Stainsby, J., Wright, G., Brittain, J., & Beaulieu, C. (2004). Musculoskeletal MRI at 3.0 T: relaxation times & image contrast. *American Journal of Roentgenology*, 183(2), 343-351.

[35] Wilhelm, M. J., Ong, H. H., Wehrli, S. L., Li, C., Tsai, P. H., Hackney, D. B., & Wehrli, F. W. (2012). Direct magnetic resonance detection of myelin & prospects for quantitative imaging of myelin density. *Proceedings of the National Academy of Sciences*, 109(24), 9605-9610.

[36] Mancini, M., Karakuzu, A., Cohen-Adad, J., Cercignani, M., Nichols, T. E., & Stikov, N. (2020). An interactive meta-analysis of MRI biomarkers of myelin. *Elife*, 9, e61523.

[37] Mezer, A., Rokem, A., Berman, S., Hastie, T., & Wandell, B. A. (2016). Evaluating quantitative proton-density-mapping methods. *Human brain mapping*, 37(10), 3623-3635.

[38] Le Bihan, D., & Breton, E. (1985). Imagerie de diffusion in vivo par résonance magnétique nucléaire. *Comptes rendus de l'Académie des sciences. Série 2, Mécanique, Physique, Chimie, Sciences de l'univers, Sciences de la Terre*, 301(15), 1109-1112.

[39] White, N. S., McDonald, C. R., Farid, N., Kuperman, J., Karow, D., Schenker-Ahmed, N. M., ... & Dale, A. M. (2014). Diffusion-weighted imaging in cancer: physical foundations and applications of restriction spectrum imaging. *Cancer research*, 74(17), 4638-4652.

[40] Basser, P. J., Mattiello, J., & LeBihan, D. (1994). Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance*, Series B, 103(3), 247-254.

[41] Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001). Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 13(4), 534-546.

[42] Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., & Wedeen, V. J. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 48(4), 577-582.

[43] Tournier, J. D., Calamante, F., & Connelly, A. (2013). Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR in Biomedicine*, 26(12), 1775-1786.

[44] Tournier, J. D., Calamante, F., Gadian, D. G., & Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3), 1176-1185.

[45] Tuch, D. S. (2004). Q-ball imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 52(6), 1358-1372.

[46] Behrens, T. E., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., ... & Smith, S. M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 50(5), 1077-1088.

[47] Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., & Alexander, D. C. (2012). NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61(4), 1000-1016.

[48] Kaden, E., Kruggel, F., & Alexander, D. C. (2016). Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magnetic resonance in medicine*, 75(4), 1752-1763.

[49] Kaden, E., Kelm, N. D., Carson, R. P., Does, M. D., & Alexander, D. C. (2016). Multi-compartment microscopic diffusion imaging. *NeuroImage*, 139, 346-359.

[50] Barkhof, F. (1999). MRI in multiple sclerosis: correlation with expanded disability status scale (EDSS). *Multiple Sclerosis Journal*, 5(4), 283-286.

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

[52] Ge, Y. (2006). Multiple sclerosis: the role of MR imaging. *American Journal of Neuroradiology*, 27(6), 1165-1176.

[53] Neema, M., Stankiewicz, J., Arora, A., Dandamudi, V. S., Batt, C. E., Guss, Z. D., ... & Bakshi, R. (2007). T1-and T2-based MRI measures of diffuse gray matter and white matter damage in patients with multiple sclerosis. *Journal of Neuroimaging*, 17, 16S-21S.

[54] Glasser, M. F., & Van Essen, D. C. (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI. *Journal of Neuroscience*, 31(32), 11597-11616.

[55] Ganzetti, M., Wenderoth, N., & Mantini, D. (2014). Whole brain myelin mapping using T1-and T2-weighted MR imaging data. *Frontiers in human neuroscience*, 8, 671.

[56] Filippi, M., & Agosta, F. (2007). Magnetization transfer MRI in multiple sclerosis. *Journal of Neuroimaging*, 17, 22S-26S.

[57] Schmierer, K., Scaravilli, F., Altmann, D. R., Barker, G. J., & Miller, D. H. (2004). Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain. *Annals of neurology*, 56(3), 407-415.

[58] Gass, A., Barker, G. J., Kidd, D., Thorpe, J. W., MacManus, D., Brennan, A., ... & Miller, D. H. (1994). Correlation of magnetization transfer ration with clinical disability in multiple sclerosis. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 36(1), 62-67.

[59] Vavasour, I. M., Laule, C., Li, D. K., Traboulsee, A. L., & MacKay, A. L. (2011). Is the magnetization transfer ratio a marker for myelin in multiple sclerosis?. *Journal of Magnetic Resonance Imaging*, 33(3), 710-718.

[60] Yarnykh, V. L. (2012). Fast macromolecular proton fraction mapping from a single off-resonance magnetization transfer measurement. *Magnetic resonance in medicine*, 68(1), 166-178.

[61] Mezer, A., Yeatman, J. D., Stikov, N., Kay, K. N., Cho, N. J., Dougherty, R. F., ... & Wandell, B. A. (2013). Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nature medicine*, 19(12), 1667-1672.

[62] Laule, C., Vavasour, I. M., Moore, G. R. W., Oger, J., Li, D. K., Paty, D. W., & MacKay, A. L. (2004). Water content and myelin water fraction in multiple sclerosis. *Journal of neurology*, 251(3), 284-293.

[63] Laule, C., Kozlowski, P., Leung, E., Li, D. K., MacKay, A. L., & Moore, G. W. (2008). Myelin water imaging of multiple sclerosis at 7 T: correlations with histopathology. *NeuroImage*, 40(4), 1575-1580.

[64] Kolind, S., Matthews, L., Johansen-Berg, H., Leite, M. I., Williams, S. C., Deoni, S., & Palace, J. (2012). Myelin water imaging reflects clinical variability in multiple sclerosis. *NeuroImage*, 60(1), 263-270.

[65] Liu, C., Li, W., Tong, K. A., Yeom, K. W., & Kuzminski, S. (2015). Susceptibility-weighted imaging and quantitative susceptibility mapping in the brain. *Journal of magnetic resonance imaging*, 42(1), 23-41.

[66] Lakhani, D. A., Schilling, K. G., Xu, J., & Bagnato, F. (2020). Advanced multicompartment diffusion MRI models and their application in multiple sclerosis. *American Journal of Neuroradiology*, 41(5), 751-757.

[67] Bagnato, F., Franco, G., Li, H., Kaden, E., Ye, F., Fan, R., ... & Xu, J. (2019). Probing axons using multi-compartmental diffusion in multiple sclerosis. *Annals of clinical and translational neurology*, 6(9), 1595-1605.

[68] Martínez-Heras, E., Solana, E., Prados, F., Andorrà, M., Solanes, A., López-Soley, E., ... & Llufriu, S. (2020). Characterization of multiple sclerosis lesions with distinct clinical correlates through quantitative diffusion MRI. *NeuroImage: Clinical*, 28, 102411.

[69] Filippi, M., & Rocca, M. A. (2013). Present and future of fMRI in multiple sclerosis. *Expert review of neurotherapeutics*, 13(sup2), 27-31.

[70] Bonnet, M. C., Allard, M., Dilharreguy, B., Deloire, M., Petry, K. G., & Brochet, B. (2010). Cognitive compensation failure in multiple sclerosis. *Neurology*, 75(14), 1241-1248.

[71] Inglese, M., Madelin, G., Oesingmann, N., Babb, J. S., Wu, W., Stoeckel, B., ... & Johnson, G. (2010). Brain tissue sodium concentration in multiple sclerosis: a sodium imaging study at 3 tesla. *Brain*, 133(3), 847-857.

[72] Paling, D., Solanky, B. S., Riemer, F., Tozer, D. J., Wheeler-Kingshott, C. A., Kapoor, R., ... & Miller, D. H. (2013). Sodium accumulation is associated with disability and a progressive course in multiple sclerosis. *Brain*, 136(7), 2305-2317.

[73] Gandini Wheeler-Kingshott, C. A., Riemer, F., Palesi, F., Ricciardi, A., Castellazzi, G., Golay, X., ... & D'Angelo, E. U. (2018). Challenges and perspectives of quantitative functional sodium imaging (fNaI). *Frontiers in neuroscience*, 12, 810.

[74] Bakshi, R., Thompson, A. J., Rocca, M. A., Pelletier, D., Dousset, V., Barkhof, F., ... & Filippi, M. (2008). MRI in multiple sclerosis: current status and future prospects. *The Lancet Neurology*, 7(7), 615-625.

[75] Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., ... & Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding & visualization. *Human brain mapping*, 38(11), 5391-5420.

[76] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Springer Series in Statistics The Elements of Statistical Learning The Elements of Statistical Learning*.

[77] Young, A. L., Marinescu, R. V., Oxtoby, N. P., Bocchetta, M., Yong, K., Firth, N. C., ... & Alexander, D. C. (2018). Uncovering the heterogeneity and temporal

complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nature communications*, 9(1), 1-16.

[78] Eshaghi, A., Young, A. L., Wijeratne, P. A., Prados, F., Arnold, D. L., Narayanan, S., ... & Ciccarelli, O. (2021). Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. *Nature communications*, 12(1), 1-12.

[79] Tan, P. N., Steinbach, M., & Kumar, V. (2016). *Introduction to data mining*. Pearson Education India.

[80] Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml* (Vol. 30, No. 1, p. 3).

[81] Phung, V. H., & Rhee, E. J. (2019). A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets. *Applied Sciences*, 9(21), 4500.

[82] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing & computer-assisted intervention* (pp. 234-241). Springer, Cham.

[83] Bendfeldt, K., Klöppel, S., Nichols, T. E., Smieskova, R., Kuster, P., Traud, S., ... & Borgwardt, S. J. (2012). Multivariate pattern classification of gray matter pathology in multiple sclerosis. *NeuroImage*, 60(1), 400-408.

[84] Wottschel, V., Alexander, D. C., Kwok, P. P., Chard, D. T., Stromillo, M. L., De Stefano, N., ... & Ciccarelli, O. (2015). Predicting outcome in clinically isolated syndrome using machine learning. *NeuroImage: Clinical*, 7, 281-287.

[85] Eshaghi, A., Wottschel, V., Cortese, R., Calabrese, M., Sahraian, M. A., Thompson, A. J., ... & Ciccarelli, O. (2016). Gray matter MRI differentiates neuromyelitis optica from multiple sclerosis using random forest. *Neurology*, 87(23), 2463-2470.

[86] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

[87] ImageNet, http://image-net.org/index

[88] Zhang, Y. D., Pan, C., Sun, J., & Tang, C. (2018). Multiple sclerosis identification by convolutional neural network with dropout & parametric ReLU. *Journal of computational science*, 28, 1-10.

[89] Brosch, T., Tang, L. Y., Yoo, Y., Li, D. K., Traboulsee, A., & Tam, R. (2016). Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. *IEEE transactions on medical imaging*, 35(5), 1229-1239.

[90] Lee, J., Lee, D., Choi, J. Y., Shin, D., Shin, H. G., & Lee, J. (2020). Artificial neural network for myelin water imaging. *Magnetic resonance in medicine*, 83(5), 1875-1883.

[91] Alexander, D. C., Zikic, D., Ghosh, A., Tanno, R., Wottschel, V., Zhang, J., ... & Criminisi, A. (2017). Image quality transfer and applications in diffusion MRI. *NeuroImage*, 152, 283-298.

[92] Fiorini, S., Verri, A., Tacchino, A., Ponzio, M., Brichetto, G., & Barla, A. (2015, August). A machine learning pipeline for multiple sclerosis course detection from clinical scales and patient reported outcomes. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4443-4446). IEEE.

[93] Pardini, M., Sudre, C. H., Prados, F., Yaldizli, Ö., Sethi, V., Muhlert, N., ... & Chard, D. T. (2016). Relationship of grey and white matter abnormalities with distance from the surface of the brain in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 87(11), 1212-1217.

[94] Clare, S., & Jezzard, P. (2001). Rapid T1 mapping using multislice echo planar imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 45(4), 630-634.

[95] Cardoso, M. J., Modat, M., Wolz, R., Melbourne, A., Cash, D., Rueckert, D., & Ourselin, S. (2015). Geodesic information flows: spatially-variant graphs and their application to segmentation and fusion. *IEEE transactions on medical imaging*, 34(9), 1976-1988.

[96] Modat, M., Ridgway, G. R., Taylor, Z. A., Lehmann, M., Barnes, J., Hawkes, D. J., ... & Ourselin, S. (2010). Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine*, 98(3), 278-284.

[97] Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *NeuroImage*, 62(2), 782-790.

[98] Andersson, J. L., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*, 20(2), 870-888.

[99] MATLAB, https://uk.mathworks.com/

[100] Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org

[101] Tofts, P. S. (2003). PD: proton density of tissue water. *Quantitative MRI of the brain: Measuring changes caused by disease*, 85-110.

[102] Rydberg, J. N., Riederer, S. J., Rydberg, C. H., & Jack, C. R. (1995). Contrast optimization of fluid-attenuated inversion recovery (FLAIR) imaging. *Magnetic resonance in medicine*, 34(6), 868-877.

[103] Volz, S., Nöth, U., Jurcoane, A., Ziemann, U., Hattingen, E., & Deichmann, R. (2012). Quantitative proton density mapping: correcting the receiver sensitivity bias via pseudo proton densities. *NeuroImage*, 63(1), 540-552.

[104] McPhee, K. C., & Wilman, A. H. (2018). Limitations of skipping echoes for exponential T2 fitting. *Journal of Magnetic Resonance Imaging*, 48(5), 1432-1440.

[105] JIM v6.0, Xinapse systems, Aldwincle, UK, http://www.xinapse.com

[106] Prados, F., Cardoso, M. J., Kanber, B., Ciccarelli, O., Kapoor, R., Wheeler-Kingshott, C. A. G., & Ourselin, S. (2016). A multi-time-point modality-agnostic patch-based method for lesion filling in multiple sclerosis. *NeuroImage*, 139, 376-384.

[107] Modat, M., Cash, D. M., Daga, P., Winston, G. P., Duncan, J. S., & Ourselin, S. (2014). Global image registration using a symmetric block-matching approach. *Journal of Medical Imaging*, 1(2), 024003.

[108] Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R., & Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, (47), S102.

[109] XNAT, https://www.xnat.org/about/xnat-implementations.php

[110] Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of statistics*, 1165-1188.

[111] McPhee, K. C., & Wilman, A. H. (2019). T1 and T2 quantification from standard turbo spin echo images. *Magnetic resonance in medicine*, 81(3), 2052-2063.

[112] Anderson, S. W., Sakai, O., Soto, J. A., & Jara, H. (2013). Improved T2 mapping accuracy with dual-echo turbo spin echo: Effect of phase encoding profile orders. *Magnetic resonance in medicine*, 69(1), 137-143.

[113] Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., ... & Lerer, A. (2017). Automatic differentiation in pytorch.

[114] Riba, E., Mishkin, D., Ponsa, D., Rublee, E., & Bradski, G. (2020). Kornia: an open source differentiable computer vision library for pytorch. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 3674-3683).

[115] Pytorch-msssim, https://github.com/VainF/pytorch-msssim

[116] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[117] Reddi, S. J., Kale, S., & Kumar, S. (2019). On the convergence of adam & beyond. *arXiv preprint arXiv:1904.09237*.

[118] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision & pattern recognition* (pp. 770-778).

[119] Zhu, H., Jones, C. K., Van Zijl, P. C., Barker, P. B., & Zhou, J. (2010). Fast 3D chemical exchange saturation transfer (CEST) imaging of the human brain. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 64(3), 638-644.

[120] SMT, https://github.com/ekaden/smt

[121] Prados, F., Solanky, B. S., Alves Da Mota, P., Cardoso, M. J., Brownlee, W. J., Riemer, F., ... & Ourselin, S. (2016). Automatic sodium maps reconstruction using PatchMatch algorithm for phantom detection. In *Scientific Meeting and Exhibition of the International Society for Magnetic Resonance in Medicine — ISMRM 2016*. https://archive.ismrm.org/2016/4330.html

[122] Pedregosa, F., Gramfort, A., Gael, V., Michel, V., Thirion, B., Grisel, O., Blondel, M., ... Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

[123] Lemaître, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *The Journal of Machine Learning Research*, 18(1), 559-563.

[124] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.

[125] Ramentol, E., Caballero, Y., Bello, R., & Herrera, F. (2012). SMOTE-RS B*: a hybrid preprocessing approach based on oversampling & undersampling for high imbalanced data-sets using SMOTE & rough sets theory. *Knowledge & information systems*, 33(2), 245-265.

[126] Eshaghi, A., Marinescu, R. V., Young, A. L., Firth, N. C., Prados, F., Jorge Cardoso, M., ... & Ciccarelli, O. (2018). Progression of regional grey matter atrophy in multiple sclerosis. *Brain*, 141(6), 1665-1677.

[127] Sastre-Garriga, J., Pareto, D., Battaglini, M., Rocca, M. A., Ciccarelli, O., Enzinger, C., ... & Rovira, À. (2020). MAGNIMS consensus recommendations on the use of brain & spinal cord atrophy measures in clinical practice. *Nature Reviews Neurology*, 16(3), 171-182.

[128] Ontaneda, D., Raza, P. C., Mahajan, K. R., Arnold, D. L., Dwyer, M. G., Gauthier, S. A., ... & Azevedo, C. J. (2021). Deep grey matter injury in multiple sclerosis: A NAIMS consensus statement. *Brain*.

[129] Carne, R. P., Vogrin, S., Litewka, L., & Cook, M. J. (2006). Cerebral cortex: an MRI-based study of volume & variance with age & sex. *Journal of Clinical Neuroscience*, 13(1), 60-72.

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

[131] Zivadinov, R., Havrdová, E., Bergsland, N., Tyblova, M., Hagemeier, J., Seidl, Z., ... & Horáková, D. (2013). Thalamic atrophy is associated with development of clinically definite multiple sclerosis. *Radiology*, 268(3), 831-841.

[132] Roosendaal, S. D., Geurts, J. J., Vrenken, H., Hulst, H. E., Cover, K. S., Castelijns, J. A., ... & Barkhof, F. (2009). Regional DTI differences in multiple sclerosis patients. *NeuroImage*, 44(4), 1397-1403.

[133] Peterson, J. W., Bö, L., Mörk, S., Chang, A., & Trapp, B. D. (2001). Transected neurites, apoptotic neurons, & reduced inflammation in cortical multiple sclerosis lesions. *Annals of Neurology: Official Journal of the American Neurological Association & the Child Neurology Society*, 50(3), 389-400.

[134] Johnson, D., Ricciardi, A., Brownlee, W., Kanber, B., Prados, F., Collorone, S., ... & Grussu, F. (2021). Comparison of Neurite Orientation Dispersion and Density Imaging and Two-Compartment Spherical Mean Technique Parameter Maps in Multiple Sclerosis. *Frontiers in Neurology*, 12, 944.

[135] Deppe, M., Krämer, J., Tenberge, J. G., Marinell, J., Schwindt, W., Deppe, K., ... & Meuth, S. G. (2016). Early silent microstructural degeneration & atrophy of the thalamocortical network in multiple sclerosis. *Human brain mapping*, 37(5), 1866-1879.

[136] Muthuraman, M., Fleischer, V., Kolber, P., Luessi, F., Zipp, F., & Groppa, S. (2016). Structural brain network characteristics can differentiate CIS from early RRMS. *Frontiers in neuroscience*, 10, 14.

[137] Calabrese, M., Rinaldi, F., Seppi, D., Favaretto, A., Squarcina, L., Mattisi, I., ... & Gallo, P. (2011). Cortical diffusion-tensor imaging abnormalities in multiple sclerosis: a 3-year longitudinal study. *Radiology*, 261(3), 891-898.

[138] Liu, W., Yang, J., Burgunder, J., Cheng, B., & Shang, H. (2016). Diffusion imaging studies of Huntington's disease: a meta-analysis. *Parkinsonism & related disorders*, 32, 94-101.

[139] Poonawalla, A. H., Hasan, K. M., Gupta, R. K., Ahn, C. W., Nelson, F., Wolinsky, J. S., & Narayana, P. A. (2008). Diffusion-tensor MR imaging of cortical lesions in multiple sclerosis: initial findings. *Radiology*, 246(3), 880-886.

[140] Rocca, M. A., Mezzapesa, D. M., Falini, A., Ghezzi, A., Martinelli, V., Scotti, G., ... & Filippi, M. (2003). Evidence for axonal pathology & adaptive cortical

reorganization in patients at presentation with clinically isolated syndromes suggestive of multiple sclerosis. *NeuroImage*, 18(4), 847-855.

[141] Whittall, K. P., MacKay, A. L., Li, D. K., Vavasour, I. M., Jones, C. K., & Paty, D. W. (2002). Normal-appearing white matter in multiple sclerosis has heterogeneous, diffusely prolonged T2. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 47(2), 403-408.

[142] Vrenken, H., Geurts, J. J., Knol, D. L., Van Dijk, L. N., Dattola, V., Jasperse, B., ... & Pouwels, P. J. (2006). Whole-brain T1 mapping in multiple sclerosis: global changes of normal-appearing gray and white matter. *Radiology*, 240(3), 811-820.

[143] Aquino, D., Bizzi, A., Grisoli, M., Garavaglia, B., Bruzzone, M. G., Nardocci, N., ... & Chiapparini, L. (2009). Age-related iron deposition in the basal ganglia: quantitative analysis in healthy subjects. *Radiology*, 252(1), 165-172.

[144] Acosta-Cabronero, J., Betts, M. J., Cardenas-Blanco, A., Yang, S., & Nestor, P. J. (2016). In vivo MRI mapping of brain iron deposition across the adult lifespan. *Journal of Neuroscience*, 36(2), 364-374.

[145] Maarouf, A., Audoin, B., Pariollaud, F., Gherib, S., Rico, A., Soulier, E., ... & Zaaraoui, W. (2017). Increased total sodium concentration in gray matter better explains cognition than atrophy in MS. *Neurology*, 88(3), 289-295.

[146] Petracca, M., Fleysher, L., Oesingmann, N., & Inglese, M. (2016). Sodium MRI of multiple sclerosis. *NMR in Biomedicine*, 29(2), 153-161.

[147] Armstrong, C. M. (2003). The Na/K pump, Cl ion, and osmotic stabilization of cells. *Proceedings of the National Academy of Sciences*, 100(10), 6257-6262.

[148] Dutta, R., & Trapp, B. D. (2011). Mechanisms of neuronal dysfunction and degeneration in multiple sclerosis. *Progress in neurobiology*, 93(1), 1-12.