ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

SCIENZE BIOMEDICHE E NEUROMOTORIE

Ciclo 36

Settore Concorsuale: 05/E3 - BIOCHIMICA CLINICA E BIOLOGIA MOLECOLARE CLINICA

Settore Scientifico Disciplinare: BIO/12 - BIOCHIMICA CLINICA E BIOLOGIA MOLECOLARE CLINICA

QUANTITATIVE SUSCEPTIBILITY MAPPING AS BIOMARKER OF NEURODEGENERATION: METHODOLOGICAL MODELS AND CLINICAL APPLICATIONS

Presentata da: Cristiana Fiscone

Coordinatore Dottorato

Prof.ssa Matilde Yung Follo

Supervisore

Prof.ssa Caterina Tonon

Co-supervisore

Prof. Daniel Remondini

Esame finale anno 2024

Abstract

Quantitative Susceptibility Mapping (QSM) is an advanced magnetic resonance technique that can identify and quantify *in vivo* biomarkers of pathology, such as alteration in iron and myelin concentration. Its quantitative nature allows for the comparison of magnetic susceptibility properties within and between different subject groups. Following the introduction of the acquisition and processing pipeline, I will explore potential clinical and methodological applications of QSM to neurodegeneration.

The first project focuses on the investigation of cortical regions in the development of amyotrophic lateral sclerosis. By examining various histogram susceptibility properties, a distinct pattern of increased iron content was revealed in patients with amyotrophic lateral sclerosis compared to controls and other neurodegenerative disorders. Moreover, there was a notable correlation between susceptibility and upper motor neuron impairment, particularly in patients experiencing rapid disease progression.
Similarly, in the second application, QSM was used to examine cortical and sub-cortical areas in individuals with myotonic dystrophy type 1. This analysis identified the thalamus and the brainstem as structures of interest from a susceptibility perspective, with relevant correlations with clinical and laboratory data such as neurological evaluation and sleep records.
In the third project, a robust pipeline for assessing radiomic susceptibility-based features reliability was implemented within a cohort of patients with multiple sclerosis and healthy controls. This study provided non-invasive and reliable measurements that hold promise for future clinical applications in disease characterization and phenotyping.
Lastly, a deep learning super-resolution model was applied to QSM images of healthy controls. The employed model demonstrated excellent generalization abilities and outperformed traditional up-sampling methods, without requiring a customized re-training.

In designing these studies, significant emphasis was placed on results reproducibility by implementing automated pipelines and by using, as much as possible, open-source software and toolboxes. Interpretability was also a key consideration, as it is crucial to maintain a strong connection between imaging data and pathophysiology when searching for biomarkers of pathology.
Across the three disorders investigated in this thesis, it was evident that QSM is capable of distinguishing between patient groups and healthy controls while establishing correlations between imaging measurements and clinical data. These studies lay the foundation for future research, with the ultimate goal of achieving earlier and less invasive diagnoses of neurodegenerative disorders within the context of personalized medicine.

Contents

Introduction

This thesis aims to explore quantitative susceptibility-based Magnetic Resonance (MR) images within the context of personalized medicine and neurodegeneration. Specifically, I am going to describe an advanced MR technique, Quantitative Susceptibility Mapping (QSM), from its sequence optimization to its clinical and methodological applications. I have carried out my research at the Functional and Molecular Neuroimaging Laboratory, IRCCS Istituto delle Scienze Neurologiche di Bologna (IT), with access to a clinical scanner for both clinical and research activity.
QSM is an MR method that produces an anatomical and voxel-wise map of magnetic susceptibility χ within the brain. This advanced technique enables the *in vivo* identification of various types of signal alterations, allowing for the quantification of calcification, blood-product-based deposits, iron and myelin concentration changes, serving as important biomarkers for pathologies like cerebrovascular conditions, neuroinflammatory processes, and neurodegenerative disorders. Its inherent quantitative nature offers the unique ability to compare magnetic properties within and between different groups of subjects. I will discuss and propose various applications, using both traditional methods and those that more effectively exploit the quantitative aspects of QSM, such as deep learning and radiomic techniques.

In Chapter 1, I am going to provide a theoretical overview of QSM, as a product of the study of susceptibility-based techniques. Its importance in the context of personalized medicine and in the study of neurodegenerative disorders, focus of this thesis, is going to be explained. I will delineate the mathematical process that allows us to measure magnetic susceptibility distribution from MR measurements, and describe the main sources of susceptibility in the human brain, with particular attention to relevant biomarkers within the context of neurodegeneration. Then, I will illustrate the acquisition setting implemented and optimized in our center, together with the stages to perform QSM reconstruction, providing examples in different conditions and pathologies. In three years, QSM has been incorporated into over 1700 MR scans, becoming a standard component of the imaging protocol for both neurodegenerative and non-degenerative conditions. I have been working on ensuring the quality of these images and on developing automated workflows, which I will describe in the following chapters about applications carried out during my PhD.

The first application, discussed in Chapter 2, pertains to the utilization of QSM in the study of Amyotrophic Lateral Sclerosis (ALS). ALS is a progressive idiopathic neurodegenerative disorder, typically characterized by initial muscle weakness symptoms that gradually intensify over time.

A well-established hallmark in the development of ALS is the accumulation of iron in the motor cortex, a phenomenon that can be readily detected and quantified through QSM. Previous research has identified elevated susceptibility values in the precentral gyrus, which directly correlate with iron concentration. Given that ALS patients experience a range of symptoms beyond motor impairment, the proposed application of QSM seeks to explore whether iron accumulation is observable in other cortical regions. This research aims to enhance our comprehension of the disease, especially in relation to other motor neuron disorders.
To achieve this, a cohort of ALS patients was compared with healthy controls and individuals with other neurodegenerative disorders. A comprehensive evaluation of the disease included the analysis of various susceptibility and structural properties within the frontal, parietal, temporal lobes, and cingulate cortex gyri. Neuroimaging data were combined with clinical assessments measuring the functional status and progression of ALS, as well as the impairment of the upper motor neuron. The diagnostic accuracy of quantitative findings was compared with qualitative neuroradiological evaluation normally used in clinical practice. Iron accumulation in the development of ALS occurs in motor and extra-motor cortical structures, with a direct link to the disease process and progression.

In the third chapter (Ch. 3), I will delve into the examination of iron accumulation in cortical and sub-cortical structures within the context of Myotonic Dystrophy type 1 (DM1), a dominantly inherited disorder that affects both the musculoskeletal and central nervous systems. While structural, diffusion, spectroscopy, and functional MR measurements have revealed alterations in the brains of individuals with DM1, there is a gap in our understanding of this disorder from a susceptibility perspective.
Similar to the previous application, χ measurements were derived from QSM images and integrated with structural data, along with neurological assessments, pneumological and cardiological evaluations, as well as sleep records from polysomnography. A group of DM1 patients was compared to healthy controls matched for demographic characteristics. Multiple automated segmentation tools were employed to pinpoint the regions of interest. QSM helps to identify the brainstem and thalamus as target structures for progressive impairment in DM1 patients, showing correlations with clinical and laboratory data. The use of QSM may contribute to shedding light on the ongoing debate regarding the origin of DM1 symptomatology, whether it is central or peripheral, as discussed in the existing literature.

In the last chapter (Ch. 4), I will provide explanations of two innovative methods that I employed on QSM images, to improve their utility and facilitate data analysis.
The first method involves the implementation of a robust and accurately designed pipeline for the extraction and evaluation of radiomic features from QSM images. This pipeline specifically focuses on the normal appearing white matter and its tracts (arcuate fasciculus, cortico-spinal tract, frontal aslant tract, inferior fronto-occipital fasciculus, optic radiation and uncinate fasciculus) and it is employed within a study conducted on a cohort of patients with multiple sclerosis and healthy controls. Multiple sclerosis is an autoimmune demyelinating disease characterized by

changes in iron and myelin content, detectable by QSM; when analyzed with radiomic techniques that exploit its intrinsic quantitative nature, QSM may facilitate early diagnosis and timely assessment of progression. The objective of this approach is to systematically extract and assess radiomic features, offering valuable insights into the structural and textural characteristics of the normal appearing white matter within the context of multiple sclerosis and beyond. This study validated the processing pipeline for robustness analysis and established the reliability of QSM-based radiomic features, providing a consistent number of descriptors that may be used for disease characterization and phenotyping.
The second methodology discussed in this chapter regards the application of a super-resolution model on susceptibility images. Super resolution includes a class of techniques aiming to enhance image spatial resolution. They have been applied to biomedical images and, more specifically, to MR brain ones; however, to date, QSM is quite unexplored in this kind of application, making this work an innovative advancement in the field. The approach used was peculiar itself, since the employed model (Enhanced Deep Super Resolution) was previously trained with natural images and then directly applied to brain MR images, without *re-training* or *fine-tuning*, to test its inherent generalization ability over new data, not obvious in deep learning applications. The pipeline I proposed was applied over multiple MR modalities, possibly offering new perspectives and enhanced capabilities for the study of neurological and medical conditions. To date, only data from healthy subjects were used, aiming to extend the validation over pathological tissues.

The final section at the end of the thesis will serve to summarize the results obtained in the different applications, highlighting the advantages of the use of QSM in the study of neurodegenerative pathologies.

Chapter 1

Quantitative Susceptibility Mapping

The purpose of this thesis is to analyze quantitative susceptibility-enhanced brain images from Magnetic Resonance (MR) measurements, namely Quantitative Susceptibility Mapping (QSM) sequence, describing its implementation at IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit (Bellaria Hospital), where I carried out my PhD research, and its clinical and methodological applications. In this first chapter, I am going to provide a theoretical overview of the technique, including important aspects from a mathematical and physiological profile, and explain the pipeline implementation in terms of processing and acquisition.

More specifically, in Section 1.1 QSM is introduced as an innovative quantitative technique[(1)] product of the study of susceptibility-based techniques, introduced in the 1980s to be investigated in human subjects. I am going to illustrate its importance in the context of personalized medicine and, specifically, in the study of neurodegenerative disorders[(2)] (Sec. 1.1.1), which I will focus on in this work. The mathematical process[(1)(3)] which allows us to measure magnetic susceptibility distribution from MR measurements is going to be explained (Sec. 1.1.2), followed by the description of the main sources of susceptibility in the human brain, with particular attention to relevant biomarkers within the context of neurodegeneration, such as iron and myelin (Sec. 1.1.3).
In the second section, details of the sequence acquisition setting are given[(4)] (Sec. 1.2.1), together with the stages to perform QSM reconstruction (Sec. 1.2.2,1.2.3); QSM images in different conditions and pathologies are provided as examples, comparing the technique with other susceptibility-based sequences (Sec. 1.2.4).
Issues and possible new developments are going to be highlighted, including technical limitations and issues related to data availability (Sec. 1.3).

The first use of abbreviations includes their full expansion; for ease of reference, a glossary is provided at the end of the chapter.

1.1 Background

Quantitative Susceptibility Mapping (QSM)[1][3] is an advanced Magnetic Resonance (MR) technique measuring and mapping the magnetic susceptibility χ of structures within the body, particularly in the brain; it is often used to study and visualize conditions linked with the magnetic properties of tissues, such as iron deposition, hemorrhages or calcifications. Exploiting advancements in MR Imaging (MRI) technology and image processing techniques, it gained prominence in the early 2010s, driven by the need for more accurate and quantitative measurements of magnetic properties in the context of neuroimaging.
The development of QSM builds upon earlier work in MRI and χ imaging techniques; multiple adult and pediatric neuropathologies demonstrate abnormal accumulation of blood products or mineral deposition, which led scientists to study χ and its effects on MRI signals. Susceptibility-Based Imaging (SBI) has been introduced in the clinical practice[5], usually using magnitude and phase data from GRadient Echo (GRE) T_2^*-weighted (T_2^*w) sequences. Already in 1984, the effect of tissue susceptibility had been investigated in human subjects[6]. Nowadays, Susceptibility Weighted Imaging (SWI)[7][8], whose processing is implemented on scanners of the main vendors, is the most commonly used among those techniques: it has been recognize to be useful in the diagnosis of a wide range of Central Nervous System (CNS) disorders of various etiology and it is acquired together with conventional morphological MR T_1- and T_2- weighted (T_1w and T_2w) images.

Although conventional SBI techniques (e.g. T_2^*w, R_2^*, SWI) provide valuable information, they are less capable of quantifying the magnetic properties of tissues. In a recent work from Harada and colleagues, 2022[3], they summarize the advantages and drawbacks of imaging sequences in the evaluation of χ; QSM offers several benefits. First, it provides quantitative measurements of χ, yielding signal intensities directly related to tissue magnetic properties; this enables the comparison of χ distributions intra- and inter- groups, performing longitudinal evaluations as well. Additionally, QSM provides anatomical information: it improves tissue differentiation - especially in some structures such as the basal ganglia, not clearly identifiable using T_1w and T_2w images - and helps to distinguish between different types of lesions, hemorrhages, and tissue structures with high sensitivity. An example of QSM reconstruction, together with T_1w, T_2w and SWI images, is shown in Fig. 1.1.
Furthermore, QSM allows discrimination between dia- and para- magnetic substances, appearing respectively hypo- and hyper- intense. The accumulation of both types of biomarkers leads to a phase shift of the spins and in T_2^*w sequences this phenomenon leads to a signal drop; considering that SWI does not actually measure susceptibility distribution $\chi(r)$ but returns a magnitude image weighted on the phase signal[7] - which is mathematically linked to $\chi(r)$ (Sec. 1.1.2) - results of both dia- and para- magnetic accumulation is hypo-intense in SWI maps. In Sec. 1.2.4, I am going to continue the discussion by providing some examples.
Another advantage of using QSM with respect to other SBI sequences is the reduction of blooming artifacts, commonly occurring in GRE images near the surface perpendicular to magnetic field[9].

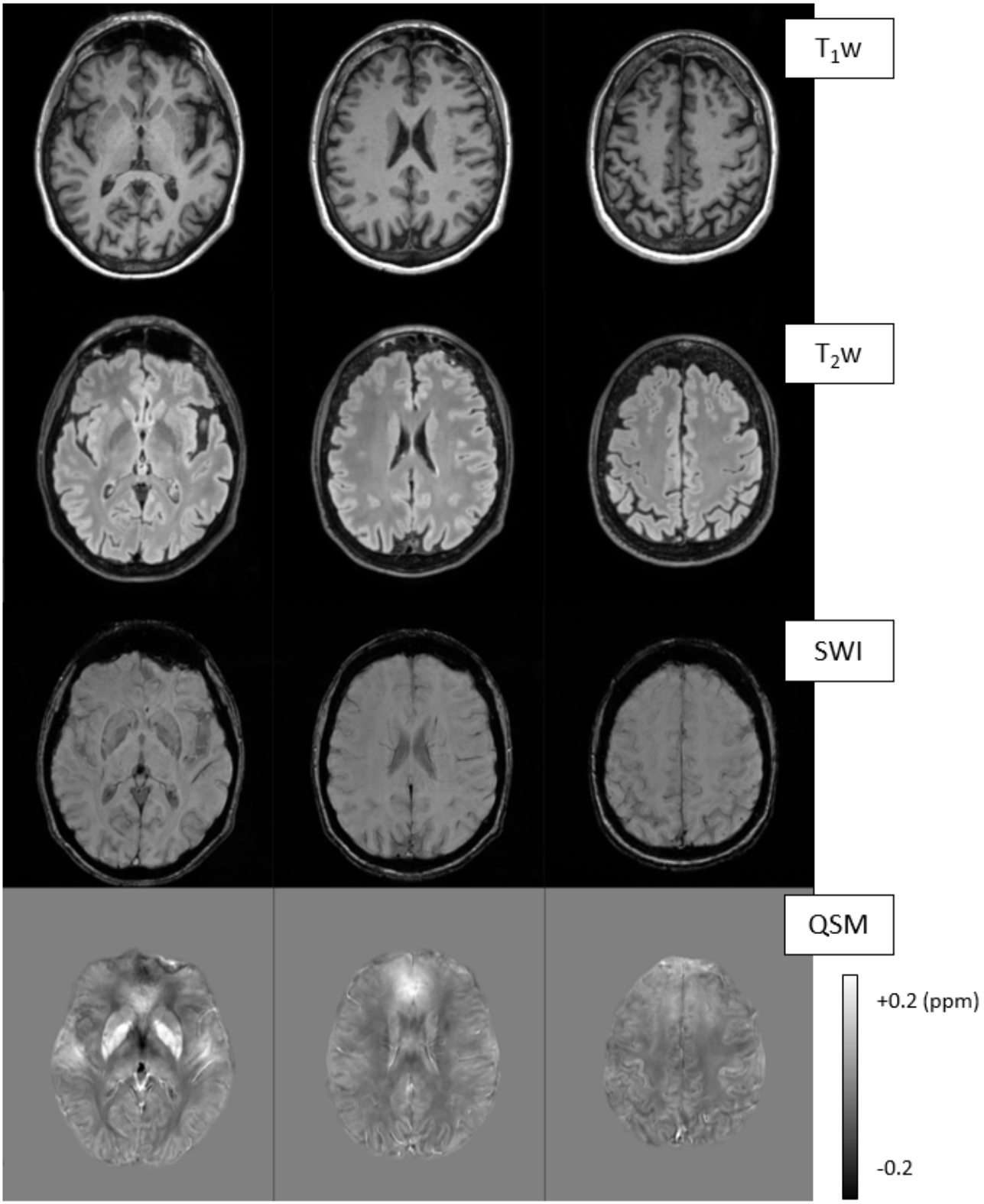

Figure 1.1: Example of conventional morphological (T_1w and T_2w) and susceptibility-based (SWI and QSM) MR sequences (healthy control, F/58 years old). Note that, since QSM is a quantitative technique, the color bar is given, while the other sequences return signal intensity in arbitrary units. The exam was performed on a 3-T clinical scanner at Bellaria Hospital, Bologna, IT (Neuroimaging Laboratory, IRCCS Istituto Neuroscienze Neurologiche di Bologna, IT). Acquisition parameters can be found in the Section 1.2.1

1.1.1 Precision medicine and neurodegeneration

In the early 1990s, the beginning of the human genome mapping project gave a fundamental boost to what is called Personalized Medicine (PM), used for the first time in 1999 in a short article of *The Wall Street Journal*[10]. Also referred to as *precision medicine*, PM is a health-care approach that tailors medical decisions, interventions and therapies to the individual characteristics of each patient[11]. PM strategies have been developed recently, aiming to consider human variability and to identify the disease-associated variants through the study of the genotype and phenotype of patients and their diseases. The use of genetic information is central in PM approach, to understand the individual genetic make-up and reveal mutations, variations and predispositions to diseases. Although, the subject's profile is evaluated adding other variables, such as environment and lifestyle, to carry out a more comprehensive evaluation. By understanding an individual's risk factors, PM may help in disease prevention and early detection. For example, individuals with a higher risk of certain cancers may undergo more frequent screening to detect the disease at an early, more treatable stage. PM often involves the development and use of targeted therapies and diagnostics, as treatments designed to specifically target the underlying causes of a disease, often based on the patient's genetic profile. They can be more effective than *one-size-fits-all* treatments. While PM holds great promise for improving healthcare outcomes and minimizing adverse effects, it also presents challenges related to privacy, data security, and ethical considerations. Additionally, the availability and affordability of genetic testing and targeted therapies can vary widely, making widespread implementation of personalized medicine a complex endeavor. Nonetheless, it continues to be an area of active research and development in modern health services.

Biomedical imaging is a key component of PM[12]; exploiting imaging information, coming from *in vivo* and non-invasive examinations, can lead to more effective diagnosis, treatment, and management of diseases. As a quantitative technique, QSM has the potential to play a significant role in PM, especially in the fields of neurology and oncology. Thanks to its sensitivity to the effects of magnetization, some areas of application may be[3]: a) brain tumor characterization, distinguishing between types, providing information about their grade, and assessing treatment response; b) evaluation of the extent and location of hemorrhages and micro-bleeds in patients with traumatic brain injuries, crucial for tailoring treatment and monitoring recovery in a personalized manner; c) identifying and characterizing hemorrhagic and ischemic strokes, providing information about the extent of brain damage. In all of these applications, QSM provides quantitative data that can help clinicians and researchers to tailor treatments to individual patients' needs. It can guide treatment decisions and monitor treatment responses, ultimately improving patient outcomes. However, it's important to note that QSM is still a relatively specialized technique and may not be available at all healthcare institutions. Its adoption in personalized medicine may vary depending on the specific clinical context and available resources.

Personalized treatment programs have been applied to neurodegeneration[2] [13], which is one of the most explored QSM applications[14], that I will focus on for this thesis. Neurodegenerative

disorders affect millions of people worldwide and they are one of the most recurrent contributors to morbidity and mortality in the elderly; it has been established that the combination of genes and environment contributes to the risk of developing a neurodegenerative disease.
Dysfunctions of cerebral iron regulation contribute to the pathogenesis of various neurodegenerative disorders[15]; QSM can help to quantify the amount of iron deposition in specific brain structures and, considering that, it is a promising tool to characterize neurodegeneration, where abnormal iron accumulation, associated with oxidative stress, is a hallmark. QSM can provide insights into the basal ganglia's iron content, which is implicated in Parkinson's Disease (PD) pathophysiology, particularly in the substantia nigra[16], monitoring iron levels in these structures and helping to track disease progression and treatment response. Iron concentration in the basal ganglia also occurs in the development of Huntington's Disease (HD)[17], which can consequently be studied using χ-based techniques such as QSM. Another example is Alzheimer's Disease (AD): while QSM is not as commonly used for AD - since other imaging techniques like positron emission tomography scans and cerebrospinal fluid analysis are often more informative in AD research - it can potentially provide information about iron accumulation in certain brain regions[18].
There are other biomarkers influencing χ distributions: paramagnetic transition metal ions, such as copper, manganese and cobalt[1], or calcium, responding diamagnetically[19]. Thus, over iron-based disorders, QSM may help in studying other biomarkers, for example, calcium in Fahr's syndrome[3] or copper in Wilson's Disease[20]. QSM is particularly useful in studying demyelinating diseases, such as Multiple Sclerosis (MS)[21][22][23], because myelin is diamagnetic and reduction in myelin concentration directly impact measure of χ[24].
In the past three years, I have been focused on the applications of QSM to Amyotrophic Lateral Sclerosis (ALS), Myotonic Dystrophy type 1 (DM1) and MS, which I am going to describe and explain in Chapters 2, 3 and in Section 4.1.

1.1.2 From χ to MR measurements

The process of generating χ maps involves post-processing of MRI data, typically acquired with 3D GRE T_2^*w sequences[4]. By measuring the phase information of the MRI signal and performing mathematical reconstruction, QSM pipelines generate maps that represent the spatial distribution of χ values in the imaged region. These maps can then be used by radiologists and researchers to gain insights into the underlying tissue properties.

As I am going to explain shortly in this section, tissue susceptibility distribution $\chi(r)$ is related to magnetic field inhomogeneities $\Delta B(r)$ generated by the induced matter magnetization when an external magnetic field is applied. $\Delta B(r)$ is proportional to the phase data ϕ measured with MR measurements. Signal $S(TE)$ at echo time TE is equal to:

$$S(TE) \propto e^{-\frac{TE}{T_2^*}} \cdot e^{i\phi} \quad \left(\phi = \phi_0 - \frac{\gamma}{2\pi}\Delta B \cdot TE\right) \tag{1.1}$$

with ϕ_0 offset related to the experimental setup and γ gyromagnetic ratio (for hydrogen, equal to 42.58 $\frac{MHz}{T}$. $\Delta B(r)$ and $\phi(r)$ are the same up to a multiplicative factor.

Magnetic susceptibility χ refers to a material's ability to become magnetized when placed in an external magnetic field[25] and it depends on the atomic and molecular structure: availability of unpaired electrons, distribution of electron cloud, competition between spin and induction, micro-structure and so on. In MRI, an external and static magnetic field is applied, and, consequently, χ differences in tissues create contrast in MRI images. In this context, I refer to static magnetic susceptibility.

For isotropic, non-ferromagnetic materials, matter magnetization is described by the magnetization vector or magnetic dipole moment per unit volume $\vec{M}$, defined as[1]:

$$\vec{M} = \frac{d\vec{m}}{dV} \tag{1.2}$$

where $d\vec{m}$ is the elementary magnetic moment. The macroscopic magnetization $\vec{M}$, the induction vector $\vec{B}$ and the magnetic field $\vec{H}$ are related with each other:

$$\vec{M} = \chi\vec{H} \tag{1.3}$$

$$\vec{B} = \mu_0(\vec{H} + \vec{M}) \tag{1.4}$$

$$\vec{B} = \mu_0(\vec{H} + \chi\vec{H}) = \mu_0(1+\chi)\vec{H} = \mu_0\mu_r\vec{H} \tag{1.5}$$

where μ_0 is the vacuum magnetic permeability ($4\pi \cdot 10^{-7}\frac{H}{m}$) and μ_r is the relative permeability, representing the degree of the magnetization of the specific material. $\vec{M}$ and $\vec{H}$ have the same unit of measure (A/m), which makes χ dimensionless. In isotropic and homogeneous materials, the magnetic response is uniform, χ has an isotropic behavior and it can be modeled as a scalar quantity; in the brain, that happens in the gray matter structures, such as the basal ganglia. In other materials, such as the cerebral white matter, the magnetic response is not uniform, because of the asymmetry of the molecules' structures and it depends on the orientation of the field, leading to an anisotropic behavior of χ which should be modeled as a tensor quantity ($\vec{\chi} = \vec{\chi}(\vec{H})$). Susceptibility Tensor Imaging (STI) technique characterizes the anisotropic tissue magnetic susceptibility with a second-order tensor model[26]; detailing the math and application of this technique is beyond the purpose of this thesis.

In medical images, each voxel represents a particle, and each particle magnetically reacts with the external magnetic field and with the surrounding particles. Particles are modeled as dipoles; assuming that $\vec{B_0}$ is aligned along the $\vec{z}$ axis ($\vec{B_0} = [0, 0, B_0]$) and consequently $\vec{M}$ is aligned along the $\vec{z}$ axis ($\vec{M} = [0, 0, M_z]$), fair to assume in MR measurements:

$$M_z(r) = \chi(r)H_0 = \chi(r)\frac{B_0}{\mu_0\mu_r} = \chi(r)\frac{B_0}{\mu_0(1+\chi(r))} \sim \chi(r)\frac{B_0}{\mu_0} \tag{1.6}$$

with $\chi << 1$, reasonable considering susceptibility values in the human body - tipically $\pm$ 20% of $\chi_{water} = -9.05 \cdot 10^{-6}$ (SI units)[1]. The magnetic field generated by this distribution $M_z(r)$ is:

$$\Delta B_z(r) = \frac{\mu_0}{4\pi}\int \frac{1}{|r-r'|^3} \times \left(3\frac{M_z(r')(r-r')}{|r-r'|^2}(r-r') - M_z(r')\right) d^3r' = \mu_0 \cdot M_z(r) \bigotimes D(r) \tag{1.7}$$

$$\frac{\Delta B_z(r)}{B_0} = \chi(r) \bigotimes D(r) \tag{1.8}$$

resulting in a convolution between the susceptibility distribution and the expression of the unit dipole field $D(r)$, describing the sum of the dipolar fields generated by each element of the distribution $M_z(r)$. In k-space, using the convolution theorem, the expression can be simplified as a point-wise multiplication:

$$\frac{\Delta B_z(k)}{B_0} = \chi(k) \cdot D(k) \tag{1.9}$$

The unit dipole field $D(k)$ has singularities near the cone surface of the magic angle 54.7° relative to the direction of the main magnetic field, which makes the susceptibility reconstruction an ill-posed problem. Different techniques have been implemented to fix this issue, regularization strategies helping the suppression of streaking and shadowing artifacts in χ reconstruction and optimization approaches incorporating additional information. Some of them are: MEDI (Morphology Enabled Dipole Inversion)[27], HEIDI (Homogeneity Enabled Incremental Dipole Inversion)[28], iterative Least SQuaRe (iLSQR)[29], TGV (Total Generalized Variation)[30]. An exhaustive overview of those methods may be found in the work of Bilgic and colleagues[4]. Additionally, multiple-orientation methods (e.g. Calculation Of susceptibility through the Multiple Orientation Sampling [COSMOS][31]) have been introduced, very accurate but impracticable in clinical practice because of the long scan time and patient discomfort[32]. Deep learning approaches have been proposed as a solution as well[33][34][35][36].

1.1.3 χ sources: dia-, para- and ferro- magnetic materials

In almost all the materials, in the absence of an externally applied magnetic field $\vec{H} = \vec{B} = 0$, and consequently there is no magnetization ($\vec{M} = 0$). When an external magnetic field is applied ($\vec{H} \neq 0$ and $\vec{B} \neq 0$), a macroscopic magnetization vector with non-zero intensity ($\vec{M} \neq 0$) occurs, whose direction, sense and intensity are related to χ. In Fig. 1.2, the scale of magnetic susceptibility values of some elements and compounds in the human body, ranging from -7.0 to 11.0 ppm (SI units)[37], is shown. According to χ, materials can be broadly categorized into diamagnetic, paramagnetic and ferromagnetic; below, is a brief description of those phenomena.

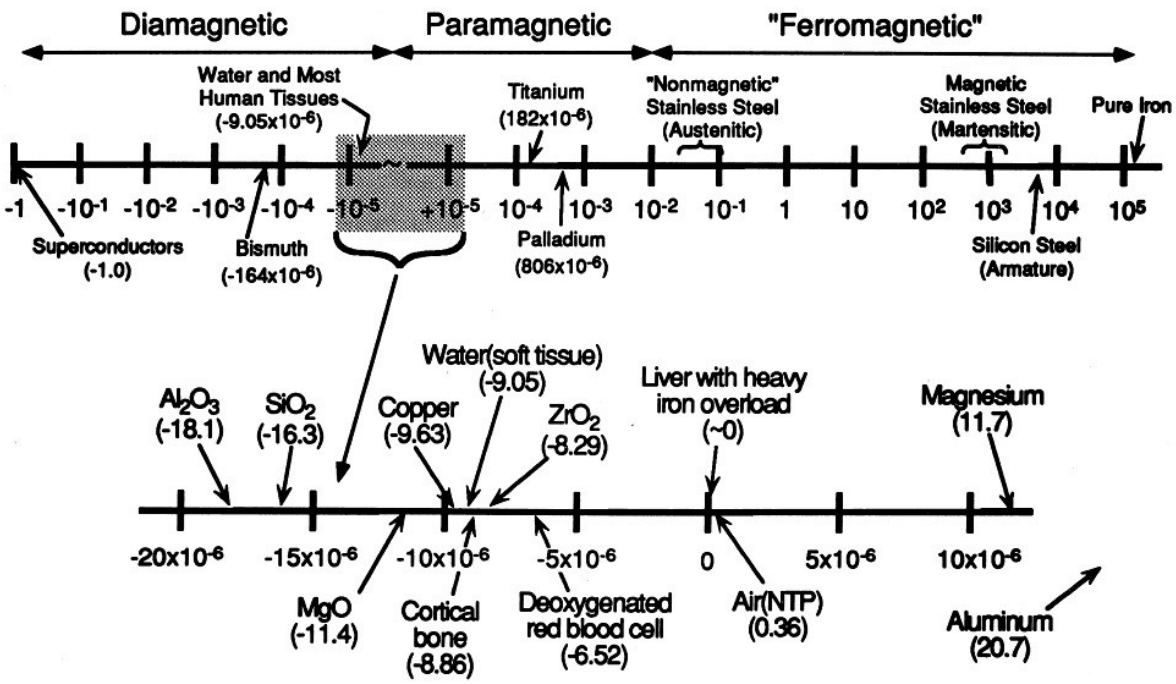

Figure 1.2: Spectrum of magnetic susceptibility values of some elements and compounds in the human body[37] (SI units). In the upper diagram, the full range of observed magnetic susceptibility values is shown in logarithmic scale, extending from -1.0 ppm for superconductors to values higher than 10^6 for ferromagnetic materials. In the bottom diagram, the susceptibility of elements and compounds in the human body, raging from -7.0 to 11.0 ppm

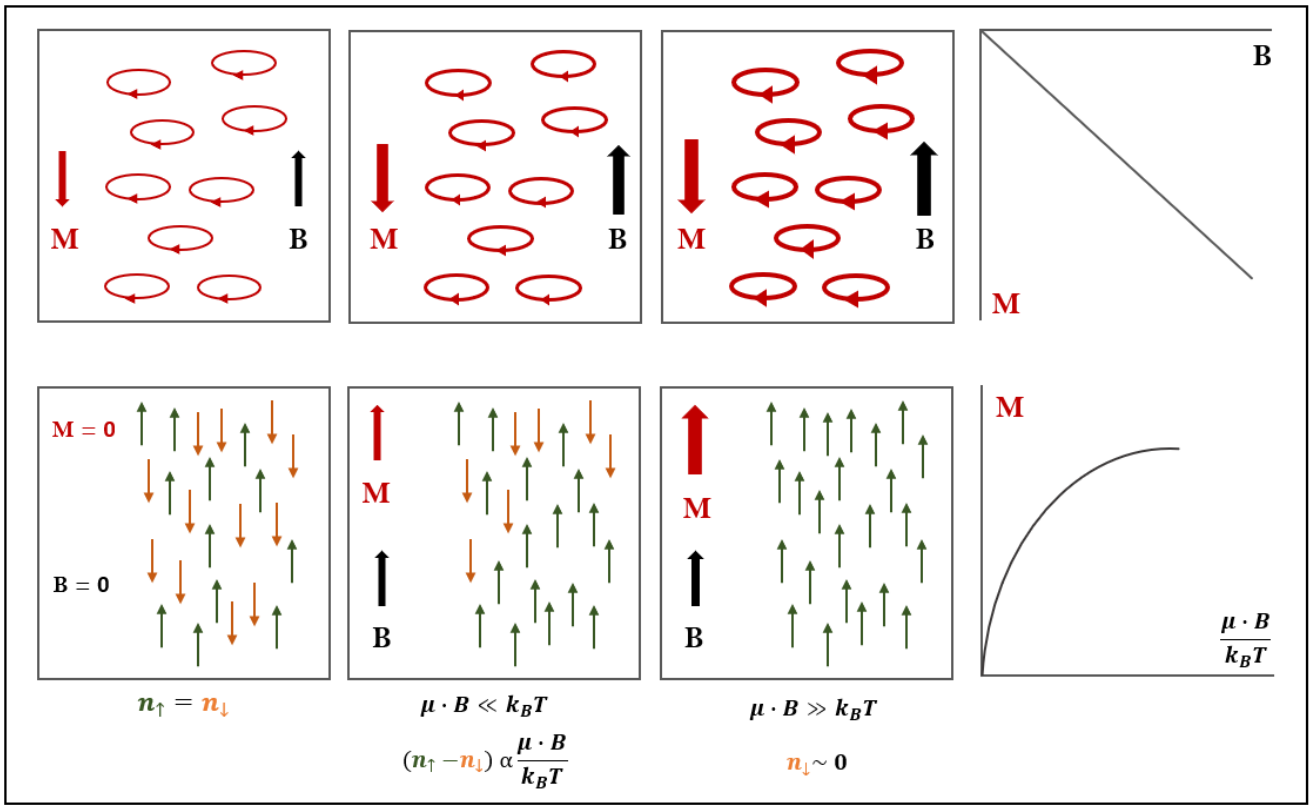

Figure 1.3: Scheme of magnetization vector $\vec{M}$ and external magnetic field $\vec{B}$ in dia- (upper diagram) and para- (bottom diagram) magnetic materials. Diamagnetic materials: in balance condition without an external magnetic field, the net magnetic moment is equal to zero; when a field $\vec{B}$ is applied, a weak magnetization vector $\vec{M}$ is formed along the same direction of the field but in the opposite sense. Paramagnetic materials: in the absence of an external magnetic field, the number of magnetic moments pointing up n_+ and the ones pointing down n_- are approximately the same; when an external field is applied, the majority of the magnetic moments start to align as the field. When the intensity of the field is strong, the number of the magnetic moments direct in the opposite sense of the field tends to zero

Diamagnetic materials

In diamagnetic materials (Fig. 1.3, upper diagram), a repulsive force is induced when a magnetic field is externally applied; magnetization $\vec{M}$ induced by the external field has a small intensity and is oppositely directed with respect to the field, because of the electron induction currents which generate a magnetic field in the opposite sense with respect of the external one (Lenz's Law). Water is slightly diamagnetic (χ_{water} = -9.035 ppm [SI units], average brain tissue susceptibility) and, because of the great abundance of its molecules ($\sim$ 70-80% of the total tissues[19]), body tissues are mainly diamagnetic and have a weak response to the magnetic field. Examples of diamagnetic biomarkers are oxy-hemoglobin, calcium and myelin.

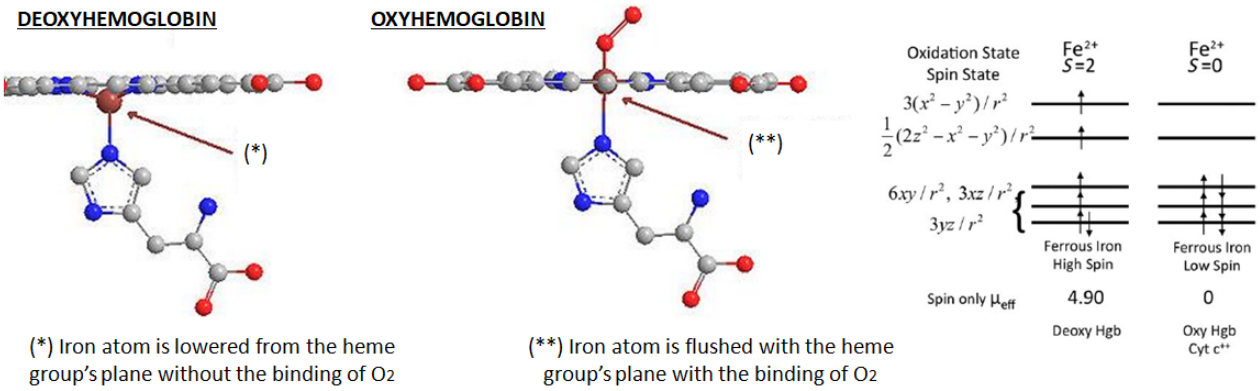

Figure 1.4: Deoxygenated and oxygenated hemoglobin structure

- **Oxy-hemoglobin**
  Oxyhemoglobin (HbO_2,Fe^{3+})[4] (Fig. 1.4) is the form of hemoglobin combined with oxygen. Hemoglobin (Hb) is the predominant protein in the erythrocytes, containing four atoms of iron; it is responsible for transporting oxygen from the lungs to the body's tissues and organs, as well as carrying carbon dioxide back to the lungs for exhalation. Hb binds and releases oxygen depending on the oxygen concentration in its environment, and it is referred to as deoxy-hemoglobin or oxy-hemoglobin respectively if it is detached or not from molecular oxygen O_2. The balance between oxyhemoglobin and deoxyhemoglobin is crucial for efficient oxygen delivery and release, and it is influenced by various factors, including pH and carbon dioxide levels. In its oxygenated form, Hb exhibits diamagnetic behavior, with iron temporarily oxidized from Fe^{2+} to Fe^{3+}.

- **Calcium**
  Ca^{2+} ions are implied into various neuronal functions, such as neuronal signaling, participating in neurotransmitter release, synaptic transmission, and the activation of various enzymes and pathways involved in learning, memory, and other cognitive functions. When calcium regulation is disrupted, it can contribute to neurodegenerative processes[38]: changes in calcium levels at synapses can modulate the strength of synaptic connections, affecting information processing in the brain. Excessive calcium influx into neurons can lead to

excitotoxicity, a process where over-stimulation of neurons by neurotransmitters like glutamate results in cell damage or death; this process is implicated in several neurodegenerative conditions, including AD and HD; disruption of calcium homeostasis can lead to neuronal dysfunction, and degeneration and dysregulation of mitochondrial calcium levels can contribute to oxidative stress and cell death. In many neurodegenerative disorders, including AD, PD, HD and ALS, abnormalities in calcium handling have been observed, and QSM may be used to detect these abnormalities since it is calcium-sensitive.

- **Myelin**
  Myelin is a spiraling sheath wrapping around nerve fibers (axons) in the CNS (Fig. 1.5), composed of 10% proteins, 40% water and 50% lipids; lipids give myelin its characteristic white appearance, which is the reason why regions of the nervous system rich in myelinated fibers are often referred to as white matter. The myelin sheath acts as an insulating layer and plays a critical role in the efficient transmission of electrical impulses along nerve cells, enhancing the speed and efficiency of electrical signal propagation (action potentials) along the axons. Myelin sheaths are not continuous along the entire length of the axon. Instead, they are periodically interrupted by small gaps called nodes of Ranvier. These nodes contain a higher concentration of ion channels and allow for the rapid propagation of the action potential.
  Myelin is involved in neurological functions, and disruptions in myelin can lead to neurological disorders and diseases, including: MS, an autoimmune disease in which the immune system mistakenly targets and damages the myelin in the CNS (Sec. 4.1); NeuroMyelitis Optica Spectrum Disorder (NMOSD), inflammatory disorder of the central nervous system characterized by severe, immune-mediated demyelination and axonal damage predominantly targeting optic nerves and the spinal cord; Guillain-Barré Syndrome (GBS), a condition where the immune system attacks the peripheral nerves and their myelin sheaths, causing muscle weakness and sometimes paralysis; Charcot-Marie-Tooth Disease (CMT), a hereditary disorder that affects the peripheral nervous system with progressive muscle weakness and sensory abnormalities due to myelin and axon damage. As already pointed out, QSM is a useful instrument to study myelin concentration *in vivo*, because of the proportionality between histopathological and imaging measurements[24] (Fig. 1.6).

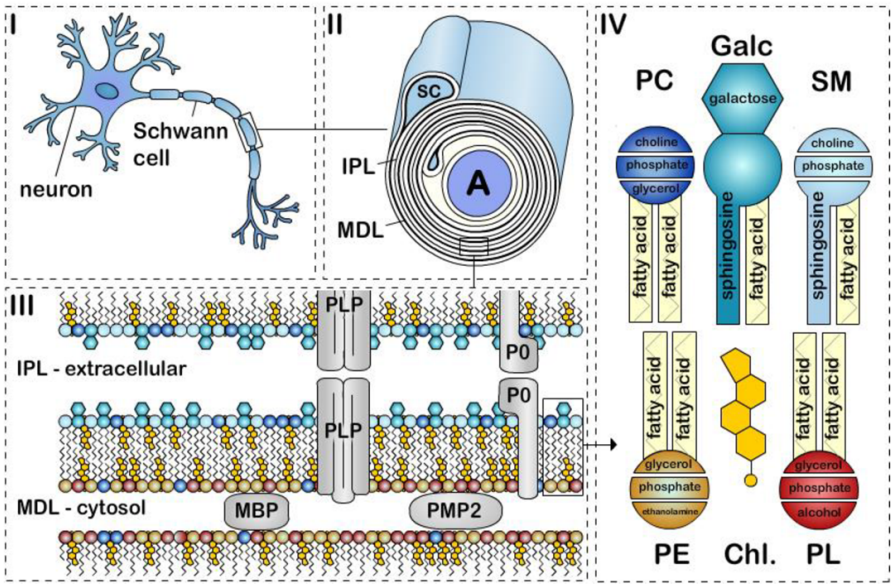

Figure 1.5: The structure of peripheral nervous system myelin sheath [39]. Schematic representation of a (i) myelinated axon, (ii) myelin sheath, (iii) bilayer membrane, and (iv) major lipids classes. Myelin is formed by apposition of the external surfaces and internal surfaces of the myelin bilayer that constitute the intraperiodic line (IPL) and the major dense line (MDL), respectively (ii, iii). The myelin bilayer has an asymmetric lipid composition (iii, iv). Myelin proteins are also asymmetrically distributed with for example PLP and P0 in the IPL and MBP and PMP2 at the MDL (iii). Cholesterol (Chl.), Galactosylceramide (Galc, cyan), Plasmalogen (PE, yellow), Phosphatidylcholine (PC, dark blue), Sphingomyelin (SM, light blue) and other phospholipids (PL, red). P0, PMP2 proteins and the enrichment of sphingomyelin in the myelin are specific to PNS myelin

Paramagnetic materials

Paramagnetic (Fig. 1.3, bottom diagram) materials respond positively when an external magnetic field is applied. The magnetic moments $\vec{\mu}_i$ of the single particles are directed randomly and have a non-zero intensity value even if there is not an external field $\vec{B}$ applied. Considering only the $\vec{z}$ direction, in a balanced situation ($\vec{H} = 0 \rightarrow \vec{M} = 0$) $n_+ = n_-$, where n_+ is the number of the magnetic moments pointing up, n_- the number of magnetic moments pointing down. This balance is broken when an external field is applied ($\vec{M} \neq 0$). The imbalance $n_+ - n_-$ is proportional to $\frac{\mu B}{k_B T}$ when $|\vec{\mu}||\vec{B}| << k_b T$ - where $\vec{\mu}$ is the total magnetic moment, k_B the Boltzmann constant ($1.23 \cdot 10^{-23} JK^{-1}$) and T is the temperature. When $\vec{B}$ has a strong intensity ($|\vec{\mu}||\vec{B}| >> k_b T$), the relationship stops to be linear and n_- tends to zero. All the single magnetic moments tend to line up in the same direction of the field, so the external magnetic field is reinforced. Unpaired electrons produce strong paramagnetic effects. Their spins tend to align themselves along the direction of the field. In this way, they make it stronger in intensity. In paramagnetic materials, χ assumes small and positive values; the relative permeability μ_r is slightly greater than 1. As paramagnetic materials in the brain, I am going to focus on iron-storage substances, heme and non-heme compounds.

- **Iron-storage substances**
  Iron is a transition metal with the following electron configuration: $[Ar]3d^6s^2$; the four unpaired electrons in the d-shell cause its magnetic response. In the human body there are 3 to 5 g of iron, 2/3 in hemoglobin and 1/3 in non-heme compounds (ionic form). Iron is involved in oxidative phosphorylation, DNA synthesis, transport of oxygen in the

blood, myelin synthesis and neurotransmitter production[14]. Iron pathogenesis is modified in abnormal situations; oxidative damage leads to the production of ferroptosis and disease-specific mechanisms: iron interacts with hallmark features of the disease (e.g. α-synuclein proteins), leading to the increase of the aggregation of these proteins, incorporations of iron in these compounds and then increase of oxidizing capacity and neuronal cell death.
An increase of χ values measured with QSM, directly proportional to iron concentration[24] (Fig. 1.6) was observed in brain regions associated with specific pathology: motor cortex in ALS, substantia nigra in PD, etc. In some disorders, the increase of susceptibility is also correlated with disease duration or severity in clinical evaluation[14].

- **Deoxy-hemoglobin**
  The majority of human iron is stored in hemoglobin; in Fig. 1.4, the structure that the molecule assumes when the oxygen is released during metabolic consumption is shown (Hb,Fe^{2+}). Deoxy-Hb is strongly paramagnetic because it has four unpaired electrons per heme; it may further oxidize in methemoglobin, which has five unpaired electrons per heme. The observation of hemoglobin magnetic effects allows us to study blood-oxygen-level-dependent signals and all disorders connected with the blood flow thanks to the oxygenation-related susceptibility changes occurring in blood products. QSM has been proven to be useful in investigating χ of intra-cerebral hemorrhages at various stages[40].

Ferromagnetic materials

Ferromagnetic materials react in a constructively when an external field is applied; they are characterized by a net magnetic moment, also in the absence of an external field. When a field is externally applied, their reaction is stronger than paramagnetic materials: their magnetization persists thanks to a very favorable energy exchange between the material and the field. Ferromagnetic materials have magnetic moments very high in intensity, so the magnetic susceptibility values are positive and very large.
There are no elements stored in known ferromagnetic structures in the human body, even if there are ferromagnetic elements - for example, iron is a ferromagnetic element, but ferritin and hemosiderin, which store iron, are not.

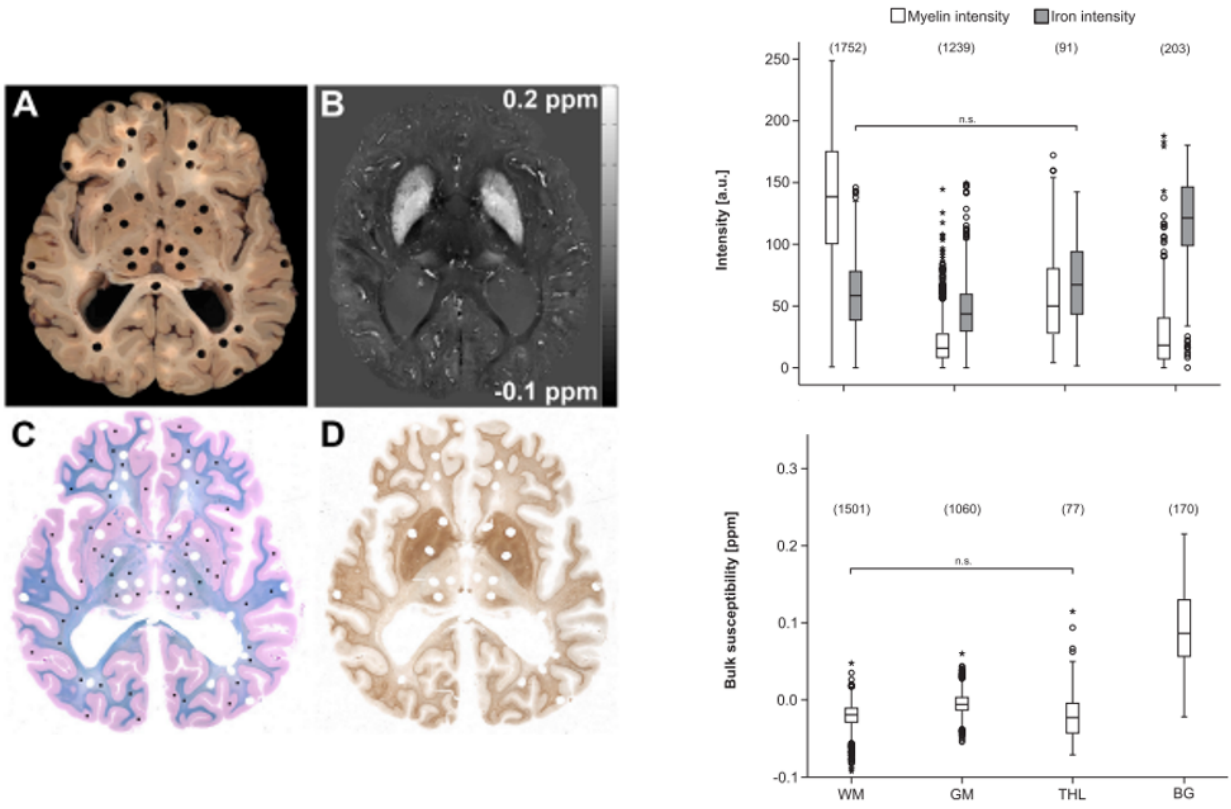

Figure 1.6: Figures from [24], comparing susceptibility measurements with iron and myelin concentration from histopathological measurements. Left panels: punched brain slice, corresponding QSM image, and myelin and iron stainings; (A) bottom cutting plane of a 5 mm thick formalin-fixed punched axial brain slice of case ID3 containing BG; (B) corresponding axial susceptibility ma; (C) corresponding myelin and (D) iron stainings. Right panel: histological intensities (upper diagram) and corresponding susceptibilities derived from MRI (bottom diagram) of white matter, grey matter, thalamus and basal ganglia. Susceptibility values from QSM positively correlate with quantitative ferrozine iron values in the basal ganglia, confirming paramagnetic effect of iron on susceptibility, while they negatively correlate with myelin intensities within the white matter, showing a diamagnetic effect. Myelin intensities were higher in the thalamus than in the basal ganglia: a significant relationship was nonetheless observed between quantitative iron values and QSM, confirming the applicability of the latter in this brain region for iron quantification

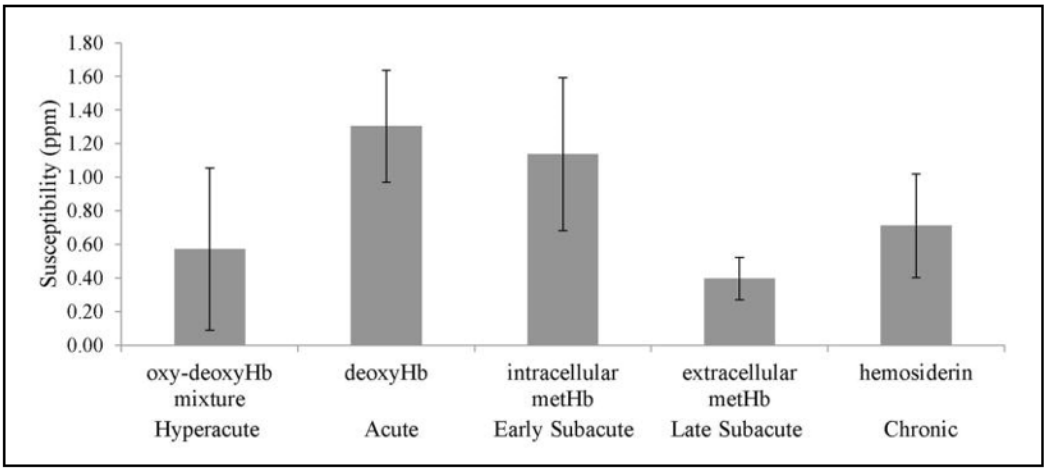

Figure 1.7: From [40], χ values of intracerebral hematomas at various stages and corresponding blood products, demonstrating positive χ at the hyperacute stage, followed by a significant increase at the acute stage from hyperacute, a significant decrease in the late subacute stage from early subacute, and an eventual significant increase at the chronic stage

1.2 Processing pipeline

An overview of QSM possible acquisition setting and processing options may be found in[4]. At Bellaria Hospital in Bologna (Functional and Molecular Neuroimaging Laboratory, IRCCS Istituto delle Scienze Neurologiche di Bologna, IT), where I have carried out my PhD research, there is a 3-T clinical scanner (Magnetom Skyra; Siemens Healthineers, Erlangen, Germany) equipped with a whole-body transmit and 64-channel Head/Neck receiver coil. QSM acquisition protocol and χ reconstruction were optimized from November 2019 to February 2020; nowadays, this sequence has been inserted in clinical routine, particularly in neurodegenerative protocols, reaching $\sim$ 1700 QSM examinations in September 2023. Together with patients, healthy volunteers have been scanned for comparison.

1.2.1 MR protocol and software

Details for QSM acquisition in our research group are (TE = Time of Echo, TR = Time of Repetition, SR = Super Resolution, FOV = Field Of View, FA = Flip Angle): 3D T_2^*w GRE, axial, 5 TEs, $TE_1/\Delta TE = 9.42/9.42$ ms, TR = 53 ms, SR = 0.5x0.5x1.5 mm^3, FOV = 196x224x144 mm^3, FA = 9°, scan time 8' 45"; ADAPTIVE[41] is used as coil combination method, automatically performed by the scanner. An example of raw data, magnitude and phase for each echo time, are shown in Fig. 1.8.

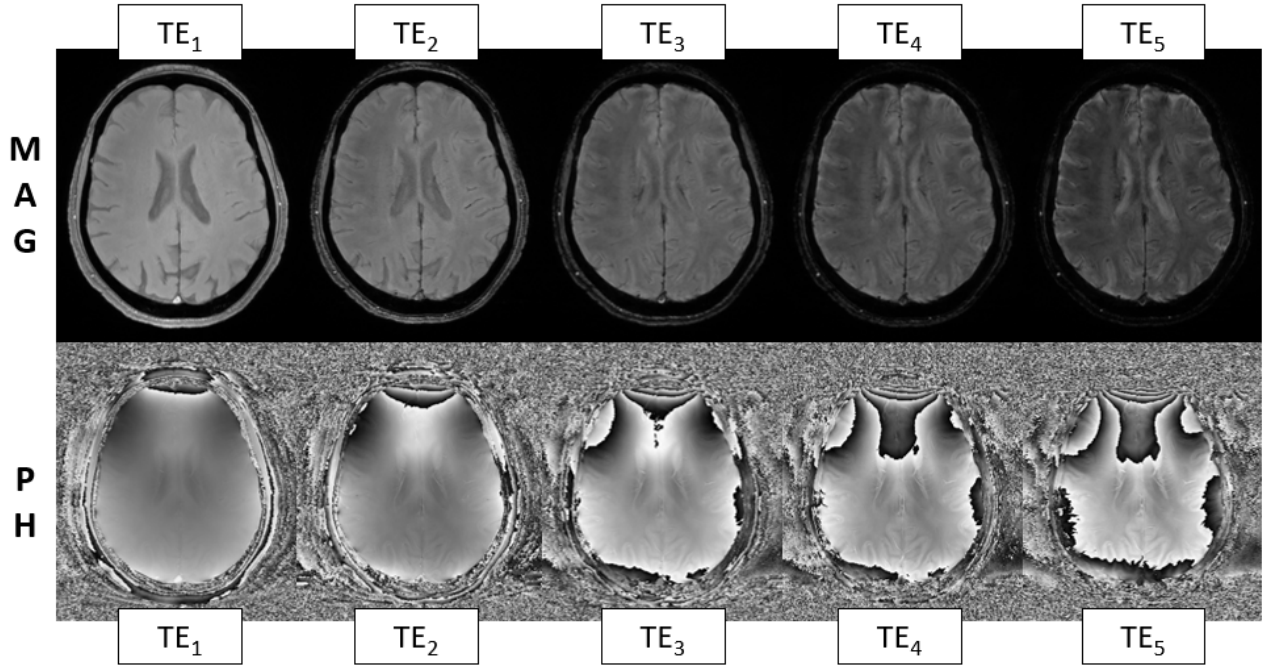

Figure 1.8: Magnitude and phase maps from the 5TE-sequence implemented in Bellaria Hospital to reconstruct QSM (healthy control, F/60 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit)

Other MR sequences are acquired during the exams, supporting QSM processing with respect to data analysis of various pathologies; for example, automated segmentation tools working on conventional morphological T_1w may be exploited and overlaid on QSM to study χ values. In

Tab. 1.1, parameters of the sequences acquired with QSM and used in the clinical applications are described. Details about their processing are going to be provided in the specific sections.

	MPRAGE (3D T_1w)	**FLAIR** (3D SPACE T_2w)	**HARDI DWI** (2D EPI single-shot)	**SWI** (3D GRE T_2*w)
Plane	sagittal	sagittal	axial	axial
TR (ms)	2300	5000	4300	28
TE (ms)	2.98	428	98	20
SR (mm^3)	1x1x1	1x1x1	2x2x2	0.5x0.5x1.5
FOV (mm^3)	176x256x256	176x256x256	220x220x176	196x224x144
FA (°)	9	120	90	15
Scan time	5' 21"	5' 55"	∼9'	4' 28"

Table 1.1: Acquisition parameters of T_1w, T_2w, DWI and SWI in Bellaria Hospital, IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit, Bologna (IT), with a 3-T clinical scanner (Magnetom Skyra; Siemens Healthineers, Erlangen, Germany) equipped with a whole-body transmit and 64-channel Head/Neck receiver coil (MPRAGE = Magnetization Prepared RApid Gradient Echo, FLAIR = FLuid Attenuated Inversion Recovery, SPACE = Sampling Perfection with Application optimised Contrast using different flip angle Evolution, HARDI = High Angular Resolution Diffusion Imaging, DWI = Diffusion Weighted Imaging, EPI = Echo Planar Imaging, SR = Spatial Resolution, FOV = Field Of View, FA = Flip Angle)

For DWI measurements, the sequence has changed in early 2021 from single- (5 unweighted volumes and 64 volumes b = 2000 s/mm^2) to multi- (8 unweighted volumes, 12 volumes b = 300 s/mm^2, 30 volumes b = 1000 s/mm^2 and 64 volumes b = 2000 s/mm^2) shell. Both sequences are acquired with Anterior-Posterior (AP) phase encoding. An additional sequence with inverted phase encoding (PA) (∼ 4' scan time) was acquired to correct EPI distortions artifacts in the EPI volumes: for single-shell measurements, 3 unweighted volumes were acquired, while for multi-shell the sequence was re-acquired repeating each volume acquisition with weighting 0-1000 s/mm^2. In the projects described in this thesis (Sec. 4.1), the DWI sequence was used to only tractography reconstructions, using volumes with b-value 2000 s/mm^2 for both single- and multi- shell measurements, without changing the pipeline. The volumes of the reconstructed tracts are consistent between the two sequences.

Most of the software and tools used to implement project pipelines illustrated in the following chapters are open source. Below there is a list of the ones used the most; specific functions used in the studies will be presented in the corresponding chapter.

- FSL: Functional magnetic resonance imaging of the brain (FMRIB) Software Library [42] (version 6.0.4) - comprehensive library analysis tools for MR brain imaging data
  - BET: Brain Extraction Tool [43], for skull-stripping
  - FLIRT: FSL's Linear Image Registration Tool [44] [45], automated robust and accurate tool for linear (affine) brain image registration, both intra- and inter- modalities
  - FNIRT (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FNIRT), performing registration with non-linear method when linear transformations are not sufficient to achieve good registration

  * FIRST[46], model-based segmentation tool; running on T_1w images, it returns the segmentation of deep brain gray matter nuclei

- ANTs: Advanced Normalization Tools (http://stnava.github.io/ANTs/) (version 2.3.5), software package for normalizing data to a template
- MRtrix3[47] (version 3.0.2), providing a set of tools for MR analysis
- FreeSurfer[48] (version 6.0.0) for the analysis and visualization of MR neuroimaging data from cross-sectional and longitudinal studies

1.2.2 χ map reconstruction

Processing stages are shown in Fig. 1.9. Phase maps were processed by Laplacian unwrapping[49] and Variable kernel Sophisticated Harmonic Artifact Reduction for Phase data (V-SHARP) as background field removal[50]. V-SHARP requires a brain mask, obtained skull-stripping TE_1 magnitude map using BET[43]. Multi-echo phase data were combined into a single-phase image through a weighted sum over echoes of processed phase maps[51]:

$$w(TE_i) = \frac{TE_i e^{-\frac{TE_i}{T_2^*}}}{\sum_{j=1}^{N} TE_j e^{-\frac{TE_j}{T_2^*}}} \tag{1.10}$$

with N is the number of echo times, equal to 5 in this pipeline. The iLSQR technique was used as dipole inversion method[29]. STI Suite[52] was used for the processing.

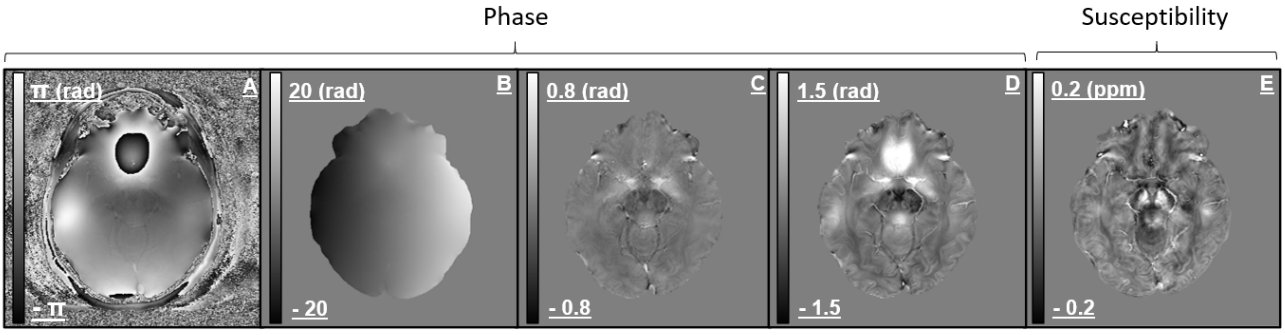

Figure 1.9: Phase and quantitative susceptibility maps from a 3D GRE 5-TEs T_2*w acquisition (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit, Ospedale Bellaria, Bologna): A) Raw phase data (1^{st} TE); B-C) phase map after laplacian unwrapping and V-SHARP background field removal (1^{st} TE); D) total phase map as result of multi-echo map combination through weighted sum; E) QSM, evaluated with iLSQR (healthy control, F/46 years old)

During the preliminary stages, different combinations of parameters and strategies have been attempted on a few subjects[53], changing spatial resolution, number and values of echo times and coil combination method. Those were the competitive configurations:

a. 2 TEs ($TE_1/\Delta TE/TR$=9.42/9.42/27 ms) and spatial resolution 1x1x1 mm^3, 0.8x0.8x1.5 mm^3, 0.5x0.5x1.5 mm^3, 0.5x0.5x1 mm^3, A Simple Phase Image Reconstruction for multi-Echo data (ASPIRE)[54] as coil combination method (scan time $\sim$ 3,4,5,7')

b. 5 TEs and spatial resolution 0.5x0.5x1.5 mm^3 (TE$_1$/ΔTE/TR=9.42/9.42/53 ms, $\sim$9 mins), using both ASPIRE, manually implemented, and the ADAPTIVE coil combination methods[41], automatically performed by the scanner

c. 8 TEs (TE$_1$/ΔTE/TR=5.6/5.6/51 ms) and spatial resolution 0.8x0.8x1.5 mm^3 (scan time $\sim$7 halving the number of axial slices)

An overall qualitative evaluation of the reconstructed images was conducted, together with quantitative evaluation of contrast of small structures, specifically of the Nigrosome-1 (N1), substructure of the Substantia Nigra *pars compacta* (SNpc) enriched by dopaminergic neuron, which loss in PD and Atypical Parkinsonian Syndromes (APS) (α-synucleinopathies) is related to iron accumulation. N1 changes are detectable using QSM and are supportive MR markers for differential diagnosis of these neurodegenerative disorders[55]. However, its detection is not that easy and requires at least the use of a high-strength magnetic field and high spatial resolution setting. In Fig.1.10, examples of N1 detection in the different configurations.

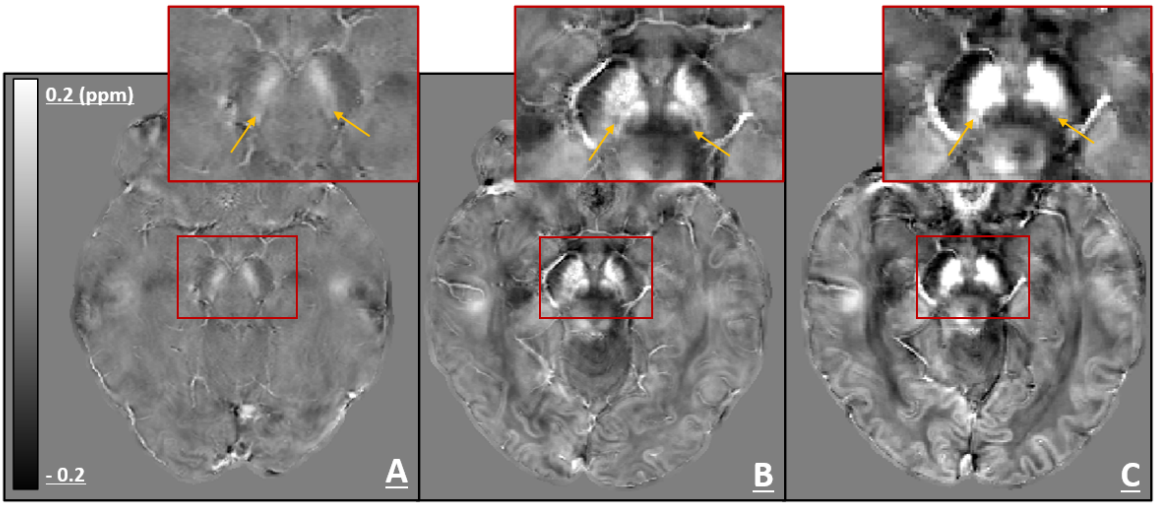

Figure 1.10: Axial view with SN and N1 of QSM maps[53]: A) 2TEs, spatial resolution 0.5x0.5x1.5 mm^3, ASPIRE; B) 5TEs, spatial resolution 0.5x0.5x1.5 mm^3, ADAPTIVE; C) 8TEs, spatial resolution 0.8x0.8x1.5 mm^3, ADAPTIVE (healthy control, F/46 years old) (Neuroimaging Laboratory, IRCCS Istituto Neuroscienze Neurologiche di Bologna, IT)

The impact of N1 recognition becomes difficult as the number of echo times decreases; of course, this is not the only factor to take into account. N1 and SN were manually segmented and the Contrast-to-Noise Ratio (CNR) N1-SN in susceptibility χ maps was evaluated, with 0.5 as estimated error:

$$CNR_{N1-SN} = \frac{| < x >_{N1} - < x >_{SN} |}{\sigma^2_{SN}} \tag{1.11}$$

where $< x >$ is the mean intensity value within N1 and SN and σ is the standard deviation. In Fig. 1.11, CNR_{N1-SN} of the different configurations is shown. Maximizing the contrast of N1 in SN, ADAPTIVE was selected as coil combination method, 5 echo times (TE$_1$/ΔTE/TR=9.42/9.42 ms) and 0.5x0.5x1 mm^3 as spatial resolution.

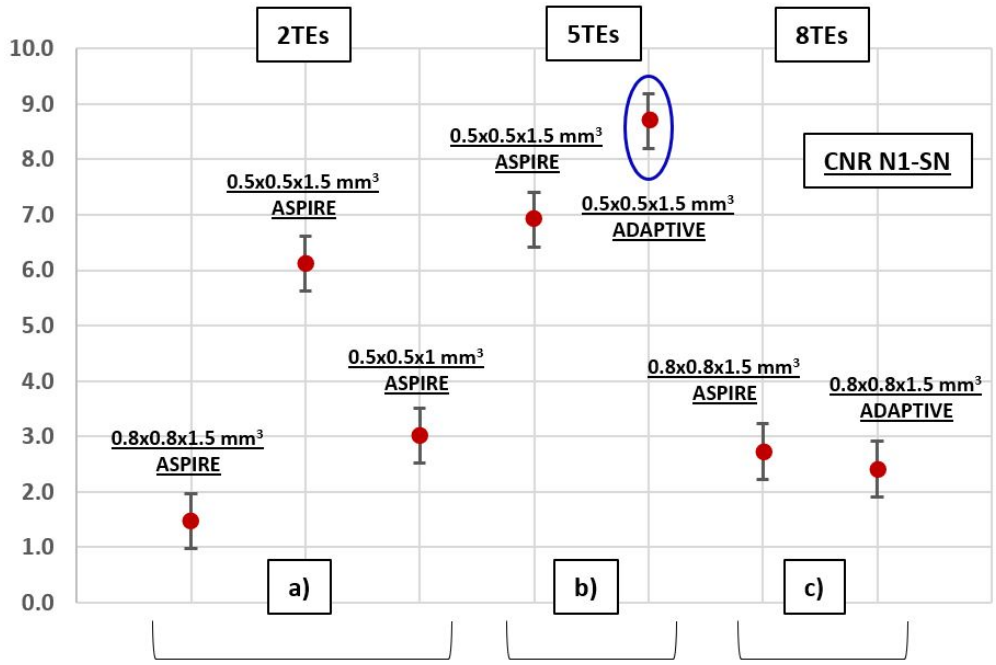

Figure 1.11: Scatter plot with CNR_{N1-SN} in the different configurations tested to optimize QSM reconstruction, with 0.5 set as error. In this graph, the number of echo times, spatial resolution and coil combination method are reported for each configuration; $TE_1/\Delta TE/TR$ (ms) are = 9.42/9.42/27 for 2TE-configurations, 9.42/9.42/53 for 5TE-configurations and 5.6/5.6/51 for 8TE-configurations

1.2.3 Cerebrospinal fluid as reference tissue: atlas-based method

QSM enables the evaluation of $\chi(r)$ up to an additive factor χ_0, called reference value[56] [57]:

$$\chi(r)^* = \chi(r) + \chi_0 \tag{1.12}$$

χ_0 changes from one acquisition to another, thus its evaluation is necessary to compare χ distributions intra- and inter- subjects. The most common method is called *zero-referencing*[58]: χ of a specific area, presuming to have a null response to the magnetic field, is set equal to zero, and the entire $\chi(r)$ is shifted by the same factor. Usually, the cerebrospinal fluid (CSF) is chosen, because its susceptibility value is not affected by the subject's age or pathological condition and there is no orientation dependence. Normally, the average CSF signal over ventricles is considered.
However, since the size of the ventricles varies and CSF often appears non-uniform in QSM because of the presence of blood vessels and/or calcifications, the QSM signal inside ventricles is not uniform. Instead of considering the entire ventricles, in our pipeline three small volumes were selected following the work of Straub and colleagues, 2017 (atrium, anterior horns and central part)[57]. An original atlas-based method was implemented to automatically identify the three volumes on each exam. MPRAGE maps of 60 subjects, mix of healthy controls and patients, were non-linearly registered using FNIRT (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FNIRT), after brain extraction with BET[43], to the space of a selected reference subject. The magnitude of the first echo time of the GRE T_2^*w sequence was registered via affine transformation to the corresponding MPRAGE using FLIRT[44] [45]; this operation was combined with the deformation matrix obtained from the non-linear transformation, to have all the 60 QSM-magnitude maps

averaged in the same space. The resulting 60-subject atlas is shown in Fig. 1.12. The three small CSF volumes were manually drawn on the atlas on both sides, selecting three adjacent slices for each. The volumes were back-registered into the original QSM map for each subject and the mean susceptibility value was extracted and considered as reference. Because of the artifacts in QSM near air-bone interfaces[59], the weighted sum was considered instead of the arithmetic one, giving less weight to the values that deviate the most from the average value. The method was tested on 120 subjects, 60 used to build up the atlas and the other 60 news.
The extracted values were consistent with values from manual segmentation, considered as the gold standard. In ∼9% of the cases the back registration failed, meaning that the back-registered volumes fell out from the ventricles, thus manual segmentation was performed. The atlas-based method implemented is promising to automatically and accurately extract susceptibility reference values. The evaluated values were consistent with the gold standard within 1.5 σ, both in the 60 subjects used to build the atlas and the 60 subjects constituting the test sample. The measurement uncertainty is significantly reduced using three small volumes instead of the whole lateral ventricles: in 70% of cases, it results in less than half.

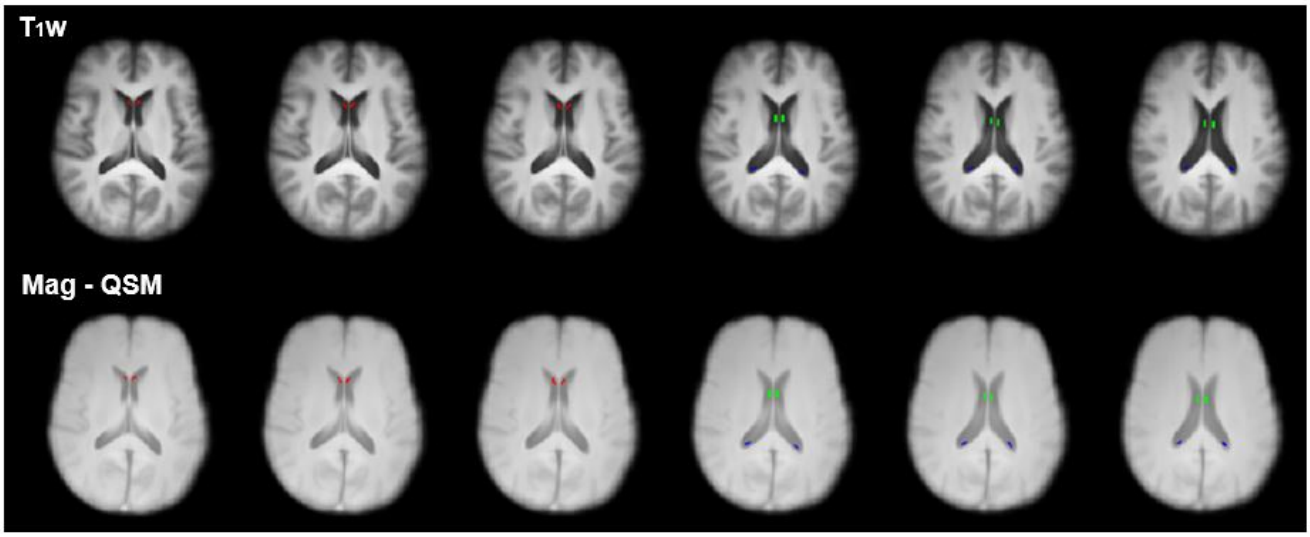

Figure 1.12: T_1w and QSM-Magnitude 60-subject atlas built-up to automatically identify the anterior horns, atrium and body of the lateral ventricles in each exam and extract the susceptibility reference value (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit, Bellaria Hospital)

1.2.4 Susceptibility-based imaging: QSM and SWI

In this section, I would like to provide a few examples of QSM compared to SWI. As aforementioned, there are quantitative and qualitative advantages in using QSM[3], including visual benefits thanks to its ability to distinguish dia- and para- magnetic materials (Fig. 1.13). The presence of dia- and para- magnetic substances leads to a de-phasing of the spins and dropping signal in T_2^*w images: while SWI considers this information qualitatively, QSM measures the susceptibility distribution from phase data.

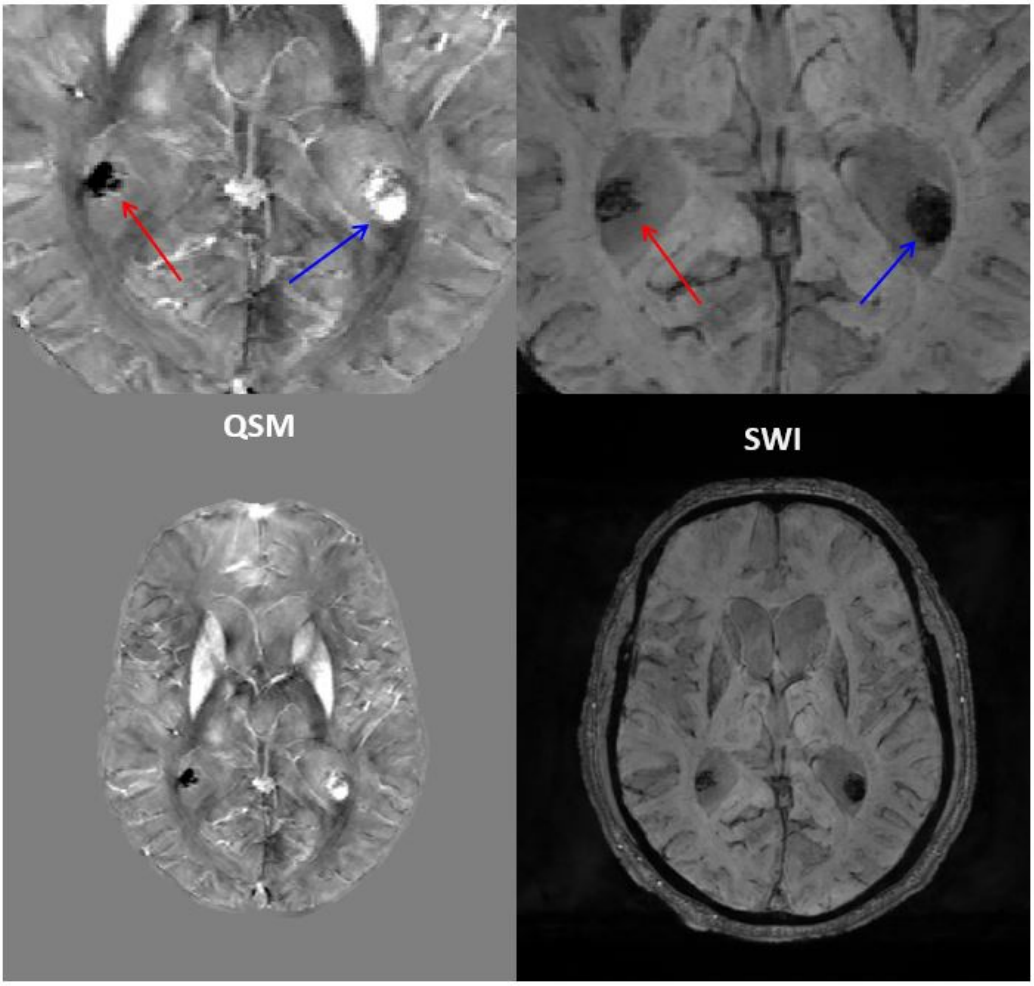

Figure 1.13: QSM and SWI of a patient with Lewy Body Dementia (M/79 years old, IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit). Diamagnetic (red arrow) and paramagnetic (blue arrow) materials appear different in QSM while they have the same intensity and contrast in SWI. Acquisition parameters are in Section 1.2.1 (Tab. 1.1)

In Fig. 1.14, an example of QSM and SWI reconstructions from the same 3D GRE T_2^* sequence: note that they can ideally be obtained using the same raw data and changing the post-processing. In our center, as it commonly happens, two different sequences were optimized for the two methods (Sec. 1.2.1, Tab. 1.1); examples shown in this section refer to those measurements.

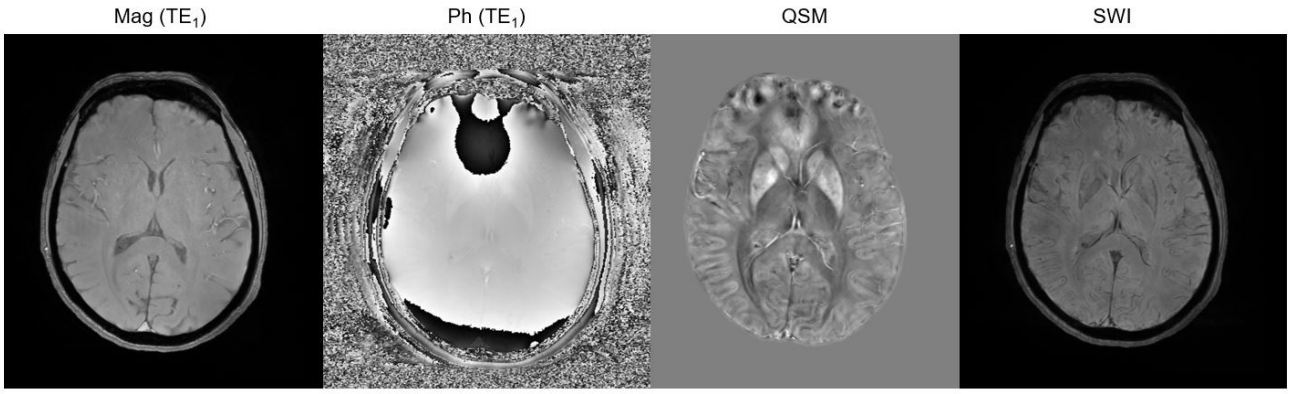

Figure 1.14: From the left: Magnitude and phase (1^{st} TE) raw data from QSM sequence, QSM reconstruction from the 5 TEs combination and SWI obtained with the same measurements (HC F/60 years old, IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit)

Brain iron accumulation in deep brain nuclei has been described in APS, mainly in Multiple System Atrophy (MSA) and Progressive Supranuclear Palsy (PSP)[60]. In Fig. 1.15, an example of a patient with MSA is shown, displaying basal ganglia appearance in both SWI and QSM: iron deposition is detectable by the two sequences as hypo- and hyper- intense respectively. However, with SWI susceptibility is not actually measured; quantitative SBI techniques have helped in the identification of disease-related patterns.

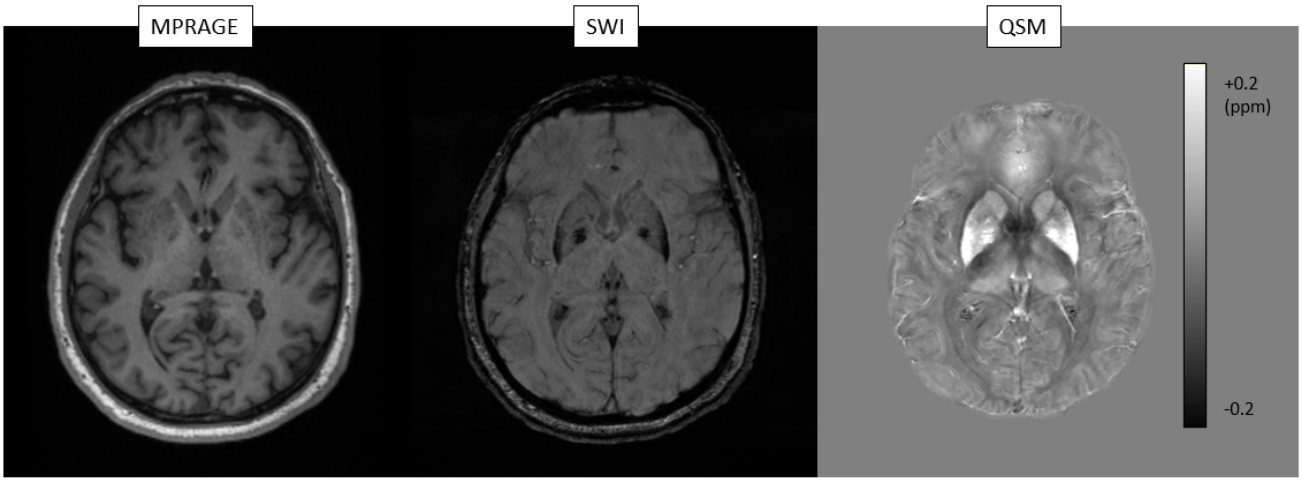

Figure 1.15: MPRAGE, SWI and QSM in a patient with MSA-C (cerebellar type) (M/65 years old, IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit). Iron accumulation is detectable using susceptibility-based techniques - as hypo- and hyper-intensity respectively in SWI and QSM - in putamen and globus pallidus. QSM provides quantitative measurements, thus its color bar is shown, while MPRAGE and SWI are in arbitrary units. Details acquisition parameters are in Section 1.2.1 (Tab. 1.1)

QSM has been used to detect paramagnetic rims of iron-loaded active microglia and macrophages in patients with MS[61]. They are also referred to as *chronic active* lesions and their negative impact on the disease course has been suggested in literature[62], making their exploration crucial to understand and characterize this disorder. In Fig. 1.16, an example is given: note that rim+ lesions may gain contrast in SWI image but without the rim. Together with the detection, with QSM it is possible to quantify χ lesions and monitor their values longitudinally[63].

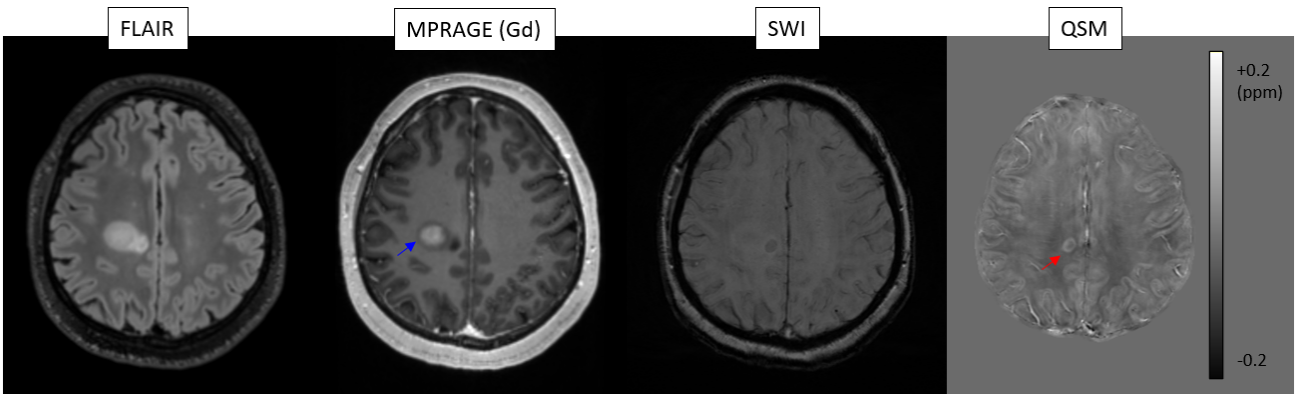

Figure 1.16: FLAIR, MPRAGE-Gd (gadolinium-enhanced), SWI and QSM in a patient with relapsing-remitting MS (F/33 years old, IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit). QSM allows the detection *chronic active* lesions (red arrow) reflecting an ongoing inflammatory process not Gd-enhanced - note that the lesion gains contrast in SWI image but without the rim. Instead, enhanced-Gd lesions (blue arrow) are not detectable in QSM. QSM provides quantitative measurements, thus its color bar is shown, while MPRAGE and SWI are in arbitrary units. Details acquisition parameters are in Section 1.2.1 (Tab. 1.1)

In Figs. 1.17 and 1.18, an example of brain calcification and Transient Ischemic Attack (TIA); as it may be observed, in SWI calcium and hemorrhages products appear both hypo-intense while in QSM respectively hypo- and hyper- intense.

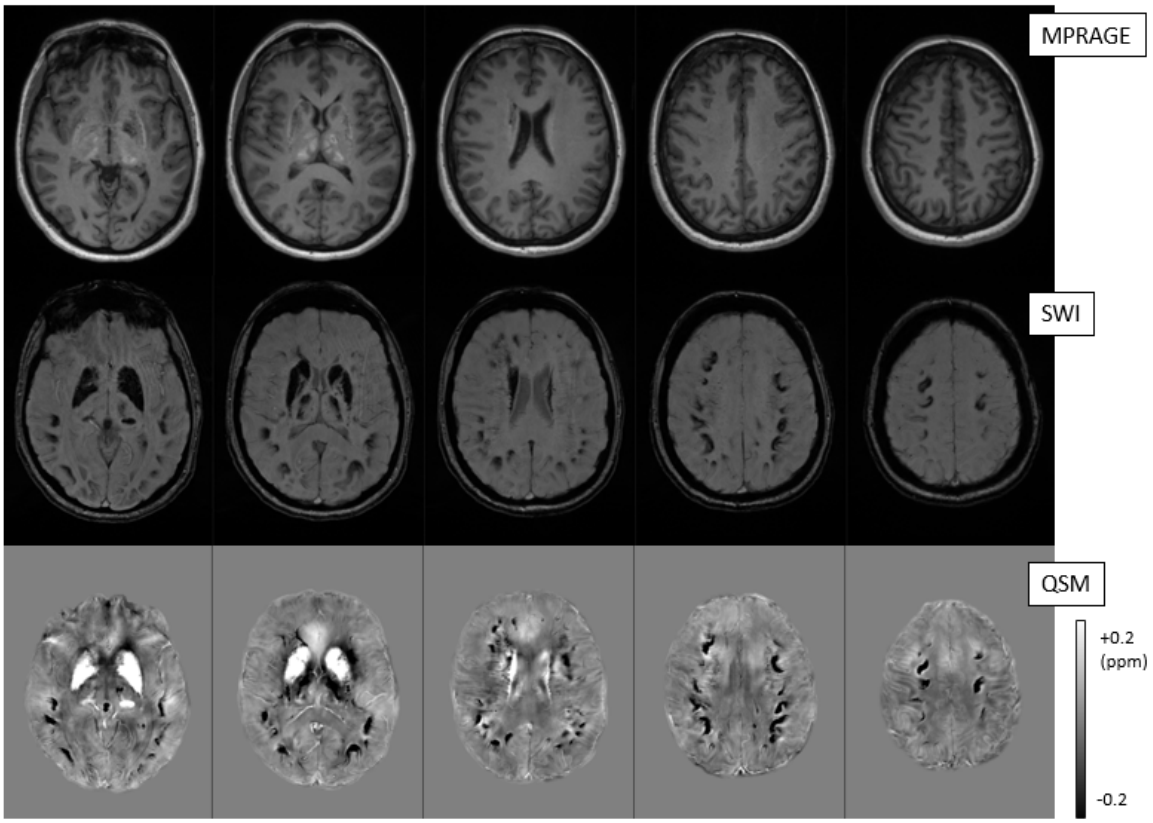

Figure 1.17: MPRAGE, SWI and QSM in a patient with brain calcification (F/44 years old, IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit). Calcium accumulation is detectable in cortical and subcortical regions both in SWI (hypo-intensity) and QSM (hypo-intensity); iron deposition is detectable in the basal ganglia (putamen, globus pallidus, thalamus) both in SWI (hypo-intensity) and QSM (hyper-intensity). Note dia- and para- magnetic contributions in the two sequences, both hypo-intense in SWI while hypo- and hyper- intense respectively in QSM. QSM provides quantitative measurements, thus its color bar is shown, while MPRAGE and SWI are in arbitrary units. Details acquisition parameters are in Section 1.2.1 (Tab. 1.1)

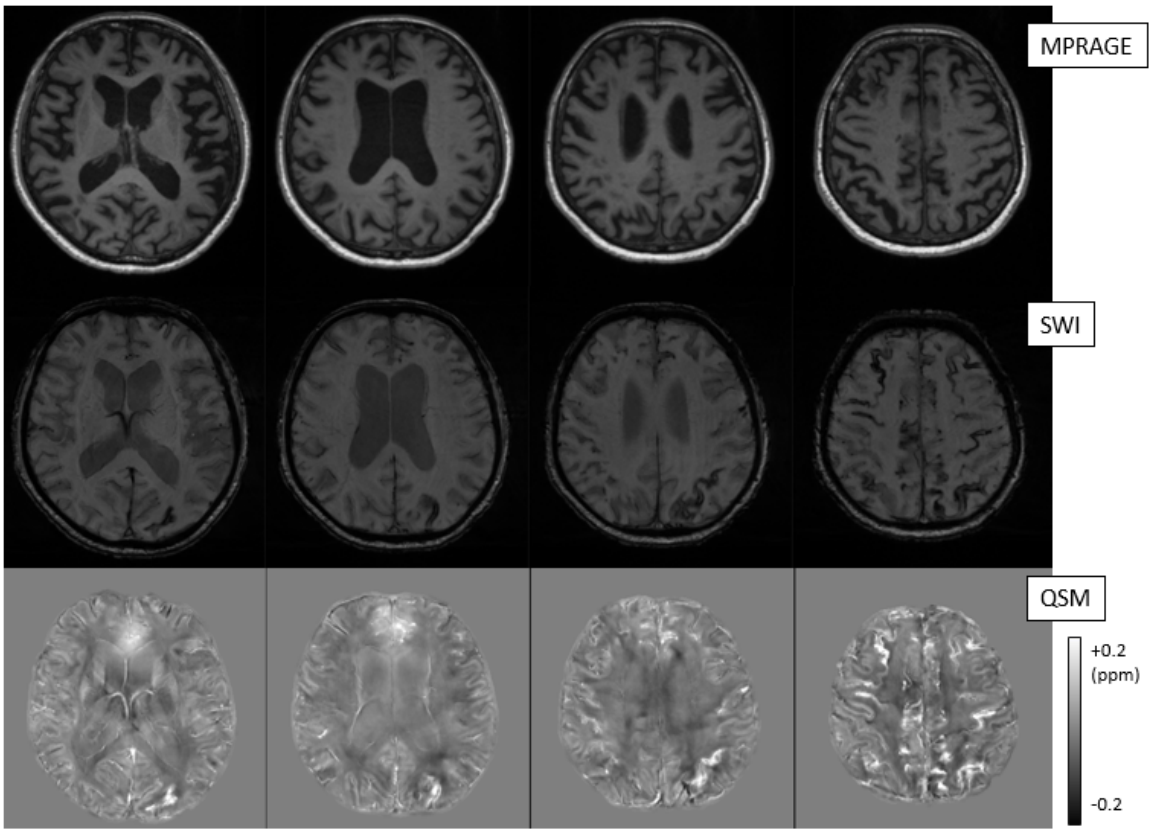

Figure 1.18: MPRAGE, SWI and QSM in a patient with TIA (transient ischemic attack) (M/83 years old, IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit). QSM hyper-intensities are observed within the old cortical hemorrhages. The hemoglobin contained in the blood has weakly diamagnetic behavior when it is in its oxygenated form, while it assumes paramagnetic behavior in its de-oxygenated form (v. Section 1.1.3). In the early stages of a hemorrhage, the blood hemoglobin is de-oxygenated and the iron ions (Fe^{2+}) within it have a paramagnetic response to the magnetic field, which translates into a hyper-intensity in QSM. In the subsequent stages, iron ions are transformed into methemoglobin and hemosiderin, the paramagnetic response is accentuated and the hyperintensities become more marked. QSM provides quantitative measurements, thus its color bar is shown, while MPRAGE and SWI are in arbitrary units. Details acquisition parameters are in Section 1.2.1 (Tab. 1.1)

1.3 Issues and new developments

Since its introduction, QSM has undergone further refinement, and various research groups and medical institutions have contributed to its development and application. Of course, while it offers advantages as MR sequence and has become an important tool in the field of neuroimaging, QSM comes with its own challenges, such as the need for complex post-processing and potential limitations in cases of complex tissue geometries. Below, some issues, requiring further developments, are listed.

χ **artifacts near air-bone interfaces**: Regions of the brain near air-bone interfaces, such as the sinuses or skull, are susceptible to susceptibility artifacts. These artifacts can manifest as streaking or distortions in the images, making it challenging to accurately interpret the data [59]. One of the techniques proposed to fix this issue is using ROMEO (Removal of Macroscopic Field Effects by Oscillating Field Gradients) as unwrapping algorithm [64]; it partially corrects for these artifacts, but it does not entirely eliminate the problem.

Sensitivity to motion: Phase measurements are highly sensitive to motion artifacts, considering that longer echo times are the ones containing more information about the transversal spin de-phasing. Any motion during the acquisition can lead to errors in the phase measurements, which consequently affect the accuracy of the susceptibility maps. This is particularly demanding for patients with neurodegenerative disorders and movement dysregulation, considering that QSM sequences usually last $\sim 8'$. 3D EPI sequences have been proposed to reduce acquisition time and motion artifacts[65], but they may not be widely adopted due to various trade-offs in image quality, acquisition time, and susceptibility to other artifacts. In fact, they are way faster than GRE sequences and can achieve better spatial resolution, but they are more liable to χ artifacts near air-tissue interfaces and regions with magnetic field inhomogeneities, which may lead to distortions and artifacts, especially in areas near the sinuses or skull. Additionally, they tend to have lower SNR than GRE sequences. Ongoing research is still striving to find an appropriate compromise.

Heterogeneity of χ sources: Contributions of dia- and para- magnetism can overlap in QSM images, leading to issues in interpreting QSM signal: when an increase in χ is observed, it may result from either an increase in paramagnetic materials or a decrease in diamagnetic materials concentration. Consider, for instance, the thalamus, where both gray matter and white matter coexist, and susceptibility values are linked to both iron and myelin concentration[24]. Recently, a new technique has been developed, called χ-separation[66], helping in distinguishing the contribution of negative and positive χ values; it has been applied, for example, to explore MS lesions[67].

Standardized atlas and sharing data: Another noteworthy issue to highlight is that certain biomarkers of pathology, such as iron accumulation, are associated with both pathological conditions and natural aging processes[68]; MR techniques such QSM certainly helps in distinguishing those processes - I will further explore this topic talking about iron accumulation in the motor cortex in ALS patients and healthy controls (Ch. 2). A more comprehensive characterization of the patterns of susceptibility is often very helpful, and, as I will explore later, radiomics can play a pivotal role in advancing this direction (Sec. 4.1).
However, it is necessary to establish a robust control group and to have a standardized atlas to provide a reference standard and develop correction methods for susceptibility values. Correction methods are nontrivial since different regions exhibit varying susceptibility changes with age and require different fitting models[69] [70]. Recruiting suitable healthy control subjects, particularly elderly individuals, poses its own set of challenges, and, since QSM is a relatively new technique, individual centers struggle to reach an adequate number of subjects to build correction models. Thus, sharing data from both control subjects and patients is essential to facilitate collaborative efforts aimed at refining harmonization techniques and developing more accurate fitting models[71].

χ = magnetic susceptibility
AD = Alzheimer's Disease
ALS = Amyotrophic Lateral Sclerosis
ANTs = Advanced Normalization Tools
APS = Atypical Parkinsonian Syndrome
ASPIRE = A Simple Phase Imaging Reconstruction for multi-Echo data
BET = Brain Extraction Tool
CMT = Charcot-Marie-Tooth Disease
CNR = Contrast-to-Noise Ratio
CNS = Central Nervous System
COSMOS = Calculation Of susceptibility through the Multiple Orientation Sampling
CSF = CerebroSpinal Fluid
DM1 = Myotonic Dystrophy type 1
DWI = Diffusion Weighted Imaging
EPI = Echo Planar Imaging
FA = Flip Angle
FLAIR = FLuid Attenuated Inversion Recovery
FLIRT = FLS's Linear Image Registration Tool
FOV = Field Of View
FSL = Functional magnetic resonance imaging of the brain (FMRIB) Software Library
GBS = Guillain-Barré Syndrome
GRE = GRadient Echo
HARDI = High Angular Resolution Diffusion Imaging
HD = Huntington's Disease
HEIDI = Homogeneity Enabled Incremental Dipole Inversion
iLSQR = iterative Least SQuaRe
MEDI = Morphology Enabled Dipole Inversion
MPRAGE = Magnetization Prepared RApid Gradient Echo
MR = Magnetic Resonance
MRI = MR imaging
MS = Multiple Sclerosis
MSA = Multiple System Atrophy
N1 = Nigrosome-1
NMOSD = NeuroMyelitis Optica Spectrum Disorder
PD = Parkinson's Disease
PM = Personalized Medicine
PSP = Progressive Supranuclear Palsy
QSM = Quantitative Susceptibility Mapping
ROMEO = Removal of Macroscopic Field Effects by Oscillating Field Gradients
SBI = Susceptibility-Based Imaging
SNpc = Substantia Nigra *pars compacta*
SPACE = Sampling Perfection with Application optimised Contrast using different flip angle Evolution
STI = Susceptibility Tensor Imaging
SWI = Susceptibility Weighted Imaging
T_1w, T_2w, T_2^*w = T_1-, T_2-, T_2^*- weighted
TE = Time of Echo
TGV = Total Generalized Variation
TIA = Transient Ischemic Attack
TR = Time of Repetition
V-SHARP = Variable kernel Sophisticated Harmonic Artifact Reduction for Phase data

Chapter 2

Amyotrophic Lateral Sclerosis

Amyotrophic Lateral Sclerosis (ALS) is a progressive idiopathic neurodegenerative disorder affecting nerve cells in the brain and spinal cord; symptoms generally start with muscle weakness and subsequently progress, becoming more severe over time[72]. A well-established hallmark of ALS development is the accumulation of iron in the motor cortex, a phenomenon easily detectable and quantifiable using QSM[73]. Previous studies have identified an increase of susceptibility values in the precentral gyrus, which directly correlate with iron concentration[74]. Considering that motor impairment is not the only symptom experienced by ALS patients, the proposed QSM application aims to investigate whether iron accumulation is observed in other cortical regions, to enhance the understanding of the disease, particularly in comparison with other motor neuron disorders. Background and aims of the study are further described in Sec. 2.1.

Materials and methods of this application are in Sec. 2.2. The study cohort consisted of 47 ALS patients, 26 patients with other neurodegenerative conditions, and 20 healthy controls. The brain MR protocol included morphological T_1w and T_2w, and QSM. ALS Functional Rating Scale (ALSFRS) and Penn Upper Motor Neuron (UMN) scores were available as clinical data.
First, qualitative evaluation was blindly performed by three neuroradiologists assessing radiological signs of UMN impairment in a group mix of patients and controls. Then, quantitative analysis was carried out, considering several χ properties (mean, median, 90^{th} percentile, skewness, and kurtosis) from QSM and structural properties (volume and cortical thickness) from T_1w images. Several gyri in the frontal, parietal, temporal lobes and in the cingulate cortex were automatically segmented and analyzed. Extracted χ and structural metrics were compared among the groups and correlated with clinical data.

Outcomes are shown and discussed in Sec. 2.3 and 2.4. First, quantitative parameters showed better accuracy than qualitative image evaluation. Additionally, mean, median, 90^{th} percentile, and skewness of susceptibility in several gyri within the frontal lobe, the parietal lobe and the cingulate

cortex were significantly higher in ALS patients compared to healthy controls. χ 90^{th} percentile emerged as the most reliable indicator for distinguishing ALS patients from healthy controls and patients with other neurodegenerative diseases. Reductions in volume and cortical thickness in the ALS group were consistent with findings from previous studies. Notably, susceptibility values positively correlated with the Penn UMN score in patients with rapid disease progression. These findings suggest that iron accumulation during the development of ALS may affect multiple cortical regions, offering valuable insights into the neuropathological changes of ALS (Sec. 2.5).

The first use of abbreviations includes their full expansion; for ease of reference, a glossary is provided at the end of the chapter.

2.1 Background and aims

Amyotrophic Lateral Sclerosis (ALS) is a progressive neurodegenerative condition that leads to a gradual impairment of both Upper and Lower Motor Neuron (UMN and LMN), resulting in the eventual loss of control over voluntary and involuntary movements[74]. ASL exhibits an annual incidence rate ranging from 1 to 2 cases per 100000 people per year, with some variations across continents and higher incidence rates observed in Europe and North America than Africa and Asia[75]. On average, ALS occurrence is higher in men than women in sporadic cases and it is approximately equal in familial cases (5-10% of ALS patients). When considering the overall population-based lifetime risk, it stands at approximately 1 in 400 for women and 1 in 350 for men. The peak age at which ALS typically manifests is around 58 to 63 years for sporadic cases and 47 to 52 years for familial cases. ALS is characterized by its relentless progression, with approximately 50% of patients succumbing to the disease within 30 months of experiencing symptoms; about 20% of patients manage to survive for a period ranging from 5 to 10 years following the onset of symptoms.

The neurodegeneration in ALS arises from a multifaceted set of mechanisms that are interconnected at the molecular and genetic levels[72][75]. The neurodegeneration associated with ALS results from a complex interplay of factors, such as glutamate excitotoxicity, generation of free radicals, presence of cytoplasmic protein aggregates, involvement of SOD1 enzymes, mitochondrial dysfunction, and disruptions in axonal transport processes due to the accumulation of intracellular neurofilament aggregates that prove detrimental to neurons. Mutations in TARDBP and FUS lead to the formation of intracellular aggregates in neurons. The activation of microglia results in the release of proinflammatory cytokines, worsening the toxicity. Ultimately, the degeneration of motor neurons is driven by the activation of calcium-dependent enzymatic pathways. Pathological studies conducted *ex vivo* have revealed that the activation of microglia in the deep layer of the primary motor cortex (M1)[76] is associated with iron accumulation in the same area[77], suggesting that iron buildup could potentially serve as a neuropathological hallmark of the disease.

ALS manifests in different ways[78][79]: 1) limb-onset, exhibiting both UMN and LMN signs in the limbs; 2) bulbar-onset, beginning with difficulties in speech and swallowing, with limb symptoms appearing later in the disease progression; 3) primary lateral sclerosis, less common and primarily involving UMNs; 4) progressive muscular atrophy, primarily affecting LMNs. Classical ALS involves both UMN and LMN signs, affecting the brainstem and multiple regions of the spinal cord; Pure UMN (PUMN) and Pure LMN (PLMN) patients show only UMN and LMN signs respectively. Patients can initially present with bulbar-onset disease ($\sim 25\%$), limb-onset disease ($\sim 70\%$), or other areas like the trunk or respiratory system ($\sim 5\%$). Some atypical presentations include weight loss, muscle cramps and twitching without weakness, emotional instability, and cognitive dysfunction resembling frontal lobe impairment.

There are issues in identifying ALS biomarkers of disease progression and therapeutic benefit; thus formal diagnostic criteria for ALS and Motor Neuron Diseases (MND) in general have been developed and modified several times; nowadays, the revised El Escorial criteria[80] are the most used. Conventional MRI radiological findings can provide support for ALS diagnosing[81][82] (Fig. 2.1): hyperintensity in the Cortico-Spinal Tracts (CST) is often detectable in T_2w images, as well as the "motor band sign", which corresponds to hypo-intensity in SWI due to iron accumulation. However, these features are not specific imaging characteristics[83]. As an advanced MR technique, QSM[1] has helped in finding imaging biomarkers and understanding the mechanisms of pathology. In contrast to SWI, QSM is a quantitative method that provides voxel-wise magnetic susceptibility χ values, influenced by both diamagnetic (e.g. myelin) and paramagnetic (e.g. iron) properties of the tissue. This enables the study of iron concentration and its spatial distribution within *in vivo* high-resolution MRI. Additionally, QSM provides qualitative information since the "motor band sign" appears as a hyper-intensity in QSM images in patients with ALS (Fig. 2.2), assessable by a standard scale assigned looking into the intensity similarity to the post-central gyrus[82].

QSM enables precise exploration of magnetic properties, facilitating intra- and inter-subject quantitative comparisons. Significant alterations have been observed in ALS patients compared to healthy controls in the motor cortex. Measures of PrCG bulk susceptibility, such as mean and median, indicate higher values in ALS patients than in healthy controls[14][73][82], suggesting an increase in iron concentration. Skewness and kurtosis, which respectively measure the symmetry and tail weight relative to a normal distribution, have been introduced in previous studies[84][85][86], showing promise as potential biomarkers for characterizing diseases and subgroups. Furthermore, atrophy of the PrCG is associated with alterations in susceptibility properties[87][84].

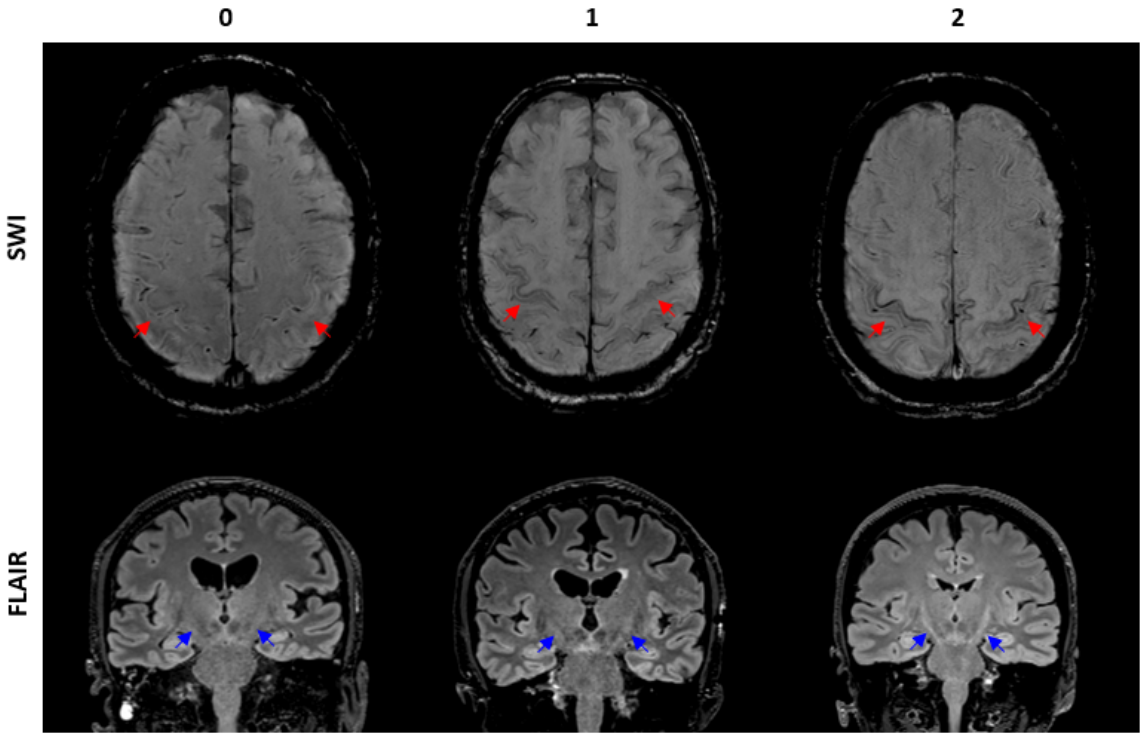

Figure 2.1: Examples of radiological signs in ALS patients: in the first row, SWI hypo-intensity in the precentral gyrus ("motor band sign"); in the second row, T_2w FLAIR hyper-intensity in the cortico-spinal tract (protocol details are in Sec. 1.2.1). Radiological evaluation was performed using a 3-score ranking as suggested from previous studies [81]: 0 = absent (first column, M/64 years old), 1 = present, mild (second column, M/68 years old), 2 = present, marked (third column, F/55 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit, IT)

The primary objective of this study [88] [89] is to employ QSM to identify more precise imaging biomarkers capable of distinguishing ALS patients from both healthy controls (HC) and patients with Other Neurodegenerative Disorders (OND). While ALS neuroimaging studies have predominantly focused on motor-related signs, various MR sequences have revealed differences between healthy controls and ALS patients in frontal and temporal cortical regions [90]. To gain a deeper understanding of the role of iron accumulation in ALS degeneration and to explore ways to characterize the disease comprehensively, this study investigates susceptibility and structural properties across multiple gyri throughout the entire cortex.

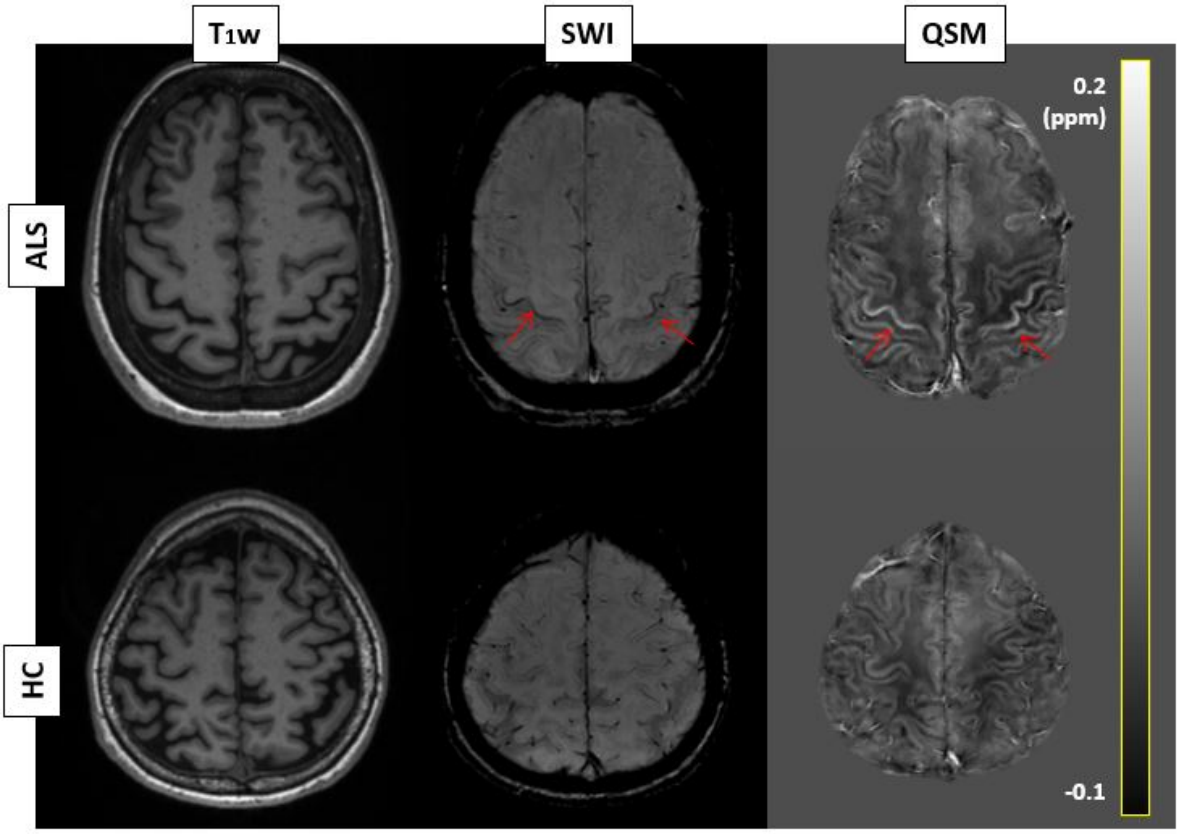

Figure 2.2: T_1w MPRAGE, SWI and QSM from F/55 years old ALS patient and F/53 years old healthy control (protocol and processing details are in Sec. 1.2.1). Iron accumulation in left and right PrCG in ALS patient (red arrows) appears hypo-intense in SWI and hyper-intense in QSM. In QSM images the gray-level scale is shown since it is a quantitative map; T_1w MPRAGE and SWI are in arbitrary units (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit, IT)

2.2 Materials and methods

2.2.1 Study sample and clinical evaluation

Between February 2020 and January 2022, 47 consecutive patients with a diagnosis of ALS (39 classic, 2 PUMN, 6 PLMN), according to the revised El Escorial criteria[80], were recruited by Clinical Neurologica at IRCCS Istituto delle Scienze Neurologiche di Bologna. All participants provided written informed consent for their involvement in the study. Brain MRI data from healthy volunteers were retrospectively collected from the Neuroimaging Laboratory's database at IRCCS Istituto delle Scienze Neurologiche di Bologna, serving both clinical and research functions. Additionally, a second group of 25 patients was included in the study, mostly subjects who could potentially be misclassified as ALS patients, known as MND mimics and chameleons[91]. Demographic characteristics of the participants are summarized in Tab. 2.1. OND patients were categorized as shown in Tab. 2.2.

The entire ALS group consisted of 47 subjects; however, for comparative analysis with HC (20 subjects) and OND (25 subjects), a subgroup of 26 subjects (22 classic, 1 PUMN, 3 PLMN) was considered. The complete ALS cohort was utilized for intra-group measurements.

	ALS (n=47)	sub-ALS (n=26)	OND (n =25)	HC (n=20)	*p-values*
F:M	18:29	13:13	12:13	13:07	.474
Age (y)	65.2 ± 12.1 (36-83)	59.5 ± 8.8 (40-69)	53.9 ± 9.2 (40-69)	56.5 ± 7.2 (41-65)	.123
DD (mo)	16.8 ± 13.5 (1-48)	18.4 ± 12.1 (1-46)	*na*	*na*	*na*
ALSFRS	38.7 ± 7.4 (13.5-47.0)	39.0 ± 7.4 (13.5-47.0)	*na*	*na*	*na*
Penn UMN	11.7 ± 6.5 (1.0-27.0)	13.5 ± 6.6 (1.0-27.0)	*na*	*na*	*na*

Table 2.1: Study demographic details; sex, age, diagnostic delay (DD), ALSFRS and Penn UMN scores are given. Where appropriate, values are given as mean ± standard deviation, with ranges in parenthesis. *p-values* are from the comparison between sub-ALS, other neurodegenerative disease (OND) and HC

	F:M	Age (y)
OND (n=25)	12:13	53.9 ± 9.2 (40-69)
HSP (n=11)	7:4	52.6 ± 8.9 (40-65)
PD (n=4)	2:2	54.4 ± 8.1 (47-67)
Parkinsonism (n=3)	1:2	57.6 ± 12.9 (43-69)
RBD (n=5)	1:4	54.5 ± 9.3 (43-64)
Hyperkinetic disorder (n=1)	1:0	40
Postural tremor (n=1)	0:1	65

Table 2.2: Age and sex of patients with other neurodegenerative diseases (OND), with details for the different pathologies (HSP = Hereditary Spastic Paraplegia, PD = Parkinson's Disease, RBD = REM Behavior Disorder)

The clinical evaluation included:

- ALSFRS (ALS Functional Rating Scale) score[92], clinical rating scale used to assess the functional status and progression of ALS. This scale consists of a series of questions that cover various aspects of a patient's daily functioning, such as speech, swallowing, fine motor skills, walking, and breathing. Each question is rated on a scale from 0 to 4, with higher scores indicating better function; individual scores for each question are summed to calculate the overall ALSFRS score, which can range from 0 (indicating complete loss of function in all areas) to 48 (indicating normal function in all areas)
- Penn UMN score[93], reliable instrument to measure UMN signs. It is a 28-item scale ranging from 0 to 32, divided into bulbar sub-score (e.g. jaw-jerk reflex and palmomental sign), upper extremity sub-score (e.g. Hoffman's sign and spasticity) and lower extremity sub-score (e.g. increased patellar reflex and Babinski's sign)

The rate of progression was calculated as $RP = \frac{48-ALFRS}{DD}$ [94], where DD is the diagnostic delay.

2.2.2 MR protocol: radiological and quantitative evaluation

MR protocol included MPRAGE, FLAIR, SWI and QSM sequences; details of acquisition and processing, as the description of software and toolboxes used, are in Sec. 1.2.1. Images did not present massive movement artifacts and were considered suitable for the analysis.

Conventional neuroradiological signs from FLAIR and SWI images were evaluated, following standard scores[81], in ALS patients and healthy controls from three neuroradiologists with 2, 5, and 15 years of experience. They were blinded to the subjects' clinical condition. For FLAIR hyperintensity in CST and SWI hypo-intensity in PrCG scores were assigned as follows: 0 = absent, 1 = present, mild and 2 = present, marked. An example is shown in Fig. 2.1.

To extract quantitative measurements, MPRAGE images were automatically segmented using FreeSurfer[48] and QSM images were linearly registered to their corresponding MPRAGE using FLIRT[44][45], to overlay the segmentation on susceptibility maps. The cortical parcellation process from FreeSurfer identified a total of 35 cortical gyri; however, not all of these areas were suitable for analysis in QSM images due to their overlap with areas prone to air-tissue and bone artifacts commonly found in QSM images[59][95]. Therefore, 10 specific cortical gyri were selected for this study: PrCG, Caudal Middle Frontal Gyrus (CMFG), Paracentral Gyrus (PCG), and Pars Triangularis (PT) in the frontal lobe; PostCentral Gyrus (PoCG), Inferior Parietal Gyrus (IPG), and SupraMarginal Gyrus (SMG) in the parietal lobe; Transverse Temporal Gyrus (TTG) in the temporal lobe; and Posterior Cingulate Cortex (PCC) and Isthmus of the Cingulate Gyrus (ICG) in the cingulate cortex. For each, various χ histogram metrics (mean, median, 90^{th} percentile, skewness, and kurtosis) as well as measurements of volume and cortical thickness were extracted. Acquisition and processing workflow is illustrated in Fig. 2.3.

2.2.3 Statistical analysis

For the assessment of radiological signs, the intraclass correlation coefficient (ICC)[96] was calculated to measure the agreement among the three radiological evaluations, ranging from 0 (= complete disagreement) to 1 (= perfect agreement among raters). The clinical accuracy of FLAIR and SWI was evaluated using the area under the Receiver Operating Characteristic (ROC) (AUC) curve[97].

For the quantitative analysis, mean, median, 90^{th} percentile[98], skewness and kurtosis[84][85] were considered as histogram χ properties, volume and cortical thickness as structural properties, for each gyrus, averaging left and right sides. To account for individual variations in head size, volume and cortical thickness were adjusted for Total Intracranial Volume (TIV) using proportional and residual methods respectively[99]. χ values of ALS and OND groups were age-corrected, assuming linear growth based on data from healthy controls.

Volume-of-interest (VOI)-based analysis was conducted for each property. To investigate potential correlations between χ and volume, Spearman's test was employed, and no significant correlations ($p\text{-}values > .05$) were found. Data were assessed for normality using the One-Sample Kolmogorov-Smirnov test and were found to be non-normally distributed. Therefore, non-parametric Kruskal-

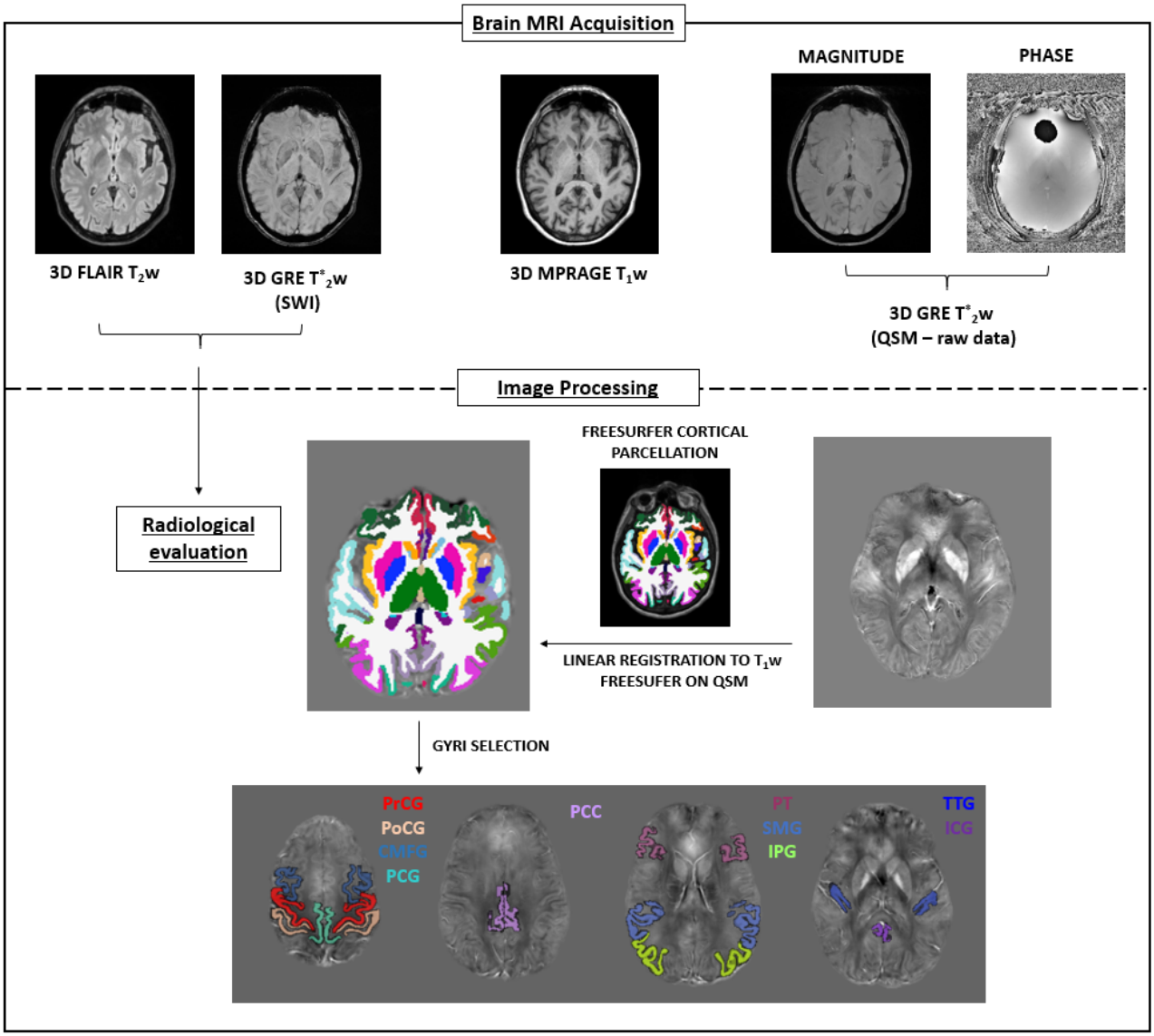

Figure 2.3: Workflow of the MR protocol acquisition and processing. Acquisition and pre-processing steps are detailed in Sec. 1.2.1. MR protocol included MPRAGE and QSM images, linearly registered to the corresponding MPRAGE. The automatic cortical parcellation from FreeSurfer was overlaid to QSM images. Ten cortical gyri were considered for the analysis (G = Gyrus, PrCG = PreCentral G, CMFG = Caudal Middle Frontal G, PCG = ParaCentral G, PT = Pars Triangularis, PoCG = PostCentral G, IPG = Inferior Parietal G, SMG = SupraMarginal G; TTG = Transverse Temporal G, PCC = Posterior Cingulate Cortex, ICG = Isthmus of the Cingulate Cortex)

Wallis was selected for comparison, followed by *post-hoc* multiple comparisons with Bonferroni correction, considering *p-values* $< .005$ as statistically significant.

Furthermore, ROC curve analysis was performed to evaluate the clinical accuracy of the measures under consideration. For correlation measurements, Spearman's test was used and *p-values* $< .05$ were considered as indicative of statistical significance.

The analysis was conducted using Matlab (version 2020b).

2.3 Results

2.3.1 Radiological evaluation

The degree of agreement among the three neuroradiologists was carefully evaluated, to ensure the consistency and reliability of the evaluation made by each rater. An ICC value of 0.74 was obtained for the evaluation of PrCG hypo-intensity in SWI images, indicating a relatively strong agreement among the raters. For CST hyper-intensity in FLAIR images, the ICC value was 0.65, demonstrating a moderately strong level of agreement.
For SWI measurements, the AUC was 0.547, with 35.3% of the findings false positives and 36.5% false negatives; similarly, for FLAIR measurements, the AUC was 0.543, with 42.9% of the results false positives, and 23.4% false negatives. These outcomes show that both SWI and FLAIR have issues in distinguishing ALS patients from healthy controls: while the inter-rater agreement was reasonably strong, the diagnostic accuracy of these imaging modalities, as indicated by the AUC values and false positive/negative rates, suggests room for improvement and the need for further investigation or the consideration of complementary diagnostic tools. This result confirms the importance of utilizing quantitative information from QSM measurements for better disease characterization.

2.3.2 χ and structural properties

Inter-group comparison

χ and structural characteristics were compared among three groups: sub-ALS, OND, and HC (Tab. 2.1). Histograms of the different metrics (mean, median, 90^{th} percentile, and skewness of χ, volume and cortical thickness) in the 10 gyri are depicted in Fig. 2.4. The analysis revealed several noteworthy differences in the assessed parameters between the HC and ALS groups, summarized in Tab. 2.3. Specifically, the following differences in the median χ values ($\pm$ mean absolute deviation) (ppm) between the HC and ALS groups were observed:

- mean χ in the frontal lobe (precentral gyrus [HC: -0.023 $\pm$ 0.011; ALS: -0.007 $\pm$ 0.010; p = .002], caudal middle frontal gyrus [HC: -0.022 $\pm$ 0.009; ALS: -0.012 $\pm$ 0.007; p = .001], pars triangularis [HC: -0.015 $\pm$ 0.012; ALS: 0.005 $\pm$ 0.011; p = .001]), parietal lobe (inferior parietal [HC: -0.012 $\pm$ 0.012; ALS: 0.002 $\pm$ 0.010; p = .004] and supramarginal gyri [HC: -0.014 $\pm$ 0.011; ALS: 0.006 $\pm$ 0.010; p = .002]), cingulate cortex (isthmus [HC: -0.010 $\pm$ 0.013; ALS: 0.006 $\pm$ 0.009; p = .004]), higher in ALS group;
- median χ in the frontal lobe (precentral gyrus [HC: -0.022 $\pm$ 0.011; ALS: -0.007 $\pm$ 0.010; p = .004], caudal middle frontal gyrus [HC: -0.021 $\pm$ 0.009; ALS: -0.010 $\pm$ 0.007; p = .001], pars triangularis [HC: -0.014 $\pm$ 0.012; ALS: 0.005 $\pm$ 0.010; p = .001], parietal lobe [supramarginal gyri (HC: -0.013 $\pm$ 0.011; ALS: 0.007 $\pm$ 0.010; p = .002]), cingulate cortex (posterior-cingulate [HC: -0.037 $\pm$ 0.010; ALS: -0.014 $\pm$ 0.012; p = .003] and isthmus [HC: -0.014 $\pm$ 0.014; ALS: 0.007 $\pm$ 0.009; p = .002]), higher in ALS group;

- 90^{th} percentile χ in the frontal lobe (precentral gyrus [HC: 0.003 ± 0.014; ALS: 0.024 ± 0.011; $p = .004$], caudal middle frontal [HC: 0.00002 ± 0.00001; ALS: 0.012 ± 0.008; $p = .001$] and pars triangularis [HC: 0.005 ± 0.014; ALS: 0.027 ± 0.013; $p = .002$]), higher in ALS group;

- skewness χ in the precentral gyrus (HC: -0.167 ± 0.120; ALS:-0.025 ± 0.149; $p = .004$).

Notably, kurtosis did not display any significant alterations and is therefore not presented in the figure. Volume and cortical thickness did not result significantly impaired, but a common reduction trend was observed in cortical thickness for almost all the gyri.
ROC analysis was performed and returned AUC values were higher than 0.70 for all the parameters that resulted significant from Kruskal-Wallis test, showing an acceptable diagnostic accuracy in classifying HC/ALS status. In PrCG, ROC curves for χ properties (mean, median, 90^{th} percentile and skewness) were 0.758, 0.738, 0.775 and 0.733 respectively. Comparison analysis did not show any significant difference between OND and HC or ALS groups.

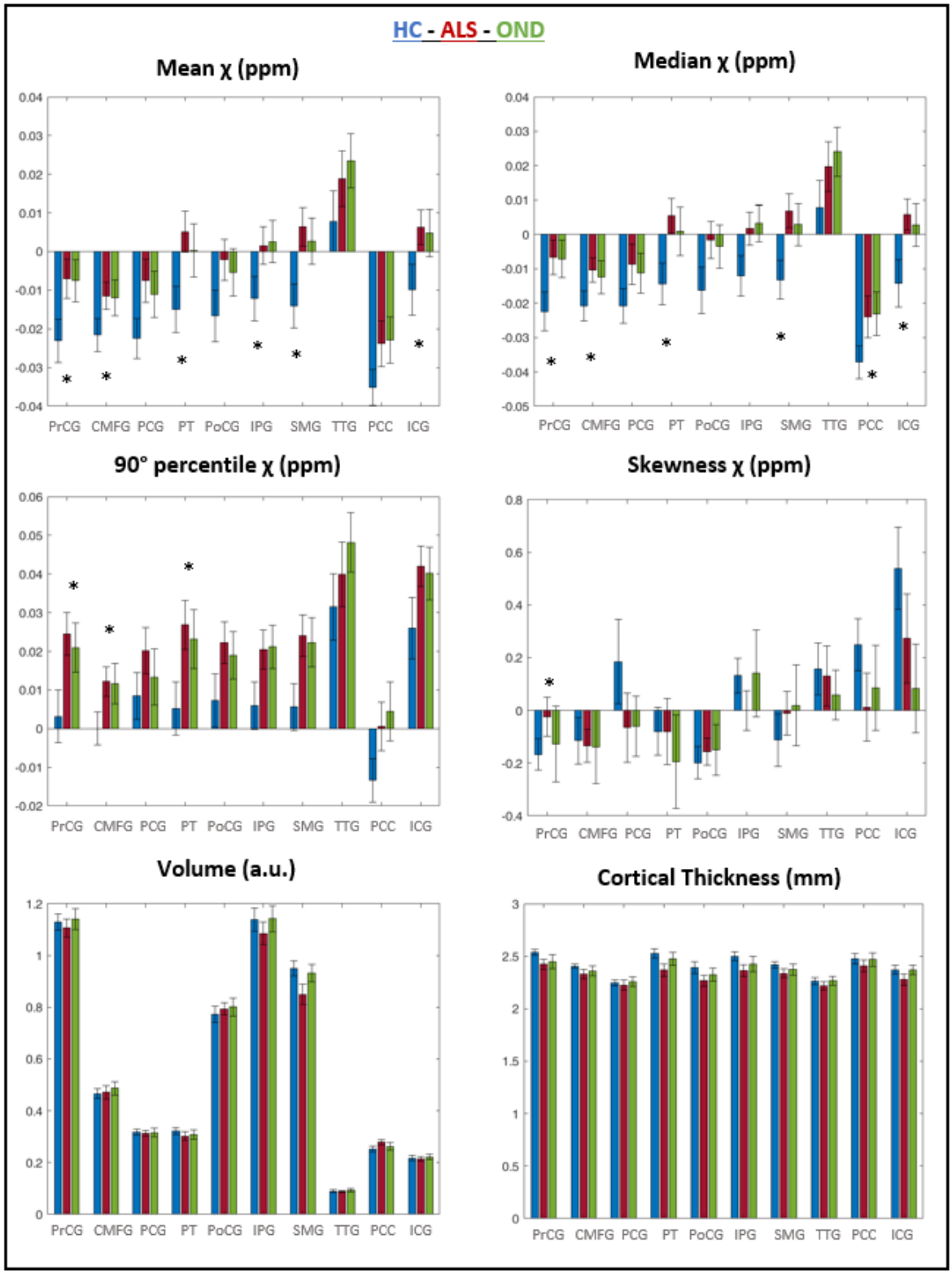

Figure 2.4: Median ± mean absolute deviation of χ (mean, median, 90^{th} percentile) and structural (volume and cortical thickness) properties within HC (blue), ALS (red) and OND (green) samples in the analyzed gyri. Significant differences between HC and ALS groups (*p-values* $< .005$, Kruskal-Wallis test, *post-hoc* multi-comparison test with Bonferroni correction) were highlighted (*). The analysis did not show any statistical differences between OND and HC or ALS groups (OND = Other Neurodegenerative Disease, G = Gyrus, PrCG= PreCentral G, CMFG = Caudal Middle Frontal G, PCG = ParaCentral G, PT = Pars Triangularis, PoCG = PostCentral G, IPG = Inferior Parietal G, SMG = SupraMarginal G, TTG = Transverse Temporal G, PCC = Posterior Cingulate Cortex, ICG = Isthmus of the Cingulate Cortex)

		χ properties					Structural properties	
		Mean	Median	90° P	Skewness	Kurtosis	Volume	Thickness
FL	PrCG	**.002 ***	**.004 ***	**.004 ***	**.004 ***	*ns*	*ns*	*ns*
	CMFG	**.001 ***	**.001 ***	**.001 ***	*ns*	*ns*	*ns*	*ns*
	PCG	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
	PT	**.001 ***	**.001 ***	**.002 ***	*ns*	*ns*	*ns*	*ns*
PL	PoCG	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
	IPG	**.004 ***	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
	SMG	**.002 ***	**.002 ***	*ns*	*ns*	*ns*	*ns*	*ns*
TL	TTG	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
CC	PCC	*ns*	**.003 ***	*ns*	*ns*	*ns*	*ns*	*ns*
	ICG	**.004 ***	**.002 ***	*ns*	*ns*	*ns*	*ns*	*ns*

Table 2.3: Comparison analysis between ALS and HC in cortical gyri. *p-values* are from Kruskal-Wallis test (*post-hoc* multi-comparison test with Bonferroni correction) from the comparison between ALS and HC. *P-values* $< .005$ were considered as significant (*) (FL = Frontal Lobe, PL = Parietal Lobe, TL = Temporal Lobe, CC = Cingulate Cortex, G = Gyrus, PrCG = PreCentral G, CMFG = Caudal Middle Frontal G, PCG = ParaCentral G, PT = Pars Triangularis, PoCG = PostCentral G, IPG = Inferior Parietal G, SMG = SupraMarginal G, TTG = Transverse Temporal G, PCC = Posterior Cingulate Cortex, ICG = Isthmus of the Cingulate Cortex)

Clinical correlation analysis

Potential correlations within the entire ALS sample (Tab. 2.1) were considered, aiming to identify any connections between χ neuroimaging measurements and clinical scores.
A significant correlation emerged between RP and cortical thickness in the precentral gyrus, with a correlation coefficient ρ of -0.302 and a corresponding *p-value* of .041, suggesting that as RP increases the cortical thickness tends to decrease. This correlation provides valuable insights into the relationship between this clinical measure and the structural characteristics of the brain. Correlations between cortical thickness and RP in the other gyri show a negative but not significant trend, indicating that the relationship between RP and cortical thickness may be specific to the motor cortex.
Another aspect of this investigation aimed to establish a connection between iron accumulation (assessed through mean, median, and 90^{th} percentile χ) and UMN impairment, measured by the Penn UMN score. Although positive correlations were observed between susceptibility properties and the Penn UMN score, these correlations did not reach statistical significance. However, these trends suggest that there may be some association between iron accumulation and UMN dysfunction that warrants further exploration.

2.3.3 Progression rate: slow, intermediate and fast

One of the aims of the study was to gain insights into potential distinct behaviors related to the speed at which ALS progresses. The ALS patient group was divided into three subgroups based on the rate of disease progression: slow, intermediate, and fast progression; 25^{th} (RP = 0.25) and the 75^{th} (RP = 1.30) were used as cut-off. Demographic and clinical characteristics of the three

subgroups are summarized in Tab. 2.4. They are consistent with each other in terms of sex, age, clinical phenotype and Penn UMN score; this consistency is crucial to ensure that the eventual observed differences are primarily related to the progression rate rather than other demographic or clinical factors.

The three subgroups are different in terms of disease delay and ALSFRS: patients in the fast progression subgroup had a notably shorter disease duration compared to those in the other subgroups - this finding aligns with expectations, as faster disease progression typically leads to a shorter duration of illness - and the ALSFRS score was lower in patients with fast progression, as anticipated, reflecting greater functional impairment, which is often associated with more rapid disease progression.

	Slow RP (n=12) (RP>0.25)	**Int RP** (n=23) (0.25<RP<1.30)	**Fast RP** (n=12) (RP >1.30)	*p-value*
F:M	3:9	10:13	5:7	.544
Age (y)	68.2 ± 6.2 (47-77)	65.7 ± 11.9 (36-82)	65.3 ± 13.3 (33-83)	.945
Cl ph (class:pumn:plmn)	8:1:3	21:0:2	1:1:1	.357
DD (mo)	25.3 ± 13.4 (4-46)	17.0 ± 11.8 (5-48)	8.0 ± 3.4 (1-13)	.003
ALSFRS	44.3 ± 2.3 (41.0-47.0)	39.5 ± 4.6 (26.0-44.0)	31.0 ± 8.8 (13.5-44.0)	<.000
Penn UMN	11.1 ± 6.3 (1.0-21.0)	10.4 ± 6.1 (1.0-21.0)	15.0 ± 6.9 (1.0-27.0)	.161

Table 2.4: Demographic details of ALS sub-groups (slow, intermediate and fast progression rate). Where appropriate, values are given as mean ± standard deviation, with ranges in parenthesis. *p-values* are from the comparison between slow, intermediate and fast progression

χ and structural properties: inter-group comparison

The analysis of χ properties did not reveal any statistically significant differences or discernible trends when comparing the three ALS subgroups. These results suggest that the susceptibility measurements represented by χ properties did not exhibit any systematic variations related to the different rates of disease progression among the subgroups.

Similarly, the analysis of structural properties did not yield any significant differences among the subgroups, although it shows a clear trend in the context of atrophy progression: as RP increased, there appeared to be a noticeable trend towards greater atrophy corresponding to thinner cortical thickness, which can be observed in Fig. 2.5. This observation suggests that as the disease progresses more rapidly, there may be a corresponding increase in the degree of brain atrophy. While these trends did not reach statistical significance, they do hint at potential associations between disease progression speed and structural changes in the brain.

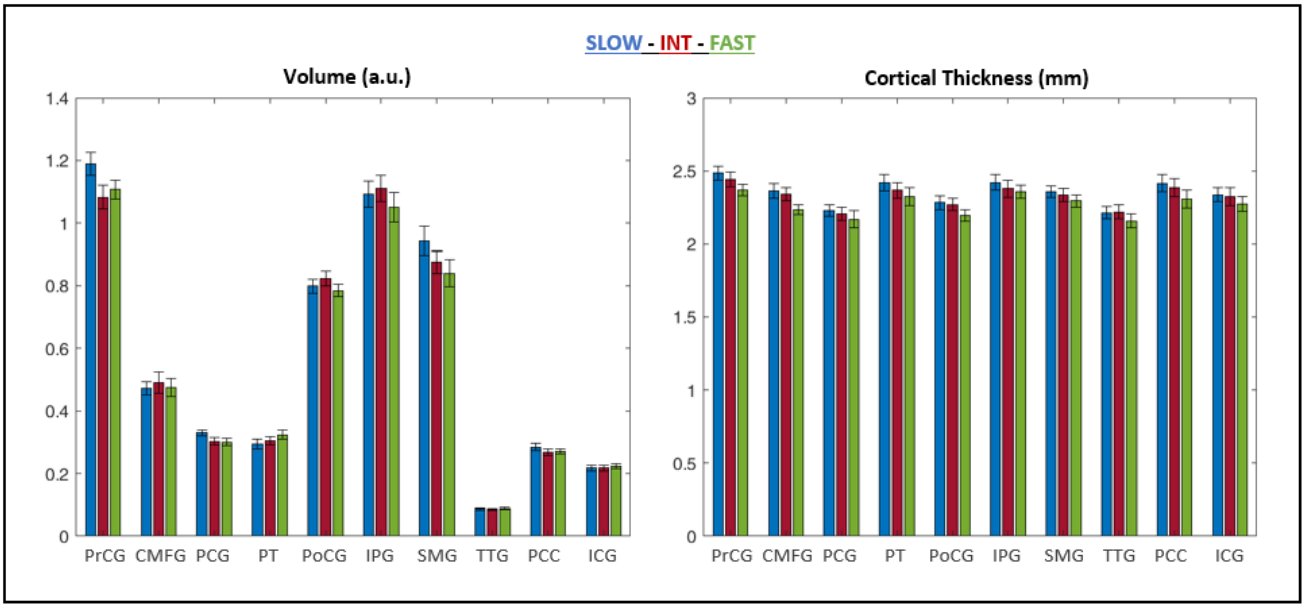

Figure 2.5: Volume and cortical thickness in ALS patients with slow (blue), intermediate (red) and fast (green) progression. Median ± MAD are shown in the analyzed gyri (G = Gyrus, PrCG = PreCentral G, CMFG = Caudal Middle Frontal G, PCG = ParaCentral G, PT = Pars Triangularis, PoCG = PostCentral G, IPG = Inferior Parietal G, SMG = SupraMarginal G, TTG = Transverse Temporal, PCG = Posterior Cingulate G, ICG = Isthmus of the Cingulate Cortex)

Clinical correlations analysis

Spearman's test was employed independently on each of the three groups to explore potential correlations between the Penn UMN score and susceptibility properties. The results revealed noteworthy findings in the group characterized by a faster rate of disease progression: within this subgroup, significant correlations were observed between iron accumulation (measured through mean, median, and 90^{th} percentile χ values) and the Penn UMN score. Findings are outlined in detail in Tab. 2.5. These significant correlations suggest a relationship between the degree of iron accumulation in the brain and the severity of UMN impairment, as measured by the Penn UMN score. This connection is particularly prominent in individuals with a faster rate of disease progression, highlighting the clinical relevance of iron accumulation in the context of ALS progression.

		Mean	Median	90^{th} P
FL	PrCG	*ns*	**(0.623, .041) ***	**(0.607, .482) ***
	CMFG	*ns*	*ns*	*ns*
	PCG	**(0.728, .011) ***	**(0.816, .002) ***	*ns*
	PT	**(0.843, .001) ***	**(0.780, .002) ***	*ns*
PL	PoCG	**(0.651, .030) ***	**(0.606, .048) ***	*ns*
	IPG	*ns*	*ns*	*ns*
	SMG	**(0.628, .039) ***	**(0.628, .039) ***	*ns*
TL	TTG	*ns*	*ns*	*ns*
CC	PCC	**(0.628, .038) ***	**(0.668, .025) ***	*ns*
	ICG	*ns*	*ns*	*ns*

Table 2.5: Correlations between susceptibility properties and Penn UMN score in ALS patients with fast disease progression. Correlation coefficients ρ with corresponding *p-values* from Spearman's test to find correlations between susceptibility properties (mean, median and 90^{th} percentile χ) and Penn UMN Score. *P-values* $< .05$ were considered significant. Positive correlations between Penn UMN Score and those susceptibility properties (mean, median and 90^{th} percentile) mean that iron accumulation corresponds to more severe impairment. (G = Gyrus, PrCG = PreCentral G, CMFG = Caudal Middle Frontal G, PCG = ParaCentral G, PT = Pars Triangularis, PoCG = PostCentral G, IPG = Inferior Parietal G, SMG = SupraMarginal G, TTG = Transverse Temporal, PCC = Posterior Cingulate Cortex, ICG = Isthmus of the Cingulate Cortex, UMN = Upper Motor Neuron)

2.4 Discussion

ALS is a neurodegenerative disorder occurring with various clinical manifestations, making it challenging to gain a deep understanding of the disease and to identify reliable *in vivo* imaging markers for effective categorization and prognosis prediction [78] [79]. Qualitative observations in brain MRI scans, such as hyper-intensity in the corticospinal tract in T_2-weighted images and hypo-intensity in the motor cortex in SWI, are more frequently observed in ALS patients compared to healthy individuals [100] [81]. However, these signs lack high specificity for ALS diagnosis [83]. The introduction of QSM has significantly advanced the ability to investigate the role of iron accumulation in the motor cortex during the progression of ALS, paving the way for a deeper understanding of the underlying pathology [101] [74].

In this study, a comprehensive exploration of various aspects of magnetic susceptibility distributions in both motor and extra-motor cortical regions was conducted. A group of ALS patients was compared with both healthy controls and patients with other neurodegenerative disorders. Additionally, structural characteristics, specifically volume and cortical thickness, were incorporated into the analysis. Results are discussed below, highlighting the key points of the study.

Qualitative *vs* quantitative evaluation

First, radiological assessment of both the motor cortex and the corticospinal tract in a sample mix of ALS patients and healthy controls was conducted blindly by three different neuroradiologists. The inter-rater agreement among them demonstrated good reliability, as evidenced by ICC of 0.74 (SWI) and 0.65 (FLAIR). However, when it came to their diagnostic accuracy in distinguishing ALS patients from controls, their performance was found to be lacking, with AUC values of 0.547

and 0.543, respectively [83]. QSM analysis was employed to examine the magnetic susceptibility of the precentral gyrus: quantitative measurements provided more valuable insights. Notably, there was a significant increase in both mean and median susceptibility observed in the ALS group when compared to healthy controls (Fig. 2.4, Tab. 2.3). This increase in susceptibility is indicative of higher iron concentrations in the motor cortex of ALS patients, as previously reported [81] [73]. Furthermore, the diagnostic accuracy of M1 susceptibility surpassed the one of the qualitative "motor band sign" with AUC values of 0.758 and 0.738 for mean and median susceptibility, respectively, underscoring the superiority of quantitative analysis in this context. Reduction in both volume and cortical thickness in the precentral gyrus was observed, consistently with prior research [87].

M1 iron accumulation: multiple χ properties

It has been established that iron accumulation in ALS primarily affects the deep layers of the motor cortex [76] [82] [102] and specific regions linked to impaired functions during the progression of the disease [101]. Additionally, iron accumulation is a natural part of the aging process [103], occurring in both healthy individuals and ALS patients, which can lead to increased susceptibility values. This complexity may result in misinterpretation when analyzing M1 mean/median susceptibility, particularly in elderly cohorts. To address these complexities, it was appropriate to focus on multiple properties of the susceptibility histogram [86] [84], specifically (Fig. 2.4, Tab. 2.3): the 90^{th} percentile, holding information related to the right part of the histogram and so to high-iron content regions – high percentiles have been studied to explore iron distribution in the substantia nigra during the development of Parkinson's Disease (PD) [98] [104] -, skewness and kurtosis [84] [85] [86]. According to previous studies, differences were observed in PrCG skewness, higher in ALS patients than controls. The 90^{th} percentile also showed a strongly significant difference and was the parameter with the highest diagnostic accuracy in our sample (AUC = 0.775).

ALS iron accumulation beyond the motor cortex

ALS was initially categorized as primarily affecting the neuromuscular domain. However, recent advances in imaging and neuropathological assessments have revealed the complexity of ALS pathophysiology beyond motor impairment [72]. Given the association between increased susceptibility and iron accumulation in the motor cortex, I investigated whether this phenomenon extended to other cortical regions. The comparisons between ALS and control groups (Fig. 2.4, Tab. 2.3) revealed increased susceptibility, measured in terms of mean and median, not only in the primary motor cortex but also in other cortical areas, including the caudal-middle frontal gyrus, pars triangularis in the frontal lobe, inferior parietal gyrus, supramarginal gyrus in the parietal lobe, and the posterior cingulate cortex. Analyzing multiple features from the susceptibility histogram allowed us to pinpoint areas of greater involvement. The 90^{th} percentile, in particular, highlighted significant differences, predominantly in the frontal lobe, while susceptibility skewness exhibited significant differences in the precentral gyrus. Following previous studies [87], a trend of reduction was observed in cortical thickness in motor and extra-motor gyri.

ALS *vs* other neurodegenerative disease

One of the objectives of this study was to explore susceptibility properties when comparing ALS and patients with other neurodegenerative diseases, which remains challenging, particularly in the early stages of the disease. Any significant difference was identified between ALS patients and those with other neurodegenerative conditions (Fig. 2.4), suggesting that the process of iron accumulation extends to multiple cortical areas in ALS development and may also be associated with other pathological conditions. However, the comparison among the three groups (HC, ALS, OND) showed significant differences between ALS and HC, but not between OND and HC; it's interesting to note that OND had lower susceptibility values in terms of mean, median, and 90^{th} percentile compared to ALS patients almost everywhere. Additionally, a consistent trend of cortical thickness reduction occurs across all gyri, with controls exhibiting the highest thickness, ALS patients having the lowest, and other patients falling in between.

Further investigation, perhaps focusing on advanced susceptibility features, could yield valuable biomarkers for disease characterization.

χ measurements and clinical data

Quantitative measurements as susceptibility properties are feasible to be correlated with clinical scores; considering heterogeneity in phenotypes and clinical conditions, previous studies have not shown agreement in linking susceptibility and clinical data[74]. Here, susceptibility measurements did not show any significant relationship/trend with Penn UMN Score and rate of progression, defined considering ALS-FRS score and the diagnostic delay[94]. A significant negative correlation resulted between the rate of progression and the cortical thickness in the motor cortex, which led us to split the group into three sub-populations, with slow, intermediate and fast progression (Tab. 2.4), with ALS-FRS and Penn UMN Score worse with the increase of the progression.

Within the three groups, susceptibility and structural measurements were compared. Susceptibility properties did not highlight any significant difference or trend, suggesting that iron accumulation is a process occurring regardless of the speed of the progression. Cortical thickness showed an interesting trend of decreasing from slow to fast progression in the analyzed gyri (Fig. 2.5).

Zooming in on the three groups individually, there were significant correlations between mean, median and 90^{th} percentile and Penn UMN score in multiple cortical gyri, including the precentral gyrus, only in ALS patients with fast progression rate (Tab. 2.5). The outcome suggests that, with the rapid evolution of the disease, iron concentration grows as the rate of progression. This result has to be strengthened by expanding the sample.

2.5 Conclusions

This study delves into the potential role of iron accumulation, which can be detected *in vivo* through QSM, as a potential hallmark of ALS, not only in the motor cortex but also in extra-motor cortical regions. The research uncovers that, apart from the motor cortex, multiple cortical gyri exhibit a significant increase in iron levels in ALS patients with respect to control subjects. Furthermore, ALS patients display a pattern of increasing iron content and decreasing cortical thickness when compared to individuals affected by other neurodegenerative disorders. These findings contribute to the characterization of ALS as a disease. While increased cortical iron concentration is a phenomenon that occurs in the majority of neurodegenerative conditions, this study suggests that it may hold particular relevance to ALS.

A range of susceptibility-based descriptors - namely mean, median, 90^{th}, skewness and kurtosis - were used in the investigation, to capture nuances in iron distribution patterns. They revealed significant variations, particularly in frontal brain regions.

By studying the clinical profiles of the participants, a noteworthy correlation between Penn UMN scores and susceptibility values across multiple cortical gyri was established, specifically in patients with rapid disease progression. This discovery underscores the potential of QSM as a valuable tool in the identification of biomarkers for ALS, particularly in cases characterized by fast disease progression.

χ = magnetic susceptibility
ALS = Amyotrophic Lateral Sclerosis
ALSFRS = ALS Functional Rate Scale
AUC = Area Under the ROC Curve
CMFG = Caudal Middle Frontal
CST = Cortico-Spinal Tract
DD = Diagnostic Delay
FLAIR = FLuid Attenuate Inversion Recovery
HC = Healthy Control
HSP = Hereditary Spastic Paraplegia
ICC = Intra-class Correlation Coefficient
ICG = Isthmus of the Cingulate Gyrus
IPG = Inferior Parietal Gyrus
LMN = Lower Motor Neuron
M1 = primary motor cortex
MND = Motor Neuron Disease
MPRAGE = Magnetization Prepared RApid Gradient Echo
MR = Magnetic Resonance
MRI = Magnetic Resonance Imaging
OND = Other Neurodegenerative Diseases
PCC = Posterior Cingulate Cortex
PCG = Paracentral Gyrus
PD = Parkinson's Disease
PLMN = Pure Lower Motor Neuron
PoCG = PostCentral Gyrus
PrCG = PreCentral Gyrus
PT = Pars Triangularis
PUMN = Pure Upper Motor Neuron
RBD = REM sleep Behavior Disorder
ROC = Receiver Operating Characteristic
ROI = Region of Interest
RP = Rate of Progression
SMG = SupraMarginal Gyrus
SWI = Susceptibility Weighted Imaging
T_1w, T_2w, T_2^*w = T_1-, T_2-, T_2^*- weighted
TIV = Total Intracranial Volume
TTG = Transverse Temporal Gyrus
UMN = Upper Motor Neuron
VOI = Volume of Interest

Chapter 3

Myotonic Dystrophy Type 1

Myotonic Dystrophy type 1 (DM1) is a dominantly inherited disorder, affecting the musculoskeletal and central nervous systems and causing a variety of symptoms including muscle weakness and cognitive deficits. Structural, diffusion, spectroscopy and functional MR measurements have shown alterations in DM1 brains[105]; to date, only one study has examined susceptibility and identified alterations in the thalamus[106]. Here, I explored iron concentration distribution in cortical and subcortical brain areas through QSM, linking imaging measurements with disability scores, pneumological and cardiological evaluations and sleep records data from polysomnography. Background and aims are described in Sec. 3.1.

Materials and methods are in Sec. 3.2. 34 DM1 participants and 35 age- and sex- matched controls went under MRI examinations, including 3D morphological T_1w and QSM sequences. Cortical and subcortical structures were automatically segmented, χ distributions within the two groups were compared and DM1 χ values were correlated with clinical and laboratory data.
Results are described in Sec. 3.3 and 3.4. Several cortical gyri showed a significant increase of χ in the patient group, as a non-specific sign of neurodegeneration. Among subcortical structures, χ of the thalamus (specifically, ventral and pulvinar nuclei) and brainstem (specifically, pons and medulla) was significantly higher in DM1 group than HC, and negatively correlated with the onset age, linking higher iron accumulation to earlier onset. Additionally, thalamic χ values significantly correlated with disability scores and cardiological parameters; χ in both the thalamus and brainstem correlated with the number of central apneas. The outcomes show a progressive impairment of thalamus and brainstem, core of autonomic functions, during the development of DM1 disease. Different methods to assess the origins of DM1 symptomatology, if central or peripheral, have recently been developed, and QSM may help in shedding light on this debate (Sec. 3.5).

The first use of abbreviations includes their full expansion; for ease of reference, a glossary is provided at the end of the chapter.

3.1 Background and aims

Myotonic Dystrophy type 1 (DM1) is one of the most frequent forms of muscular dystrophy in adults with an estimated prevalence of 1/8000, described with clarity for the first time in 1909 by Steinert and his colleagues - hence the name Steinert's disease[107][108]. It is an autosomal dominant disease, caused by the unstable expansion of trinucleotide cytosine-thymine-guanine (CTG) repeats in the 3 untranslated region (UTR) of the myotonic dystrophy protein kinase (DMPK), coding for a myosin kinase expressed in skeletal muscle, leading to an accumulation of mutant RNA aggregates[109].

DM1 patients may be grouped with respect to the onset age (congenital, childhood/juvenile, adult and late-onset/asymptomatic)[110] as well as the number of CTG repeats; in this work, I am going to refer to the genetic classification[111]. Healthy individuals have between 5 and 37 CTG repeats, while 38-50 repeats are considered permutation alleles; with more than 50 CTG repeats, patients are classified as follows, with CTG repeats predictive of clinical severity and onset age: mild (E1 class), with 50-150 CTG-repeat; classic (E2 class), with 50-1000 CTG-repeats; congenital (E3 class), with $>$ 1000 CTG-repeats. The variability of genetic mutations, which may increase or decrease in size as occurs in different tissues, causes a wide spectrum of possible muscular and extra-muscular symptoms[112][109]:

- **Muscular impairment**: muscle weakness and atrophy, especially in distal regions of limbs, and central and peripheral fatigue often occur[21]. Myotonia is one of the main symptoms, as suggested by the name of the disease, consisting in delayed relaxation of the skeletal muscles after voluntary contraction
- **Respiratory involvement**: respiratory insufficiency and pulmonary complications are the leading cause of death in patients with DM1[113]; many mechanisms are involved, such as muscle weakness, chest mechanisms and abnormal central respiratory control (central hypoventilation)
- **Cardiovascular involvement**: it can occur as dilated cardiomyopathy, heart failure, conduction and myocardial abnormalities[114]. Both respiratory and cardiovascular symptoms can be explained by muscular or Autonomic Nervous System (ANS) dysfunction (central or peripheral)
- **CNS involvement**: additionally, Central Nervous System (CNS) dysfunctions are observed[115]. DM1 is the neuromuscular disease with the most prominent sleep disorders, including excessive daytime sleepiness (EDS), sleep breathing disorders (nocturnal hypoxemia, nocturnal hypoventilation, obstructive apnoeas and hypopnoeas, central apneas), periodic leg movements during sleep and rapid eye movement sleep dysregulations. Cognitive impairment and psychiatric disorders occur[116][117], as well as the involvement of neuropsychological functioning regarding attention, executive functions, visuo-spatial and constructive abilities.

To improve the understanding of DM1 pathogenesis, particularly regarding CNS, MR brain imaging techniques have been used to detect abnormalities in DM1 brains[105], confirming clinical, molecular and histopathological brain alterations. Generalized brain atrophy was found, together with gray matter (GM) volume reduction within the cortex (frontal, temporal, parietal and occipital cortex)[118] and deep structures[119]; subcortical and periventricular white matter (WM) hyperintensities increased in DM1 patients with respect to healthy controls (HC) (Fig. 3.1)[120] and microstructural parameters from diffusion measurements show widespread alterations in normal appearing WM[121]. Other techniques have been used, such as functional MRI[122] and MR spectroscopy[123], supporting findings in alterations in DM1 brains.

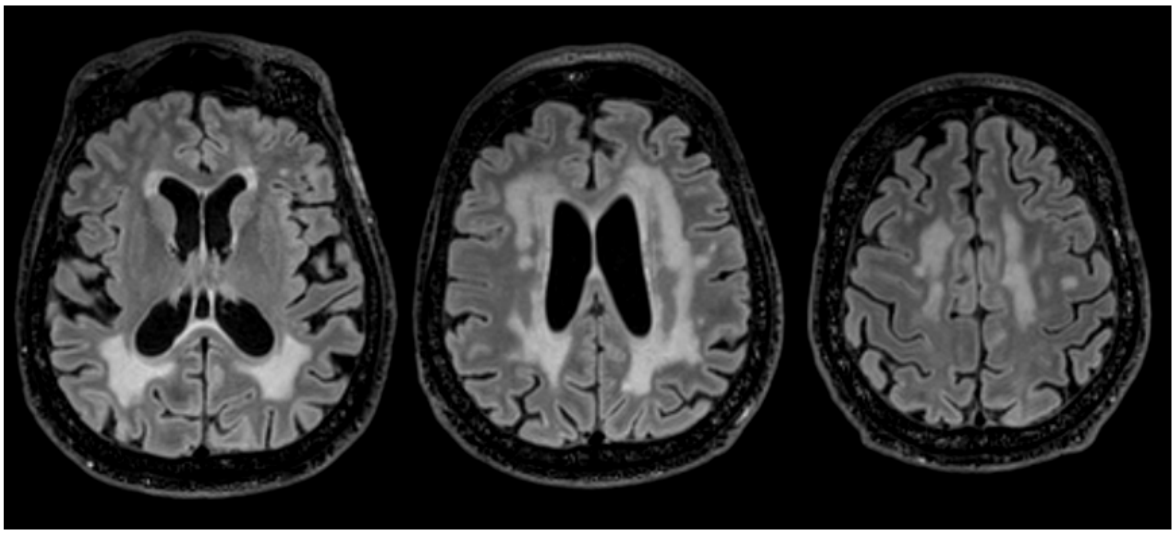

Figure 3.1: White matter hyper-intensities (T_2w 3D FLAIR) (DM1, E2 genetic class, M/57 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit, Ospedale Bellaria, Bologna)

As in[106], the purpose of this work[124] is to explore the role of iron accumulation in DM1 development using QSM, both cortical and subcortical structures, aiming to find more specific biomarkers to characterize the disease and distinguish different genetic classes, helping the understanding of DM1 pathogenesis. The analysis proposed is original, as analyzing multiple brain structures within a sample homogeneous in terms of phenotype abundance and with a wide clinical characterization.

3.2 Materials and methods

3.2.1 Study sample and clinical evaluation

34 DM1 patients participated, recruited in a prospective study at the Neuromuscular Centre of the UOC Neurological Clinic of the IRCCS Istituto delle Scienze Neurologiche di Bologna. MR and clinical examinations were performed between May and November 2021; all patients signed informed consent. The patient cohort included E1 and E2 genetic classes; originally, the sample included other three patients (1 E1 and 2 E3) which were excluded because was not possible to include QSM sequence in the MR protocol. Data were collected from the database of the

Neuroimaging Laboratory (IRCCS Istituto delle Scienze Neurologiche di Bologna). 35 age- and sex- healthy controls were included for comparisons; MR exams were performed between February 2020 and June 2022, using the same protocol. Demographic characteristics are shown in Tab. 3.1, with details about E1 and E2 genetic classes in Tab. 3.2.

	DM1 (n=34)	**HC (n=35)**	***p-values***
F:M	20:14	20:15	.888
Age (y)	46.8 ± 12.0 (20-71)	50.5 ± 17.4 (24-86)	.230
Onset age (y)	30.0 ± 15.7 (4-68)	*na*	*na*
DD (y)	16.8 ± 9.7 (0.6 - 36.1)	*na*	*na*

Table 3.1: Main characteristics of DM1 patients and healthy controls. Where appropriate, values are given as mean ± standard deviation, with ranges in parenthesis. *p-values* are from the comparison between the two groups, considered statistically significant (*) at $p < .05$ (*na* = not available, DD = Diagnostic Delay)

	DM1-E1 (n=14)	**DM1-E2 (n=20)**	***p-values***
F:M	7:7	13:7	.389
Age (y)	50.3 ± 10.0 (28-71)	44.3 ± 12.9 (22-67)	.230
Onset age (y)	36.4 ± 14.8 (16-68)	25.6 ± 15.0 (4-55)	.044 *
DD (y)	13.9 ± 9.2 (0.6 - 29.3)	18.8 ± 9.8 (6.3 - 36.1)	.152

Table 3.2: Main characteristics of DM1 patients, divided into E1 (50-150 CTG repeats) and E2 (150-1000 CTG repeats) genetic classes. Where appropriate, values are given as mean ± standard deviation, with ranges in parenthesis. *p-values* are from the comparison between the two groups, considered statistically significant (*) at $p < .05$ (*na* = not available, DD = Diagnostic Delay)

Neurological, pneumological and cardiological evaluations, together with sleep recordings study, were performed on the patient cohort together with MR scans, allowing the combination of imaging and clinical data. More specifically[125]:

- Neurological evaluation included the Muscular Impairment Rating Scale (MIRS)[126], the Neuromuscular Impairment Function and Disability Scale (NIFDS)[127], divided into four domains (neuropsychological, motor, myotonia and daily life activity), and the Epworth Sleepiness Scale (ESS) to evaluate excessive daytime sleepiness
- Pneumological evaluation included Forced Vital Capacity (FVC), arterial blood gas analysis with recording of arterial carbon dioxide partial pressure ($PaCO_2$) and arterial oxygen partial pressure (PaO_2)
- Cardiological assessment included ECG, to evaluate conduction abnormalities (PR interval, QRS duration) and systolic (left ventricular) dysfunction, as the reduction of Ejection Fraction (EF)

- Sleep breathing disorders were studied through polysomnography; the Oxygen Desaturation Index (ODI), the Apnea Hypopnea Index (AHI) and the number of central apneas (NCA) per hour of sleep were scored according to the AASM guidelines

In Tab. 3.3, indices of clinical data distribution are reported, distinguishing E1 and E2 classes.

	DM1	**E1**	**E2**	***p-val***
MIRS ↑ (*na*)	3.15 ± 0.78	2.79 ± 0.89	3.40 ± 0.60	.022 *
NIFDS (NP) ↑ (*na*)	3.34 ± 2.70	2.91 ± 2.30	3.61 ± 2.95	.507
NIFDS (Mo) ↑ (*na*)	9.41 ± 4.01	6.82 ± 3.34	11.00 ± 3.60	.004 *
NIFDS (My) ↑ (*na*)	5.48 ± 2.71	3.91 ± 2.66	6.44 ± 2.31	.012 *
NIFDS (DL) ↑ (*na*)	2.10 ± 1.80	0.91 ± 0.94	2.83 ± 1.83	.003 **
NIFDS (tot) ↑ (*na*)	20.34 ± 8.80	14.55 ± 5.18	23.89 ± 8.75	.003 **
ESS ↑ (≥ 10)	8.25 ± 2.92	8.54 ± 2.93	8.05 ± 2.97	.651
FVC ↓ (≤ 80 %)	84.24 ± 17.80	90.29 ± 15.32	80.00 ± 18.53	.098
$PaCO_2$ ↑ (≥ 45 mmHg)	45.45 ± 4.96	43.86 ± 4.54	46.76 ± 5.01	.122
PaO_2 ↓ (≤ 80 mmHg)	85.98 ± 10.15	87.47 ± 10.63	84.92 ± 9.99	.516
PR ↑ (≥ 200 ms)	194.54 ± 34.37	190.50 ± 41.20	197.80 ± 28.90	.593
QRS ↑ (≥ 100 ms)	116.25 ± 25.15	118.58 ± 28.16	120.00 ± 22.91	.391
EF ↓ (≤ 50 %)	60.72 ± 6.39	58.57 ± 4.21	62.19 ± 7.27	.117
ODI ↑ (*na* %)	11.46 ± 11.10	8.58 ± 6.80	13.76 ± 13.41	.235
AHI ↑ (≥ 30 N/h)	11.36 ± 10.49	8.80 ± 6.43	13.40 ± 12.73	.266
NCA ↑ (≥ 5 N/h)	1.26 ± 1.69	1.01 ± 1.57	1.46 ± 1.81	.500

Table 3.3: Clinical data of DM1 patients (as a whole cohort, and divided into E1 and E2 genetic classes), divided into: neurological (MIRS, NIFDS, ESS), pneumological (FVC, $PaCO_2$, PaO_2), cardiological (PR, QRS, EF) and sleep (ODI, AHI, NCA) evaluations. Arrows in the first column refer to the direction of the severity of the corresponding scale; pathological cut-off has been indicated in brackets, where it is not available *na* has been indicated. Values are given as mean ± standard deviation, significant (* $p < .05$) and highly significant (** $p < .01$) differences between E1 and E2 classes are highlighted. Scores of MIRS, NIFDS and ESS are reported; FVC is reported as % of the predicted value; AHI and NCA are reported as number of apnoeas, hypopnoeas and central apneas per hour of sleep (MIRS = Muscular Impairment Rating Scale, NIFDS = Neuromuscular Impairment Function and Disability Scale, NP = Neuropsychological, Mo = Motor, My = Myotonia, DL = Daily Life activity, ESS = Epworth Sleepiness Scale, FVC = Forced Vital Capacity, $PaCO_2$ = arterial carbon dioxide partial pressure, PaO_2 = arterial oxygen partial pressure, EF = Ejection Fraction, ODI = Oxygen Desaturation Index, AHI = Apnea Hypopnea Index, NCA = Number of Central Apneas)

3.2.2 MR measurements and VOIs segmentation

The MR protocol included MPRAGE and QSM sequences, processed as explained in Sec. 1.2.1. Images did not show movement artifacts and were considered suitable for the analysis.
The workflow pipeline is illustrated in Fig. 3.2. QSM images were linearly registered to the corresponding MPRAGE using FLIRT[44][45]. Automated segmentation tools were applied to MPRAGE images and the resulting segmentation was overlaid to QSM. More details about software and toolboxes used can be found in Sec. 1.2.1. As Volumes of interest (VOIs), the following structures were selected:

- Ten cortical gyri from FreeSurfer[48] cortical parcellation: PreCentral Gyrus (PrCG), Caudal Middle Frontal Gyrus (CMFG), ParaCentral Gyrus (PCG) and Pars Triangularis (PT) in the

frontal lobe, PostCentral Gyrus (PoCG), Inferior Parietal Gyrus (IPG) and SupraMarginal Gyrus (SMG) in the parietal lobe, Transverse Temporal Gyrus (TTG) in the temporal lobe and Posterior Cingulate Cortex (PCC) and Isthmus of the Cingulate Gyrus (ICG) in the cingulate cortex. Not all the 35 available gyri were considered because some of them overlap artifacts in QSM due to air-tissue boundaries and cortical bones[59] [95].

- Caudate nuclei, accumbens, putamen, globus pallidus, thalamus, hippocampus and amygdala from FIRST segmentation[46].
- Thalamic subnuclei from an atlas-based method implemented in Brun et al., 2022[128]. Using 30 7T- MP2RAGE images of controls, they built up an atlas in the Montreal Neurological Institute's 152 (MNI152) space and a template with 24 deep grey nuclei, 14 of which are thalamic. As suggested, their template was registered to MPRAGE images using ANTs (http://stnava.github.io/ANTs/) and nuclei were grouped into four main units (anterior, medial, ventral and pulvinar). In Fig. 3.3, thalamic nuclei are shown overlaid to MPRAGE and QSM of a representative subject.
- Brainstem and its substructures (midbrain, pons and medulla) from FreeSurfer[129]; in Fig. 3.4, brainstem sub-segmentation is shown overlaid to MPRAGE and QSM of a representative subject.
- Substantia Nigra (SN), Red Nuclei (RN) and Dentate Nuclei (DN) were identified on an original implemented χ-enhanced atlas, built up in the MNI152 space. QSM images were registered to the corresponding MPRAGE using ANTs and MPRAGE to the MNI152 space using FNIRT (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FNIRT). SN, RN and DN were manually drawn on the atlas, and the ROIs were back-registered to the native space for each subject. In Fig. 3.5, these structures are shown on a representative subject of the cohort.

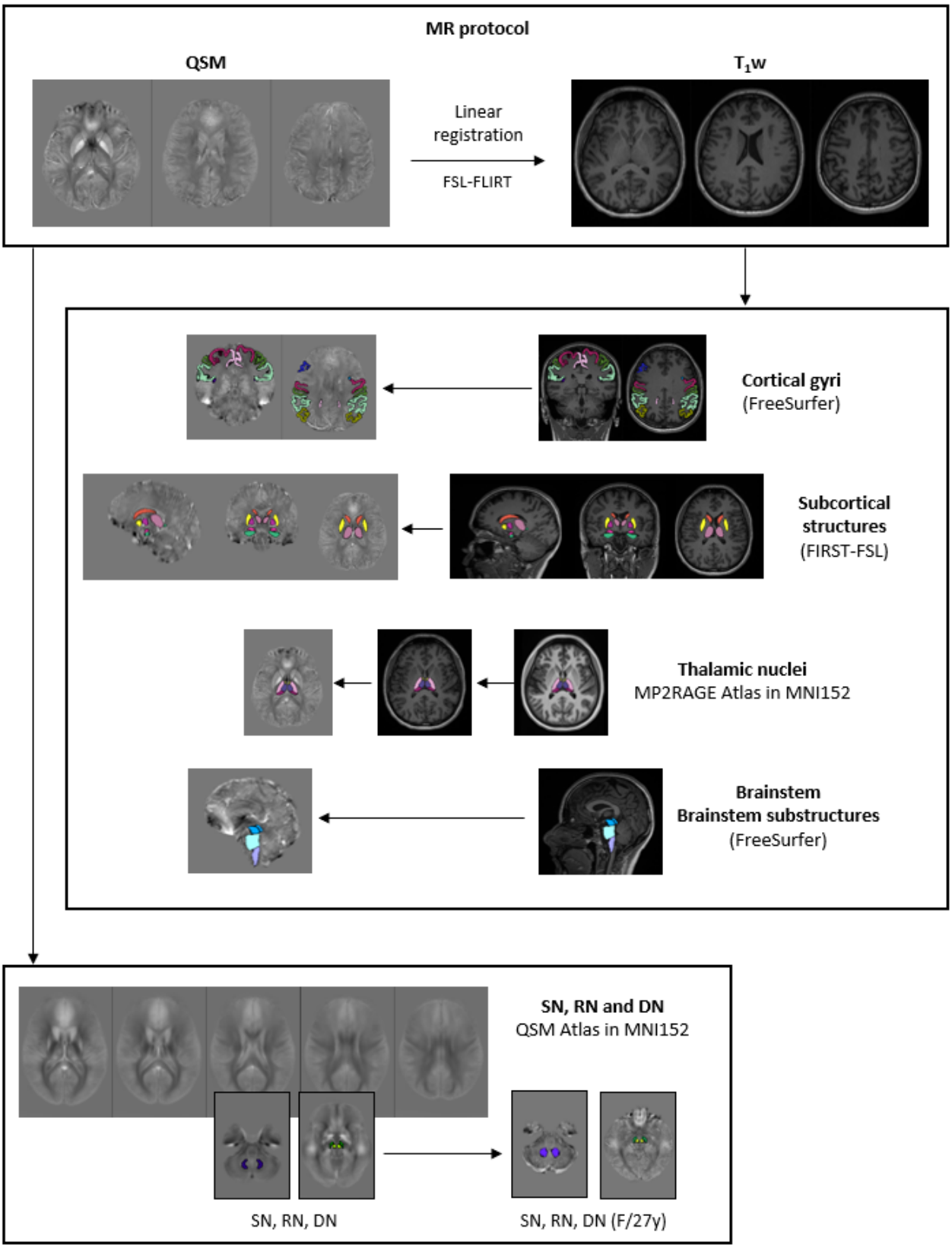

Figure 3.2: Workflow pipeline, including MR protocol and different segmentation methods used. Ten cortical gyri (precentral gyrus, caudal middle frontal gyrus, paracentral gyrus and pars triangularis in frontal lobe, postcentral gyrus, inferior parietal gyrus and supramarginal gyrus in parietal lobe, transverse temporal gyrus in temporal lobe and posterior cingulate cortex and isthmus of the cingulate gyrus in cingulate cortex) were selected from FreeSurfer cortical parcellation; the brainstem and its substructures (midbrain, pons and medulla) were selected from FreeSurfer segmentation; caudate, accumbens, putamen, globus pallidus, thalamus, hippocampus and amygdala were selected from FIRST segmentation; subthalamic nuclei (anterior, medial, ventral and pulvinar) were selected from a MP2RAGE-enhanced atlas developed in [128]; Substantia Nigra (SN), Red (RN) and Dentate Nuclei (DN) were selected from a susceptibility-enhanced atlas originally implemented. Images are from a representative healthy control (F/27 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit)

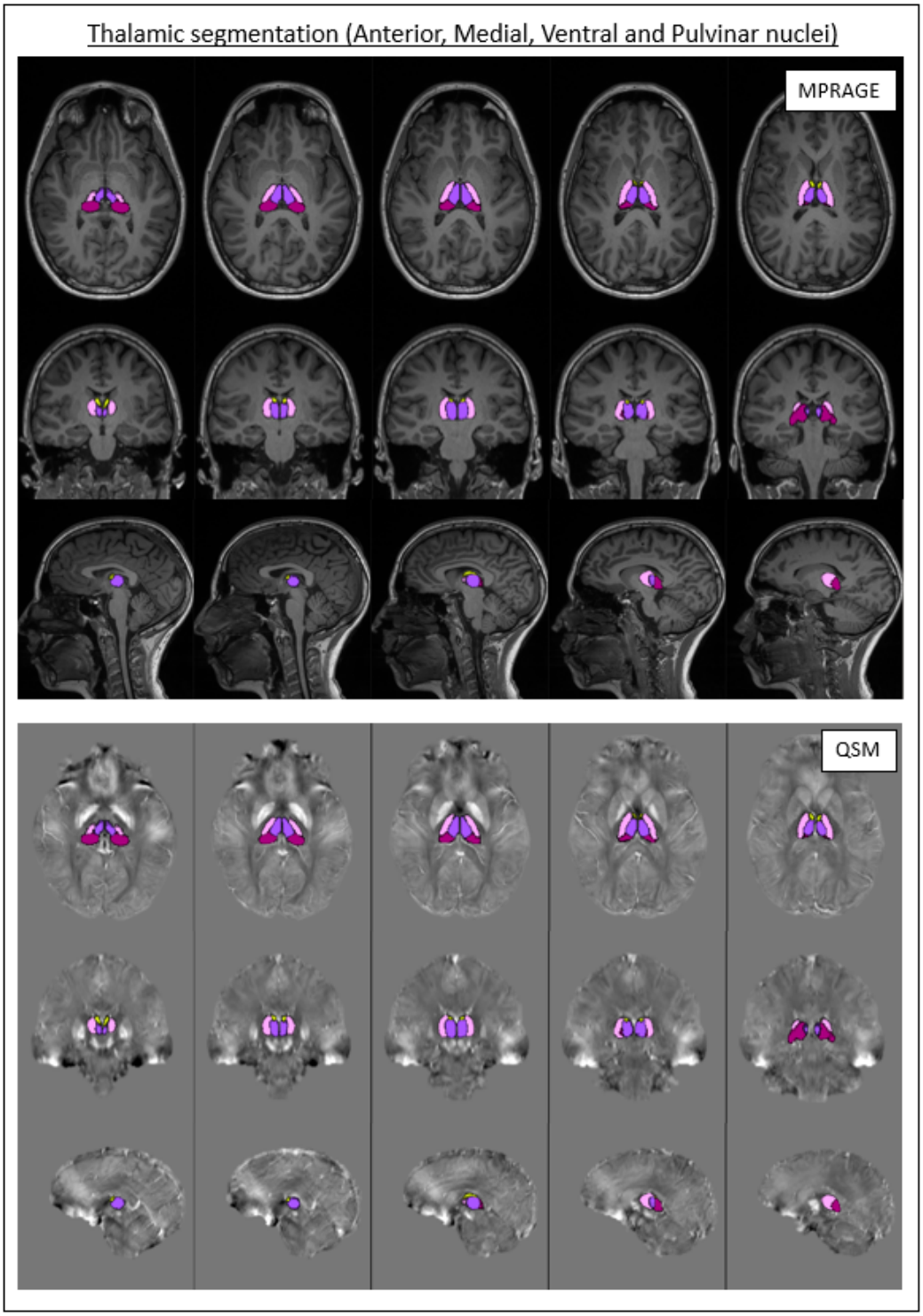

Figure 3.3: Thalamic segmentation of anterior, medial, ventral and pulvinar nuclei overlaid on MPRAGE and QSM images of a representative healthy control of the study cohort (F/27 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit). Thalamic nuclei segmentation was performed using the template from [128]

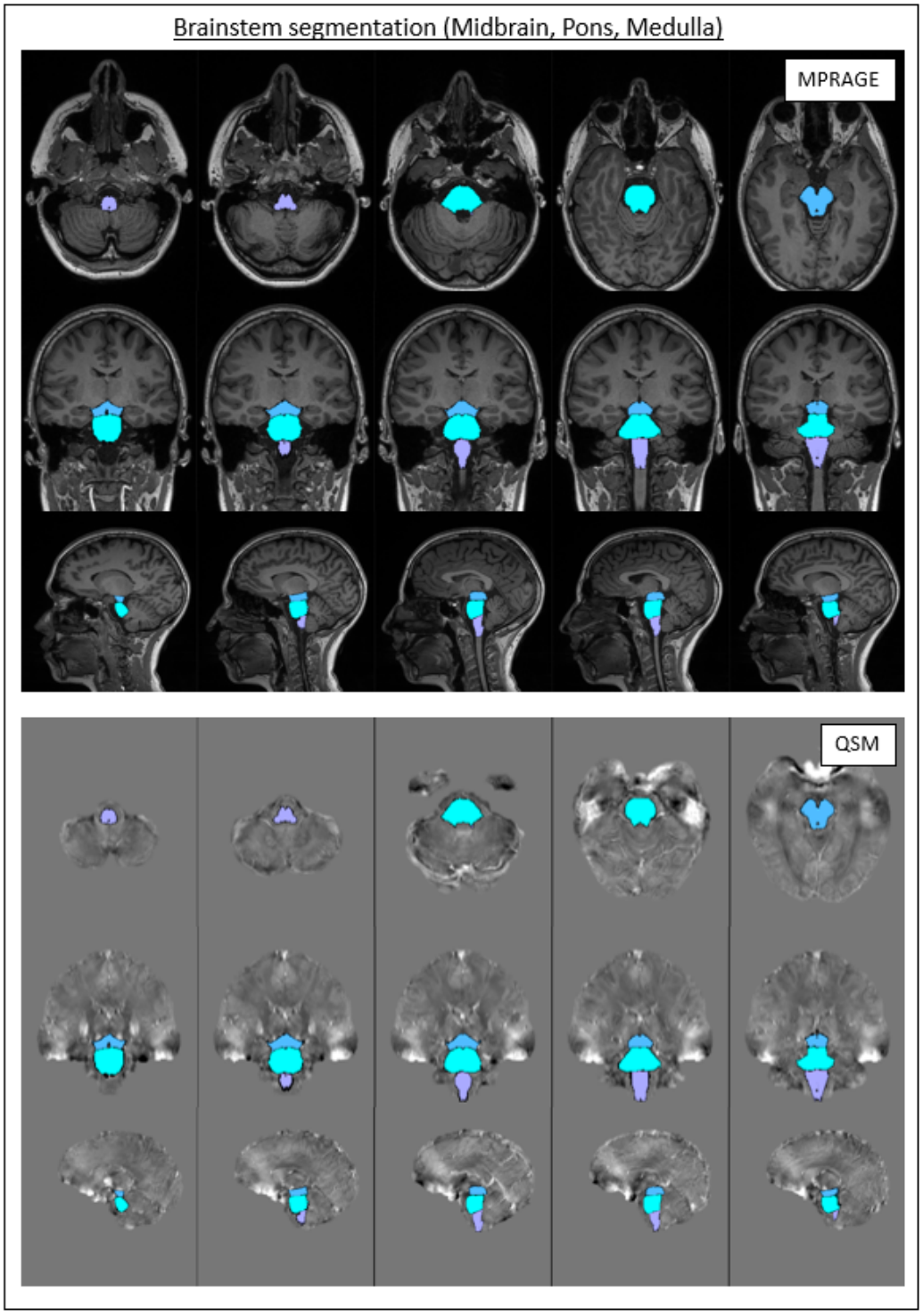

Figure 3.4: Brainstem segmentation of midbrain, pons and medulla from FreeSurfer, overlaid on MPRAGE and QSM images of a representative healthy control of the study cohort (F/27 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit)

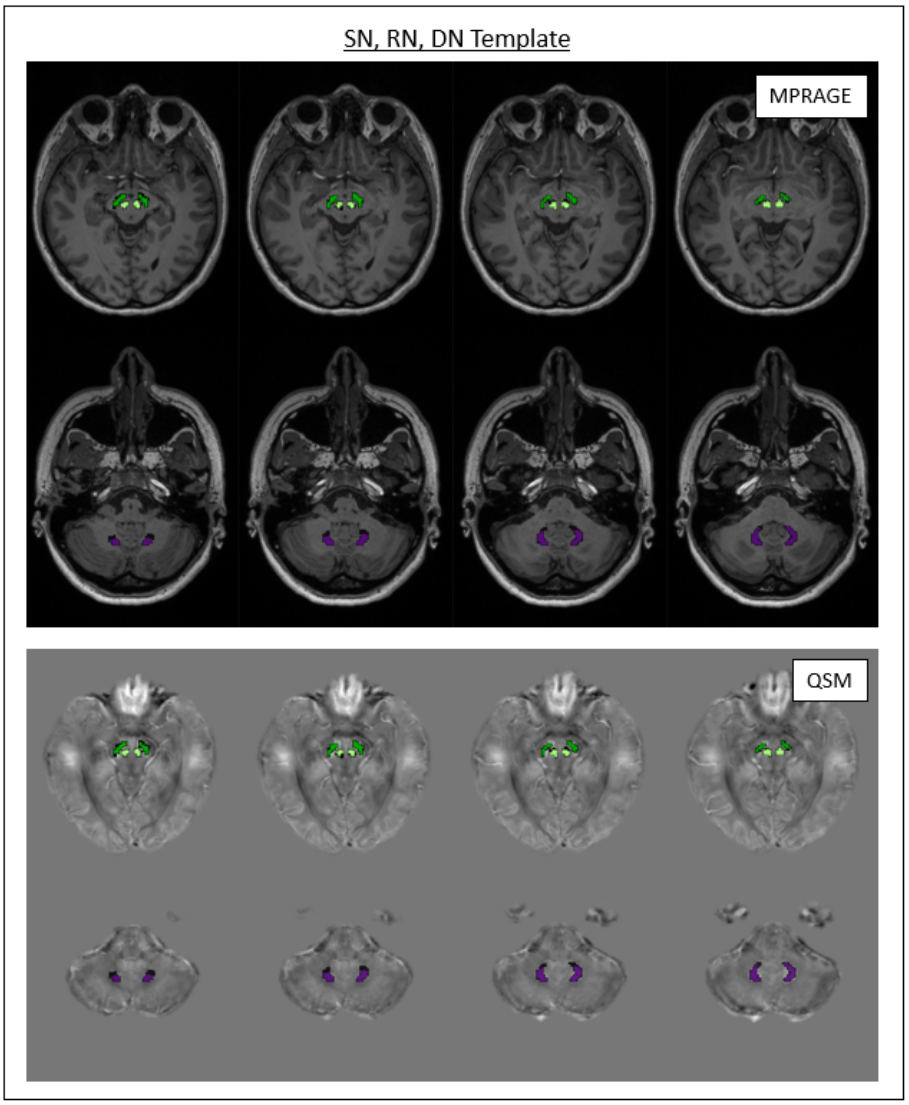

Figure 3.5: Substantia nigra (SN), red and dentate nuclei (RN and DN) from the QSM atlas-based method implemented, overlaid on MPRAGE and QSM images of a representative healthy control of the study cohort (F/27 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit)

3.2.3 Statistical analysis

ROI-based analysis was performed comparing the median χ and volume distributions between HC and DM1 and between E1 and E2 genetic classes, averaging the left and the right side. Median χ values were corrected by age assuming a linear increase in the control groups[103], and volume was corrected by Total Intracranial Volume (TIV) using the proportional method[99]. From Spearman's test, correlations between volume and χ values did not result. One-sample Kolmogorov-Smirnov test was used to check the normality of the distributions; χ and volumes were not normally distributed in any of the structures, thus the non-parametric Kruskal-Wallis test was used (* *p-values* $< .05$ and ** *p-values* $< .01$). Correlations were evaluated between χ values and clinical data using Spearman's test (* *p-values* $< .05$ and ** *p-values* $< .01$).
The analysis was carried out using Python (version 3.7.6).

3.3 Results

3.3.1 χ and volume: cortical and subcortical structures

Susceptibility and volume distributions were evaluated in the selected gyri, comparing HC and DM1 groups; results are summarized in Tab. 3.4. Median χ is significantly higher in multiple cortical gyri in DM1 group, in particular in PrCG ($p = .016$), CMF ($p = .001$), PT ($p = .005$), PoCG ($p = .044$), IPG ($p = .005$), SMG ($p = .002$), TTG ($p = .002$) and ICG ($p = .006$), together with significant atrophy in some of them (PrCG, $p = .005$; PCG, $p = .010$; PoCG, $p = .012$; IPG, $p = .003$; SMG, $p = .002$; TTG, $p = .001$). Note that, despite the increase in susceptibility corresponding to a decrease in volume, there are no significant correlations between the two properties in any structures. The same analysis was performed to compare, within the patient group, E1 and E2 classes. There were no significant differences between the two genetic classes exploring those measurements.

	Median χ		**Volume**	
PrCG	*	↑	**	↓
CMFG	**	↑	*ns*	↓
PCG	*ns*	↑	*	↓
PT	**	↑	*ns*	↓
PoCG	*	↑	*	↓
IPG	**	↑	**	↓
SMG	**	↑	**	↓
TTG	**	↑	**	↓
PCC	*ns*	↑	*ns*	↑
ICG	**	↑	*ns*	↑

Table 3.4: Susceptibility and volume distribution comparison between HC and DM1 groups (↑ = χ increase in DM1 group, ↓ = χ decrease of in DM1 group) in cortical gyri. ROI-based analysis was performed (Kruskal-Wallis test, * $p < .05$, ** $p < .01$) (G = Gyrus, PrCG = PreCentral G, CMFG = Caudal Middle Frontal G, PCG = ParaCentral G, PT = Pars Triangularis, PoCG = PostCentral G, IPG = Inferior Parietal G, SMG = SupraMarginal G, TTG = Transverse Temporal G, PCC = Posterior Cingulate Cortex, ICC = Isthmus of the Cingulate Gyrus)

χ distributions in subcortical structures between controls and DM1 were compared, at first considering the patients' group as a whole cohort and then distinguishing the two genetic classes. The first outcomes are shown in Tab. 3.5. Bar plots with χ distributions within the two groups are shown in Fig. 3.6: there is a significant and highly significant increase of bulk χ respectively in the thalamus ($p = .020$) (Fig. 3.7) and in the brainstem ($p = .003$) (Fig. 3.8), without detecting significant volume changes.

	Median χ		Volume	
Cau	ns	↑	ns	↑
Acc	ns	↓	ns	↓
Put	ns	↓	ns	↓
GP	ns	↓	ns	↑
SN	ns	↓	ns	↑
RN	ns	↑	ns	↑
Th	*	↑	ns	↑
Hipp	ns	↑	ns	↑
Amy	ns	↑	ns	↑
DN	ns	↑	ns	↑
Brainstem	**	↑	ns	↑

Table 3.5: Susceptibility and volume distribution comparison between HC and DM1 groups (↑ = χ increase in DM1 group, ↓ = χ decrease of in DM1 group) in the analyzed subcortical structures. ROI-based analysis was performed (Kruskal-Wallis test, * $p < .05$, ** $p < .01$) Cau = Caudate, Acc = Accumbens, Put = Putamen, GP = Globus Pallidus, SN = Substantia Nigra, RN = Red Nucleus, Th = Thalamus, Hipp = Hippocampus, Amy = Amygdala, DN = Dentate Nucleus)

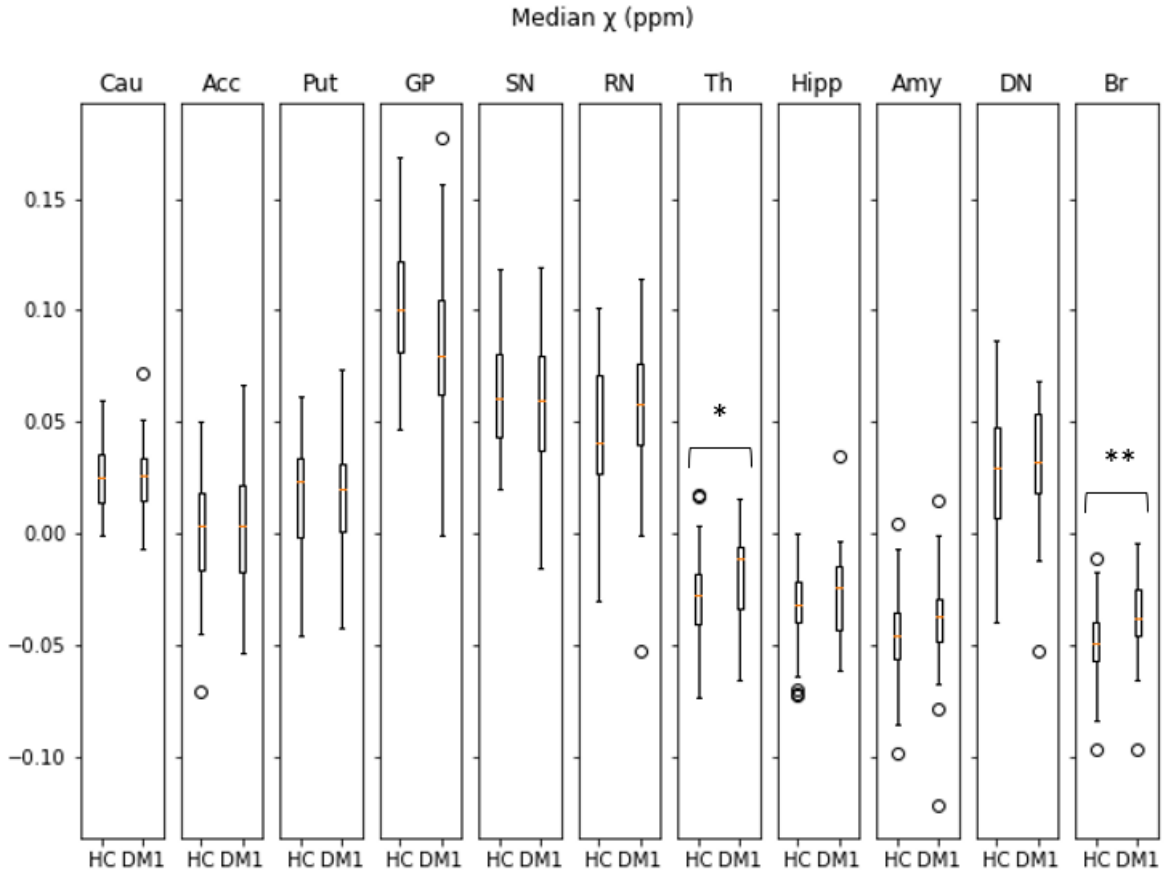

Figure 3.6: Bar plots of χ distributions of the analyzed subcortical structures. VOI-based analysis was performed (Kruskal-Wallis test, * $p < .05$, ** $p < .01$) (Cau = Caudate, Acc = Accumbens, Put = Putamen, GP = Globus Pallidus, SN = Substantia Nigra, RN = Red Nucleus, Th = Thalamus, Hipp = Hippocampus, Amy = Amygdala, DN = Dentate Nucleus)

No significant difference was found between E1 and E2 groups in χ and volume distributions. An increase of χ (E2 > E1) occurred in the caudate, thalamus and amygdala; specifically in the thalamus, susceptibility is higher in patients than healthy controls and in E2 than in E1 group

(Fig. 3.7, panel b), which may indicate a marker of worsening of clinical conditions. Since the thalamus and brainstem showed significant differences in DM1 patients, those structures were better analyzed by studying their sub-units.

3.3.2 Thalamic nuclei and brainstem sub-units

Anterior, medial, ventral and pulvinar nuclei of the thalamus were analyzed (Figures 3.2 and 3.3). Results about the comparisons HC *vs* DM1 and E1 *vs* E2 are reported in Tab. 3.6 (upper part) and Fig. 3.7: there is a significant increase in susceptibility in the ventral and pulvinar nuclei, leading, as mentioned in the previous section, to a significant difference considering the whole thalamus. Regarding the comparison between the two genetic classes, there were no significant differences.

		HC *vs* **DM1**		**E1** *vs* **E2**	
Thalamus	**Tot**	*	↑	*ns*	↑
	Anterior	*ns*	↑	*ns*	↑
	Medial	*ns*	↑	*ns*	↑
	Ventral	**	↑	*ns*	↓
	Pulvinar	*	↑	*ns*	↑
Brainstem	**Tot**	**	↑	*ns*	↓
	Midbrain	*ns*	↑	*ns*	↓
	Pons	**	↑	*ns*	↓
	Medulla	*	↑	*ns*	↑

Table 3.6: Comparison of χ distribution in sub-units in the thalamus and the brainstem. On the left side, HC and DM1 ($\uparrow = \chi$ increase in DM1 group, $\downarrow = \chi$ decrease of in DM1 group); on the right side, E1 and E2 (genetic classes of DM1) ($\uparrow = \chi$ increase in E2 class, $\downarrow = \chi$ decrease in E2 class). VOI-based analysis was performed (Kruskal-Wallis test, * $p < .05$, ** $p < .01$) (ns = not significant)

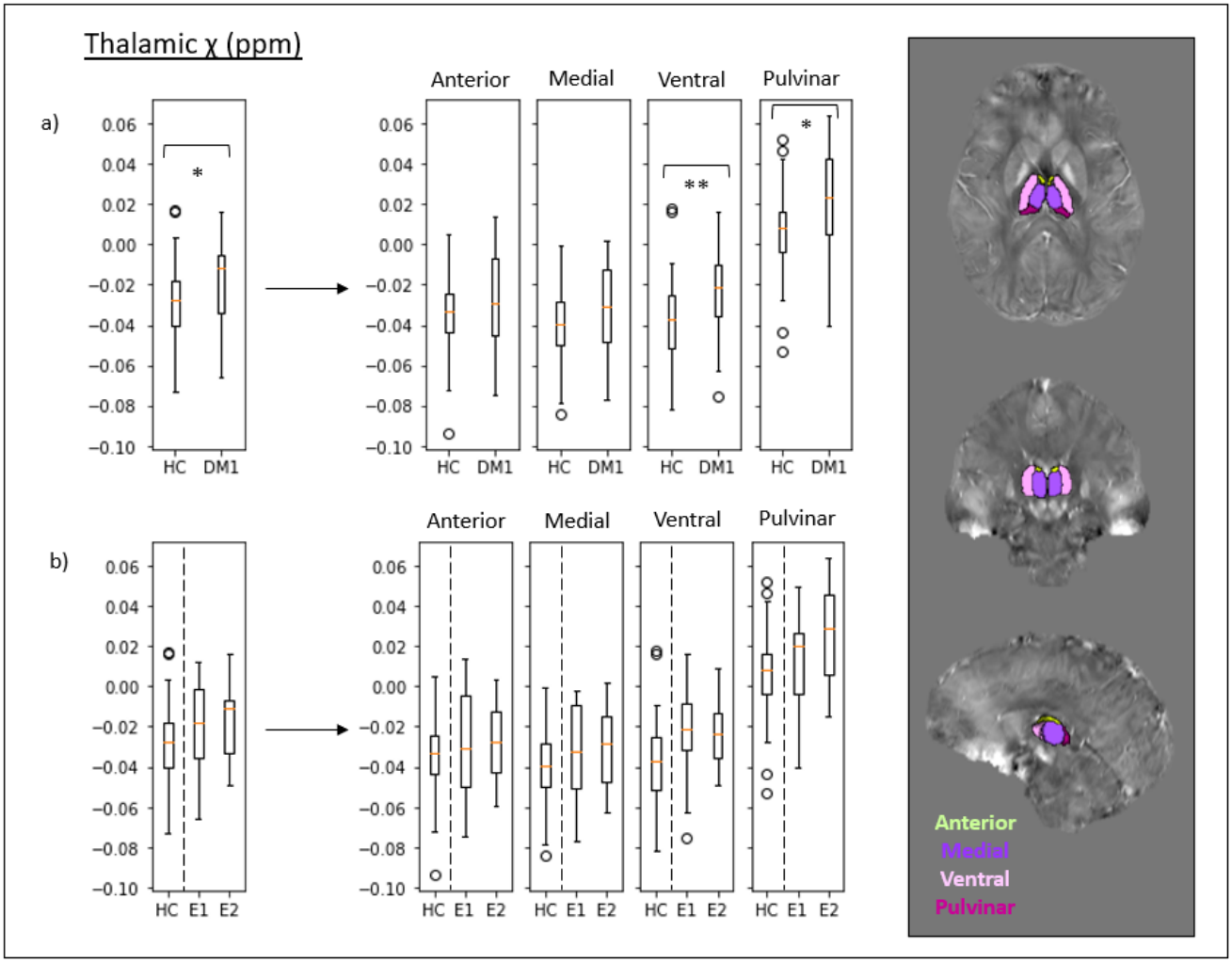

Figure 3.7: Thalamic χ distributions in HC and DM1 (E1 and E2 classes). Distributions were evaluated by comparing the control cohort with the entire DM1 group (panel a) and distinguishing E1 and E2 classes (panel b). On both panels, the first plot on the left is the one regarding the entire structure, while on the right the plots of the thalamic nuclei (anterior, medial, ventral, pulvinar) are shown. VOI-based analysis was performed (Kruskal-Wallis test, * $p < .05$, ** $p < .01$). On the right, representative images of thalamic nuclei segmentation (HC, F/26 years old)

Midbrain, pons and medulla were individually analyzed as parts of the brainstem (Figures 3.2 and 3.4). In Tab. 3.6 (bottom part) and in Fig. 3.8, results are reported: pons, medulla and the entire brainstem showed a significant increase in DM1, major in the pons; there is no trend of increases looking into E1/E2 differentiation.

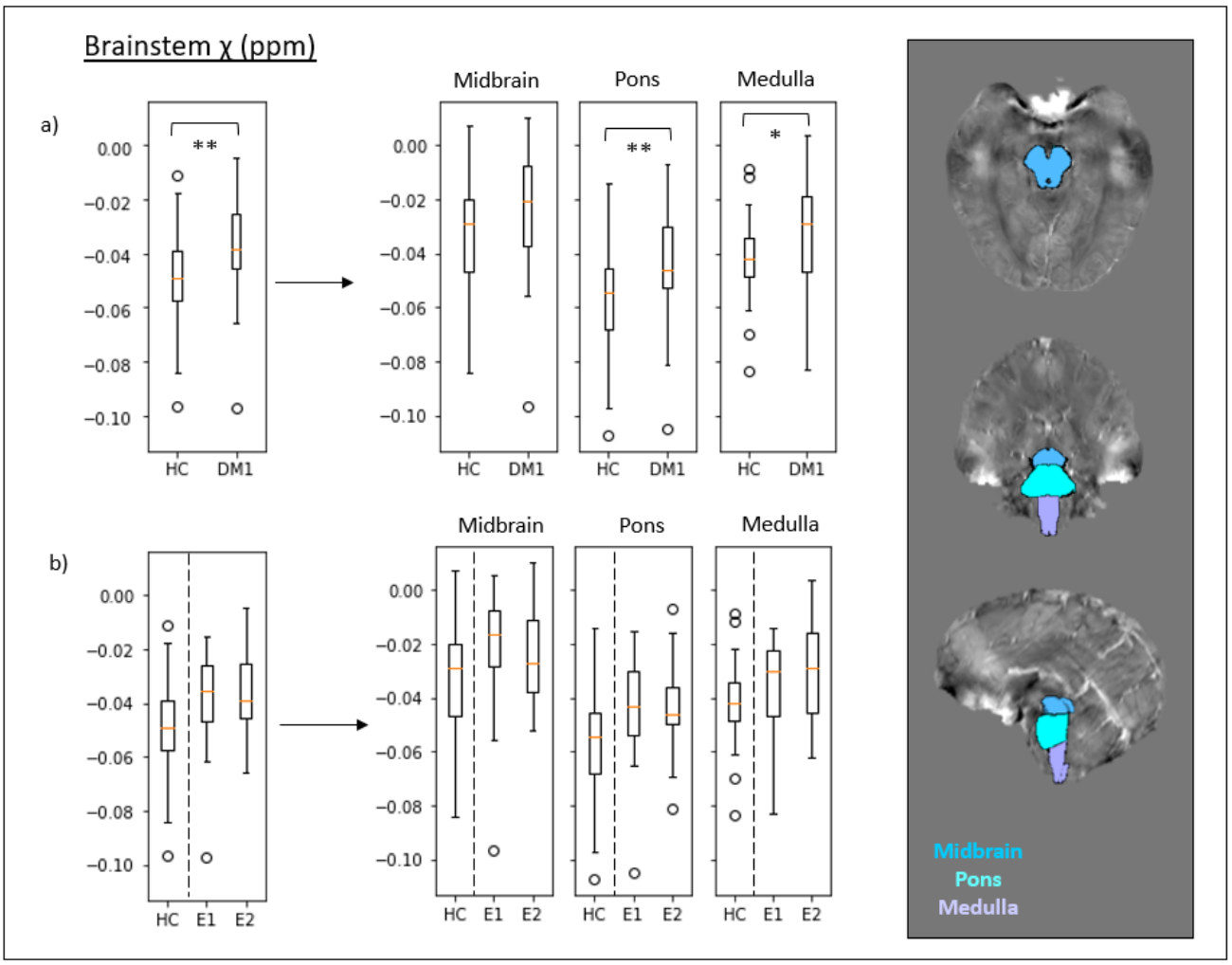

Figure 3.8: χ distributions of the brainstem in HC and DM1 (E1 and E2 classes). Distributions were evaluated by comparing the control cohort with the entire DM1 group (panel a) and distinguishing E1 and E2 classes (panel b). On both panels, the first plot on the left is the one regarding the entire structure, while on the right the plots of its substructures (midbrain, pons and medulla) are shown. VOI-based analysis was performed (Kruskal-Wallis test, * $p < .05$, ** $p < .01$). On the right, representative images of brainstem segmentation (HC, F/26 years old)

3.3.3 χ correlations with clinical and laboratory data

Available data in the study included clinical evaluation. Since subcortical regions, particularly the thalamus and brainstem, showed a higher specificity than cortical gyri, possible correlations between χ values and clinical data in those structures, including the sub-segmentation of the thalamus and the brainstem, were explored. In Tab. 3.7, significant correlations in the DM1 groups are shown; within the text below, eventual significances within the genetic classes considered individually are reported. In the table the direction of severity for each parameter is reported; please consider χ values in those structures correlate with iron accumulation [24] [130], as a possible biomarker of disease severity. The main findings were the following.

- χ values negatively correlated with the onset age in the thalamus – considering the sub-segmentation, in medial and pulvinar nuclei – and the brainstem.
- Thalamus and amygdala are linked with disability scores. Thalamic χ in the pulvinar positively correlated with three out of four domains of NIFDS (motor, myotonic and daily life

activities); the correlation in the motor domain remained true in the E2 class individually. Amygdala χ correlated with motor and daily life domain of NIFDS, as well as in the total counts; as before, this resulted also in the individual E2 class, where, additionally, the correlation between amygdala χ and the neuropsychological domain occurred.

- Pneumological data did not correlate with any structures considering the entire cohort of patients; there was in the E2 class a positive correlation with the $PaCO_2$ within the anterior and medial thalamic nuclei and the medulla oblongata.

- Within the cardiological data, correlations occur between the PR interval length and brain structures: thalamus, as a whole and in the medial, ventral and pulvinar nuclei; medulla oblongata; amygdala and putamen. Thalamic, medial and pulvinar χ positively correlated within E1 and E2 classes individually; amygdala χ correlated also in E2 class.

- There are correlations between χ values and the index of central apnea in the thalamic medial nucleus, the brainstem, the hippocampus and the putamen

	Thalamus	**Brainstem**	**Others**
Onset age $\downarrow$	(-) Th (*)	(-) Br (*)	
MIRS $\uparrow$			
NIFDS (NP) $\uparrow$			
NIFDS (Mo) $\uparrow$	(+) Th (*)		(+) Amy (*)
NIFDS (My) $\uparrow$	(+) Th (**)		
NIFDS (DL) $\uparrow$	(+) Th (*)		(+) Amy (*)
NIFDS (tot) $\uparrow$			(+) Amy (*)
ESS $\uparrow$			
FVC $\downarrow$			
$PaCO_2$ $\uparrow$			
PaO_2 $\downarrow$			
PR $\uparrow$	(+) Th (**)	(+) Br (*)	(+) Amy (*) Put (**)
QRS $\uparrow$			
EF $\downarrow$			
ODI $\uparrow$			
AHI $\uparrow$			
NCA $\uparrow$	(+) Th (*)	(+) Br (*)	(+) Hipp (*) Put (*)

Table 3.7: Correlations between χ values and clinical data of DM1 patients; data are divided into: neurological (MIRS, NIFDS, ESS), pneumological (FVC, $PaCO_2$, PaO_2), cardiological (PR, QRS, EF) and sleep (ODI, AHI, NCA) evaluations. Arrows in the first column refer to the direction of the severity of the corresponding scale. Spearman's test (* $p < .05$, ** $p < .01$) was used, evaluating the following structures: thalamus, brainstem and the other subcortical structures (Th = Thalamus, Br = Brainstem, Cau = Caudate, Acc = Accumbens, Put = Putamen, GP = Globus Pallidus, SN = Substantia Nigra, RN = Red Nucleus, Hipp = Hippocampus, Amy = Amygdala, DN = Dentate Nucleus, MIRS = Muscular Impairment Rating Scale, NIFDS = Neuromuscular Impairment Function and Disability Scale, NP = Neuropsychological, Mo = Motor, My = Myotonia, DL = Daily Life activity, ESS = Epworth Sleepiness Scale, FVC = Forced Vital Capacity, $PaCO_2$ = arterial carbon dioxide partial pressure, PaO_2 = arterial oxygen partial pressure, EF = Ejection Fraction, ODI = Oxygen Desaturation Index, AHI = Apnea Hypopnea Index, NCA = Number of Central Apneas)

3.4 Discussion

DM1 is a hereditary multisystem disorder that affects both the musculoskeletal system and the CNS. Previous research has utilized MR brain imaging techniques to identify anomalies in the brains of individuals with DM1[105]. Building on this prior work, the current study aims to investigate χ distributions in several brain structures. To provide context, it's important to note that there is only one previous study, conducted by Ates and colleagues in 2019[106], that examined susceptibility-based brain images in DM1 patients. In this research, their findings were confirmed and expanded, including additional brain structures in the analysis and integrating imaging biomarkers with a comprehensive neurological assessment. This investigation focused on gray matter structures, including both cortical and subcortical regions, to explore potential correlations between neurological observations and the data derived from susceptibility-based imaging. The primary objective was to gain a deeper understanding of the pathological progression of DM1, specifically distinguishing between central and peripheral origins of DM1 symptoms. Furthermore, we aimed to identify specific biomarkers for the disease since neurodegenerative conditions often exhibit quantitative and qualitative indicators that are sensitive but lack specificity. Below, I am going to delve into the results of the study, emphasizing the key findings.

χ in cortical gyri

Prior research has underscored disparities in the cortex when comparing individuals with DM1 to healthy controls[105], specifically widespread cortical atrophy involving reductions in both cortical volume and thickness in regions of the frontal, temporal, parietal, and occipital cortex has been revealed. Hence, the first part of the analysis has been about magnetic susceptibility within cortical gyri. Notably, outcomes indicated a significant increase in susceptibility among DM1 patients when compared to healthy controls (Tab. 3.4). This increase was observed in regions such as the precentral gyrus, caudal middle frontal gyrus and pars triangularis within the frontal lobe, as well as the post-central, inferior parietal and supramarginal gyri within the parietal lobe. Similar increases in susceptibility were noted in the transverse-temporal region of the temporal lobe and the mid-posterior cingulate cortex. These observations aligned with the alterations revealed through structural analysis.

It is important to note that an increase in susceptibility in cortical gyri is typically associated with the accumulation of iron and is commonly linked to a more severe clinical condition in various neurodegenerative disorders, such as amyotrophic lateral sclerosis[15] (Ch. 2). It is plausible to hypothesize that a similar association might hold for DM1. The affected gyri are closely related to essential functions, including motor and language skills within the frontal lobe, sensory perception within the parietal lobe, auditory function in the temporal lobe, and memory and awareness in the posterior cingulate cortex. These functions may be compromised in DM1 patients, who often exhibit myotonia and muscle weakness[21], cognitive difficulties in both written and spoken languages[131], and a lack of awareness regarding the burden and progression of the disease[116].

χ in subcortical structures

Cortical iron accumulation is indeed present in DM1 patients, and it may be associated with some of the clinical symptoms they experience. However, this process is a common characteristic of many neurodegenerative conditions and it is not specific to DM1. Because of that, the analysis was extended to include sub-cortical regions, namely: caudate, accumbens, putamen, globus pallidus, substantia nigra, red nuclei, thalamus, hippocampus, amygdala, dentate nuclei and brainstem. Among those, both the thalamus and the brainstem displayed a significant increase in susceptibility in DM1 patients when compared to healthy controls (Tab. 3.5 and Fig. 3.6). Alterations in susceptibility measurements in the thalamus have also been noted in previous research [106], where they were identified as part of the pathophysiological changes seen in DM1 patients. In both the thalamus and brainstem, as it happens for cortical gray matter, a positive correlation between susceptibility values and iron concentration has been identified by combining imaging and histopathological measurements [24] [130]. Below, I am going to focus on the analysis of these two structures, considering their sub-units.

χ in thalamic nuclei

The thalamus plays a pivotal role as a central structure in the transmission of information from the body to the brain. Given its strategic position and its extensive connections within the CNS, alterations in the thalamus can potentially lead to a wide range of functional impairments, spanning from motor deficits to disturbances in consciousness. A more in-depth investigation into thalamic nuclei can provide valuable insights into which specific aspects are most affected, consequently offering a deeper understanding of the corresponding clinical spectrum. In this study, I focused on four distinct thalamic nuclei (Fig. 3.3): anterior, ventral, medial and pulvinar. The analysis revealed an increase of χ in the entire thalamus and the ventral and pulvinar nuclei independently (Tab. 3.6 and Fig. 3.7). The ventral thalamic nuclei are responsible for relaying information to the somatosensory and motor cortex, playing a vital role in transmitting data related to movement, tremors and the sensory system. On the other hand, the pulvinar nuclei project to the visual cortex, carrying crucial information related to visual skills. Note that the three corresponding cortical areas (motor, sensory, and visual) displayed χ alterations in the examined cohort of patients.

Moreover, a comparative analysis of thalamic χ, both considering the thalamus as a whole structure and breaking it down into individual nuclei, was conducted across the E1 and E2 classes within the DM1 patient group. No statistically significant difference was found; however, a noticeable trend of increased thalamic susceptibility was observed, progressing from healthy controls to E1 and E2 classes. This trend - remaining consistent when considering the anterior, medial, and pulvinar nuclei individually - led to infer that the increase in thalamic susceptibility is associated with the worsening of clinical conditions (Fig. 3.7).

Furthermore, this research revealed significant correlations between thalamic susceptibility and clinical data (Tab. 3.7). Notably, a significant negative correlation was observed between susceptibility values and the age of disease onset. This negative correlation suggests that an earlier onset

is linked to a higher concentration of iron in the entire thalamus and specifically in the medial and pulvinar nuclei. Additionally, thalamic pulvinar susceptibility exhibited a significant positive correlation with DM1 disability scores, especially with motor function, myotonia, and activities of daily living as assessed by the NIFDS. These correlations remained significant within the E2 class, but not in the E1 class. Moreover, thalamic susceptibility (both as a whole and in the medial, ventral, and pulvinar nuclei) was strongly correlated with the PR internal length within the DM1 group, and this correlation held for both E1 and E2 classes. Finally, a significant correlation was observed between the number of central apneas and the susceptibility of the medial thalamic nuclei, which are also involved in the regulation of sensory information related to respiration[132]. This outcome suggests that CNS alterations may correspond to breathing dysfunction during sleep. Importantly, it's worth noting that no such correlations were found, whether within the thalamus or other structures, with the other two apnea indices, ODI and AHI, which consider both central and non-central episodes.

χ in brainstem sub-units

Another significant target for differentiation between DM1 and healthy controls is the brainstem, which holds a central role in the processing of information within the CNS and serves as a crucial intermediary for signals traveling from the body to the brain. It is involved in regulating many vital functions, including breathing, heart rate, blood pressure, and sleep. The brainstem is traversed by numerous white matter tracts, such as the corticospinal and corticobulbar tracts. Any alterations or damage to the brainstem can impact an individual's clinical condition in diverse ways, contingent upon the location of the affected areas. In the analysis, the three primary divisions (midbrain, pons, medulla) of the brainstem were considered (Fig. 3.4). Findings revealed that the most significant differences between individuals with DM1 and healthy controls were predominantly observed in the pons (Tab. 3.6 and Fig. 3.8). The pons contains various nuclei responsible for functions like sleep and respiration. Additionally, differences were noted in the medulla, which comprises the lower half of the brainstem responsible for managing information related to cardiac and respiratory function.

This observation is particularly interesting in light of the associations found between brainstem χ values and cardiological and respiratory evaluations (Tab. 3.7). For instance, susceptibility in the medulla positively correlated with the PR internal length, and susceptibility of the brainstem as a whole structure correlated with the number of central apneas. Sleep disorders are a common feature in the development of DM1, and previous research in the literature has indicated that these anomalies originate in the CNS[133], corroborating our study's findings. A recent study by Nepozitek et al. in 2023[134] also linked magnetic susceptibility in the brainstem to sleep anomalies in a cohort of patients with REM sleep Behavior Disorder. Furthermore, χ in the brainstem displayed a negative correlation with the age of disease onset. An earlier onset was associated with a greater accumulation of iron in the brainstem when considered as a unified structure. No significant correlations were observed when examining the individual sub-structures within the brainstem (midbrain, pons, and medulla) or when considering the individual genetic classes.

Additional considerations and limitations

It is worth noting that other deep gray matter structures seem to be involved in the development of DM1. First, susceptibility in the amygdala exhibited correlations with motor and daily life items in the NIFDS within the DM1 group and the E2 class individually. Additionally, it correlated with the PR internal length within the entire DM1 group and the E2 class. Second, putaminal susceptibility was linked with the PR internal length and the number of central apneas. Last, susceptibility within the hippocampus correlated with the number of central apneas; this structure has been already associated with sleep breathing disorders [125].

There are a few limiting factors to consider: our sample did not include DM1 patients belonging to E3 class; even if the number of patients is consistent considering previous studies, analysis on a larger sample would be useful to confirm the outcomes; it would be interesting to analyze correlations between susceptibility values and neurological, cardiological and respiratory evaluations in the group of controls as well.

3.5 Conclusion

In conclusion, this study helps in better characterization of DM1 as a neurodegenerative disorder, focusing in particular on the role of iron accumulation during the development of the disease. The thalamus and brainstem appeared to be the target structures in QSM analysis: they show a significant increase in susceptibility and correlations with clinical and laboratory data. Specifically, it is interesting the correlation with the number of central apneas in both structures, in addition to the putamen and the hippocampus, already linked with sleep breathing disorder. The thalamus is also involved with motor, myotonia and daily life items of the NIFDS, disability scale implemented for DM1 patients, and, together with the brainstem, amygdala and putamen, it shows significant correlations with PR interval length, associating indeed conduction abnormalities with iron accumulation. Additionally, thalamus and brainstem χ negatively correlates with the onset age, associating iron accumulation with earlier disease onset. Last, the amygdala is also involved in a few items of NIFDS, namely motor, daily life and total score.

These results certainly bring us one step closer to understanding the mechanisms that govern the pathology, and they deserve further investigation. Some differences have been observed between the two genetic classes analyzed, both in the comparisons of distributions and in the trends of clinical data, even though a clear clustering of the two groups using susceptibility data was not identified. Further analyses, with a larger sample or the incorporation of additional features extracted from QSM images, could assist in distinguishing the two genetic classes more precisely.

χ = magnetic susceptibility
AHI = Apnea Hypopnea Index
ANS = Autonomic Nervous System
CMFG = Caudal Middle Frontal
CNS = Central Nervous System
CTG = Cytosine-Thymine-Guanine
DD = Diagnostic Delay
DM1 = Myotonic Dystrophy type 1
DMPK = Myotonic Dystrophy Protein Kinase
DN = Dentate Nuclei
EDS = Excessive Daytime Sleepiness
EF = Ejection Fraction
ESS = Epworth Sleepiness Scale
FCV = Forced Vital Capacity
GM = Gray matter
HC = Healthy Controls
ICG = Isthmus of the Cingulate Gyrus
IPG = Inferior Parietal Gyrus
MIRS = Muscular Impairment Rating Scale
MNI152 = Montreal Neurological Institute's 152
MPRAGE = Magnetization-Prepared Rapid Acquisition Gradient Echo
MR = Magnetic Resonance
MRI = Magnetic Resonance Imaging
NCA = Number of Central Apneas
NIFDS = Neuromuscular Impairment Function and Disability Scale
ODI = Oxygen Desaturation Index
$PaCO_2$ = arterial carbon dioxide partial pressure
PaO_2 = arterial oxygen partial pressure
PCC = Posterior Cingulate Cortex
PCG = Paracentral Gyrus
PoCG = PostCentral Gyrus
PrCG = PreCentral Gyrus
PT = Pars Triangularis
RN = Red Nuclei
ROI = Region of Interest
SMG = SupraMarginal Gyrus
SN = Substantia Nigra
TIV = Total Intracranial Volume
TTG = Transverse Temporal Gyrus
UTR = UnTranslated Region
VOI = Volume of Interest
WM = White Matter

Chapter 4

Innovative Methods to QSM images

In this chapter, I am going to elaborate on two innovative approaches I applied to QSM images, aiming to enhance the use and analysis of those data.
The first method involves the implementation of a robust and accurately designed pipeline for the extraction and evaluation of radiomic features from QSM images. This pipeline specifically focuses on the normal appearing white matter and it is employed within a study conducted on a cohort of patients with multiple sclerosis and healthy controls. The objective of this approach is to systematically extract and assess radiomic features, offering valuable insights into the structural and textural characteristics of the normal appearing white matter within the context of multiple sclerosis and beyond. Section 4.1 refers to this study, whose results have been published in a research paper[23].
The second methodology discussed regards the application of a super-resolution model aiming to enhance the spatial resolution of susceptibility images. To date, QSM is quite unexplored in this kind of application, and this work represents an innovative advancement in the field. The approach used was peculiar, since a model already trained with natural images was directly applied to brain MR images, without *re-training* or *fine-tuning*. The model was validated on morphological T_1w and T_2w images of a large sample from the Cam-CAN database at first[135], and then applied to QSM images from Neuroimaging Laboratory of IRCCS Istituto delle Scienze Neurologiche di Bologna, possibly offering new perspectives and enhancing capabilities for the study of neurological and medical conditions. Results are shown in Section 4.2.

4.1 QSM radiomic features in Multiple Sclerosis

Multiple Sclerosis (MS) is an autoimmune demyelinating disease characterized by changes in iron and myelin content, biomarkers detectable using QSM[1][24]. When analyzed with radiomic techniques that exploit its intrinsic quantitative nature, QSM may furnish biomarkers to facilitate early diagnosis and timely assessment of progression.
In this work, the robustness of QSM radiomic features was explored by varying the number of grey levels and echo times, in a sample of healthy controls and patients with MS. Normal appearing white matter was analyzed in total and within six clinically relevant tracts, namely: arcuate fasciculus, cortico-spinal tract, frontal aslant tract, inferior fronto-occipital fasciculus, optic radiation and uncinate fasciculus. The background and aims are described in Sec. 4.1.1 and 4.1.2, the pipeline implementation and database in Sec. 4.1.3.
Results and discussion are in Sec. 4.1.4 and 4.1.5. After optimizing the number of gray levels (n = 64), at least 65% of features were robust for each volume of interest, with no difference ($p > .05$) between the left and right hemispheres. Different outcomes in feature robustness among the analyzed regions depend on their characteristics, such as volume and variance of susceptibility values.
This study[23] validated the processing pipeline for robustness analysis and established the reliability of QSM-based radiomic features against gray levels and echo times (Sec. 4.1.6), providing important insights for future radiomic studies using QSM in clinical applications, which I am going to pursue in future work.

The first use of abbreviations includes their full expansion; for ease of reference, a glossary is provided at the end of the section.

4.1.1 Radiomics and biomarkers

Radiomics has been defined as the conversion of digital images into mineable high-dimensional data[136]. Traditionally, images are intended for visual interpretation, but there is more beyond signal intensity and contrast: focusing on a specific Volume of Interest (VOI) in a 3D image, it is possible to analyze descriptors of the shape, intensity histogram and texture. Over the past two decades, radiomic applications to medical imaging have emerged as quantitative tools for personalized medicine[137][138], using non-invasive instruments to assess aspects of anatomy, physiology and pathology and to accurately define diseases' diagnosis, prognosis and suitable treatments. Quantitative measurements, such as quantitative MRIs (qMRIs), are particularly suited to this type of analysis since they go beyond conventional MRI measurements and carry information about tissue microstructure and pathophysiological properties, complementing - and possibly, in the future, taking over - traditional histology examinations[139], providing benefit to patients care.

Multiple measurements or *radiomic features* can be extracted from images, quantifying their characteristics in an automated high-throughput manner[140] and returning a potentially mas-

sive source of *biomarkers* of pathology. It is important to distinguish image measurements from biomarkers, accurately defined: in 1998, the National Institutes of Health Biomarkers Definition defined a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention"[141]. Indeed, they are objective and quantifiable descriptors of biological processes, able to consistently predict clinical outcomes and endpoints: interpretability and reproducibility are two crucial aspects. First, current hardware, software and computing developments allow us to manipulate images, and data in general, in infinite ways, which may be intriguing on one side but chaotic and confusing on the other. In medical applications, it is important to keep a clinical perspective driving the research[142]; validity and relevance of a biomarker, as its utility in clinical practice and ability to provide useful information, need to be confirmed[143]. Second, whenever an informative or relevant feature is identified, outcomes generalizability and replicability need to be guaranteed; as we are referring to descriptors representing physiological processes, they should be the same independently from the acquisition system or processing workflow. Stages from the scanner to the feature extraction are several and complex, and multiple parameters may influence the reliability of the results. Thus, to preserve the outcomes from one system to another, robustness analysis is crucial in radiomic applications[144][140].

Radiomics has been used to explore different medical image modalities, such as computed tomography[145][146][147] or photoacoustic imaging[148]. In brain MRI applications, it has been adopted in a few studies mainly to characterize brain tumors, using anatomical T_1w and T_2w images as support[149][150], and it has been shown that images contain information associated with treatment response, tumor grade and clinical outcomes[151]. However, radiomics is still a new method in susceptibility-based imaging and neurodegeneration applications. This work aims to analyze susceptibility-based features from QSM in a cohort of patients with MS.

4.1.2 Multiple sclerosis: background and aims

MS is an acquired autoimmune disorder characterized by demyelination in the Central Nervous System (CNS), affecting both the brain and the spinal cord. This condition presents a wide array of symptoms, including fatigue, sensory loss in the limbs, tingling sensations (paraesthesias), cognitive decline, and visual disturbances[152]. MS is a global health concern, impacting over 2 million individuals worldwide[153][152]; the disease typically emerges between the ages of 15 and 60, with a higher prevalence in women (accounting for three-quarters of MS cases). The course of MS can span decades, leading to a life expectancy that is 5 to 10 years shorter than that of healthy individuals. The triggers for MS can be both environmental and genetic[153]: factors such as geographic latitude, obesity, gender, early exposure to infections, smoking, and air pollution can contribute to the development of the condition. Genetic predisposition also plays a role, with approximately a 20% familial recurrence rate. There are over 200 genetic variants associated with an increased risk of MS, with the most significant one being the HLA DRB*1501 haplotype.

In the progression of MS, CNS exhibits focal lymphocyte infiltration, leading to damage to myelin and axons. In the initial stages, inflammation is transient, and the CNS can undergo remyelination following demyelination. However, over time, pathological changes driven by microglial activation result in chronic neurodegeneration, contributing to the progression of disability[153]. MS lesions can manifest in both White and Gray Matter (WM and GM), with cortical lesions typically being less inflammatory and displaying lower blood-brain barrier permeability. Depending on the location and timing of lesion appearance, clinical symptoms can vary, affecting systems such as motor, sensory, visual, and autonomic functions. Lesions are categorized into types, including periventricular WM, juxtacortical WM, corpus callosum, infratentorial brain (pons and cerebellum), optic nerve, and spinal cord[154][155].
The precise pathogenesis of MS remains an object of ongoing investigation, with several key questions that still need to be answered. Firstly, it is unclear whether the primary cause of the disease is intrinsic or extrinsic to CNS. Secondly, the role of neurodegeneration in MS remains enigmatic. In the course of MS, both inflammatory and neurodegenerative processes occur concurrently, but there is no consensus on whether they are independent of each other or if one precedes the other[153]. While MS was initially classified as a demyelinating disease, there is now growing support for the theory that acute and chronic axonal loss plays a central role in neurodegeneration. The progression of the disease is not only dependent on the occurrence of new lesions but also on the accumulation of axonal degeneration. In fact, individuals with MS often exhibit a significantly reduced parenchymal fraction in both the brain and spinal cord[152].

Currently, the diagnosis of MS is based on the McDonald diagnostic criteria, that combine clinical, imaging, and laboratory biomarkers[154]. Conventional MRI has been a cornerstone in the McDonald diagnostic criteria and in monitoring disease activity; however, conventional techniques often demonstrate established lesions lacking insights into the pathophysiological mechanisms leading to the demyelination[156]: focal or confluent alteration in WM, detectable with MRI, can be found in more than 95% of patients with MS, but lesion presence is not a necessary condition to diagnose the disease, since non-specific WM lesions often occur in $>$ 50yo healthy people. The recent development of different innovative qMRI techniques, capable of *in vivo* quantification of imaging biomarkers, is promising to explore brain microstructure[157] and metabolism[158] to unravel the pathophysiology of MS potentially revealing pre-clinical inflammatory demyelination providing new therapeutic windows. QSM, as a quantitative MR sequence, may help us explore both lesions and unlesioned parenchyma. Main findings achieved up to this point may be summarized as follows: a) chronic active or *smoldering* lesions, signs of an ongoing inflammatory process, are detected with QSM and show increased susceptibility at the edges because of iron-laden microglia and macrophages[61]; b) susceptibility values increase, meaning that iron concentration increases, in the basal ganglia and decrease in the thalamus, where they have been associated with higher disability[159]; c) differences have been highlighted in the Normal Appearing WM (NAWM)[21][22], analyzing the non-lesioned parenchyma in patients with MS and healthy controls; in WM, changes in susceptibility are related to ongoing demyelination processes[160].

This work[161][162] focuses on the NAWM[21][22], analyzing the non-lesioned parenchyma in patients with MS and healthy controls. In literature, differences have already been highlighted in this tissue, even in the absence of lesions, associated with the presence of activated microglia as a sign of inflammation[163][164]. Studying NAWM may help in the early diagnosis of the disease, the characterization of different clinical phenotypes of MS, and the distinction from other demyelinating diseases. Here, QSM-based radiomic features were extracted from NAWM and clinically relevant WM tracts in healthy controls and patients with MS, and a robustness analysis was performed to suggest potential features for use in classification, characterization, and prognosis. The reliability of susceptibility features was assessed against the number of gray levels (GLs) and echo times (TEs); all details of the implemented pipeline are described in the following sections. The dataset generated and analysed during the study[23] is available in the Zenodo repository (https://zenodo.org/record/8271881)

4.1.3 Materials and methods

Study sample and MR examinations

The analyzed sample included 121 MS patients (F:M 71:50, 48.7 ± 12.6 year old [23-75]) and 30 Healthy Controls (HCs) (F:M 17:13, 53.4 ± 17.9 year old [24-86]), totaling of 151 exams acquired at the IRCCS Istituto delle Scienze Neurologiche di Bologna (Bellaria Hospital) (IT) between February 2020 and September 2022. The sample included various MS clinical phenotypes (relapsing-remitting, primary and secondary progressive) to prevent any phenotype-biased results. Throughout the study, there was no substantial modification to the acquisition and processing software or the data analysis pipeline, ensuring uniformity and comprehensiveness of the sample. The sample size aligns with previous studies that have investigated susceptibility-based radiomic features[165][104][166]. The MR protocol included: MPRAGE, FLAIR, DWI and QSM. QSM acquisition and processing details are illustrated in Chapter 1 (Section 1.2.1). FLAIR, DWI, and QSM images were linearly registered with FLIRT[44][45] to the corresponding MPRAGE images. All the images did not present significant movement artifacts and were considered suitable for the analysis.

Pre-processing pipeline and ROI segmentation

WM tissue was segmented using the MRtrix tool `5ttgen`, based on FreeSurfer segmentation[48]. WM lesions were automatically segmented by the Lesion Prediction Algorithm[167] as implemented in the open-source Lesion Segmentation Tool (LST www.statistical-modelling.de/lst.html); FLAIR images were used as input data[168][169]. The toolbox provided an estimate for the lesion probability map, used to obtain a binary map of lesions: for each exam, the inverse of this map was multiplied by the WM mask to identify the NAWM.

Diffusion and tractography processing pipeline was previously implemented as described in[170][171][172].

DWI images were skull-stripped using BET[43] and denoised using `dwidenoise`, function of MRtrix3[47], with a principal component analysis approach. Susceptibility-related distortion was estimated using `topup` function of FSL, and correction for susceptibility, eddy currents effects, and signal dropout was performed (`eddy_openmp`[173] function of FSL).
For WM tract reconstructions, diffusivity was modeled along the spatial eigenvector, employing the tensor model, a high-order fiber modeling technique and a probabilistic streamline approach for evaluating crossing fibers. Diffusion images were used to estimate fiber orientation distributions by the single tissue, single shell spherical deconvolution algorithm 'csd' implemented by the `dwi2fod` function from MRtrix3[174]. Six WM tracts were defined: arcuate fasciculus (AF), cortico-spinal tract (CST), frontal aslant tract (FAT), inferior fronto-occipital fasciculus (IFOF), optic radiation (OR), uncinate fasciculus (UF). For each tract, seed and inclusion regions of interest for streamline generation and selection were defined on the Montreal Neurological Institute's 152 (MNI152) standard brain. Non-linear registration between DWI and MNI spaces was performed using FNIRT. Streamlines were generated using the iFOD1 method `tckgen` from MRtrix3[47], using criteria such as entering inclusion regions, FOD amplitude, and deviation angle. Tracts were individually reconstructed for the two hemispheres and then combined into a single tract, converted into a NIfTI image with voxel intensity representing the streamline count. To obtain a relative fiber count estimate image, a threshold was applied (referring to the maximum value, 10% for AF and 2% for the others). As NAWM, tract VOIs were analyzed excluding the lesions.

Details of software used are in Chapter 1 (Section 1.2.1). The pipeline was entirely automated and it is illustrated in Figure 4.1.

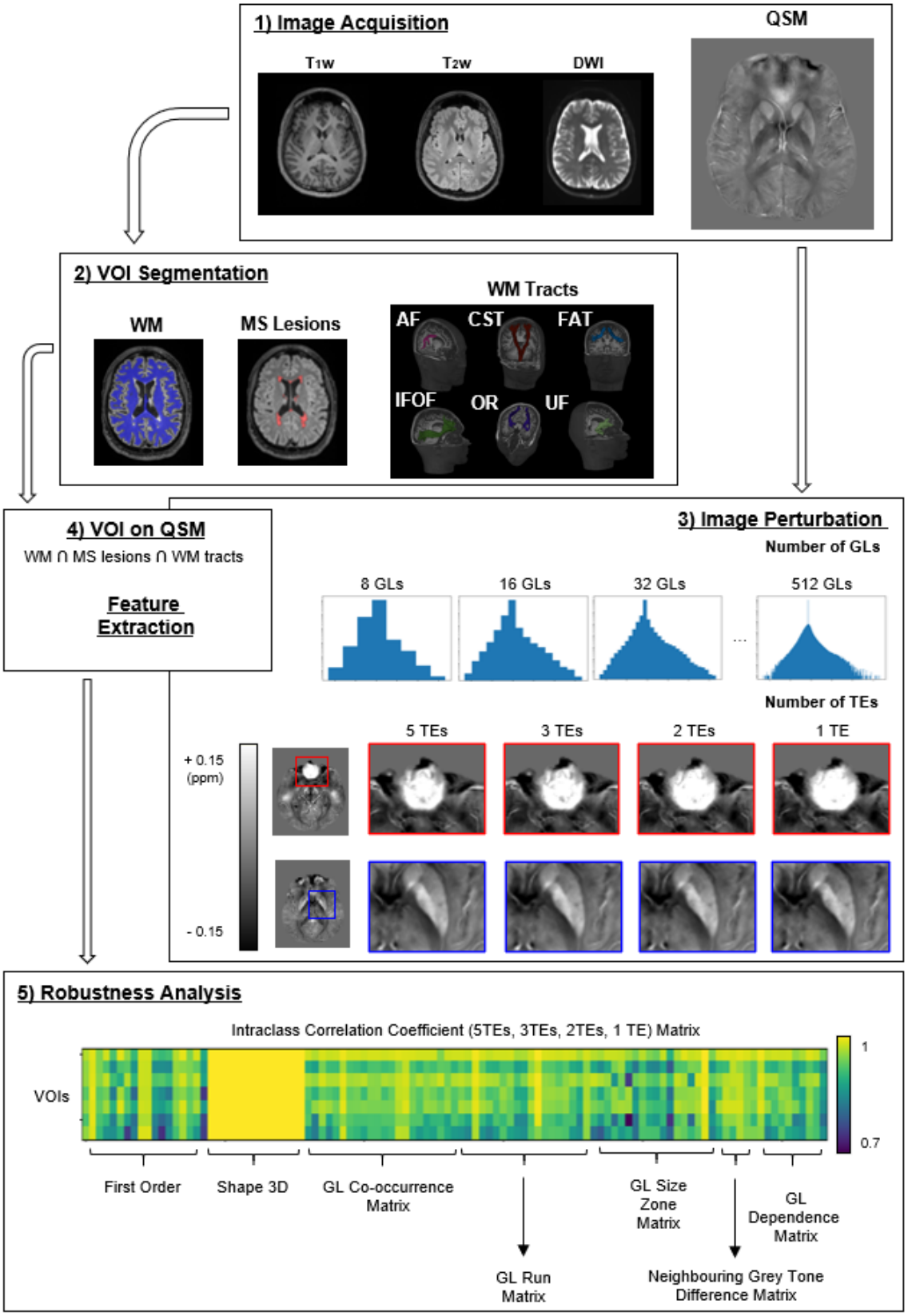

Figure 4.1: Scheme of the workflow[23]: 1) MPRAGE, FLAIR, DWI and QSM acquisition (M/25y, HC); 2) WM and lesions segmentation from FLAIR (M/51y, MS) and WM tracts reconstruction from DWI (M/25y, HC); 3) image perturbation, changing the number of Gray Levels (GLs) and of Echo Times (TEs) (please note changes in artifacts (red boxes) and contrast (blue boxes) moving from 5TEs to 1TEs) 4) VOIs overlap on QSM images and feature extraction; 5) Robustness analysis (ICC evaluation); note that ICC values for 3D shape features are equal to 1 for all the ROIs and all the measurements, as the shape of the regions was not change during the work and they depend only for it. For this reason, this class of features is not considered for the next steps of the work (AF = Arcuate Fasciculus, CST = Cortico-Spinal Tract, FAT = Frontal Aslant Tract, IFOF = Inferior Fronto-Occipital Fasciculus, OR = Optic Radiation, UF = Uncinate Fasciculus)

Radiomic features extraction

Because the study data were collected using a consistent acquisition protocol, clinical scanner, and processing pipeline, all within a single center, the sample exhibited a high degree of homogeneity, and the need for histogram normalization[151] was deemed unnecessary.
PyRadiomics (version 3.0.1)[175], running on Python (version 3.7.6), was used to extract features from 3D VOIs overlapped to QSM images; the features extracted were 107, divided into the following categories:

- **First-order features (FO, # 18)**: descriptors of the histogram intensity through commonly used metrics, such as mean, median, 10^{th} and 90^{th} percentile, skewness, kurtosis, and so on; they are independent of the number of GLs
- **Shape 3D (S3D, # 14)**: descriptors of the 3D size and shape of the ROI (e.g. volume, surface, minimum and maximum axes); they are independent of the number of GLs and their intensity distributions
- **Gray Level Co-occurrence Matrix (GLCM, # 24)**: they describe the second-order joint probability function of an image region constrained by the mask.
- **Gray Level Run Length Matrix (GLRLM, # 16)**: they quantify GL runs (number of consecutive pixels that have the same grey level value)
- **Gray Level Size Zone Matrix (GLZM, # 16)**: they quantify GL zones in an image (number of the connected voxels that share the same grey level intensity)
- **Neighboring Gray Tone Difference Matrix (NGTDM, # 5)**: they quantify the difference between a GL value and the average grey value of its neighbors
- **Gray Level Dependence Matrix (GLDM, # 14)**: they quantify GL dependencies (number of connected voxels within a distance that are dependent on the central voxel)

Categories 3 to 7 are referred to as texture, and they give information about the spatial distribution of intensity levels in the image; a complete list of features can be found in PyRadiomics documentation (https://pyradiomics.readthedocs.io/en/latest/features.html) and in[23].

Robustness of radiomic features

The reliability of radiomic features was assessed by evaluating: a) different numbers of GLs in the images, which affects texture features, and b) different numbers of echo times considered to reproduce the susceptibility maps. The number and values of TE change a lot from pipeline to pipeline and their setting affects contrast and intensities in the resulting map. Normally, a compromise needs to be found between high (major contrast, both in well-reconstructed tissue and in artifacts areas) and low (more details) values and the number of echoes, which increases as the signal-to-noise ratio and the acquisition time.

Seven different re-binning (2^n , $n \in [3, 4, \ldots, 9] = 8, 16, 32, 64, 128, 256, 512$ GLs) and four different QSM reconstructions were considered, using 5 (TE_i, $i \in$ [1,2,3,4,5]), 3 (TE_i, $i \in$ [1,3,5]), 2 (TE_i, $i \in$ [1,5]) and 1 (TE_5) echo times (TEs = 9.42, 18.84, 28.26, 37.68 and 47.10 ms). To evaluate robustness, the Intraclass Correlation Coefficient ICC[96] was measured: its higher value is equal to 1 when there is perfect repeatability, while decreasing values from 1 to 0 show a decrease in resemblance. Features considered with good reliability were the ones corresponding to ICC > 0.85, set as threshold. Measurements were evaluated considering the left and right hemispheres together at first and then individually; eventual differences between the two sides were assessed using the two-sample t-test, considering significant $p - values < .05$.
First, the optimal number of GLs was estimated and then changes in ICC between the different susceptibility maps were evaluated. Shape 3D features are independent of the intensity values; ICC values are equal to 1 in all the measurements (Fig. 4.1), thus this class of features was not considered further in this study.

4.1.4 Results

The optimal number of GLs in QSM images was evaluated as the one returning the highest number of robust features. For each binning, from 8 to 512, ICC values between the 4 QSM reconstructions were measured in all the VOIs (NAWM Tracts: AF, CST, FAT, IFOF, OR, UF, and NAWM; 7 VOIs · 107 features = 749 features in total). Different threshold levels were explored at this stage, from 0.85 to 0.90; results are in Tab. 4.1. At all threshold levels tested, the maximum number of robust features was obtained using 64 GLs. The same analysis was repeated extracting features by side (2 sides · 7 VOIs · 107 features = 1498) and considering patients and control groups independently. The results confirmed 64 GLs as the optimal binning.

N° GLs	**ICC threshold**					
	0.90	**0.89**	**0.88**	**0.87**	**0.86**	**0.85**
8	461	489	517	543	556	568
16	525	570	605	624	636	648
32	545	574	603	624	640	659
64	555	589	613	626	652	667
128	533	584	607	622	635	654
256	531	568	592	612	623	640
512	540	576	593	608	621	631

Table 4.1: Number of robust features *vs* number of gray levels used to quantize the images[23]. Features were evaluated for all the VOIs (6 NAWM tracts and NAWM), leading to a total number of 749 (= 7 VOIs · 107 features). A feature was considered robust when the ICC over the 4 QSM reconstructions was higher than a set threshold, from 0.85 to 0.90

Robustness analysis

In Fig. 4.2, robust features in NAWM Tracts and NAWM are shown divided into different categories. In NAWM, ∼65% of features were robust; tracts generally work better in terms of robust features, between 75% and 100%. There are differences in the different tracts: AF, FAT, and OR

(∼ 95 to 100% robust features) showed better results than CST, IFOF and UF (∼ 75 to 90%). Thus, further analysis was carried out to explain those differences, and it is illustrated in the subsequent subsection. To highlight eventual differences between the left and right hemispheres, the analysis was carried out by dividing the left and right sides for each ROI. The left hemisphere showed slightly better results; however, differences between the two sides can be considered negligible ($p > .05$). The same results were obtained for the entire NAWM ($p > .05$).

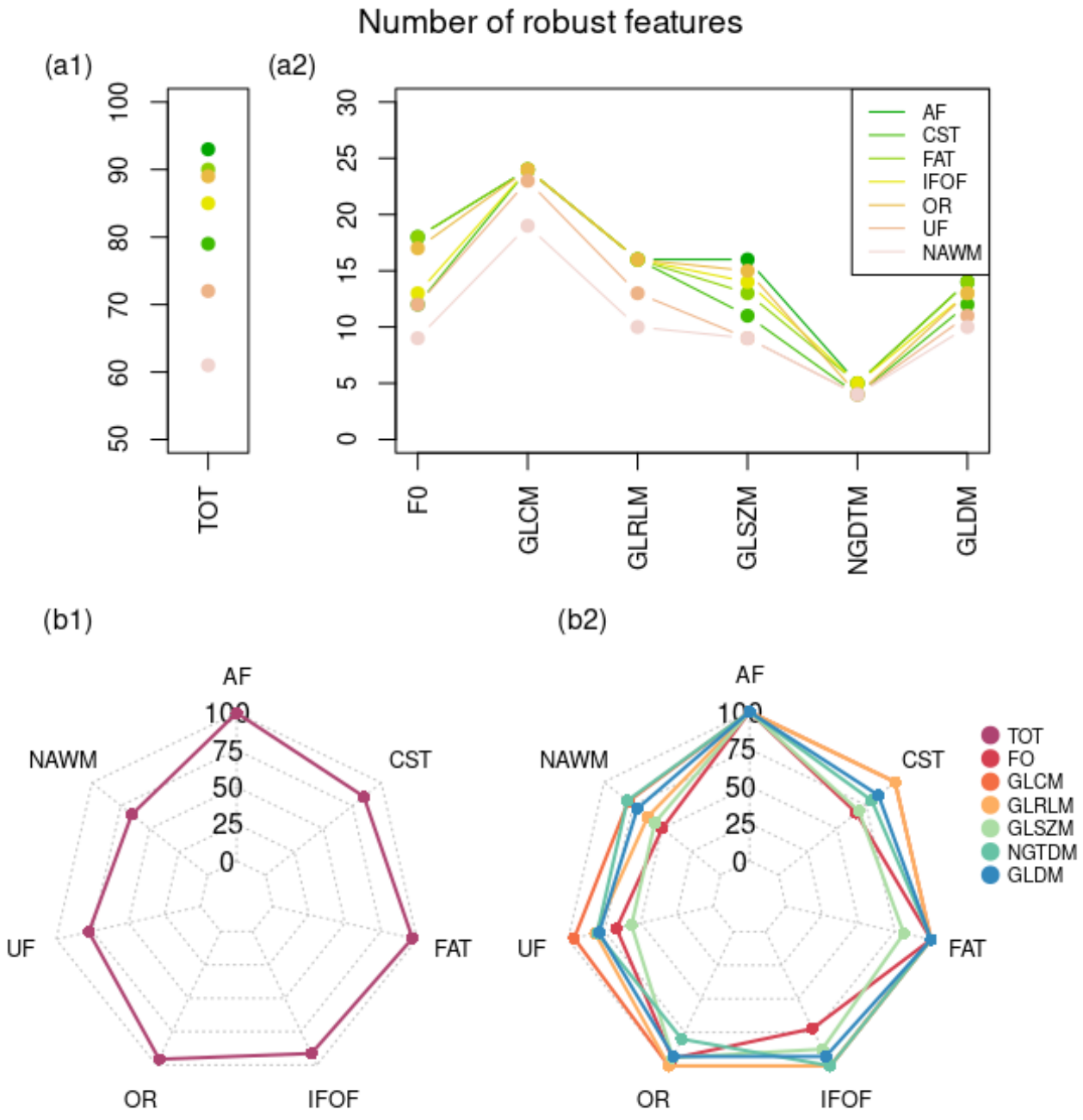

Figure 4.2: Robust features in NAWM and NAWM tracts[23]. Features were extracted setting 64 as binning, and ICC was evaluated between the 4 QSM reconstructions, setting 0.85 as threshold. First row: line plot with number of robust features, total (a1) and considering the different feature categories (a2); second row: spider graph with the percentage of the number of robust features, total (b1) and considering the different feature categories (b2) (AF = Arcuate Fasciculus, CST = Cortico-Spinal Tract, FAT = Frontal Aslant Tract, IFOF = Inferior Fronto-Occipital Fasciculus, OR = Optic Radiation, UF = Uncinate Fasciculus, FO = First-Order, GL = Gray Level, GLCM = GL Co-occurrence Matrix, GLRLM = GL Run Length Matrix, GLZM = GL Zone Matrix, NGTDM = Neighboring Gray Tone Difference Matrix, GLDM = GL Dependence Matrix)

Different NAWM VOIs were analyzed in this study, considering individually left and right hemispheres. Below, a list of features reliable in all the considered regions is given:

- FO (50%, 9/18): 10^{th} and 90^{th} percentile, energy and total energy, mean, median and root mean squared, skewness and kurtosis
- GLCM (∼80%, 19/24): cluster shade, contrast, correlation, difference average, difference entropy, difference variance, inverse difference, inverse difference moment, inverse different moment normalized, inverse difference normalized, information measure of correlation 1 and 2, inverse variance, joint average, joint energy, joint entropy, maximal correlation coefficient, maximum probability and sum average
- GLRLM (∼65%, 10/16): GL non uniformity, long run emphasis, long run high GL emphasis, run entropy, run length non uniformity, run length non uniformity normalized, run percentage, run variance, short run emphasis, short run high GL emphasis
- GLSZM (25%, 4/16): GL non uniformity, size zone non uniformity, zone entropy and zone percentage
- NGDTM (80%, 4/5): busyness, coarseness, contrast, strength
- GLDM (∼65%, 9/14): dependence entropy, dependence non-uniformity, dependence non uniformity normalized, dependence variance, GL non uniformity, large dependence emphasis, large dependence high GL emphasis

First-order statistics analysis

Volume may play a significant role in robustness, indeed, NAWM, which has the largest volume, showed less robust features. Fig. 4.3 shows the volume distribution of the NAWM tracts: CST and IFOF, which showed fewer robust features compared to other tracts (e.g. AF and FAT), showed greater volumes, suggesting that the assumption that volume may be influencing the number of robust features is maintained. Volumes of the structures of interest were corrected by using the proportional adjustment method[99]. UF also showed a lower number of robust features, but as opposed to CST and IFOF, its volume distribution was comparable with those of AF, FAT, and OR. Hence, I focused on the first-order features to further investigate the lower robustness of the UF.

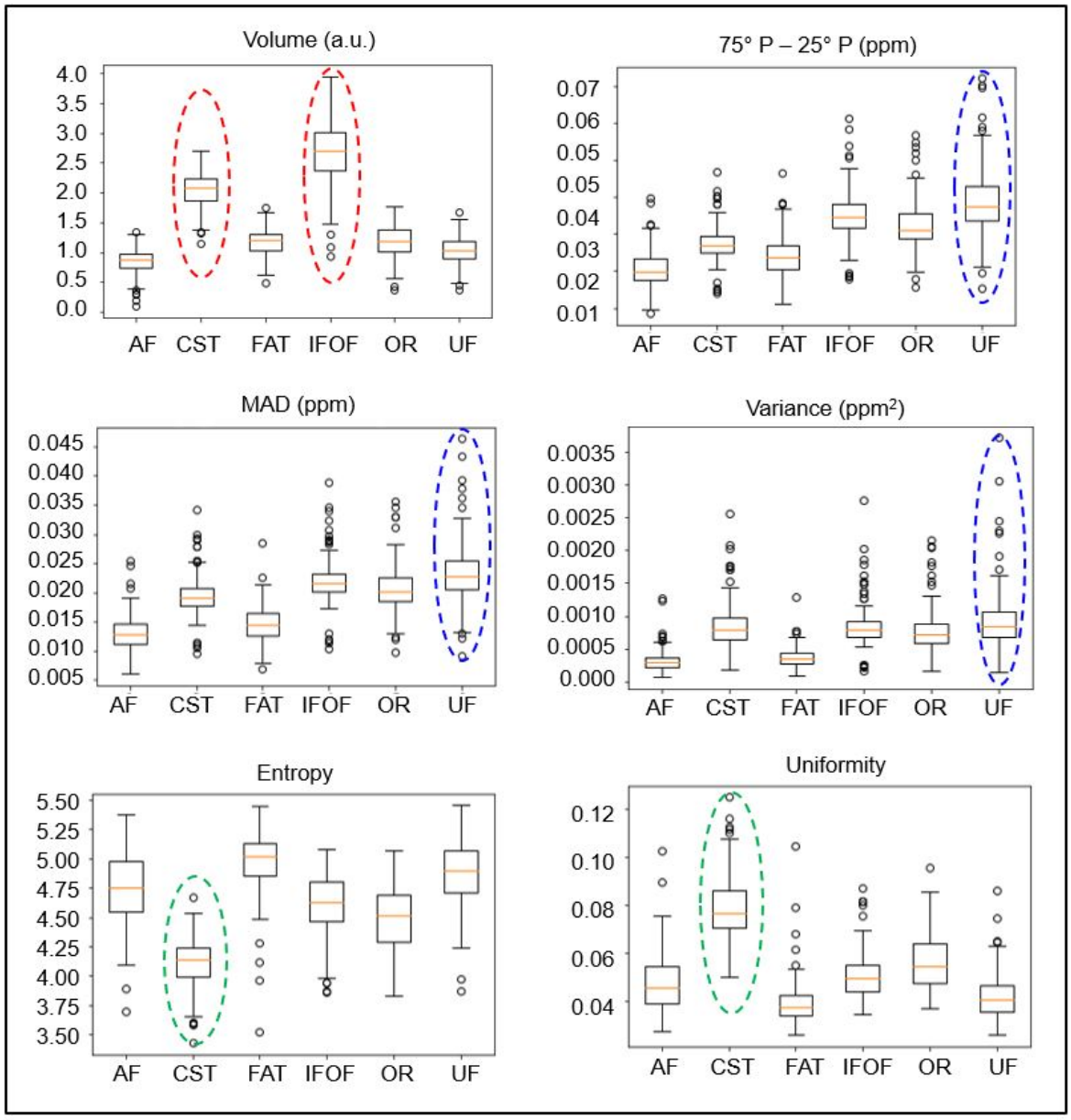

Figure 4.3: Box plots of volume (corrected by the total brain volume, proportional method), interquartile range (= 75^{th} percentile – 25^{th} percentile), MAD (Mean Absolute Deviation), variance, entropy and uniformity (first-order statistics features) distributions in the NAWM tracts (5TE-QSM) among the analyzed sample [23]. CST and IFOF are the more extended tracts (red circle); UF is the one more inhomogeneous in terms of range, MAD, and variance (blue circles); CST is the tract with lower entropy and higher uniformity (green circles), thus showing more uniformity in terms of texture despite its wide extension (AF = Arcuate Fasciculus, CST = Cortico-Spinal Tract, FAT = Frontal Aslant Tract, IFOF = Inferior Fronto-Occipital Fasciculus, OR = Optic Radiation, UF = Uncinate Fasciculus)

ICC values for first-order features in the NAWM VOIs are reported in Tab. 4.2. Review of individual values confirmed that, as mentioned in the previous section, AF, FAT, and OR showed higher robustness, while CST, IFOF, and UF, were more variable, similar to the overall NAWM. Most of the less-robust features in those VOIs (75^{th} P – 25^{th} P, Mean Absolute Deviation, Range, Robust Mean Absolute Deviation, and Variance) are related to the degree of spread of the intensity histogram, which is generally higher in UF distributions (Fig. 4.3, '75° - 25° P', 'Mean Absolute Deviation' and 'Variance' panels). Hence, a QSM-contrasted atlas was built in the MNI152 space using QSM images of healthy controls to explore spatial relationships that could influence intensity values. In Fig. 4.4, the variance of the susceptibility values among the sample is shown voxel-by-

voxel; the regions with more variability were areas prone to streak artifacts, i.e., closer to air-tissue boundaries or cortical bone, and the basal ganglia. Fig. 4.4 also shows individual variance maps for each tract: CST, IFOF, and UF include areas with higher variance.

First-Order Features	AF	CST	FAT	IFOF	OR	UF	NAWM
10^{th} P	**0.962**	**0.946**	**0.948**	**0.919**	**0.942**	**0.909**	**0.947**
90^{th} P	**0.985**	**0.982**	**0.984**	**0.988**	**0.986**	**0.986**	**0.982**
Energy	**0.936**	**0.935**	**0.940**	**0.915**	**0.922**	**0.906**	**0.892**
Entropy	**0.976**	**0.908**	**0.968**	**0.928**	**0.952**	**0.888**	*0.839*
75^{th} P – 25^{th} P	**0.906**	*0.830*	**0.899**	*0.810*	**0.883**	*0.815*	*0.834*
Kurtosis	**0.986**	**0.907**	**0.874**	**0.956**	**0.982**	**0.921**	**0.881**
Maximum	**0.944**	**0.916**	**0.962**	**0.903**	**0.914**	**0.852**	*0.729*
MAD	**0.893**	*0.842*	**0.887**	*0.783*	**0.870**	*0.791*	*0.811*
Mean	**0.979**	**0.973**	**0.973**	**0.969**	**0.974**	**0.961**	**0.983**
Median (50^{th} P)	**0.979**	**0.972**	**0.972**	**0.970**	**0.973**	**0.960**	**0.982**
Minimum	**0.906**	*0.832*	**0.909**	**0.855**	**0.863**	*0.814*	*0.802*
Range	**0.893**	*0.833*	**0.914**	*0.833*	**0.853**	*0.780*	*0.717*
Robust MAD	**0.902**	*0.828*	**0.896**	*0.805*	**0.880**	*0.811*	*0.830*
RMS	**0.964**	**0.947**	**0.953**	**0.930**	**0.942**	**0.925**	**0.903**
Skewness	**0.988**	**0.943**	**0.978**	**0.943**	**0.981**	**0.958**	**0.956**
Total Energy	**0.936**	**0.935**	**0.940**	**0.915**	**0.922**	**0.906**	**0.892**
Uniformity	**0.979**	**0.919**	**0.969**	**0.936**	**0.956**	**0.893**	*0.845*
Variance	**0.872**	*0.848*	**0.860**	*0.724*	*0.843*	*0.741*	*0.732*

Table 4.2: ICC values (n° GLs = 64) for the First-Order features in the analyzed NAWM tracts and overall VOIs [23]. Considering 0.85 as the threshold, the robust features are highlighted in bold, the others in italic (P = Percentile, MAD = Mean Absolute Deviation, RMS = Root Mean Squared, AF = Arcuate Fasciculus, CST = Cortico-Spinal Tract, FAT = Frontal Aslant Tract, IFOF = Fronto-Occipital Fasciculus, OR = Optic Radiation, UF = Uncinate Fasciculus)

Texture analysis

Among the first-order features, distributions of uniformity, which measures the homogeneity of the intensity values, and entropy, which measures the randomness in the image values, showed lower and higher median values in CST than other tracts (Fig. 4.3). Details in the intensity values distributions are better described by the texture features, among which several measurements refer to fine or coarse texture in the images. Many of those agree on the fact that CST, one of the wider white matter tracts analyzed, is the one reporting the less fine texture in GLCM, GLRLM, GLSZM, NGTDM and GLDM categories, followed by IFOF and OR. In Tab. 4.3, a list of features supporting this observation.

GLRLM, GLSZM and GLDM describe the distribution of the pixel measuring similar properties of different matrices, evaluating respectively length runs, size zones and dependences. There are correspondences between the three classes with respect to the texture homogeneity of CST: long run emphasis, long area emphasis and large dependence emphasis all show higher values for CST, run and zone percentage both show higher values for CST; GL variance, measuring the variance in

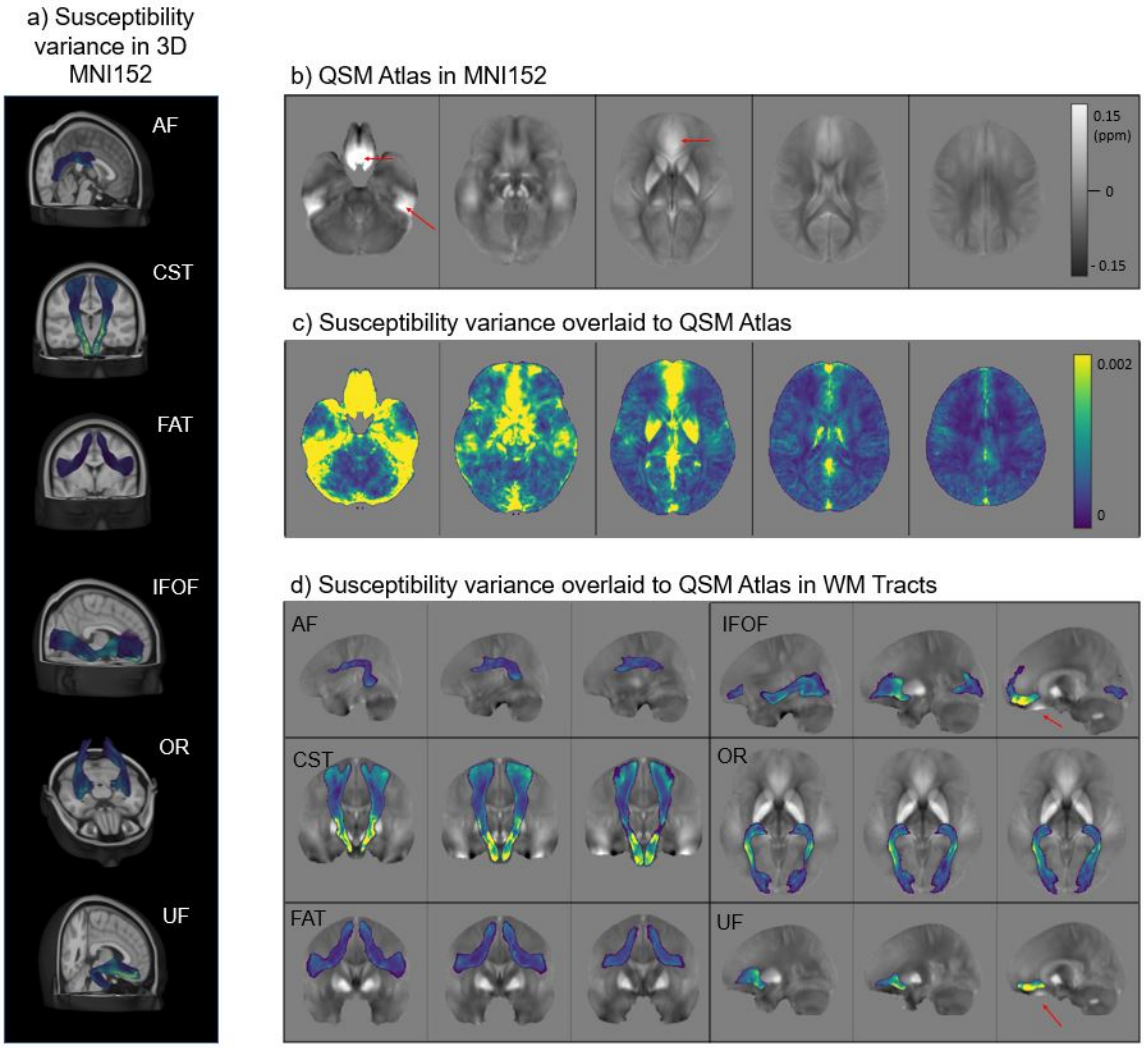

Figure 4.4: QSM atlas was obtained non-linearly registering 30 healthy control exams to the MNI152 space and a pixel-wise variance of the susceptibility measurements map was evaluated [23] - a) susceptibility variance map in the 3D MNI 152 in the white matter tracts; b) QSM atlas; c) susceptibility variance map overlaid to QSM atlas; d) susceptibility variance map reported individually for each WM tract, overlaid to QSM atlas, with the most representative projection for each of them (sagittal for AF, IFOF, and UF; coronal for CST and FAT; axial for OR). Red arrows in a) and c) point to the hyperintensities artifacts areas in QSM reconstructions (AF = Arcuate Fasciculus, CST = Cortico-Spinal Tract, FAT = Frontal Aslant Tract, IFOF = Inferior Fronto-Occipital Fasciculus, OR = Optic Radiation, UF = Uncinate Fasciculus)

GL intensity respectively for the runs, the sizes and the dependences, is lower in CST distributions for the three categories. CST is wide and so the variance in terms of intensity values is high. GL Non-Uniformity, which measures the similarity in intensity values, is higher for CST and IFOF for the three categories of GLRLM, GLSZM and GLDM. Other tracts, such as AF and FAT, which are more reliable in terms of number of robust features, return measurements corresponding to a more fine texture: for example, Contrast in GLCM is higher in AF and FAT; Difference Entropy in GLCM, lower in CST, is higher in AF and FAT; Complexity in NGDTM, lower in CST and IFOF, is higher in FAT.

Texture category	Feature	Trend
GLCM	Difference Entropy	↓ CST
	Maximal Correlation Coefficient	↓ CST
	Joint Entropy	↓ CST
	Joint Energy	↑ CST
GLRLM	Short Run Emphasis	↓ CST
	Long Run Emphasis (*)	↑ CST
	Run Percentage (**)	↑ CST
	Gray Level Variance (***)	↓ CST
	Run Entropy	↓ CST, OR
	Run Length Non-Uniformity Normalised (****)	↓ CST
GLSZM	Long Area Emphasis (*)	↑ CST
	Zone Percentage (**)	↑ CST
	Gray Level Variance (***)	↓ CST, IFOF, OR
	Zone Entropy	↓ CST, OR
NGTDM	Complexity	↓ CST, IFOF
GLDM	Small Dependence Emphasis	↓ CST
	Large Dependence Emphasis (*)	↑ CST
	Dependence Non-Uniformity Normalised (****)	↓ CST
	Gray Level Variance (***)	↓ CST

Table 4.3: Second-order features measuring fineness, coarseness and texture homogeneity [23]; their trend shows that the CST is the one reporting less fine texture and more homogeneity in terms of texture, despite its wider histogram distribution (CST = Cortico-Spinal Tract, GL = Gray Levels, GLCM = GL Co-occorrence Matrix, GLRLM = GL Run Length Matrix, GLSZM = GL Size Zone Matrix, NGDTM = Neighboring Gray Difference Tone Matrix, GLDM = GL Dependance Matrix). * were signed for corresponding features in different categories.

4.1.5 Discussion

This study has illustrated the reliability of radiomic features derived from QSM images when evaluated within NAWM of both patients with MS and controls, both overall and within selected tracts. Nevertheless, some variability was noted among different structures. In the investigation, a few factors were identified as potentially influencing this robustness, to aid in the future selection of anatomical volumes of interest. Furthermore, we pinpointed a subset of these robust radiomic features that merit further exploration as potential tools for clinical decision-making support.
The analysis incorporated a mixed sample of patients and healthy controls; thus the reliability of radiomic features was evaluated regardless of subjects' clinical condition. Future work will help find differences between controls and patients and different clinical phenotypes. Results are discussed below, highlighting the key aspects of the study.

Intensity normalization

Conventional MR images exhibit arbitrary, non-calibrated grey intensity values, normally necessitating image normalization before conducting feature extraction. Various methods have been proposed for this purpose, and the chosen normalization method can directly impact the pre-

dictability of features extracted from the normalized images[176]. QSM is an intrinsically quantitative sequence, and, in this study, all examinations were conducted using the same scanner and processed with a standardized pipeline. In this way, an extremely homogeneous dataset was generated, eliminating the need for intensity histogram normalization[151].

Optimal number of gray levels in QSM

The intensity values within QSM images directly represent the magnetic susceptibility of a specific voxel. The manipulation of visualization parameters influences how these images are perceived, primarily due to the limited capacity of the human eye to perceive gray levels. However, these adjustments do not alter the quantitative values representing the magnetic susceptibility of the examined structures. Consequently, the first stage of the analysis was to examine how different numbers of GLs (bin number) affect the robustness of radiomic features while keeping the bin size constant. This method aligns with the recommendations of the Image Biomarker Standardisation Initiative[140], particularly when working with quantitative data, as it preserves a direct connection with the original intensity values. The number of robust features reaches its peak at 64 GLs (Tab. 4.1), that was selected as the optimal quantization level for QSM and used for subsequent assessments of robustness. Interestingly, a recent study exploring the robustness of radiomic features in computed tomography images of the liver and muscles, which also falls under the category of quantitative imaging modalities, arrived at similar conclusions. It identified the optimal range of GLs to be between 32 and 64[147], underscoring the need for a sufficient number of GLs to account for variations among neighboring pixels and produce meaningful results.

QSM robust features in NAWM

After fixing the number of gray levels in QSM reconstructions, it was evaluated that at least 55 descriptors for each volume of interest are robust and potentially informative for future clinical study (Fig. 4.2); a full list is given in the previous section. Features were evaluated first by considering together the left and right hemispheres: 65% of radiomic features were robust within the entire NAWM with no fewer than 50% in each class. Texture classes showed higher robustness (NGTDM 80%, GLCM 79.2%, GLDM 71.4%, GLRLM 62.5%, GLSZM 56.3%), compared to first-order statistics (50%). Analysis of WM tracts showed that more than 75% of features were robust within all the analyzed NAWM tracts. Texture classes GLCM (99.3%) and GLRLM (96.9%) showed a higher percentage of robust features compared to GLDM (91.7%), NGTM (90.0%), FO (83.3%) and GLSZM (81.3%).

It is noteworthy to consider that the ISMRM Electro-Magnetic Tissue Properties Study Group is working on a consensus paper about QSM acquisition and processing[4], but still, the number and values of TEs change from one acquisition system to another. This study revealed a significant level of agreement among the results obtained from susceptibility maps reconstructed using different echo times, even when considering maps generated with a single TE of 47.1 ms. However, while this observation might imply that conducting radiomics with a limited number of TE values would be adequate, such an approach would entail a loss of valuable QSM-derived information,

potentially offering distinct insights. For example, shorter echo times enhance the visualization of fine details, proving particularly beneficial in identifying and visualizing small demyelinating lesions in patients with multiple sclerosis. Additionally, a combination of short and long echo times serves to balance image contrast and minimize image degradation due to streak artifacts. Furthermore, using multiple TE values enhances the signal-to-noise ratio for various tissues and structures in the brain[177]. Therefore, the utilization of multiple TE values not only contributes to the stability of radiomic analysis but also aids in the conventional evaluation of the resulting radiological images.

Radiomics and QSM: brief overview

There is a limited number of investigations of the application of radiomics to QSM. For instance, Xiao et al. in 2019[165] developed a machine learning algorithm capable of distinguishing patients with Parkinson's disease (PD) from healthy controls by using radiomic features extracted from the substantia nigra. More recently, Kang et al. in 2022[166] compared radiomic features extracted from the substantia nigra, the head of the caudate nucleus, and the putamen in patients with PD and controls, confirming that features derived from the substantia nigra were more effective in diagnosing PD. Both studies concentrated on small but clinically relevant regions of interest and did not formally evaluate the reproducibility or repeatability of their findings. Zhang et al. in 2022[104] employed a deep convolutional neural network that integrated lesion-level radiomic and convolutional image features for the automated identification of chronic active MS lesions (QSMRim-Net), defined by the presence of a peripheral rim of iron-laden activated microglia/macrophages. QSMRim-Net demonstrated promising results in detecting these lesions, although the training and validation dataset was relatively limited, and rim-positive lesions were relatively scarce. Similarly, Yan et al. in 2021[178] constructed a machine learning model that combined radiomic features extracted from deep grey matter regions and demographic characteristics to differentiate between patients with MS and those with neuromyelitis optica spectrum disorder (NMOSD) with a high degree of accuracy. However, the model solely relied on image information from deep grey matter structures not typically associated with NMOSD, which raised concerns about the possibility of overestimating its ability to distinguish between the two disorders[179]. Additionally, the choice of using deep grey matter regions may have been influenced by the limited presence of lesions with rim-like paramagnetic phase changes in NMOSD patients[180].

Despite the limitations of these studies, there is a growing potential for applying sophisticated analytical approaches to QSM data in the context of neurodegenerative and neuroinflammatory disorders. The robustness assessment outlined in this study could significantly enhance future investigations. For example, since both NMOSD and MS affect the corticospinal tract, this study identified a set of reliable features that could be employed for assessing changes. This could pave the way for computer-assisted differentiation between these two disorders. Furthermore, even though this study did not assess microstructural information, the capability to evaluate measurements from DWI sequence is integrated into the pipeline, enabling the incorporation in the future of susceptibility and diffusion measurements.

Radiomic features influenced by volume

Prior research has highlighted the significant impact of the volume on radiomic features [181] [182]; specifically, first-order features are influenced by larger volumes due to the increased number of voxels contained within the VOI [183]. Consequently, given that NAWM constituted the largest VOI in our analysis, a higher degree of variability in terms of robustness was expected, which was indeed confirmed with the analysis. A recent study investigating the reproducibility of QSM-derived radiomic features when examining intramural hematomas and atherosclerotic calcifications of intracranial vertebral arteries also observed a dependence on volume for certain features. In this study, the smaller volume of the VOIs jeopardized the robustness of these features [184].

This information assumes particular significance when comparing the NAWM between patients with MS and healthy controls. Brain volume loss, which occurs at an accelerated rate in patients with MS, can result in different sizes of NAWM between the two groups. This discrepancy introduces a potential confounding factor due to the volume difference in the NAWM. Ideally, dissimilarly sized VOIs should not be compared when the chosen metrics are sensitive to volume effects. If, for some specific reason, metrics sensitive to volume effects cannot be avoided, as is the case in this context, a normalization to account for volume dependency must be integrated into the pipeline when VOIs have varying sizes [147] [181] [185].

Despite their sensitivity to volume effects, tract-based analysis within the analyzed cohort indicated that most first-order statistic descriptors (including the 10^{th} and 90^{th} percentiles, mean, median, skewness, and kurtosis) exhibited robustness in all the examined tracts, even during side-based sub-analysis. Additionally, these first-order statistic descriptors outperformed texture descriptors. These statistic descriptors are associated with tissue bulk susceptibility and are directly proportional to the concentration of biomarkers. Notably, susceptibility values correlate positively and negatively with the concentration of iron and myelin, both of which are altered in various neurological disorders, including MS, as indicated by [24]. This robustness of first-order statistic descriptors aligns with recent studies that have shown high reproducibility of susceptibility values across different sites [186] [187], highlighting the suitability of this technique for conducting multicenter, cross-vendor, quantitative analyses of brain iron and myelin.

Radiomic features influenced by χ variance

Upon examining the statistical characteristics of individual tracts, our analysis divided the NAWM tracts into two distinct clusters: AF, FAT and OR exhibited a higher degree of robustness, with 100%, 96.8%, and 95.7% of robust features, respectively; in contrast, CST, IFOF and UF demonstrated less robustness, with 84.9%, 91.4%, and 77.4% of robust features, respectively. As discussed earlier, the volume of the VOIs has a noticeable influence on radiomics features. Indeed, an examination of the volume distribution among these tracts (Fig. 4.3) affirmed that CST and IFOF had the largest volume among the assessed white matter tracts. However, this volume-dependence doesn't elucidate the performance of UF, which had a volume that was not greater than that of AF or FAT.

Compelling evidence from a susceptibility variance map generated using QSM images of healthy volunteers (Fig. 4.4) indicated that voxels along the CST, IFOF, and UF exhibited greater variability, implying that radiomic features in these areas would be less reliable. Furthermore, the variance map demonstrated that regions with higher variability were situated adjacent to areas prone to streak artifacts[59][95]. These artifacts were typically found near air-tissue boundaries, cortical bone, or close to high-susceptibility sources, such as the basal ganglia, which is known to accumulate susceptibility-inducing compounds with age[188][70]. While reconstructed tracts should not overlap with these artifact-prone areas, the stability of features based on gray level intensity may be affected by their proximity. In particular, the susceptibility variance of both IFOF and UF peaked in the anterior cranial fossa, which is particularly susceptible to streaking artifacts. The UF may be more adversely affected due to its smaller volume, resulting in a higher overall proportion of unreliable pixels.

CST: microstructural properties and χ features

Of distinct interest in the clinical context of MS, texture features within the cortico-spinal tract unveiled the inherent microstructural characteristics of the tract itself (Tab. 4.3). Notably, compared to other tracts (Fig. 4.3), the CST exhibited lower entropy, indicating a reduced level of randomness in susceptibility values. Additionally, the distribution of uniformity within the CST was higher than that in other tracts, implying a greater level of homogeneity.

A plausible explanation for these findings lies in the fact that white matter fibers in the CST exhibit a cranio-caudal orientation, whereas other tracts consist of a mixture of fibers with varying directions. Consequently, it is possible to infer that the uniform orientation within the CST leads to the uniformity observed in its texture features. Motor symptoms are highly prevalent in patients with MS and significantly impact their quality of life and independence. The robustness of radiomics features in the CST holds promise for the development of a radiomic signature specific to the CST. Such a signature could potentially be employed for prospective monitoring, enabling the early detection of changes in the pyramidal tract before the onset of symptoms and before these changes become evident through morphological imaging.

Limitations

This study is subject to a few limitations to be considered. Firstly, despite the overall large sample size, the patient subgroup significantly outnumbered the control group due to the hybrid clinical and research focus of our institution. To address this imbalance, a sub-group analysis within the patient and control groups was conducted, which confirmed that the independent results were consistent with the overall pooled analysis. This information may provide valuable guidance for determining sample sizes in future studies, particularly those focused on disease characterization, as achieving an equal number of patients and controls may be challenging.

Secondly, the sample studied in this study is from a single site, with all scans performed on the same clinical scanner. While this approach enhanced data homogeneity, it raises the possibility

that the radiomic features identified as robust in this study may be specific to datasets acquired using similar scanners and in similar settings. Assessing the robustness of a more heterogeneous dataset would offer a more comprehensive evaluation, decoupling the results from the specific acquisition and post-processing techniques. Nonetheless, now that robust radiomic features have been identified within this particular setting, the proposed pipeline can be employed in similar studies to verify the reproducibility of the outcomes across different hardware manufacturers and settings.

Lastly, an automated tool for lesion segmentation was used to maintain consistent accuracy and minimize potential bias throughout the analysis, avoiding the inherent variability that arises from manual segmentation[189]. This approach assured that the segmentation did not impact our results; however, it is important to note that automatic lesion segmentation tools may have limited accuracy, particularly when it comes to identifying small lesions (those smaller than 5mm), potentially leading to their inclusion within what was labeled as normal-appearing tissue.

4.1.6 Conclusions

In conclusion, to the best of my knowledge, this study represents the first attempt to assess the robustness of QSM-derived radiomic features in normal appearing white matter. Image perturbations - against the number of grey levels in the image and adjusting the number of echo times used to reconstruct quantitative maps - were conducted. Findings reveal that over 65% of the features demonstrated robustness across the entire NAWM. Moreover, the white matter tracts exhibited a higher level of robustness, with over 75% of features proving to be robust in all the white matter tracts. Differences among these tracts are likely attributed to the volume of the examined structure and the distribution of susceptibility variability, which can be influenced by neighboring anatomical features. This research sets the stage for future investigations that can leverage the set of robust features identified here to non-invasively characterize patients with multiple sclerosis, rapidly detect therapy responses, and monitor the progression of the disease.

AF = Arcuate Fasciculus
BET = Brain Extraction Tool
CNS = Central Nervous System
CST = Cortico-Spinal Tract
DWI = Diffusion Weighted Imaging
FAT = Frontal Aslant Tract
FLAIR = FLuid Attenuate Inversion Recovery
FLIRT = FSL's Linear Image Registration Tool
FO = First-order
FSL = FMRIB Software Library
GL = Gray Levels
GLCM = Gray Level Co-occurrence Matrix
GLDM = Gray Level Dependence Matrix
GLRLM = Gray Level Run Length Matrix
GLZM = Gray Level Size Zone Matrix
GM = Gray Matter
HC = Healthy Controls
ICC = Intraclass Correlation Coefficient
IFOF = Inferior Fronto-Occipital Fasciculus
LPA = Lesion Prediction Algorithm
LST = Lesion Segmentation Tool
MAD = Mean Absolute Deviation
MNI = Montreal Neurological Institute
MPRAGE = Magnetization Prepared RApid Gradient Echo
MR = Magnetic Resonance
MRI = MR Imaging
MS = Multiple Sclerosis
NAWM = Normal Appearing White Matter
NGTDM = Neighbouring Gray Tone Difference Matrix
NMOSD = NeuroMyelitis Optica Spectrum Disorder
OR = Optic Radiation
qMRI = quantitative MRI
QSM = Quantitative Susceptibility Mapping
RMS = Root Mean Squared
S3D = Shape 3D
T_1w, T_2w, T_2^*w = T_1-, T_2-, T_2^*- weighted
TE = Time of Echo
UF = Uncinate Fasciculus
VOI = Volume of Interest
WM = White Matter

4.2 QSM and Deep Learning: Super Resolution

Super resolution includes a class of techniques aiming to enhance image spatial resolution; they have been applied in the biomedical imaging area, including MR brain T_1w and T_2w images[190]. In this study, the performance of a super-resolution deep learning algorithm was evaluated on QSM, which has not been previously explored using this type of application.
28 healthy controls were selected; the MR protocol provided MPRAGE (T_1w), FLAIR (T_2w) and QSM. Enhanced-Deep-Super-Resolution (EDSR) was used as a deep learning model[191], trained with general-purpose images and directly applied on MR sequences, avoiding *re-training* or *fine-tuning* stages. EDSR outcomes were compared to traditional bicubic interpolation, using pSNR and SSIM as criteria: EDSR better reproduced the original images for all the sequences, under qualitative and quantitative evaluation.
EDSR was trained with general-purpose figures and still allows better reconstructions with respect to bicubic interpolation for MPRAGE, FLAIR and QSM. In the latter reconstructions, some issues in accurately reproducing high-intensity pixels were observed; thus, arrangements may be recommended before applying EDSR to those images, such as a tailored fine-tuning of the model.

4.2.1 SR with biomedical images: background and aims

Super Resolution (SR) algorithms aim to enhance the spatial resolution of low-resolution (LR) images, improving the detection of fine details and complex structures found in their high-resolution (HR) counterparts[192]. The process of up-sampling is an ill-posed problem by definition, as it involves increasing the amount of information from poor data while accounting for various potential sources of noise impacting LR images. These sources of noise can originate from hardware factors, such as acquisition sensors and acquisition time, as well as software factors like compression artifacts and stochastic noise[193]. Consequently, there are potentially infinite solutions linking LR and HR images, and the effectiveness of up-sampling algorithms should be assessed based on their capacity to preserve the original information or enhance the initial degradation.

Various types of SR models have emerged over the years. Deep Learning (DL) SR methods were first introduced by Dong and colleagues in 2014[194], as they initiated the use of Convolutional Neural Networks (CNNs), a category of artificial neural networks well-suited for numerous visual image-related tasks. Subsequently, Residual Neural Networks (ResNet) have been proposed[195] [196] to improve the performance. ResNet refers to a category of models incorporating residual connections into the structure of CNNs, enabling the retention of information from various residual structural blocks and mitigating issues related to performance degradation. Kim and colleagues in 2016[197] were the first which integrated residual learning techniques into SR tasks, which have now become widely adopted for image restoration purposes. In this study, the Enhanced-Deep-Super-Resolution (EDSR) model, built upon these techniques, has been employed. Details about the model are given in the next section. Deep SR neural networks have been a significant development in the fields of computer vision and image processing; they have found utility

in various real-world areas, including medical imaging[198] [190] [199] [200]. In this context, the application of the EDSR model to biomedical images, particularly MR brain data, has been explored.

The use of SR algorithms in clinical practice has already been implemented in some medical centers. Their employ guarantees several advantages, primarily by offering a trade-off between the duration of the scan time and and accuracy of the image processing. High-resolution image acquisition demands lengthy scan times, impacting both the cost of the examination and patient comfort. Prolonged scan duration can lead to patient motion artifacts in the resulting medical images, thereby compromising lesion detection, disease diagnosis, and the extraction of radiomic features. Furthermore, in the MRI context, this extended acquisition time can result in a reduced Signal-to-Noise Ratio (SNR) and high Specific Absorption Rate (SAR)[190]. Consequently, post-processing spatial resolution enhancement stands as a viable strategy to unlock the full potential of medical imaging[199] [200]. They have been used to move from low-field MR images - poor in resolution but acquired with less expensive technology and faster in terms of acquisition - to high-resolution ones, obtaining interesting findings[201].
However, the application of DL SR models to biomedical images, as DL models designed for other specific tasks, is not straightforward. It requires the development of purpose-built architectures tailored to the unique characteristics of the data type at hand. Moreover, the limited availability of data can pose a challenge, often requiring the application of data augmentation techniques[202], to reduce the risk of under-fitting. Finally, DL algorithms are not universally applicable, meaning they cannot guarantee generalization to data that differs from the training dataset. As a result, *re-training* stage or even *re-initiating* the training process is frequently necessary when dealing with novel data.

In this work[203] [204], the 2D Enhanced-Deep-Super-Resolution model, with 2x as an up-sampling factor, was applied to MR brain images, specifically MPRAGE, FLAIR and QSM. The model was used to evaluate 2D multi-slice super-resolved images, and 3D MRI reconstructions obtained from them; details of the pipeline are illustrated in the following sections. EDSR application on new data was performed avoiding re-training stages, to test its inherent generalization ability. The goal was to confirm whether it was possible to reconstruct a kind of data never seen by the model during the training, and, in case of a positive outcome, to verify if the results achieved were superior to those of the traditional bicubic up-sampling method. Additionally, an analysis of the explain-ability of the model was proposed, describing the EDSR functioning with respect to the image pixel intensity.

4.2.2 EDSR-2x Model

EDSR is a deep 2D CNN based on residual learning techniques, winner of the example-based single-image SR NTIRE Challenge 2017[205] [191] [206]. The SNU Computer Vision Laboratory, from Seoul National University, promoted the model during the contest. EDSR was trained with end-to-end deep learning techniques, achieving the best reconstruction performance. Nowadays is

one of the CNNs with the best outcomes[198] [207] also processing biomedical images.
The single-scale architecture of EDSR was optimized starting from SRResNet[208], itself derived from the original ResNet architecture: batch normalization layers, which are usually inserted to reduce the risk of overfitting and to guarantee fast convergence, were removed, resulting in an improvement of model time convergence, flexibility and performance. In more detail, the input data are LR $N \times M$ images, which are first processed in a 3x3 convolutional layer extracting 256 LR feature maps. Those are used as input for the central part of the model, containing 33 residual learning blocks. Each of them is composed of two 3x3 convolutional layers and a skip connection to keep together information from one residual block to the following one. There is then an up-sampling layer which returns 256 $2N \times 2M$ maps and a final convolutional layer with a high-resolution map $2N \times 2M$ as output. Pixel loss λ_{L1} was used as loss function. A scheme of the model is proposed in Fig. 4.5. For each 2D image EDSR works separately for the 3 RGB channels; images used in this work are grayscale figures, thus the average result over the 3 channels was considered.
The model was trained with the DIV2K dataset (https://data.vision.ee.ethz.ch/cvl/DIV2K/), which consists of 1000 DIVerse 2K RGB general-purpose images collected from dozens of sites, with a large diversity of contents (people, handmade objects, environments, flora and fauna, and so on). For the challenge, high-resolution and corresponding low-resolution images for 2,3 and 4 down-sampling factors were provided, for the training (800 images), test (100 images) and validation (100 images) stages.

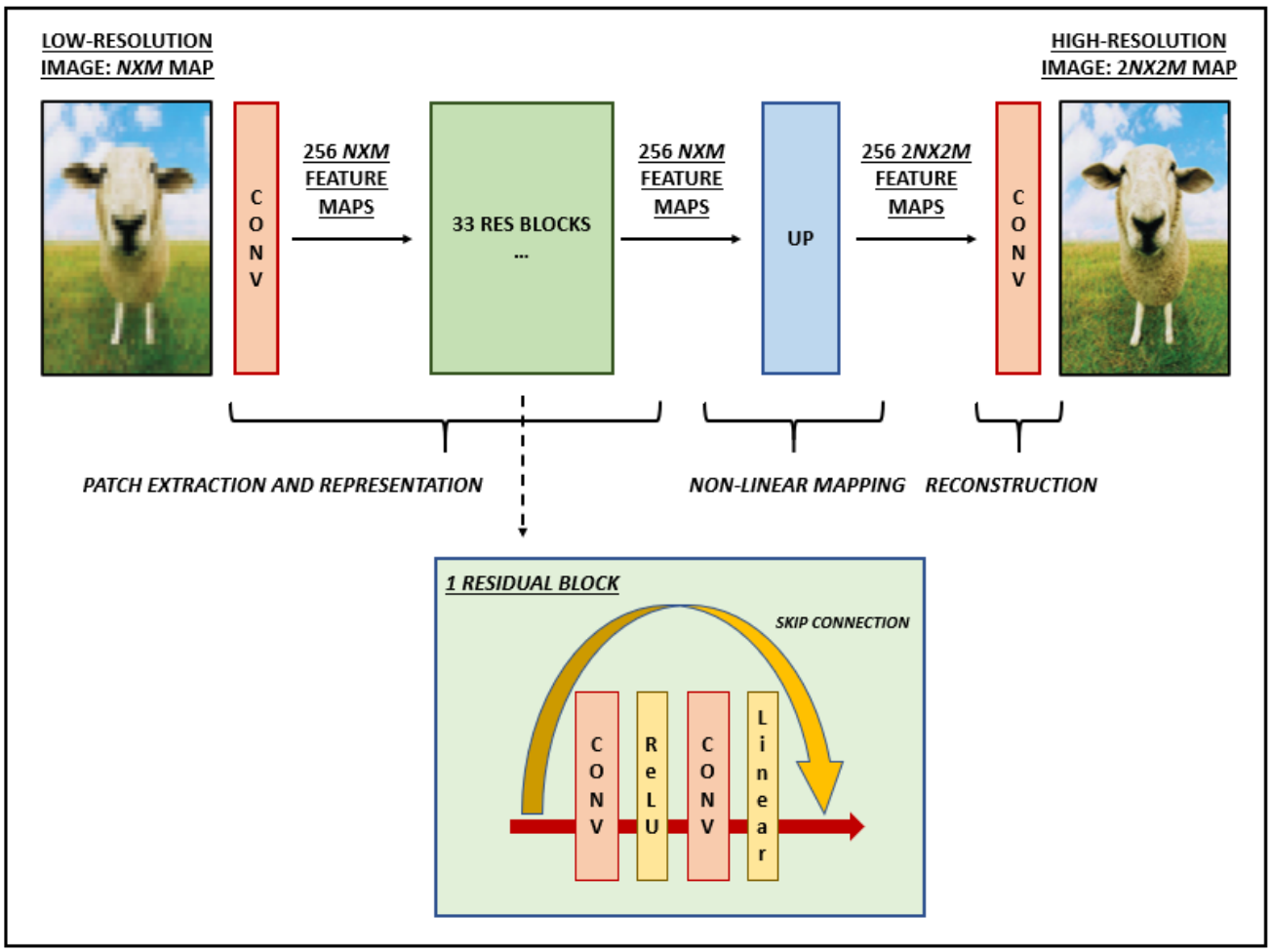

Figure 4.5: Scheme of EDSR architecture: input data are LR $N \times M$ images, initially processed in a 3x3 convolutional layer extracting 256 LR feature maps. The central part of the model contains 33 residual learning blocks, each of them composed of two 3x3 convolutional layers and a skip connection. There is then an up-sampling layer which returns 256 $2N \times 2M$ maps and a final convolutional layer with a high-resolution map $2N \times 2M$ as output

4.2.3 Methods and materials

Study sample and MR measurements

Data from 28 healthy controls (F:M 17:11, 48.1 $\pm$ 17.9 years old [24-86]) were selected from the database of Neuroimaging Laboratory (Functional and Molecular Neuroimaging Unit, IRCCS Istituto delle Scienze Neurologiche di Bologna, Bellaria Hospital).

The MR protocol included MPRAGE, FLAIR and QSM as sequences; acquisition and processing details are given in Chapter 1 (Sec. 1.2.1). QSM images, originally acquired with 0.5x0.5x1.5 mm^3 spatial resolution, were linearly registered to the corresponding MPRAGE using FLIRT [44] [45] from FSL [42], as FLAIR images, gaining 1mm-isotropic spatial resolution.

The pipeline proposed for this study was at first validated by our group [135] over a 70-subject dataset from the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) [209] [210]. The repository, based on a large-scale collaborative research project at the University of Cambridge, contains multi-modal data, including structural and functional (resting and task-based) MR images, magnetoencephalography data and several cognitive tests of nearly 700 healthy subjects, 100 per decade

from 18 to 88 years old. From there, T_1w and T_2w brain MR images of 70 healthy subjects (10 per decade from 18 to 85 years old) were considered (F:M 40:30, 53.3 ± 20.0 years old [18-85]).

Image Processing Pipeline

LR images generation: To generate LR maps, from 1x1 mm^2 to 2x2 mm^2, 2D slices extracted from the original images underwent a convolution process using a 3×3 kernel Gaussian filter with a unit standard deviation. Subsequently, these slices were down-sampled by a factor of 2 using bicubic (BC) interpolation. This degradation procedure results in MR images that closely resemble those generated by actual LR MR image acquisition[199]. This pipeline is commonly employed in SR studies to establish LR-HR image pairs for model validation. In an ideal scenario, the same subject, perfectly motionless, would undergo both LR and HR scans within the same field of view. However, this is impractical, and it would necessitate registration operations, which are preferably avoided given the need to assess interpolation methods. Additionally, doubling the number of measurements would require twice the resources in terms of scan time.

Up-sampling methods: The EDSR model and the BC interpolation were used to up-sample LR images, moving from 2x2 mm^2 to 1x1 mm^2 spatial resolution. Pipeline was designed to establish a gold-standard reference (HR images) and a benchmark (BC up-sampling, a commonly used up-sampling technique) to evaluate EDSR reconstructions.
After the down-sampling, the images exhibited more pronounced ringing or Gibbs artifacts. Initially, a Gibbs artifact removal tool[211] was experimented with LR images before applying the two up-sampling methods, but it did not yield any noticeable impact on the results. Consequently, this step was excluded from the pipeline.
As the EDSR model processes 2D images, data obtained from sagittal, axial, and coronal orthogonal planes were analyzed individually. This analysis resulted in the creation of a volume composed of 2D super-resolved slices for each direction, referred to as *2D multi-slice reconstructions*. The 3D reconstructions were then generated by averaging the 2D multi-slice reconstructions from all three planes. This pipeline's workflow is shown in Figure 4.6, and it was applied to MPRAGE, FLAIR, and QSM images. An illustrative example of EDSR and BC reconstructions can be found in Figure 4.7.

The pipeline was implemented using Python (version 3.7.6) and included the following libraries: OpenCV[212] (version 4.5.2), NumPy[213] (version 1.20.2) and NiBabel (version 3.2.1).

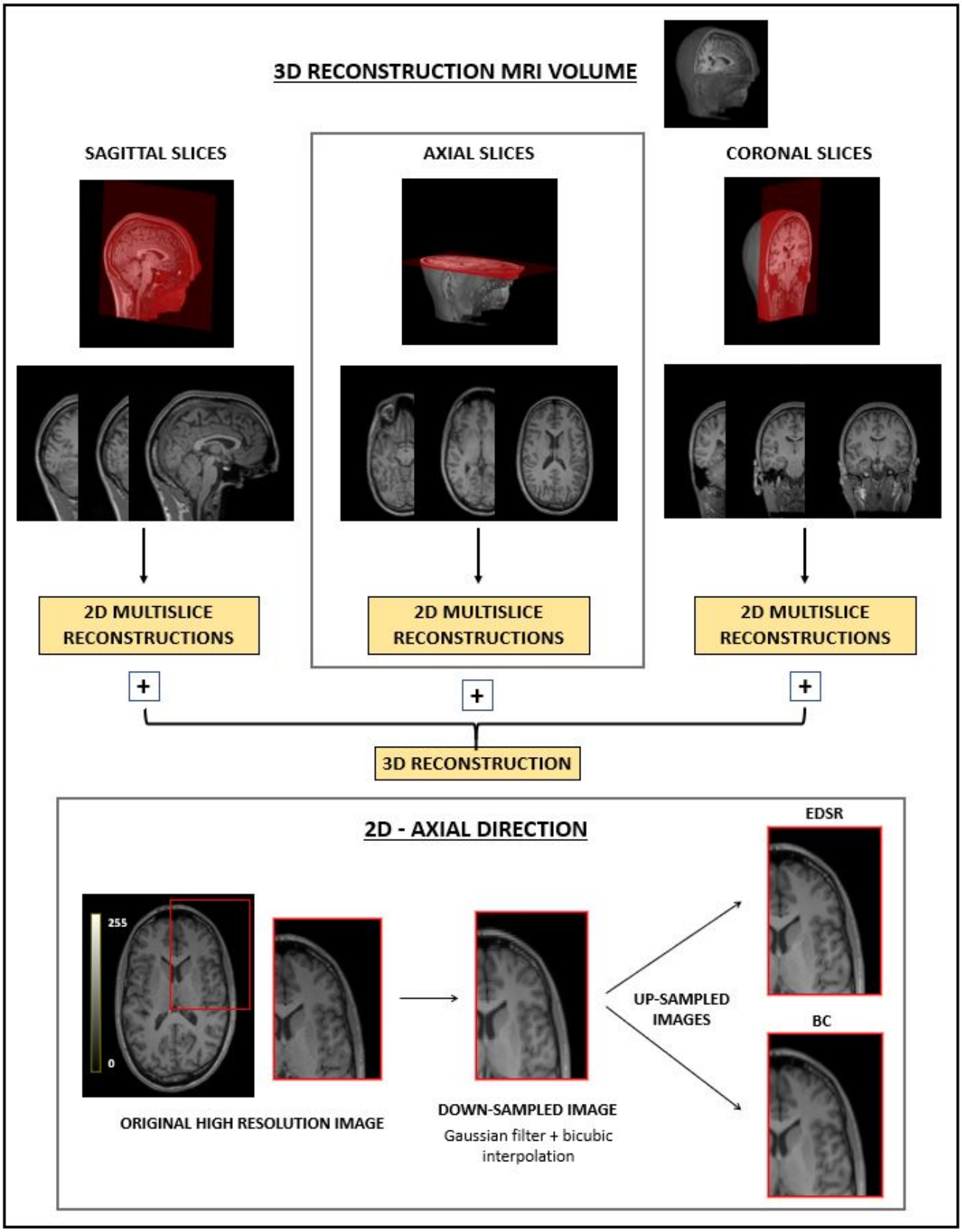

Figure 4.6: Pipeline applied to original images to perform 3D reconstructions[135]. Sagittal, axial and coronal planes were processed separately, and distinct 2D multi-slice images were achieved and then averaged to obtain the final map. The process for 2D multi-slice reconstruction is illustrated at the bottom of the panel, taking axial direction as an example: 2D images were processed with a Gaussian filter, down-sampled with BC interpolation and then up-sampled using EDSR and BC interpolation.

Image quality assessment and statistical analysis

Human perception plays of course an important role in the assessment of image quality and the similarity of images to a reference standard. To model and quantify perceived quality is challenging, and several measures have been proposed[214]. In this work, the quality of EDSR and BC images was quantified as the pixel-level fidelity to the ground truth using standard metrics. Considering N the number of pixels, G the gold-standard image (original HR), and I the up-sampled image (EDSR or BC), we measured peak Signal-to-Noise Ratio (pSNR)[191] and Structural SIMilarity Index (SSIM)[215], defined as:

$$pSNR = 10 \cdot log_{10} \frac{{MAX_G}^2}{MSE} \tag{4.1}$$

$$SSIM = l \cdot c \cdot s = \; = \frac{2\mu_G\mu_I + c_1}{\mu_G^2 + \mu_I^2 + c_1} \cdot \frac{2\sigma_G\sigma_I + c_2}{\sigma_G^2 + \sigma_I^2 + c_2} \cdot \frac{\sigma_{GI} + \frac{c_2}{2}}{\sigma_G + \sigma_I + \frac{c_2}{2}} \tag{4.2}$$

where μ_G is the average of G, μ_I is the average of I, σ_G^2 is the variance of G, σ_I^2 is the variance of I, σ_{GI} is the covariance of G and I, $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$, L is the dynamic range of the pixel-values and $k_1 = 0.01$ and $k_2 = 0.03$ by default; l is for luminance, c for contrast and s for structure.

Standard image quality metrics, such as pSNR and root mean square error (RMSE)[205] refer to the absolute error between the obtained image and reference one. However, differences in intensity values are not enough to analyze and quantify the perceptual distortion experienced by human vision. Perception-based metrics, such as SSIM, have been introduced to provide a more comprehensive analysis. SSIM considers image quality degradation as perceived changes in structural information. It is a full reference metric based on visible structures in the image, extracting 3 features - luminance, contrast and structure - and performing the measure based on those.

Images with high pSNR and high SSIM indicate high quality of reconstruction. Those metrics were evaluated with Python (version 3.7.6) (libraries: Scikit-image[216] [version 0.16.2], SciPy[217] [version 1.5.3], NumPy[213] [version 1.20.2]) for EDSR and BC reconstructed images of each subject and each sequence, using the original HR images as ground-truth, over a region from which the scalp had been excluded using BET[43] from FSL[42]. For 3D reconstructions, similarity parameters were directly evaluated over the three-dimensional matrices. The non-parametric Kruskal-Wallis test was selected for comparison between the two up-sampling methods, considering $p - values < .05$ as significant.

4.2.4 Results

On qualitative assessment, EDSR better reproduced the original images for MPRAGE, FLAIR and QSM sequences (Fig. 4.7). Quantitative analysis confirmed this evaluation.

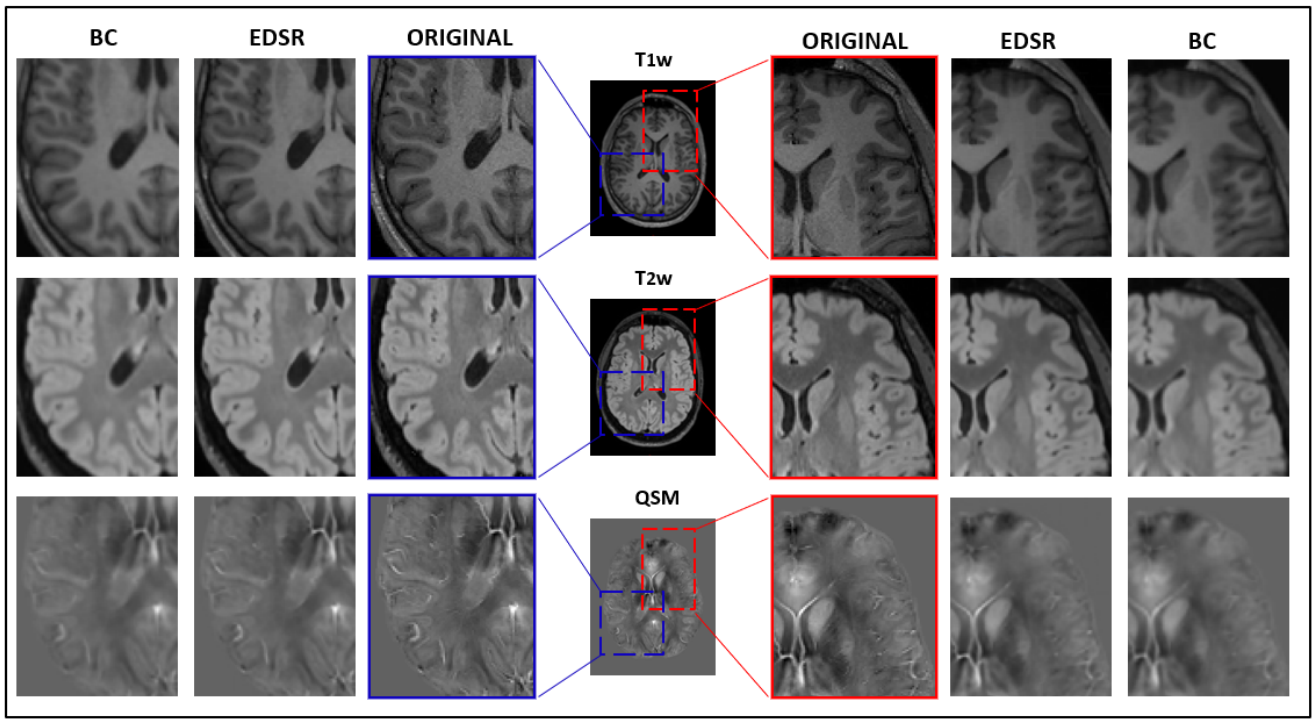

Figure 4.7: Example of MPRAGE, FLAIR and QSM images of a representative subject (healthy control, M/24 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit), comparing the original ones with EDSR and BC reconstructions. From the images, it can be observed that EDSR better reconstructs the original than BC interpolation, restoring high spatial-frequency structures in more detail

In Tab. 4.4 and Fig. 4.8, the results showed significant outperformance of EDSR in MPRAGE and FLAIR. As aforementioned, the superior performance of EDSR with 1mm-isotropic T_1w and T_2w images using this pipeline has already been confirmed[135].

	MPRAGE		**FLAIR**	
	pSNR ↑	SSIM ↑	pSNR ↑	SSIM ↑
EDSR	34.5 ± 2.8 *	0.985 ± 0.003 *	30.5 ± 2.7 *	0.989 ± 0.003 *
BC	28.8 ± 2.4	0.965 ± 0.006	28.5 ± 2.6	0.977 ± 0.004

Table 4.4: Median ± mean absolute deviation of pSNR and SSIM distributions within the sample in MPRAGE and FLAIR images. The outperforming method between EDSR and BC is highlighted (*) when there was significant difference ($p - values$ from Kruskal-Wallis test $< .05$)

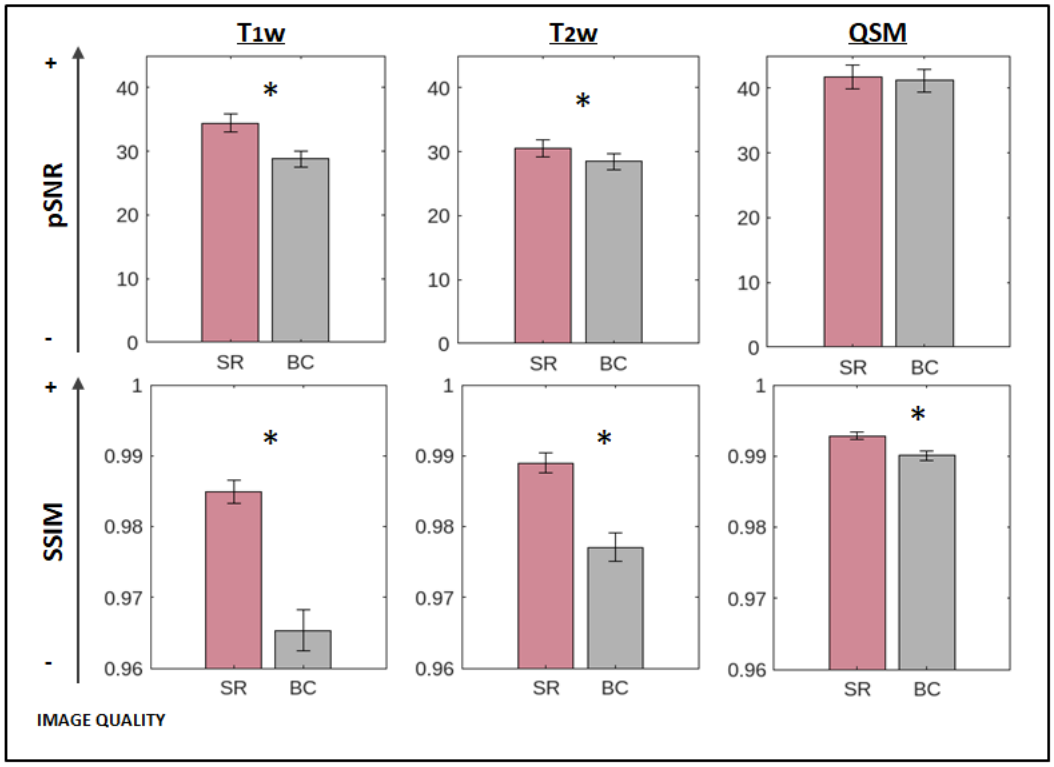

Figure 4.8: Median ± mean absolute deviation of pSNR and SSIM in MPRAGE, FLAIR and QSM images up-sampled with EDSR and BC methods, over the 28-subject dataset. Higher values of pSNR and SSIM indicate higher quality of reconstruction. Values from the 3D reconstructed images are reported. The outperforming up-sampling method between EDSR and BC is highlighted (*) when there was a significant difference ($p - values < .05$ from Kruskal-Wallis test)

Additionally, EDSR achieved better results in QSM images (Fig. 4.8 and Tab. 4.5, left side), evaluating pSNR and SSIM with respect to the original gold-standard. SSIM is a full reference metric and a reliable indicator of image quality degradation, and it was significantly higher in EDSR reconstructions. Thus, the model showed an excellent ability of generalization over new kinds of data and different MR sequences including QSM. However, there are differences in QSM images with respect to the other MR measurements looking into the pSNR metric, which is based on the absolute difference between pixel level intensity and did not show a significant difference in EDSR reconstructions compared to BC ones. Those differences are probably due to the appearance of high-intensity areas in QSM maps (Fig. 4.9). QSM maps were upper-threshold considering the 95^{th} percentile of each image as cut-off, reducing the extension of hyper-intense areas, and the comparison analysis was carried out again: with this adjustment, QSM reconstructions from EDSR resulted significantly better for both pSNR and SSIM (Tab. 4.5, right side).

	QSM		**QSM-95th**	
	pSNR ↑	SSIM ↑	pSNR ↑	SSIM ↑
EDSR	41.7 ± 3.6 *	0.993 ± 0.001 *	25.4 ± 1.0 *	0.990 ± 0.002 *
BC	41.2 ± 3.5	0.990 ± 0.001	24.1 ± 1.2	0.986 ± 0.003

Table 4.5: Median ± mean absolute deviation of pSNR and SSIM distributions among the 28-subject sample in QSM images, before (left side) and after (right side) the 95^{th} upper-thresholding. The outperforming method between EDSR and BC is highlighted (*) when there was significant difference ($p - values$ from Kruskal-Wallis test $< .05$)

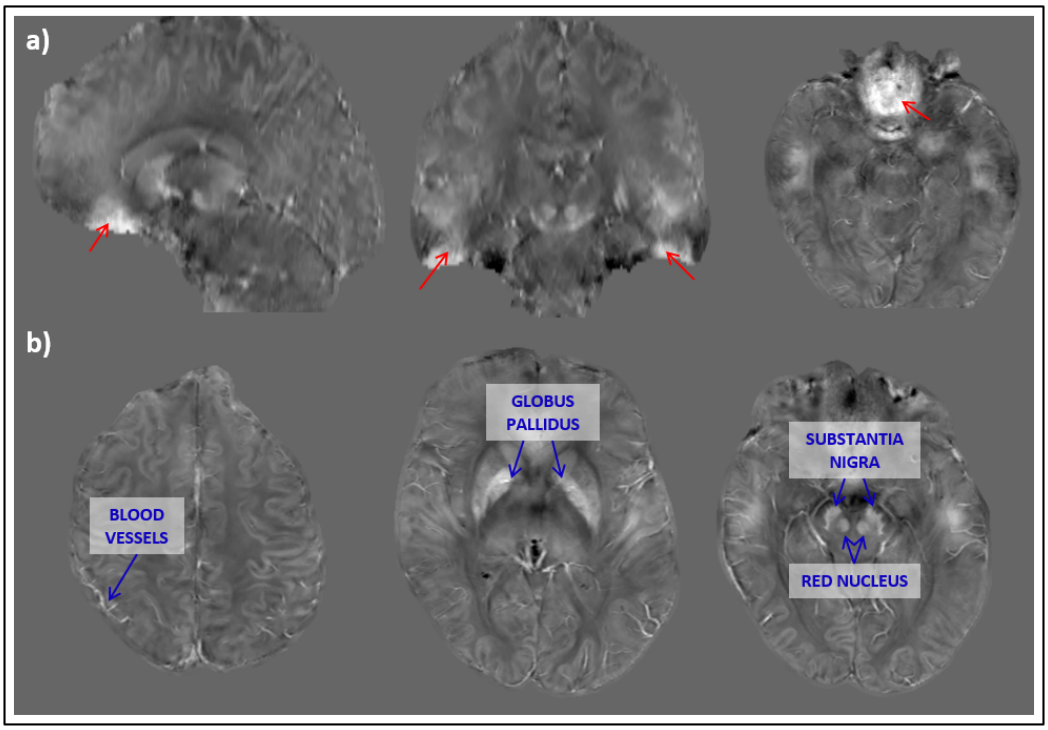

Figure 4.9: QSM (healthy control, F/57 years old) (IRCCS Istituto delle Scienze Neurologiche di Bologna, Functional and Molecular Neuroimaging Unit) - panel a) air-tissue artifacts appearing in QSM reconstructions (red arrows); panel b) examples of iron-storages structures, which appear with high-intensity pixels in susceptibility maps (blue arrows): blood vessels and deep gray matter nuclei, such as globus pallidus, substantia nigra and red nucleus

4.3 Discussion

SR techniques improve the trade-off between diagnostic accuracy and limited total scan time, supporting faster screening and follow-up exams; low-resolution images can be acquired rapidly and then have their spatial resolution enhanced at the post-processing stage. The 2D CNN EDSR model was introduced in 2017[206] and it is still one of the state-of-the-art SR DL algorithms: in a recent review of single-image SR methods[207], EDSR was among the top five best performing out of a total of 19 models. It has recently been used in biomedical applications[218], and it was adopted for the current study. The model, pre-trained on natural images, was directly tested on MR scans (MPRAGE, FLAIR, QSM), aiming to confirm its generalization ability. Results were compared with BC up-sampling. In previous work, the same pipeline was validated over a larger sample of T_1w and T_2w images, from Cam-CAN dataset[135], more frequently used in this kind of applications. In Fig. 4.7, an example of EDSR and BC MPRAGE, FLAIR and QSM up-sampled images is shown: a qualitative assessment highlights that the model better reproduces the original HR images in the analyzed sequences, restoring high spatial-frequency structures in more detail. To quantify aspects of reconstructions relevant to human perception, a robust analysis was carried out. Results are discussed below, highlighting the key points of the study.

EDSR over MR brain images: MPRAGE, FLAIR and QSM

Among MR modalities, T_1w and T_2w are the ones usually explored in SR applications[190]. EDSR, as 2D CNN model, was exploited to reproduce 3D images, combining reconstructions from the

three principal orthogonal planes (Fig. 4.6); since the MR images analyzed are 3D data, an overall assessment of the entire volume is preferred. The analysis (Tab. 4.4, Fig. 4.8) show that EDSR outperforms BC interpolation, with a significant difference in favor of EDSR model for both MPRAGE and FLAIR images. EDSR shows the superior performance of BC in QSM images as well, significant looking into SSIM metric (Tab. 4.5, Fig. 4.8).
It is well established that DL algorithms, including EDSR, work better than traditional BC upsampling over data from the training dataset[194]. In this work, the aim was to test the response of EDSR to data with characteristics unlike those of the training set: the network weights, derived in the NTIRE Challenge 2017 from a training with a variety of photographic images, were left untouched and the model was directly applied to MR brain images. The poor generalizability over different data is an issue in SR model applications: commonly train and test data are of the same thematic type, whichever is the model used. For example, Zhang et al. in 2021[198], used 3T structural T_1w images divided into train and test datasets; Sui et al. in 2021[190] trained their model using 3T T_1w and T_2w images from an online dataset and tested it on T_2w images acquired on their 3T scanner. Surprisingly from this analysis, EDSR performed excellent reconstructions of MR brain images, superior to the BC standard method, leading to the consideration that a re-training stage was not necessary, although the kind of data used for the training was very different. The high quality of EDSR reconstructions achieved in this work is not trivial, considering the transition from general-purpose images to brain MR scans. The capacity to generalize over new data is probably due to the large number of model parameters: referring to the review of 2021[207], EDSR was the model with the highest number of parameters among those analyzed, more than one-third higher than the second-ranked model. This fact does not affect the computational cost, which remains low. Furthermore, EDSR is a 2D CNN, which positively influences computational effort and memory allocation. Such models can be used to reconstruct three-dimensional data, including brain images from MR exams: Zhang et al. in 2021[198] obtained 3D reconstructions from two multi-slice super-resolution images, achieving better results than three-dimensional super resolution technology. We decided to use an analogous method as a viable alternative to 3D CNNs.

EDSR over QSM images: influence of high-intensity pixels

As mentioned before, EDSR shows superior performance of bicubic interpolation in QSM images, with differences with respect to MPRAGE and FLAIR sequences though: only SSIM shows significant differences between EDSR and BC as a metric parameter, while pSNR does not. The latter parameter is based on the intensity of the pixels, thus we hypothesized that the high-intensity pixels in susceptibility images - due to iron-storage substances and air-tissue artifacts[59] (Fig. 4.9) - may influence the outcomes. Consequently, the QSM maps were upper-threshold considering the 95^{th} percentile of each image as cut-off, reducing the extension of hyper-intense areas, and the comparison analysis was carried out again: with this adjustment, QSM reconstructions from EDSR resulted significant better for both pSNR and SSIM (Tab. 4.5).
In the previous work, we pursued a similar line of reasoning[135]. It had indeed become apparent that there were differences between T_1w and T_2w images, where white matter, gray matter,

and cerebrospinal fluid appeared with varying contrasts; in particular, in T_2w images, the cerebrospinal fluid appeared very bright. Thus, to attempt to explain EDSR functioning, we initially assessed the model's performance on each of the three tissues individually. Then, we reversed the histogram of the original images, retracing the entire down-sampling and up-sampling pipeline. The outcomes confirmed the hypothesis. EDSR performance is likely to vary in images with large hyper-intense areas, and this may be related to the fact that in natural images like landscapes and animals, used to train the model, bright and shiny elements, such as light reflected in a mirror, are less common than the darker ones, such as shadows of objects and people. Moreover, images including larger glares or shadow components are commonly excluded by training datasets like the DIV2K, since they are interpreted as undesired components in a picture.

Limitations and future work

Further analysis needs to be performed for the optimal usage of EDSR. First, data used in this work are high-quality MR images of healthy subjects, without sizeable movement artifacts, and their spatial resolution was improved from 2×2 to 1×1 mm^2. At a millimeter-centimeter scale, image intensity in these images is dominated by the presence of myelinated axons (white matter) neuronal and glial cell bodies (gray matter), and CSF. Within a 1-mm isotropic voxel, one tissue type typically predominates, and this pattern recurs in neighboring pixels over a centimeter scale, producing patches of fairly uniform pixel intensity, separated by fairly sharp boundaries. It is possible to speculate that these image characteristics are close to those found in the photographs of natural scenes used to train the EDSR. Additionally, changes in signal intensity or morphology due to neurological diseases variably introduce additional tissue types at the voxel scale. Thus, the achievement of good performance with different scales and with pathological tissue is not obvious and will have to be tested.

Second, EDSR was tested over LR images generated from HR ones, which is a standard procedure in SR studies to have a pixel-wise reference map to compare the model reconstructions, avoiding registration operations. Nevertheless, since they are simulated LR images, the superior performance of the model needs to be confirmed on LR images from real MR acquisitions.

Third, we observed that EDSR performance is negatively affected by the presence of high-intensity pixels, which may be informative in QSM images since they are, in some structures, index of iron accumulation; as pointed out in previous chapters, that may be hallmark of pathology in neurodegenerative processes. *Fine-tuning* of the model may be necessary to properly exploit the pipeline over susceptibility images.

4.4 Conclusion

The application of the EDSR-2x on MR biomedical images in this work leads to promising results: without needing tailored re-training, the model shows its ability to generalize from general-purpose images to different MR sequences of healthy subjects, achieving better performance than traditional up-sampling methods. Although the model works with two-dimensional data, the outper-

formance holds when three-dimensional reconstructions are considered. The data analyzed in this study include images from different MR sequences (MPRAGE, FLAIR, QSM) of subjects with different characteristics in terms of age and sex. The application of this model on MR brain images was satisfactorily validated. It is noteworthy that images analyzed in this work were previously unseen by the model and we can reasonably expect similar outcomes on images from the same sequences, which would appear with similar intensities and contrasts, acquired with different scanners. In QSM reconstructions, we observed some issues in accurately reproducing high-intensity pixels, some of them corresponding to the presence of iron in the underlying tissues; fine-tuning training may be needed to analyze those images.

BET = Brain Extraction Tool
BC = BiCubic
Cam-CAN = Cambridge Centre for Aging and Neuroscience
CNN = Convolutional Neural Network
DIV2K = DIVerse 2K DL = deep Learning
EDSR = Enhanced Deep Super Resolution
FOV = Field Of View
FLAIR = FLuid Attenuate Inversion Recovery
FLIRT = FSL's Linear Image Registration Tool
FSL = FMRIB Software Library
HR = High Resolution
LR = Low resolution
MPRAGE = Magnetization Prepared RApid Gradient Echo
MR = Magnetic Resonance
MRI = MR Imaging
pSNR = peak Signal-to-Noise Ratio
QSM = Quantitative Susceptibility Mapping
ResNet = Residual Neural Network
RMSE = Root Mean Squared Error
SAR = Specific Absorption Rate
SNR = Signal-to-Noise Ratio
SR = Super Resolution
SSIM = Structural SIMmilarity Index
T_1w, T_2w, T_2^*w = T_1-, T_2-, T_2^*- weighted

Conclusion

In this thesis, I explored QSM as an advanced MR technique, able to detect and quantify iron and myelin concentration as possible biomarkers of neurodegeneration. Its inherent quantitative nature offers the unique ability to compare magnetic properties within and between different groups of subjects. After the theoretical overview and the acquisition and post-processing pipeline description, optimized during my first year of PhD and described in the first chapter of the thesis, I illustrated different applications as possible approaches to QSM employment, including traditional as well as innovative techniques.

The first application I proposed is about the examination of cortical iron accumulation in patients with ALS using QSM, which is iron-sensitive and, consequently, allows to measure *in vivo* iron concentration in gray matter tissue. Various descriptors were used for the investigation, both susceptibility-related metrics and structural measurements, which revealed distinctive patterns of rising iron content of ALS patients with respect to healthy controls and patients with other neurodegenerative disorders, in the motor cortex and extra-motor cortical regions. Additionally, the analysis of the clinical profiles of the participants showed a significant correlation between susceptibility measurements and upper motor neuron impairment, especially in patients with rapid disease progression, highlighting the promising potential of QSM as a valuable tool for identifying biomarkers for ALS.

In the second application I introduced the utilization of QSM for examining the cortical and subcortical structures in a group of DM1 patients, divided into two genetic classes E1 and E2, comparing median susceptibility and volume distributions to healthy controls. Since gray matter structures are analyzed, QSM measurements revealed the presence of iron concentration as before. The results contributed to a more comprehensive characterization of DM1 patients, pinpointing the thalamus and brainstem as key structures of interest from a susceptibility perspective: these areas exhibited a significant increase in iron concentration compared to the control group and displayed correlations with clinical data, including neurological, cardiological, and pneumological evaluations, along with sleep records from polysomnography, and the onset age. These findings certainly advance our understanding of the mechanisms underlying this pathology and of the origins, central *vs* peripheral, of DM1 symptomology.

The third application focused on MS; QSM was used to explore the normal appearing white matter, in which myelin is the substance influencing susceptibility values the most, diamagnetically. For the first time, a processing pipeline to assess the robustness of QSM-derived radiomic features in white matter tracts was implemented. Multiple MR sequences were used to build up the workflow, including morphological images for tissue and lesion segmentation and diffusion measurements for tractography reconstruction. This study sets the stage for upcoming research, which can use the array of reliable features to non-invasively profile individuals with multiple sclerosis, promptly assess treatment outcomes, and track the evolution of the condition. Different outcomes in feature reliability among the analyzed regions were explained with respect to their characteristics.

Last, a deep learning super resolution model was applied to QSM images in a cohort of healthy controls. The approach was peculiar, since the adopted neural network was previously trained with general-purpose images and then directly applied to different MR techniques, achieving better performance than traditional up-sampling methods without needing a tailored re-training. The excellent generalization ability of the model over these new data suggests that similar outcomes could be achieved on images from different scanners. The analyzed sequences are certainly different in terms of intensities and contrasts; the examination by tissue type highlights variations in model performance linked to gray-level values, revealing a relative lack of out-performance in reconstructing highly intense regions.

In designing the studies, special attention has been given to the implementation of automated pipelines using as much as possible open-source software and toolboxes, to promote outcomes reproducibility and sharing of processing and analysis techniques within the scientific community. Additionally, it is important to emphasize that the interpretability of the results derived from QSM data is crucial within the context of neurodegeneration, as the other neuroimaging applications. Susceptibility maps are 3D numerical matrices that reflect the properties of underlying tissues. In the search for biomarkers of pathology, it is essential to understand the meaning of these measurements to maintain a strong connection between numerical data and the studied pathophysiology. This consideration applies to traditional analyses, exploring the average intensity of QSM maps, as well as to more innovative approaches like radiomics and deep learning, which involve less intuitive measurements.
Since its introduction, QSM has become an important tool to study neurodegenerative processes. In the three pathologies studied in this thesis, it was observed that QSM is capable of distinguishing patient groups from healthy controls and correlating imaging measurements with clinical data. In some instances, a connection between structural features (e.g. the orientation of white matter fibers in the cortico-spinal tract) and QSM measurements has even been identified. These studies pave the way for future research, aiming to achieve progressively more early and less invasive diagnoses of neurodegenerative disorders in the context of personalized medicine.

Bibliography

[1] A. Deistung, F. Schweser, and J.R. Reichenbach. Overview of quantitative susceptibility mapping. *NMR in Biomedicine*, 30(4):e3569, 2017.

[2] C. Strafella, V. Caputo, M.R. Galota, S. Zampatti, G. Marella, S. Mauriello, R. Cascella, and E. Giardina. Application of precision medicine in neurodegenerative diseases. *Frontiers in Neurology*, 9(701), 2018.

[3] T. Harada, K. Kudo, N. Fujima, M. Yoshikawa, Y. Ikebe, R. Sato, T. Shirai, Y. Bito, I. Uwano, and M. Miyata. Quantitative susceptibility mapping: basic methods and clinical applications. *Radiographics*, 42(4):1161–1176, 2022.

[4] B. Bilgic, M. Costagli, K.S. Chan, J. Duyn, C. Langkammer, J. Lee, X. Li, C. Liu, J.P. Marques, C. Milovic, S.D. Robinson, F. Schweser, K. Shmueli, P. Spincemaille, S. Straub, P. van Zijl, Y. Wang, and Group IETPS. Recommended implementation of quantitative susceptibility mapping for clinical research in the brain: a consensus of the ISMRM Electro-Magnetic Tissue Properties study group. *ArXiv [Preprint]*, 2023.

[5] S. Soman, J.A. Bregni, B. Bilgic, U. Nemec, A. Fan, Z. Liu, R.L. Barry, J. Du, K. Main, J. Yesavage, M.M. Adamson, M. Moseley, and Y. Wang. Susceptibility-based neuroimaging: standard methods, clinical applications, and future directions. *Current Radiology Reports*, 5(3):11, 2017.

[6] D. Faul, J. Abart, and P. Margosian. Quick measurements of magnetic field variations within the body. *Radiology*, 153(P):303, 1984.

[7] E.M. Haacke, S. Mittal, Z. Wu, J. Neelevalli, and Y.C.N. Cheng. Susceptibility-weighted imaging: technical aspects and clinical applications, part 1. *American Journal of Neuroradiology*, 30(1):19–30, 2009.

[8] S. Mittal, Z. Wu, J. Neelevalli, and E.M. Haacke. Susceptibility-weighted imaging: technical aspects and clinical applications, part 2. *American Journal of Neuroradiology*, 30(2):232–252, 2009.

[9] J. Li, S. Chang, T. Liu, Q. Wang, D. Cui, X. Chen, M. Jin, B. Wang, M. Pei, C. Wisnieff, B. Spincemaille, M. Zhang, and Y. Wang. Reducing the object orientation dependence of susceptibility effects in gradient echo MRI through quantitative susceptibility mapping. *Magnetic Resonance in Medicine*, 68(5):1563–1569, 2012.

[10] R. Langreth and M. Waldholz. New era of personalized medicine: targeting drugs for each unique genetic profile. *Oncologist*, 4(5):426–427, 1999.

[11] L.H. Goetz and N.J. Schork. Personalized medicine: motivation, challenges, and progress. *Fertility and Sterility*, 109(6):952–963, 2018.

[12] European Society of Radiology (ESR). Medical imaging in personalised medicine: a white paper of the research committee of the European Society of Radiology (ESR). *Insights into Imaging*, 6(2):141–155, 2015.

[13] A. Tate, M. Suárez-Calvet, M. Ekelund, S. Eriksson, M. Eriksdotter, W.M. Van Der Flier, J. Georges, M. Kivipelto, M.G. Kramberger, P. Lindgren, J.D.G. López, J. Lötjönen, S. Persson, S. Pla, A. Solomon, L. Thurfjell, A. Wimo, B. Winblad, and L. Jönsson. Precision medicine in neurodegeneration: the IHI-PROMINENT project. *Frontiers in Neurology*, 14, 2023.

[14] P. Ravanfar, S.M. Loi, W.T. Syeda, T.E. Van Rheenen, A.I. Bush, P. Desmond, V.L. Cropley, D.J.R. Lane, O.M. Opazo, B.A. Moffat, D. Velakoulis, and C. Pantelis. Systematic review: quantitative susceptibility mapping (QSM) of brain iron profile in neurodegenerative diseases. *Frontiers in Neuroscience*, 15:618435, 2021.

[15] P. Dusek, T. Hofer, J. Alexander, P.M. Roos, and J.O. Aaseth. Cerebral iron deposition in neurodegeneration. *Biomolecules*, 12(5):714, 2022.

[16] C. Langkammer, L. Pirpamer, S. Seiler, A. Deistung, F. Schweser, S. Franthal, N. Homayoon, P. Katschnig-Winter, M. Koegl-Wallner, T. Pendl, E.M. Stoegerer, K. Wenzel, F. Fazekas, S. Ropele, J.R. Reichenbach, R. Schmidt, and P. Schwingenschuh. Quantitative susceptibility mapping in Parkinson's disease. *PLoS One*, 11(9):e0162460, 2016.

[17] L. Chen, J. Hua, C.A. Ross, S. Cai, P.C.M. van Zijl, and X. Li. Altered brain iron content and deposition rate in Huntington's disease as indicated by quantitative susceptibility MRI. *Journal of Neuroscience Research*, 97(4):467–479, 2019.

[18] Y. Uchida, H. Kan, K. Sakurai, K. Oishi, and N. Matsukawa. Quantitative susceptibility mapping as an imaging biomarker for Alzheimer's disease: the expectations and limitations. *Frontiers in Neuroscience*, 16:938092, 2022.

[19] J. H. Duyn and J. Schenck. Contributions to magnetic susceptibility of brain tissue. *NMR in Biomedicine*, 30(4):e3546, 2016.

[20] S. Doganay, K. Gumus, G. Koc, A.K. Bayram, M.S. Dogan, D. Arslan, H. Gumus, S.B. Gorkem, S. Ciraci, H.I. Serin, and A. Coskun. Magnetic susceptibility changes in the basal ganglia and brain stem of patients with Wilson's disease: evaluation with quantitative susceptibility mapping. *Magnetic Resonance in Medical Sciences*, 17(1):73–79, 2018.

[21] V. Wiggermann, S. Hametner, E. Hernández-Torres, C. Kames, V. Endmayr, G. Kasprian, R. Höftberger, D.K.B. Li, A. Traboulsee, and A. Rauscher. Susceptibility-sensitive MRI of Multiple Sclerosis lesions and the impact of normal-appearing white matter changes. *NMR in Biomedicine*, 30(8), 2017.

[22] R. Sibgatulin, D. Güllmar, A. Deistung, C. Enzinger, Ropeles, and J.R. Reinchenbach. Magnetic Susceptibility anisotropy in normal appearing white matter in Multiple Sclerosis from single-orientation acquisition. *NeuroImage: Clinical*, 35:103059, 2022.

[23] C. Fiscone, L. Rundo, A. Lugaresi, D.N. Manners, K. Allinson, E. Baldin, G. Vornetti, R. Lodi, C. Tonon, C. Testa, M. Castelli, and F. Zaccagna. Assessing robustness of quantitative susceptibility-based MRI radiomic features in patients with Multiple Sclerosis. *Scientific Reports*, 13:16239, 2023.

[24] S. Hametner, V. Endmayr, A. Deistung, P. Palmrich, M. Prihoda, E. Haimburger, C. Menard, X. Feng, T. Haider, M. Leisser, U. Köck, A. Kaider, R. Höftberger, S. Robinson, J.R. Reichenbach, H. Lassmann, H. Traxler, S. Trattnig, and G. Grabner. The influence of brain iron and myelin on magnetic susceptibility effective transverse relaxation – A biomchemical and histological validation study. *NeuroImage*, 179:117–133, 2018.

[25] C. Liu, W. Li, K.A. Tong, K.W. Yeom, and S. Kuzminski. Susceptibility-weighted imaging and quantitative susceptibility mapping in the brain. *Journal of Magnetic Resonance Imaging*, 42, 2015.

[26] W. Li, C. Liu, T.Q. Duong, P.C. van Zijl, and X. Li. Susceptibility tensor imaging (STI) of the brain. *NMR in Biomedicine*, 30(4), 2017.

[27] T. Liu, W. Xu, P. Spincemaille, A.S. Avestimehr, and Y. Wang. Accuracy of the morphology enabled dipole inversion (MEDI) algorithm for quantitative susceptibility mapping in MRI. *IEEE Transactions on Medical Imaging*, 31(3):816–824, 2012.

[28] F. Schweser, K. Sommer, A. Deistung, and J.R. Reichenbach. Quantitative susceptibility mapping for investigating subtle susceptibility variations in the human brain. *NeuroImage*, 62(3):2083–2100, 2012.

[29] W. Li, N. Wang, F. Yu, H. Han, W. Cao, R. Romero, B. Tantiwongkosi, T.Q. Duong, and C. Liu. A method for estimating and removing streaking artifacts in quantitative susceptibility mapping. *NeuroImage*, 108:112–122, 2015.

[30] F. Yanez, A. Fan, B. Bilgic, C. Milovic, E. Adalsteinsson, and P. Irarrazaval. Quantitative susceptibility map reconstruction via a total generalized variation regularization. *International Workshop on Pattern Recognition in Neuroimaging, Philadelphia, PA, USA*, 2013.

[31] T. Liu, P. Spincemaille, L. de Rochefort, B. Kressler, and Y. Wang. Calculation of susceptibility through multiple orientation sampling (COSMOS): a method for conditioning the inverse problem from measured magnetic field map to susceptibility source image in MRI. *Magnetic Resonance in Medicine*, 61(1):196–204, 2009.

[32] S. Wharton and R. Bowtell. Whole-brain susceptibility mapping at high field: a comparison of multiple- and single-orientation methods. *NeuroImage*, 53(2):515–525, 2010.

[33] W. Jung, S. Bollmann, and J. Lee. Overview of quantitative susceptibility mapping using deep learning: current status, challenges and opportunities. *NMR in Biomedicine*, 35(4):e4292, 2022.

[34] C. Fiscone. A machine learning approach to QSM: quantitative susceptibility map reconstruction with convolutional autoencoders. Master's thesis, University of Bologna, Department of Physics and Astronomy, 2019.

[35] C. Fiscone, C. Testa, R. Lodi, C. Tonon, R. Bowtell, and D. Remondini. Brain quantitative susceptibility mapping: comparison between single- and multiple- orientation methods and machine learning approach. *XI Congress AIRMM - ISMRM Italian Chapter (Selected Abstract)*, 2020.

[36] C. Fiscone, C. Testa, R. Lodi, G. Castellani, R. Bowtell, and D. Remondini. A machine learning approach to QSM: susceptibility map reconstruction with convolutional autoencoders. *X Congress AIRMM - ISMRM Italian Chapter (Selected Abstract)*, 2019.

[37] J.F. Schenck. Safety of strong, static magnetic fields. *Journal of Magnetic Resonance Imaging*, 12:2–19, 2000.

[38] E. Pchitskaya, E. Popugaeva, and I. Bezprozvanny. Calcium signaling and molecular mechanisms underlying neurodegenerative diseases. *Cell Calcium*, 70:87–94, 2018.

[39] Y. Poitelon, A.M. Kopec, and S. Belin. Myelin fat facts: an overview of lipids and fatty acid metabolism. *Cells*, 9(812), 2020.

[40] S. Chang, J. Zhang, T. Liu, A.J. Tsiouris, J. Shou, T. Nguyen, D. Leifer, Y. Wang, and I. Kovanlikaya. Quantitative susceptibility mapping of intracerebral hemorrhages at various stages. *Magnetic Resonance Imaging*, 44(2):420–425, 2016.

[41] D.O. Walsh, A.F. Gmitro, and M.W. Marcellin. Adaptive reconstruction of phased array MR imagery. *Magnetic Resonance in Medicine*, 43(5):682–690, 2000.

[42] M. Jenkinson, C.F. Beckmann, T.E.J. Behrens, M.W. Woolrich, and S. Smith. FSL. *NeuroImage*, 62(2):782–790, 2012.

[43] S.M. Smith. Fast robust automated brain extraction. *Human Brain Mapping*, 17(3):143–155, 2002.

[44] M. Jenkinson and S.M. Smith. A global optimization method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143–156, 2001.

[45] M. Jenkinson, P. Bannister, M. Brady, and S. Smith. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841, 2002.

[46] B. Patenaude, S.M. Smith, D. Kennedy, and M. Jenkinson. A bayesian model of shape and appearance for subcortical brain. *NeuroImage*, 56(3):907–922, 2011.

[47] J.D. Tournier, R.E. Smith, D. Raffelt, R. Tabbara, T. Dhollander, M. Pietsch, D. Christiaens, B. Jeurissen, C.H. Yeh, and A. Connelly. MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, 202:116–137, 2019.

[48] B. Fischl, D.H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A.M. Dale. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002.

[49] F. Schweser, A. Deistung, B.W. Lehr, and J.R. Reichenbach. Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: an approach to in vivo brain iron metabolism? *NeuroImage*, 54(4):2789–2807, 2011.

[50] W. Li, B. Wu, and C. Liu. Quantitative susceptibility mapping of human brain reflects spatial variation in tissue composition. *NeuroImage*, 55(4):1645–1656, 2011.

[51] E. Biondetti, A. Karsa, F. Grussu, M. Battiston, M.C. Yiannakas, D.L. Thomas, and K. Shmueli. Multi-echo quantitative susceptibility mapping: how to combine echoes for accuracy and precision at 3 Tesla. *Magnetic Resonance in Medicine*, 88(5):2101–2116, 2022.

[52] W. Li, B. Wu, and C. Liu. A software package for quantitative susceptibility imaging. *ISMRM-ESMRMB*, 2014.

[53] C. Fiscone, C. Testa, S. Evangelisti, C. Bianchini, R. Lodi, and C. Tonon. Nigrosome-1 detection in quantitative susceptibility mapping: contrast optimization with 64-channel coil. *First Annual Meeting Rete RIN – IRCCS Network of Neuroscience and Neurorehabilitation (Selected Abstract)*, 2020.

[54] K. Eckstein, B. Dymerska, B. Bachrata, W. Bogner, K. Poljanc, S. Trattnig S, and S.D. Robinson. Computationally efficient combination of multi-channel phase data from multi-echo acquisitions (ASPIRE). *Magnetic Resonance in Medicine*, 79(6):2996–3006, 2018.

[55] M. Lancione, G. Donatelli, E. Del Prete, N. Campese, D. Frosini, M. Cencini M, M. Costagli, L. Biagi, G. Lucchi, M. Tosetti, M. Godani, D. Arnaldi, M. Terzaghi, F. Provini, C. Pacchetti, P. Cortelli, E. Bonanni, R. Ceravolo, and M. Cosottini. Evaluation of iron overload in nigrosome 1 via quantitative susceptibility mapping as a progression biomarker in prodromal stages of synucleinopathies. *NeuroImage*, 260:119454, 2022.

[56] F. Schweser, A. Deistung, and J.R. Reichenbach. Foundations of MRI phase imaging and processing for quantitative susceptibility mapping (QSM). *Zeitschrift für Medizinische Physik*, 26(1):6–34, 2016.

[57] S. Straub, T.M. Schneider, J. Emmerich, M.T. Freitag, C.H. Ziener, H.P. Schlemmer, M.E. Ladd, and F.B. Laun. Suitable reference tissues for quantitative susceptibility mapping of the brain. *Magnetic Resonance in Medicine*, 78(1):204–214, 2017.

[58] Y. Wang and T. Liu. Quantitative susceptibility mapping (QSM): decoding MRI data for a tissue magnetic biomarker. *Magnetic Resonance in Medicine*, 73(1):82–101, 2015.

[59] N. Yaghmaie, W.T. Syeda, C. Wu, Y. Zhang, T.D. Zhang, E.L. Burrows, A. Brodtmann, B.A. Moffat, D.K. Wright, R. Glarin, S. Kolbe, and L.A. Johnston. QSMART. Quantitative susceptibility mapping artifact reduction technique. *NeuroImage*, 231:117701, 2021.

[60] J.H. Lee and M.S. Lee. Brain iron accumulation in atypical parkinsonian syndromes: in vivo MRI evidences for distinctive patterns. *Frontiers in Neurology*, 10(74), 2019.

[61] M. Marcille, S. Hurtado Rúa, C. Tyshkov, A. Jaywant, J. Comunale, U.W. Kaunzner, N. Nealon, J.S. Perumal, L. Zexter, N. Zinger, O. Bruvik, Y. Wang, E. Sweeney, A. Kuceyeski, T.D. Nguyen, and S.A. Gauthier. Disease correlates of rim lesions on quantitative susceptibility mapping in Multiple Sclerosis. *Scientific Reports*, 12(1):4411, 2022.

[62] M. Absinta, P. Sati, F. Masuzzo, G. Nair, V. Sethi, H. Kolb, J. Ohayon, T. Wu, I.C.M. Cortese, and D.S. Reich. Association of chronic active Multiple Sclerosis lesions with disability in vivo. *JAMA Neurology*, 76(12):1474–1483, 2019.

[63] S. Zhang, T.D. Nguyen, S.M.Hurtado Rua, U.W. Kaunzner, S. Pandya, I. Kovanilikaya, P. Spincemaille, Y. Wang, and S.A. Gauthier. Quantitative susceptibility mapping of time-dependent susceptibility changes in Multiple Scleorosis lesions. *American Journal of Neuroradiology*, 40(6):987–993, 2019.

[64] B. Dymerska, K. Eckstein, B. Bachrata, B. Siow, S. Trattnig, K. Shmueli, and S.D. Robinson. Phase unwrapping with a rapid opensource minimum spanning tree algorithm (ROMEO). *Magnetic Resonance in Medicine*, 85(4):2294–2308, 2021.

[65] C. Langkammer, K. Bredies, B.A. Poser, M. Barth, G. Reishofer, A.P. Fan, B. Bilgic, F. Fazekas, C. Mainero, and S. Ropele. Fast quantitative susceptibility mapping using 3D EPI and total generalized variation. *NeuroImage*, 111:622–630, 2015.

[66] H.G. Shin, J. Lee, Y.H. Yun, S.H. Yoo, J. Jang, S.H. Oh, Y. Nam, S. Jung, S. Kim, M. Fukunaga, W. Kim, H.J. Choi, and J. Lee. χ-separation: magnetic susceptibility source separation toward iron and myelin mapping in the brain. *NeuroImage*, 240:118371, 2021.

[67] W. Kim, H.G. Shin, H. Lee, D. Park, J. Kang, Y. Nam, J. Lee, and J. Jang. χ-separation imaging for diagnosis of Multiple Sclerosis versus Neuromyelitis Optica Spectrum Disorder. *Radiology*, 307(1):e220941, 2023.

[68] R.J. Ward, F.A. Zucca, J.H. Duyn, R.R. Crichton, and L. Zecca. The role of iron in brain ageing and neurodegenerative disorders. *Lancet Neurology*, 13(10):1045–1060, 2014.

[69] J. Acosta-Cabronero, M.J. Betts, A. Cardenas-Blanco, S. Yang, and P.J. Nestor. In vivo MRI mapping of brain iron deposition across the adult lifespan. *Journal of Neuroscience*, 36(2):364–374, 2016.

[70] G. Li, R. Tong, M. Zhang, K.M. Gillen, W. Jiang, Y. Du, Y. Wang, and J. Li. Age-dependent changes in brain iron deposition and volume in deep gray matter nuclei using quantitative susceptibility mapping. *NeuroImage*, 269, 2023.

[71] Y. Li, S.K. Sethi, C. Zhang, Y. Miao, K.K. Yerramsetty, V.K. Palutla, S. Gharabaghi, C. Wang, N. He, J. Cheng, F. Yan, and E.M. Haacke. Iron content in deep gray matter as a function of age using quantitative susceptibility mapping: a multicenter study. *Frontiers in Neuroscience*, 14:607705, 2021.

[72] O. Hardiman, A. Al-Chalabi, A. Chio, E.M. Corr, G. Logroscino, W. Robberecht, P.J. Shaw, Z. Simmons, and L.H. van den Berg. Amyotrophic Lateral Sclerosis. *Nature Reviews Disease Primers*, 3, 2017.

[73] J. Acosta-Cabronero, J. Machts, S. Schreiber, S. Abdulla, K. Kollewe, S. Petri, N. Spotorno, J. Kaufmann, H.J. Heinze, R. Dengler, S. Vielhaber, and P.J. Nestor. Quantitative susceptibility MRI to detect brain iron in Amyotrophic Lateral Sclerosis. *Radiology*, 289(1):195–203, 2018.

[74] A. Bhattarai, G.F. Egan, P. Talman, P. Chua, and Z. Chen. Magnetic resonance iron imaging in Amyotrophic Lateral Sclerosis. *Journal of Magnetic Resonance Imaging*, 55:1283–1300, 2022.

[75] C. Wolfson, D.E. Gauvin, F. Ishola, and M. Oskoui. Global prevalence and incidence of Amyotrophic Lateral Sclerosis: a systematic review. *Neurology*, 101:e613–e623, 2023.

[76] J.Y. Kwan, S.Y. Jeong, P. Van Gelderen, H.X. Deng, M.M. Quezado L.E. Danielian, J.A. Butman, L. Chen, E. Bayat, J. Russell, T. Siddique, J.H. Duyn, T.A. Rouault, and M.K. Floeter. Iron accumulation in deep cortical layers accounts for MRI signal abnormalities in ALS: correlating 7 Tesla MRI and pathology. *PLoS One*, 2012.

[77] I.C. Nnah and M. Wessling-Resnick. Brain iron homeostasis: a focus on microglia iron. *Pharmaceutical*, 11:129, 2018.

[78] M.C. Kiernan, S. Vucic, B.C. Cheah, M.R. Turner, A. Eisen, O. Hardiman, J.R. Burrell, and M.C. Zoing. Amyotrophic Lateral Sclerosis. *Lancet*, 377(9769):942–955, 2011.

[79] J. Rosenfeld and M.J. Strong. Challenges in the understanding and treatment of Amyotrophic Lateral Sclerosis/Motor Neuron Disease. *Neurotherapeutics*, 12:317–325, 2015.

[80] B.R. Brooks, R.G. Miller, M. Swash, T.L. Munsat, and World Federation of Neurology Research Group on Motor Neuron Diseases. El Escorial revisited: revised criteria for the diagnosis of Amyotrophic Lateral Sclerosis. *Amyotrophic Lateralal Sclerosis and Other Motor Neuron Disorders*, 1(5):293–299, 2000.

[81] A.D. Schweitzer, T. Liu, A. Gupta, K. Zheng, S. Seedial, A. Shtilbans, M. Shahbazi, D. Lange, Y. Wang, and A.J. Tsiouris. Quantitative susceptibility mapping of the motor cortex in Amyotrophic Lateral Sclerosis and Primary Lateral Sclerosis. *American Journal of Roentgenology*, 2015.

[82] M. Costagli, G. Donatelli, L. Biagi, I.E. Caldarazzo, G. Siciliano, and M. Tosetti. Magnetic susceptibility in the deep layers of the primary motor cortex in Amyotrophic Lateral Sclerosis. *NeuroImage: Clinical*, 2(12):965–696, 2016.

[83] C. Zejlon, D. Nakhostin, S. Winklhofer, A. Pangalu, Z. Kulcsar, S. Lewandowski, J. Finnsson, F. Piehl, C. Ingre, T. Granberg, and B.V. Ineichen. Structural magnetic resonance imaging findings and histopathological correlations in motor neuron diseases - A systematic review and meta-analysis. *Frontiers in Neurology*, 2022.

[84] V.E. Contarino, G. Conte, C. Morelli, F. Trogu, E. Scola, S.F. Calloni, L.C. Sanmiguel Serpa, C. Liu, V. Silani, and F. Triulzi. Toward a marker of upper motor neuron impairment in Amyotrophic Lateral Sclerosis: a fully automatic investigation of the magnetic susceptibility in the precentral cortex. *European Journal of Radiology*, 124:108815, 2020.

[85] G. Conte, V.E. Contarino, S. Casale, C. Morelli, S. Sbaraini, F. Trogu, S. Siggillino, C.M. Cinnante, L. Caschera, F.M. Lo Russo, F.M. Triulzi, and V. Silani. Amyotrophic Lateral Sclerosis phenotypes significantly differ in terms of magnetic susceptibility properties of the precentral cortex. *European Radiology*, 31(7):52272–52280, 2021.

[86] M. Costagli, G. Donatelli, P. Cecchi, P. Bosco, G. Migaleddu, G. Siciliano, and M. Cosottini. Distribution indices of magnetic susceptibility values in the primary motor cortex enable to classify patients with Amyotrophic Lateral Sclerosis. *Brain Sciences*, 12(7):942, 2022.

[87] C. Schuster, E. Kasper, J. Machts, D. Bittner, J. Kaufmann, R. Benecke, S. Teipel, S. Vielhaber, and J. Prudlo. Longitudinal course of cortical thickness decline in Amyotrophic Lateral Sclerosis. *Journal of Neurology*, 261(10):1871–1880, 2014.

[88] C. Fiscone, F. Bartiromo, V. Vacchiano, M. Mitolo, G. Rizzo, C. Bianchini, D.N. Manners, C. Testa, C. Tonon, R. Liguori, and R. Lodi. In vivo investigation of susceptibility properties of cerebral gyri in Amyotrophic Lateral Sclerosis. *Italian-French International Conference on Magnetic Resonance (GIDRM – GERM) (Selected Abstract)*, 2022.

[89] C. Fiscone, F. Bartiromo, V. Vacchiano, C. Bianchini, D.N. Manners, C. Testa, F. Salvi, R. Liguori, R. Lodi, and C. Tonon. Quantitative susceptibility-based imaging in Amyotrophic Lateral Sclerosis: investigation of iron concentration in several cortical gyri. *XII Congress AIRMM - ISMRM Italian Chapter (Selected Abstract)*, 2021.

[90] T.D. Kocar, H.P. Muller, A.C. Ludolph, and J. Kassubek. Feature selection from magnetic resonance imaging data in ALS. A systematic review. *Therapeutic Advances in Chronic Disease*, 12, 2021.

[91] M.R. Turner and K. Talbot. Mimics and chamaleons in motor neuron disease. *Practical Neurology*, 13:153–164, 2013.

[92] B. Brooks, M. Sanjak, S. Ringel, J. England, and J. Brinkmann. The Amyotrophic Lateral Sclerosis Functional Rating Scale: assessment of activities of daily living in patients with Amyotrophic Lateral Sclerosis. *Archives of Neurology*, 53(2):141–147, 1996.

[93] C. Quinn, C. Edmundson, N. Dahodwala, and L. Elman. Reliable and efficient scale to assess upper motor neuron disease burden in Amyotrophic Lateral Sclerosis. *Muscle & Nerve*, 61:508–511, 2020.

[94] V. Vacchiano, A. Mastrangelo, C.Zenesini, M. Masullo, C. Quadalti, P. Avoni, B. Polischi, A. Cherici, S. Capellari, F. Salvi, R. Liguori, and P. Parchi. Plasma and CSF neurofilament light chain in Amyotrophic Lateral Sclerosis. A cross-sectional and longitudinal study. *Frontiers in Aging Neuroscience*, 13:753242, 2021.

[95] B. Kressler, L. de Rochefort, T. Liu, P. Spincemaille, Q. Jiang, and Y. Wang. Nonlinear regularization for per voxel estimation of magnetic susceptibility distributions from MRI field maps. *IEEE Transaction on Medical Imaging*, 29(2):273–281, 2010.

[96] P.R. Shrout and J.L. Fleiss. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2):420–428, 1979.

[97] F.S. Nahm. Receiver operating characteristic curve: overview and practical use for clinicians. *Korean Journal of Anesthesiology*, 75(1):25–36, 2022.

[98] K. Ghassaban, N. He, S.K. Sethi, P. Huang, S. Chen, F. Yan, and E.M. Haacke. Regional high iron in the substantia nigra differentiates Parkinson's disease patients from healthy controls. *Frontiers in Aging Neuroscience*, 11:106, 2019.

[99] C. Sanchis-Segura, M.V. Ibanez-Gual, J. Adrian-Ventura, N. Aguirre, A.J. Gomez-Cruz, C. Avila, and C. Forn. Sex differences in gray matter volume: how many and how large are they really? *Biology of Sex Differences*, 10(32), 2019.

[100] Y. Adachi, N. Sato, Y. Saito, Y. Kimura, Y. Nakata, K. Ito, K. Kamiya, H. Matsuda, T. Tsukamoto, and M. Ogawa. Usefulness of SWI for the detection of iron in the motor cortex in Amyotrophic Lateral Sclerosis. *Journal of Neuroimaging*, 25(3):443–451, 2015.

[101] K.E. Dean, B. Shen, G. Askin, A.D. Schweitzer, M. Shahbazi M., Y. Wang, D. Lange, and A.J. Tsiouris. A specific biomarker for Amyotrophic Lateral Sclerosis: quantitative susceptibility mapping. *Clinical Imaging*, 75:125–130, 2021.

[102] M. Cosottini, G. Donatelli, M. Costagli, E. Caldarazzo Ienco, D. Frosini, I. Pesaresi, L. Biagi, G. Siciliano, and M. Tosetti. High-resolution 7T MR imaging of the motor cortex in Amyotrophic Lateral Sclerosis. *American Journal of Neuroradiology*, 37(3):455–461, 2016.

[103] R. Burgetova, P. Dusek, A. Burgetova, A. Pudlac, M. Vaneckova, D. Horakova, J. Krasensky, Z. Varga, and L. Lambert. Age-related magnetic susceptibility changes in deep grey matter and cerebral cortex of normal young and middle-aged adults depicted by whole brain analysi. *Quantitative Imaging in Medicine and Surgery*, 11(9):3906–3919, 2021.

[104] H. Zhang, T.D. Nguyen, J. Zhang, M. Marcille, P. Spincemaille, Y. Wang, S.A. Gauthier, and E.M. Sweeney. QSMRim-Net: imbalance-aware learning for identification of chronic active Multiple Sclerosis lesions on quantitative susceptibility maps. *NeuroImage: Clinical*, 34, 2022.

[105] K. Okkersen, D.G. Monckton, N. Le, A.M. Tuladhar, J. Raaphorst, and B.G.M. van Engelen. Brain imaging in Myotonic Dystrophy type 1: a systematic review. *Neurology*, 89(9), 2017.

[106] S. Ates, A. Deistung, R. Schneider, C. Prehn, C. Lukas, J.R. Reichenbach, C. Schneider-Gold, and B. Bellenberg. Characterization of iron accumulation in deep gray matter in Myotonic Dystrophy type 1 and 2 using quantitative susceptibility mapping and R2* relaxometry: a magnetic resonance imaging study at 3 Tesla. *Frontiers in Neurology*, 2019.

[107] P.S. Harper. Myotonic Dystrophy. *London: W.B. Saunders*, 3^{rd} edition, 2001.

[108] G. Meola and R. Cardani. Myotonic Dystrophies: an update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1852(4):594–606, 2015.

[109] M. De Antonio, C. Dogan, D. Hamroun, M. Mati, S. Zerrouki, B. Eymard, S. Katsahian, and G. Bassez. French Myotonic Dystrophy Clinical Network. Unravelling the Myotonic Dystrophy type 1 clinical spectrum: a systematic registry-based study with implications for disease classification. *Revue neurologique (Paris)*, 172(10):572–580, 2016.

[110] M. Stokes, N. Varughese, S. Iannacone, and D. Castro. Clinical and genetic characteristics of childhood-onset Myotonic Dystrophy. *Muscle Nerve*, 60(6):732–738, 2019.

[111] S. Wenninger, F. Montagnese, and B. Schoser. Core clinical phenotypes in Myotonic Dystrophies. *Frontiers in Neurology*, 9(303), 2018.

[112] N.E Johnson and C.R Heatwole. Myotonic Dystrophy: from bench to bedside. *Seminars in Neurology*, 32(3):246–254, 2012.

[113] A.M. Hawkins, C.L. Hawkins, K. Abdul Razak, T.K. Khoo, K. Tran, and R.V. Jackson. Respiratory dysfunction in Myotonic Dystrophy type 1: a systematic review. *Neuromuscular Disorders*, 29(3), 2019.

[114] K. Wahbi and D. Furling. Cardiovascular manifestations of Myotonic Dystrophy. *Trends in Cardiovascular Medicine*, 30(4):232–238, 2020.

[115] C. Simoncini, G. Spadoni, E. Lai, L. Santoni, C. Angelini, G. Ricci, and G. Siciliano. Central nervous system involvement as outcome measure for clinical trials efficacy in Myotonic Dystrophy type 1. *Frontiers in Neurology*, 11(624), 2020.

[116] S. Baldanzi, F. Bevilacqua, R. Lorio, L. Volpi, C. Simoncini, A. Petrucci, M. Cosottini, G. Massimetti, G. Tognoni, G. Ricci, C. Angelini, and G. Siciliano. Disease awareness in Myotonic Dystrophy type 1: an observational cross-sectional study. *Orphanet Journal of Rare Diseases*, 11(34), 2016.

[117] E. Callus, E.G. Bertoldo, M. Beretta, S. Boveri, R. Cardani, B. Fossati, E. Brigonzi, and G. Meola. Neuropsychological and psychological functioning aspects in Myotonic Dystrophy type 1 patients in Italy. *Frontiers in Neurology*, 9(751), 2018.

[118] A. Giorgio, M.T. Dotti, M. Battaglini, S. Marino, M. Mortilla, M.L. Stromillo, P. Bramanti, A. Orrico, A. Federico, and S. De Stefano. Cortical damage in brains of patients with adult-form of Myotonic Dystrophy type 1 and no or minimal MRI abnormalities. *Journal of Neurology*, 253(11):1471–1477, 2006.

[119] E. van der Plas, M.J. Hamilton, J.N. Miller, T.R. Koscik, J.D. Long, S. Cumming, J. Povilaikaite, M.E. Farrugia, J. McLean, R. Jampana, V.A. Magnotta, L. Gutmann, D.G. Monckton, and P.C. Nopoulos. Brain structural features of Myotonic Dystrophy type 1 and their relationship with CTG repeats. *Journal of Neuromuscolar Diseases*, 6(3):321–332, 2019.

[120] A. Di Costanzo, F. Di Salle, L. Santoro, V. Bonavita, and G. Tedeschi. Brain MRI features of congenital- and adult- form Myotonic Dystrophy type 1: case-control study. *Neuromuscular Disorders*, 12:476–483, 2002.

[121] G. Labayru, B. Camino, A. Jimenez-Marin, J. Garmendia, J. Villanua, M. Zulaica, J.M. Cortes, A.L. de Munain, and A. Sistiaga. White matter integrity changes and neurocognitive functioning in adult-late onset DM1: a follow-up DTI study. *Scientific Reports*, 12(3988), 2022.

[122] L. Serra, G. Silvestri, A. Petrucci, B. Basile, M. Masciullo, E. Makovac, M. Torso, B. Spanò, C. Mastropasqua, N.A. Harrison, M.L. Bianchi, M. Giacanelli, C. Caltagirone, M. Cercignani, and M. Bozzali. Abnormal functional brain connectivity and personality traits in Myotonic Dystrophy type 1. *JAMA Neurology*, 71(5):603–611, 2014.

[123] S. Evangelisti, L.L. Gramegna, S. De Pasqua M.J. Rochat, L. Morandi, M. Mitolo, C. Bianchini, G. Vornetti, C. Testa, P. Avoni, R. Liguori, R. Lodi, and C. Tonon. In vivo parieto-occipital white matter metabolism is correlated with visuospatial deficits in adult DM1 patients. *Diagnostics*, 12(10):2305, 2022.

[124] C. Fiscone, M.J. Rochat, S. Evangelisti, S. De Pasqua, M. Mitolo, C. Bianchini, G. Vornetti, F. Bartiromo, L. Cirignotta, D.N. Manners, C. Testa, P. Avoni, R. Liguori, R. Lodi, and C. Tonon. Brainstem and thalamus susceptibility in Myotonic Dystrophy type 1: quantitative imaging biomarkers of disease severity. *XIV Congress AIRMM 2023 – ISMRM Italian Chapter (Selected Abstract)*, 2023.

[125] S. De Pasqua. *Myotonic Syndromes: analysis of factor with pathogenetic and prognostic significance*. PhD thesis, University of Bologna, Department of Biomedical and Neuromotor Sciences, 2022.

[126] J. Mathieu, H. Boivin, D. Meunier, M. Gaudreault, and P. Bégin. Assessment of a disease-specific muscular impairment rating scale in Myotonic Dystrophy. *Neurology*, 56(3):336–340, 2001.

[127] S. Contardi, F. Pizza, F. Falzone, R. D'Alessandro, P. Avoni, V. Di Stasi, P. Montagna, and R. Liguori. Development of a disability scale for Myotonic Dystrophy type 1. *Acta Neurologica Scandinavica*, 125(6):431–438, 2012.

[128] G. Brun, B. Testud, O.M. Girard, P. Lehmann, L. de Rochefort, P. Besson, A. Massire, B. Ridley, N. Girard, M. Guye, J.P. Ranjeva, and A. Le Troter. Automatic segmentation of deep grey nuclei using a high-resolution 7T magnetic resonance imaging atlas - Quantification of T1 values in healthy volunteers. *European Journal of Neuroscience*, 55(2):438–460, 2022.

[129] J.E. Iglesias, K. Van Leemput, P. Bhatt, C. Casillas, S. Dutt, N. Schuff, D. Truran-Sacrey, A. Boxer, and B. Fischl. Bayesian segmentation of brainstem structures in MRI. *NeuroImage*, 113:184–195, 2015.

[130] B. Bilgic, A. Pfefferbaum, T. Rohlfing, E.V. Sullivan, and E. Adalsteinsson. MRI estimates of brain iron concentration in normal aging using quantitative susceptibility mapping. *NeuroImage*, 59(3):2625–2635, 2012.

[131] S. Hanoun, Y. Sun, F. Ebrahimi, and M. Ghasemi. Speech and language abnormalities in Myotonic Dystrophy: an overview. *Journal of Clinical Neuroscience*, 96:212–220, 2022.

[132] V. Bernhardt, N. Garcia-Reyero, A. Vovk, N. Denslow, and P.W. Davenport. Tracheal occlusion modulates the gene expression profile of the medial thalamus in anesthetized rats. *Journal of Applied Physiology (1985)*, 111(1):117–124, 2011.

[133] M.J. Hamilton, A. Atalaia, J. McLean, S.A. Cumming, J.J. Evans, B. Ballantyne, R. Jampana, The Scottish Myotonic Dystrophy Consortium, C. Longman, E. Livingston, E. van der Plas, T. Koscik, P. Nopoulos, M.E. Farrugia, and D.G. Monckton. Clinical and neuroradiological correlates of sleep in Myotonic Dystrophy type 1. *Neuromuscular Disorders*, 32(5), 2022.

[134] J. Nepozitek, Z. Varga, S. Dostalova, P. Perinova, J. Keller, S. Robinson, V. Ibarburu, I. Prihodova, O. Bezdicek, E. Ruzicka, K. Sonka, and P. Dusek. Magnetic susceptibility changes in the brainstem reflect REM sleep without atonia severity in isolated REM sleep behavior disorder. *NPJ Parkinson's Disease*, 9(1), 2023.

[135] C. Fiscone, N. Curti, M. Ceccarelli, D. Remondini, C. Testa, R. Lodi, C. Tonon, D.N. Manners, and G. Castellani. Generalizing the Enhanced-Deep-Super-Resolution neural network to brain MR images: a retrospective study on the Cam-CAN dataset. *accepted for publication by eNeuro*, 2023.

[136] R. Gillies, P. Kinahan, and H. Hricak. Radiomics: images are more than pictures, they are data. *Radiology*, 276:563–577, 2016.

[137] J. Guiot, A. Vaidyanathan, L. Deprez, F. Zerka, D. Danthine, A.N. Frix, P. Lambin, F. Bottari, N. Tsoutzidis, B. Miraglio, S. Walsh, W. Vos, R. Hustinx, M. Ferreira, P. Lovinfosse, and R.T.H. Leijenaar. A review in radiomics: making personalized medicine a reality via routine imaging. *Medicinal Research Reviews*, 42:426–440, 2022.

[138] A.A. Ardakani, N.J. Bureau, E.J. Ciaccio, and U.R.Archarya. Interpretation of radiomics features - A pictorial review. *Computer Methods and Programs in Biomedicine*, 2022.

[139] N. Weiskopf, L.J. Edwards, G. Helms, S. Mohammadi, and E. Kirilina. Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. *Nature Reviews Physics*, 3:570–588, 2021.

[140] A. Zwanenbrug. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-troughput image-based phenotyping. *Radiology*, 295:328–338, 2020.

[141] Biomarkers Definition Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics*, 69:89–95, 2001.

[142] E.T.S. Smith. Clinical applications of imaging biomarkers. Part 1. The neuroradiologist's perspective. *British Journal of Radiology*, 84(S2):S196–S204, 2011.

[143] K. Strimbu and J.A. Tavel. What are biomarkers? *Current Opinion in HIV and AIDS*, 5(6):463–466, 2010.

[144] A. Zwanenbrug, S. Leger, L. Agolli, K. Pilz, E.G.C. Troost, C. Richter, and S. Löck. Assessing robustness of radiomic features by image perturbation. *Scientific Reports*, 9(1):614, 2019.

[145] M. Shafiq ul Hassan, K. Latifi, G. Zhang, G. Ullah, R. Gillies, and E. Moros. Voxel size and gray level normalization of CT radiomic features in lung cancer. *Scientific Reports*, 8(1):10545, 2018.

[146] M. Mottola, S. Ursprung, L. Rundo, L.E. Sanchez, T. Klatte, I. Mendichovszky, G.D. Stewart, E. Sala, and A. Bevilacqua. Reproducibility of CT-based radiomic features against image resampling and perturbations for tumour and healthy kindney in renal cancer patients. *Scientific Reports*, 11(1):11542, 2021.

[147] L.R. Sanchez LE, L. Rundo, A.B. Gill, M. Hoare, E.M. Serrao, and E. Sala. Robustness of radiomic features in CT images with different slice thickness, comparing liver tumor and muscle. *Scientific Reports*, 11(1):8262, 2021.

[148] L.E. Sanchez, E. Brown, L. Rundo, S. Ursprung, E. Sala, S.E. Bohndiek, and K. Partattieu. Photoacustic imaging radiomics in patient-derived xenografts: a study on feature sensitivity and model discrimination. *Scientific Reports*, 12(1):15142, 2022.

[149] J.Y. Kim, J.E. Park, Y. Jo, W.M. Shim, S.J. Nam, and J.H. Kim. Incorporating diffusion- and perfusion-weighted MRI into radiomics model improves diagnostic performance for pseudogression in glioblastoma patients. *Neuro-Oncology*, 21(3):404–414, 2019.

[150] G. Li, L. Li, Y. Li, Z. Qian, F. Wu, Y. He, H. Jiang, R. Li, D. Wang, Y. Zhai, Z. Wang, T. Jiang, J. Zhang, and W. Zhang. An MRI radiomics approach to predict survival and tumor-infiltrating macrophages in gliomas. *Brain*, 143(3):1151–1161, 2022.

[151] P. Kickingereder and O.C. Andronese. Radiomics, metabolic and molecular MRI for brain tumors. *Seminars in Neurology*, 38(1):32–40, 2018.

[152] D.S. Reich, C. F. Lucchinetti, and P. A. Calabresi. Multiple Sclerosis. *New England Journal of Medicine*, 378:169–180, 2018.

[153] A. Compston and A. Coles. Multiple Sclerosis. *Lancet*, 372:1502–1517, 2008.

[154] A.J.T. Thompson, B.L. Banwell, F. Barkhof, W.M. Carrol, T. Coetzee, G. Comi, J. Correale, F. Fazekas, M. Filippi, M.S. Freedman, K. Fujihara, S.L. Galetta, H.P. Hartung, L. Kappos, F.D. Lublin, R.A. Marrie, A.E. Miller, D.H. Miller, X. Montalban, E.M. Mowry, P.S. Sorensen, M. Tintoré, A.L. Traboulsee, M. Trojano, B.M.J Uitdehaag, S. Vukusic, E. Waubant, B.G. Weinshenker, S.C. Reingold, and J.A. Cohen. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurology*, 17(2):162–173, 2018.

[155] M.P. Wattjes, O. Ciccarelli, D.S. Reich, B. Banwell, N. de Stafano, C. Enzinger, F. Fazekas, M. Filippi, J. Frederiksen, C. Gasperini, Y. Hacohen, L. Kappos, D.K.B. Li, K. Mankad, X. Montalban, S.D. Newsome, J. Oh, J. Palace, M.A. Rocca, J. Sastre-Garriga, M. Tintoré, A. Traboulsee, H. Vrenken, T. Yousry, F. Barkhof, A. Rovira, Magnetic Resonance Imaging in Multiple Sclerosis study group, Consortium of Multiple Sclerosis Centres, and North American Imaging in Multiple Sclerosis Cooperative MRI guidelines working group. 2021 MAGNIMS-CMSC-NAIMS consensus recomandations on the use of MRI in patients with multiple sclerosis. *Lancet Neurology*, 20(8):653–670, 2021.

[156] M. Filippi, M.A. Rocca, O. Ciccarelli, N. De Stefano, N. Evangelou, L. Kappos, A. Rovira, J. Sastre-Garriga, M. Tintorè, J.L. Frederiksen, C. Gasperini, J. Palace, D.S. Reich, B. Banwell, X. Montalban, F. Barkhof, and MAGNIMS Study Group. MRI criteria for the diagnosis of Multiple Sclerosis: MAGNIMIS consensus guidelines. *Lancet Neurol*, 15(3):292–303, 2016.

[157] P. Sowa, H.F. Harbo, N.S. White, E.G. Celius, H. Bartsch, P. Berg-Hansen, S.M. Moen, A. Bjørnerud, L.T. Westlye, O.A. Andreassen, A.M. Dale, and M.K. Beyer. Restriction spectrum imaging of white matter and its relation to neurological disability in Multiple Sclerosis. *Multiple Sclerosis*, 25(5):687–698, 2019.

[158] J.T. Grist, F. Riemer, M.A. McLean, T. Matys, F. Zaccagna, S.F. Hilborne, J.P. Mason, I. Patterson, R. Slough, J. Kaggie, S.S. Deen, M.J. Graves, J.L. Jones, A.J. Coles, and F.A. Gallagher. Imaging intralesional heterogeneity of sodium concentration in Multiple Sclerosis: initial evidence from ^{23}Na-MRI. *Journal of Neurological Sciences*, 387:111–114, 2018.

[159] R. Zivadinov, R. Tavazzi, N. Bergsland, J. Hagemeier, F. Lin, M.G. Dwyer, E. Carl, C. Kolb, D. Hojnacki, D. Ramasamy, J. Durfee, B. Weinstock-Guttman, and F. Schweser. Brain iron at quantitative MRI is associated with disability in Multiple Sclerosis. *Radiology*, 289(2):487–496, 2018.

[160] M. Tranfa M, G. Pontillo, M. Petracca, A. Brunetti, E. Tedeschi, G. Palma, and S. Cocozza. Quantitative MRI in Multiple Sclerosis: from theory to application. *American Journal of Neuroradiology*, 43(12):1688–1695, 2022.

[161] C. Fiscone, L. Rundo, A. Lugaresi, D.N. Manners, K. Allinson, E. Baldin, R. Lodi, C. Tonon, C. Testa, M. Castelli, and F. Zaccagna. Robustness analysis of QSM radiomic features in patients with Multiple Sclerosis. *2023 ISMRM & ISMRT Annual Meeting & Exhibition (Selected Abstract)*, 2023.

[162] C. Fiscone, L. Rundo, A. Lugaresi, D.N. Manners, K. Allinson, E. Baldin, G. Vornetti, R. Lodi, C. Tonon, C. Testa, M. Castelli, and F. Zaccagna. Reliability of susceptibility-based features in patients with Multiple Sclerosis: methodological pipeline implementation exploring normal appearing white matter tracts. *XIV Congress AIRMM 2023 – ISMRM Italian Chapter (Selected Abstract)*, 2023.

[163] W. Chen, Y. Zhang, K. Mu, C. Pan, S.A. Gauthier, W. Zhu, and Y. Wang. Quantifying the susceptibility variation of normal-appearing white matter in Multiple Sclerosis by quantitative susceptibility mapping. *American Journal of Roentgenology*, 209(4):889–894, 2017.

[164] F.F. Yu, F.L. Chiang, N. Stephens, S.Y. Huang, B. Bilgic, B. Tantiwongkosi, and R. Romero. Characterization of normal-appearing white matter in Multiple Sclerosis using quantitative susceptibility mapping in conjunction with diffusion tensor imaging. *Neuroradiology*, 61(1):71–79, 2019.

[165] B. Xiao, N. He, Q. Wang, Z. Cheng, Y. Jiao, E.M. Haacke, F. Yan, and F. Shi. Quantitative susceptibility mapping based hybrid feature extraction for diagnosis of Parkinson's disease. *NeuroImage: Clinical*, 24, 2019.

[166] J.J. Kang, Y. chen, G.D. Xu, S.L. Bao, J. Wang, M. Ge, L.H. Shen, and Z.Z. Jia. Combining quantitative susceptibility mapping to radiomics in diagnosing Parkinson's disease and assessing cognitive impairment. *European Radiology*, 32:6992–7003, 2022.

[167] P. Schmidt, C. Gaser, M. Arsic, D. Buck, A. Förschler, A. Berthele, M. Hoshi, R. Ilg, V.J. Schmid, C. Zimmer, B. Hemmer, and M. Mühlau. An automated tool for detection of FLAIR hyperintense white-matter lesion in Multiple Sclerosis. *NeuroImage*, 59(4):3774–3783, 2021.

[168] C. Guo, D. Ferreira, K. Fink, E. Westman, and R. Granberg. Repeatability and reproducibility of FreeSurfer, FSL-SIENAX and SPM brain volumetric measurements and the effect of leson filling in Multiple Sclerosis. *European Radiology*, 29:1355–1364, 2019.

[169] C.H. Coffman, R. White, K. Subramaninan, S. Buch, E. Bernitsas, and E.M. Haacke. Quantitative susceptibility mapping of both ring and non-ring white matter lesions in relapsing remitting Multiple Sclerosis. *Magnetic Resonance Imaging*, 91:45–51, 2002.

[170] L. Talozzi, C. Testa, S. Evangelisti, L. Cirignotta, C. Bianchini, S. Ratti, P. Fantazzini, C. Tonon, D.N. Manners, and R. Lodi. Along-tract analysis of the arcuate fasciculus using the Laplacian operator to evaluate different tractography methods. *Magnetic Resonance Imaging*, 54:183–193, 2018.

[171] M. Zoli, L. Talozzi, M. Martinoni, D.N. Manners, F. Badaloni, and C. Testa. From neurosurgical planning to histopathological brain tumor characterization: potentialities of arcuate fasciculus along-tract diffusion tensor imaging tractography measures. *Frontiers in Neurology*, 12:633209, 2021.

[172] D.N. Manners, L.L. Gramegna, C. La Morgia, G. Sighinolfi, C. Fiscone, M. Carbonelli, M. Romagnoli, V. Carelli, C. Tonon, and R. Lodi. Multishell diffusion MR tractography yields morphological and microstructural information of the anterior optic pathway: a proof-of-concept study in patients with Leber's Hereditary Optic Neuropathy. *International Journal of Environmental Research and Public Health*, 19(11):6914, 2022.

[173] J.L.R. Andersson and S.N. Sotiropoulos. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, 125:1063–1078, 2016.

[174] J.D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459–1472, 2007.

[175] J.J.M. Van Griethuysen, A. Federov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R.G.H. Beets-Tan, J.C. Fillion-Robin, S. Pieper, and H.J.W.L. Aerts. Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, 77(21):e104–107, 2017.

[176] P. Salome, F. Sforazzini, G. Brugnara, A. Kudak, M. Dostal, C. Herold-Mende, S. Heiland, J. Debus, A. Abdollahi, and M. Knoll. MR intensity normalization methods impact sequence specific radiomics prognostic model performance in primary and recurrent high-grade glioma. *Cancers*, 15, 2023.

[177] E. Biondetti, A. Karsa, D.L. Thomas, and K. Shmueli. Investigating the accuracy and precision of TE-dependent versus multi-echo QSM Laplacian-based methods at 3T. *Magnetic Resonance in Medicine*, 84(6):3040–3053, 2020.

[178] Z. Yan, H. Liu, X. Chen, Q. Zheng, Y. Zheng, S. Ding, Y. Peng, and Y. Li. Quantitative susceptibility mapping-derived radiomic features in discriminating Multiple Sclerosis from Neuromyelitis Optica Spectrum Disorder. *Frontiers in Neuroscience*, 15, 2021.

[179] B.G. Dutra, A.J. de Rocha, R.H. Nunes, and M.A.C.M. Junior. Neuromyelitis Optica Spectrum Disorders: Spectrum of MR Imaging Findings and Their Differential Diagnosis. *Radiographics*, 38(1):169–193, 2018.

[180] T. Sinnecker, S. Schumacher, K. Mueller, F. Pache, P. Dusek, L. Harms, K. Ruprecht, P. Nytrova, S. Chawla, T. Niendorf, I. Kister, F. Paul, Y. Ge, and J. Wuerfel. MRI phase changes in Multiple Sclerosis vs Neuromyelitis Optica lesions at 7T. *Neurology Neuroimmunology & Neuroinflammation*, 3(4), 2016.

[181] L.J. Jensen, D. Kim, T. Elgeti, I.G. Steffen, B. Hamm, and S.N. Nagel. Stability of radiomic features across different region of interest sizes— a CT and MR phantom study. *Tomography*, 7:238–252, 2021.

[182] A. Traverso, M. Kazmierski, I. Zhovannik, M. Welch, L. Wee, D. Jaffray, A. Dekker, and A. Hope. Machine learning helps identifying volume-confounding effects in radiomics. *Physica Medica*, 71:24–30, 2020.

[183] A. Zwanenbrug, S. Leger, M. Vallieres, and S. Lock. Image biomarker standardisation initiative. *ArXiv*, (1), 2016.

[184] S.I. Park, D. Kim, and S.C. Jung. Feasibility and intra-and interobserver reproducibility of quantitative susceptibility mapping with radiomic features for intracranial dissecting intramural hematomas and atherosclerotic calcifications. *Scientific Reports*, 13, 2023.

[185] S. Roy, T.D. Whitehead, J.D. Quirk, A. Salter, F.O. Ademuyiwa, S. Li, H. An, and K.I. Shoghi. Optimal co-clinical radiomics: sensitivity of radiomic features to tumour volume, image noise and resolution in co-clinical T_1-weighted and T_2-weighted magnetic resonance imaging. *EBioMedicine*, 59, 2020.

[186] N. Naji, M.L. Lauzon, P. Seres, E. Stolz, R. Frayne, C. Label, C. Beaulieu, and A.H. Wilman. Multisite reproducibility of quantitative susceptibility mapping and effective transverse relaxation rate in deep gray matter at 3 T using locally optimized sequences in 24 traveling heads. *NMR in Biomedicine*, 2022.

[187] C. Rua, W.T. Clarke, I.D. Driver, O.Mougin, A.T. Morgan, S. Clare, S. Francis, K.W. Muir, R.G. Wise, T.A. Carpenter, G.B. Williams, J.B. Rowe, R. Bowtell, and C.T. Rodgers. Multi-centre, multi-vendor reproducibility of 7T QSM and R_2* in the human brain: results from the UK7T study. *NeuroImage*, 223, 2020.

[188] M. Liu, S. Liu, K. Ghassaban, W. Zheng, D. Dicicco, Y. Miao, C. Habib, T. Jazmati, and E.M. Haacke. Assessing global and regional iron content in deep gray matter as a function of age using susceptibility mapping. *Journal of Magnetic Resonance Imaging*, 44(1):59–71, 2016.

[189] M.P.A. Starmans MPA, S.R. van der Voort, J.M.C. Tovar, J.F. Veenland, S. Klein, and W.J. Niessen. *Chapter 18 - Radiomics: data mining using quantitative medical image features.* 2020.

[190] Y. Sui, O. Afacan, A. Gholipour, and S.K. Warfield. MRI super-resolution through generative degradation learning. *Medical Image Computing and Computer Assisted Intervention*, 12906:430–440, 2021.

[191] R. Timofte, E. Agustsson, L.V. Gool, M.H. Yang, L. Zhang, B. Lim, S. Son, H. Kim, S. Nah, K.M. Lee, X. Wang, Y. Tian, K. Yu, Y. Zhang, S. Wu, C. Dong, L. Lin, Y. Qiao, C.C. Loy, W. Bae, J. Yoo, Y. Han, J.C. Ye, J. Choi, M. Kim, Y. Fan, J. Yu, W. Han, , D. Liu, H. Yu, Z. Wang, H. Shi, X. Wang, T.S. Huang, Y. Chen, K. Zhang, W. Zuo, Z. Tang, L. Luo, S. Li, M. Fu, L. Cao, W. Heng, G. Bui, T. Le, Y. Duan, D. Tao, R. Wang, X. Lin, J. Pang, J. Xu, Y. Zhao, X. Xu, J. Pan, D. Sun, Y. Zhang, X. Song, Y. Dai, X. Qin, X. Huynh, T. Guo, , H.S. Mousavi, T.H. Vu, V. Monga, C. Cruz, K. Egiazarian, V. Katkovnik, R. Mehta, A.K. Jain, A. Agarwalla, C.V.S. Praveen, R. Zhou, H. Wen, C. Zhu, Z. Xia, Z. Wang, and Q. Guo. 2017 NTIRE 2017 challenge on single image super-resolution: methods and results. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 17(2):1110–1121, 2017.

[192] W. Yang, X. Zhang, Y. Tian, W. Wang, and J.H. Xue. Deep learning for single image super-resolution: a brief review. *IEEE Transactions on Multimedia*, 21(12):3106–3121, 2019.

[193] R.M. Umer, G.L. Foresti, and C. Micheloni. Deep generative adversarial residual convolutional networks for real-world super-resolution. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1769–1777, 2020.

[194] C. Dong, C.C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. *ECCV Lecture Notes in Computer Science Springer Cham 8692*, pages 184–199, 2014.

[195] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[196] C. Dong and K. Inoue. Super-resolution reconstruction based on two-stage residual neural network. *Machine Learning Applications*, 6:100162, 2016.

[197] J. Kim, J.K. Lee, and K.M. Lee. Accurate image super-resolution using very deep convolutional networks. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1646–1654, 2016.

[198] H. Zhang, Y. Shinomiya, and S. Yoshida. 3D MRI reconstruction based on 2D generative adversarial network super-resolution. *Sensors*, 21(9):2978, 2021.

[199] K. Zeng, H. Zheng, C. Cai, Y. Yang, K. Zhang, and Z. Chen. Simultaneous single- and multi-contrast super-resolution for brain MRI images based on a convolutional neural network. *Computers in Biology and Medicine*, 99:133–141, 2018.

[200] R. Gupta, A. Sharma, and A. Kumar. Super-resolution using GANs for medical imaging. *Procedia Computer Science*, 173:28–35, 2020.

[201] M.L. de Leeuw den Bouter, G. Ippolito, T.P.A. O'Reilly, R.F. Remis, M.B. van Gijzen, and A.G. Webb. Deep learning-based single image super-resolution for low-field MR brain images. *Scientific Reports*, 12(1):6362, 2022.

[202] P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, and A. Haworth. A review of medical image data augmentation techniques for deep learning applications. *Journal of Medical Imaging and Radiation Oncology*, 65(5):5454–563, 2021.

[203] C. Fiscone, N. Curti, M. Ceccarelli, D.N. Manners, G. Castellani, C. Tonon, D. Remondini, R. Lodi, and C. Testa. Enhanced-Deep-Super-Resolution neural network on multiple MR brain images. *2023 ISMRM & ISMRT Annual Meeting & Exhibition (Selected Abstract)*, 2023.

[204] C. Fiscone, N. Curti, M. Ceccarelli, D.N. Manners, G. Castellani, C. Tonon, D. Remondini, R. Lodi, and C. Testa. Enhanced-Deep-Super-Resolution neural network on QSM brain images. *QMR Lucca 2022 Joint Workshop on MR phase, magnetic susceptibility and electrical properties mapping (Selected Abstract)*, 2022.

[205] E. Agustsson and R. Timofte. NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1122–1131, 2017.

[206] B. Lim, S. Son H. Kim, S. Nah, and K.M. Lee. Enhanced deep residual networks for single image super-resolution. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1132–1140, 2017.

[207] S.M.A. Bashir, Y. Wang, M. Khan, and Y. Niu. A comprehensive review of deep learning-based single image super-resolution. *PeerJ Computer Science*, 2022.

[208] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, T. Johannes, Z. Wang, and W.S. Twitter. Photo-realistic single image super resolution using a generative adversarial network. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 105–144, 2017.

[209] M.A. Shafto, L.K. Tyler, M. Dixon, J.R. Taylor, J.B. Rowe, R. Cusack, A.J. Calder, W.D. Marslen-Wilson, J. Duncan, T. Dalgleish, R.N. Henson, C. Brayne, F.E. Matthews, and Cam-CAN. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurology*, 14:204, 2014.

[210] J.R., Taylor N. Williams, R. Cusack, T. Auer, M.A. Shafto, M. Dixon, L.K. Tyler, Cam-Can, and R.N. Henson. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *NeuroImage*, 144(B):262–269, 2017.

[211] E. Kellner, B. Dhital, V.G. Kiselev, and M. Reisert. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*, 76(5):1574–1581, 2016.

[212] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 120:122–125, 2000.

[213] C. R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, and D. Cournapeau et al. Array programming with NumPy. *Nature*, 585(7825):357–362, 2020.

[214] D.R. Bull and F. Zhang. *Chapter 10 - Measuring and managing picture quality*, pages 335–384. 2^{nd} edition, 2021.

[215] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[216] S. Van der Walt, J.L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J.D. Warner, N. Yager, E. Gouillart, T. Yu, and scikit-image contributors. 2014 scikit-image: image processing in Python. *PeerJ*, page e453, 2014.

[217] P. Virtanen, R. Gommers, T.E Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, I. Polat, Y. Feng, E.W. Moore, J. VanderPlas, and SciPy 1.0 contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020.

[218] B. Huang, H. Xiao, W. Liu, Y. Zhang, H. Wu, W. Wang, Y. Yang, Y. Yang, G.W. Miller, T. Li, and J. Cai. MRI super-resolution via realistic downsampling with adversarial learning. *Physics in Medicine and Biology*, 66(20):205004, 2021.