Multimodal mapping of brain structural properties by combining T1- and T2-weighted MR imaging data

Author(s):
Ganzetti, Marco

Publication Date:
2016

Permanent Link:
https://doi.org/10.3929/ethz-a-010648156

Rights / License:
In Copyright - Non-Commercial Use Permitted
Multimodal mapping of brain structural properties by combining T1- and T2-weighted MR imaging data

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH (Dr. sc. ETH Zurich)

presented by

Marco Ganzetti

M.Eng. in Biomedical Engineering - Marche Polytechnic University

born on 10.12.1986
citizen of ITALY

accepted on the recommendation of

Prof. Dr. Spyros Kollias
Prof. Dr. Dante Mantini
Prof. Dr. Nicole Wenderoth

2016
Acknowledgements

First of all, I am extremely grateful to my direct supervisor, Dr. Dante Mantini for choosing me as PhD student and for always being there for me. Dante is an excellent professional and advisor. He taught me more than I could ever imagine and gave me the confidence to explore my research interests. I am thankful to Dante for his encouragement, support and continuous optimism and in particular for being a splendid person with whom working was great.

I wish to express my most sincere gratitude to Prof. Nicole Wenderoth, principal supervisor of this thesis, for giving me the opportunity to work in her lab. I have been very fortunate to have Nici as my supervisor with her brilliant ideas, advices and support.

I would like to take this opportunity to thank Prof. Glyn Humphreys, Head of Department of Experimental Psychology (EP) at Oxford University, for giving me the opportunity to spend a significant part of this PhD in his department. Life would not have been so nice without the good friends I met at EP. I would like to warmly thank the whole laboratory especially Quanying, Shri, Rezvan, Christian, Charu, Rachel, Laura, Renata, Elisea, Alessia, Gaetan and San. Special thanks go to all the friends outside the department that accompanied me during these two unforgettable years in Oxford, especially Andrea and Karol.

I would like to extend my sincerest thanks and appreciation to all my colleagues in Zurich with whom I shared nice moments.

Finally, I would like to thank my family that through my studies has always supported and encouraged me to follow my interests and to do what makes me happy. My deepest thanks go to my girlfriend, Elizabeth, for her patience during my writing of this dissertation, and for always supporting me during these PhD years when we were far away from each other and for the wonderful year that we shared together in Switzerland.
# Index

## 1 Introduction

1.1 **Fundamental concepts** 12
   1.1.1 Nuclear magnetic resonance 12
   1.1.2 Atomic structure 12
   1.1.3 Angular momentum or spin 13
   1.1.4 Magnetic moment 14
   1.1.5 Magnetic resonance 16
   1.1.6 Spin excitation 16
   1.1.7 MR signal 17
   1.1.8 Relaxation times 18
     1.1.8.1 T1 relaxation 18
     1.1.8.2 T2 relaxation 20

1.2 **Image formation** 22
   1.2.1 Introduction to pulse sequences 22
     1.2.1.1 Spin echo (SE) pulse sequence 23
     1.2.1.2 Gradient echo (GRE) pulse sequence 24
   1.2.2 Basic concepts of image formation 25
     1.2.2.1 K space and image formation 27

1.3 **Image contrast** 29
   1.3.1 T1- and T2-weighted contrasts 29
   1.3.2 Magnetization Prepared Rapid Gradient Echo sequence 32
   1.3.3 Fast Spin Echo sequence 33

1.4 **Image properties** 35
   1.4.1 Intensity non-uniformity in MR data 35
   1.4.2 Intensity inconsistency 37

1.5 **Multimodal MR imaging** 40

1.6 **Thesis overview** 41

## 2 Quantitative evaluation of intensity inhomogeneity correction methods for structural MR images

2.1 **Introduction** 44

2.2 **Methods** 46
   2.2.1 INU correction: theory and algorithms 46
     2.2.1.1 Modelling of MR intensity inhomogeneities 46
     2.2.1.2 Methods under investigation 46
   2.2.2 Analysis of INU correction methods 48
     2.2.2.1 Simulated data 48
     2.2.2.2 Method settings 52
     2.2.2.3 Performance assessment 53

2.3 **Results** 55
3 Intensity inhomogeneity correction of structural MR images: a data-driven approach to define input algorithm parameters

3.1 Introduction 71
3.2 Methods 73
  3.2.1 Description of the data-driven approach 73
    3.2.1.1 Definition of the INU correction parameter space 73
    3.2.1.2 Selection of the indirect metric 74
    3.2.1.3 Definition of the optimal smoothing level 74
    3.2.1.4 Definition of image-specific GM and WM masks 75
    3.2.1.5 Identification of the optimal set of INU correction parameters 75
  3.2.2 Performance analysis 76
    3.2.2.1 Testing on simulated MR data 76
    3.2.2.2 Testing on actual MR images 78
3.3 Results 79
3.4 Discussion 85
3.5 Conclusion 88

4 Whole brain myelin mapping using T1- and T2-weighted MR imaging data

4.1 Introduction 91
4.2 Methods 93
  4.2.1 Description of the method 93
    4.2.1.1 Theoretical background 93
    4.2.1.2 Method implementation 94
  4.2.2 Method validation 97
    4.2.2.1 Subjects and data acquisition 97
    4.2.2.2 Similarity of image histograms across subjects 98
    4.2.2.3 Comparison of T1-w/T2-w with MTR, FA and FLAIR images 99
4.3 Results 101
4.4 Discussion 108
  4.4.1 Methodological considerations on the T1-w/T2-w approach 108
  4.4.2 Myelin-related information in T1-w/T2-w images 109
  4.4.3 Potential limitation of the method 110
4.5 Conclusion 111
Mapping pathological changes in brain structure by combining T1- and T2-weighted MR imaging data

5.1 Introduction
5.2 Methods
   5.2.1 Structural MR data
   5.2.2 Imaging analysis
   5.2.3 Statistical analysis
5.3 Results
5.4 Discussion
5.2 Conclusion

6 Main discussion

6.1 MR inhomogeneity correction
6.2 MR intensity standardization
6.3 T1-w/T2-w approach

Bibliography
Curriculum Vitae
In 1973, Paul C. Lauterbur, a chemist at the State University of New York, published the first nuclear magnetic resonance (NMR) image. Though, the image was a mere representation of two tubes of water, that moment laid the foundations of one of the most important medical advances of the 20th century. Only four years later, in 1977, Raymond Damadian, a medical doctor and research scientist, performed a full body scan of a human being to diagnose cancer. He discovered that cancerous and healthy tissue could be distinguished in-vivo because of their intrinsic tendency of emitting different response signals. Damadian also invented and patented an apparatus to use NMR safely and reliably to scan the human body, a groundbreaking innovation now well known as magnetic resonance imaging (MRI). From that moment on, the medical use of MRI has advanced rapidly. The first MRI scanners started to appear in the hospitals at the beginning of the 1980s.

Nowadays, non-invasive analyses of altered brain structures in neurological and psychiatric patients largely rely on MRI. In particular, images from conventional T1-weighted (T1-w) and T2-weighted (T2-w) MRI, being characterized by high contrast and spatial resolution, have been widely used to investigate structural properties of the brain. Since being introduced into clinics several decades ago, the interpretation of T1-w and T2-w images is presently a routine task for the majority of clinicians. Importantly, the two images differ with respect to their intrinsic sensitivities towards biophysical properties of brain tissues. This implies that, in the case of brain pathologies, structural alterations resulting from edema, inflammation, tumor infiltration, iron accumulation, or atrophy, will manifest themselves differently in the two kinds of images.

However, T1-w and T2-w images also have some common limitations, which primarily relate to: 1) the presence of spatial inhomogeneities in image intensity, produced by interactions between the subject’s body and the MR scanner, and 2) the lack of image intensity standardization, mainly related to the features of the MR instrumentation and acquisition sequence used. These factors limit the use of conventional structural MR images for the study of pathological changes in brain structure. Therefore, by combing an intensity inhomogeneity correction procedure with a retrospective calibration algorithm, we aim to address these issues. This will permit standardizing the intensity histogram of both T1-w and T2-w images, thereby allowing for across-subject statistical analyses.

Clinicians most often tend to infer brain tissue integrity through visual comparison of T1-w and T2-w images, when both are available. Also, a number of clinical studies have combined the information
from the two kinds of images to study, for instance, multiple sclerosis or mesial temporal lobe epilepsy. The parallel analysis of T1-w and T2-w images represents a first attempt to implement a multi-modal structural imaging approach for the study of brain tissue abnormalities. However, direct fusion of the two kinds of images, so as to produce a multi-modal T1-w/T2-w image could be an interesting alternative. In particular, T1-w sequences are characterized by better contrast-to-noise ratio in white matter regions, whereas T2-w sequences can be used to better discriminate structural differences in fluid-filled regions. In this study, we develop a workflow for the combined analysis of T1-w and T2-w scans from the same individual, based on their ratio of the two images. After testing the method on simulated and real MR data from healthy subjects, we also apply the T1-w/T2-w analysis to structural MRI data collected in individuals with schizophrenia and age-matched healthy controls. Accordingly, we compare the T1-w/T2-w images between groups to verify whether the technique can detect disease-related structural impairment. Since T1-w and T2-w images are most often collected in clinical scanning protocols and do not require long acquisition times, the T1-w/T2-w technique may become a widely-used tool for mapping potential pathological changes resulting from brain disease.

Oggi, la maggior parte delle analisi non invasive della struttura cerebrale in neurologia e psichiatria sono effettuate mediante RM. In particolare, le immagini T1-pesato (T1-w) e T2-pesato (T2-w), essendo caratterizzate da un elevato contrasto e risoluzione spaziale, sono state ampiamente adottate nello studio delle proprietà strutturali del cervello. Dall’introduzione in clinica diversi decenni fa l’interpretazione delle immagini T1-w e T2-w è un compito di routine per la maggior parte dei medici. È importante notare che le due immagini differiscono per quanto riguarda la loro sensibilità verso le proprietà biofisiche dei diversi tessuti cerebrali. Ciò implica che, in caso di patologie cerebrali, alterazioni strutturali derivanti da edema, infiammazione, infiltrazione tumorale, accumulo di ferro, o atrofia, si manifestano diversamente nei due tipi d’immagini.

Tuttavia, le immagini T1-w e T2-w sono soggette ad alcune limitazioni, in particolare: 1) la presenza di disomogeneità spaziali d’intensità, causate principalmente da interazioni tra soggetto e scanner, e 2) la mancanza di una standardizzazione dell’intensità, principalmente legata alle caratteristiche della strumentazione e della sequenza di acquisizione utilizzata. Questi fattori limitano l’utilizzo delle tradizionali tecniche di RM per lo studio di alterazioni patologiche nella struttura del cervello. Pertanto, combinando una procedura di correzione delle disomogeneità intensità con un algoritmo di calibrazione retrospettiva, ci proponiamo di risolvere queste problematiche. Questo permetterà di ottenere una distribuzione standardizzata delle intensità per entrambe le immagini T1-w e T2-w, consentendo in tal modo di eseguire analisi statistiche tra diversi soggetti.

Molto spesso, i medici tendono a fare una diagnosi circa l’integrità del tessuto cerebrale attraverso
il confronto visivo delle immagini T1-w e T2-w, quando entrambe sono disponibili. Inoltre, diversi studi clinici hanno provato a combinare l’informazione prodotta dalle due immagini per lo studio di patologie come la sclerosi multipla e l’epilessia del lobo temporale mediale. L’analisi parallela delle immagini T1-w e T2-w rappresenta un primo tentativo per attuare una strategia di imaging strutturale multimodale per lo studio delle anomalie del tessuto cerebrale. Tuttavia, la fusione diretta delle due immagini, in modo da produrre un’immagine multimodale T1-w/T2-w potrebbe essere un’interessante possibilità. In particolare, le sequenze T1-w sono caratterizzate da un migliore rapporto di contrasto-rumore nelle regioni della sostanza bianca, mentre sequenze T2-w possono essere utilizzati per discriminare meglio le differenze strutturali nelle regioni contraddistinte dalla presenza di liquidi. In questo studio, ci proponiamo di sviluppare un flusso di lavoro per l’analisi combinata d’immagini T1-w e T2-w dallo stesso individuo, sulla base del diretto rapporto delle due immagini. Dopo aver testato il metodo su dati simulati e in seguito validato su soggetti sani, andremo ad applicare l’analisi T1-w/T2-w su immagini strutturali raccolte in individui con schizofrenia e le confronteremo con controlli sani di pari età. Di conseguenza, metteremo a confronto le immagini T1-w/T2-w tra gruppi per verificare se la tecnica sia in grado di rilevare alterazioni strutturali dovute alla malattia. Poiché le immagini T1-w e T2-w sono normalmente acquisite in quasi tutti i protocolli clinici e non richiedono lunghi tempi di scansione, la tecnica T1-w/T2-w può diventare un potenziale strumento per la mappatura di alterazioni strutturali derivanti da patologie cerebrali.
1.1 Fundamental concepts

In this section general background information about several technical aspects of MRI physics will be provided, with a focus on the image formation and interpretation processes.

1.1.1 Nuclear magnetic resonance

Magnetic resonance (MR) imaging is a technique that uses a strong magnetic field and radio waves to create detailed images of organs and tissues. It is widely used in hospitals for medical diagnosis, staging of disease and follow-ups. It is based on a physics phenomenon discovered back in the 1940s by two researchers Edward Purcell and Felix Bloch, called nuclear magnetic resonance (NMR). Despite the word ‘nuclear’, the technique does not involve any form of radioactivity or ionizing radiations (x-rays) thus it is a safe and harmless diagnostic tool. Through a concurrent interaction between magnetic fields and radio waves, atomic nuclei in tissues are indirectly forced to give off minuscule radio signals. The signal then acquired by a receiver is recorded and analyzed by computerized systems to generate astounding anatomical images.

1.1.2 Atomic structure

Atoms are considered the tiniest neutral particles into which matter can be divided by chemical reactions. Every atom is composed of a central nucleus and one or more orbiting electrons. Each type of atom corresponds to a specific chemical element. The nucleus is composed by subatomic particles called nucleons, which are divided into protons and neutrons. The number of protons and neutrons defines two important properties of an atom: the atomic number and the mass number. The number of protons in a nucleus is called atomic number, which identifies a specific element. The mass number is the sum of both protons and neutrons in the nucleus. Usually, the number of protons and the number of neutrons is the same, so that the mass number is even. However, an atom of a certain element may have several instances varying in number of neutrons. These variants of a specific chemical element are called isotopes. In this regard, only nuclei characterized by an odd mass number are essential in MR imaging. Electrons are
subatomic particles spinning around the nucleus. They are smaller than the other two elements of the atom with a mass of almost one-thousandth the mass of a proton. They are located around the nucleus in a cloud. The electron cloud model states that it is not possible to know exactly the location of an electron at any given time, but only where the electrons are more likely to be in certain areas. This concept is also known as Heisenberg's uncertainty principle. In general, the number of electrons is the same as the number of protons. Furthermore, protons and electron are characterized by a positive and negative electric charge, respectively. On the other hand, neutrons are neutral particles. Therefore, the overall charge of the nucleus is positive.

### 1.1.3 Angular momentum or spin

In terms of physical properties, an important characteristic of atomic nuclei is that they posses an intrinsic angular momentum, also known as spin. This physical quantity may be easily perceived considering a classic spinning top evolution. As shown in Figure 1.1, a spinning top is characterized by the rotational speed or angular velocity ($\omega$) and the axis about which the object is rotating.

![Figure 1.1 – Angular momentum (L) and angular velocity ($\omega$) of a spinning top.](image)

Though, in physics, a common measure to describe the quantity of rotation of a spinning object is the angular momentum, which accounts for its mass, rotational speed and shape. It is generally expressed as a vector (L) pointing along the axis of rotation of the spinning top and with a magnitude directly proportional to the angular velocity of the object (Figure 1.1). The nuclear spin originates from the individual spin of protons and neutrons within a nucleus. Because nucleons have spin, as electrons do, their spin can pair up in nuclei that have an even mass number generating a null net spin. On the other
hand, in a nucleus with odd mass number, we assist to an imbalance between spin directions leading to a non-zero net spin or angular momentum. Only isotopes with a non-zero nuclear spin are of interest in MRI (Table 1.1). However, NMR experiments are generally performed on isotopes whose natural abundance is high enough to generate measurable signals.

<table>
<thead>
<tr>
<th>Chemical element</th>
<th>Isotope</th>
<th>Mass number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrogen</td>
<td>$^{1}\text{H}$</td>
<td>1</td>
</tr>
<tr>
<td>Carbon</td>
<td>$^{13}\text{C}$</td>
<td>13</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>$^{15}\text{N}$</td>
<td>15</td>
</tr>
<tr>
<td>Oxygen</td>
<td>$^{17}\text{O}$</td>
<td>17</td>
</tr>
<tr>
<td>Fluorine</td>
<td>$^{19}\text{F}$</td>
<td>19</td>
</tr>
<tr>
<td>Sodium</td>
<td>$^{23}\text{Na}$</td>
<td>23</td>
</tr>
<tr>
<td>Phosphorus</td>
<td>$^{31}\text{P}$</td>
<td>31</td>
</tr>
</tbody>
</table>

### 1.1.4 Magnetic moment

As stated in the Faraday's law of induction, the interaction between the positive charge of atomic nuclei and the nuclear net spin generates a magnetization, thus a nuclear magnetic moment. This is the quantity at the very basis of the NMR phenomenon. In the specific case the net angular moment is zero, the nuclear magnetic moment is zero. This happens if the number of protons and neutrons is the same. In other cases, with odd numbers, the nucleus has spin, consequently, a magnetic polarization. In terms of our organism, atoms are the building blocks of our body, with the isotope hydrogen-1 ($^{1}\text{H}$) being the most abundant in atomic percentage than any other element. In addition, due to its sole proton, $^{1}\text{H}$ possess a nuclear magnetic moment, hence it is the primary candidate in clinical MR imaging. Its magnetic moment can be denoted as a vector with a given length and direction, representing the magnitude of the magnetic moment and its direction, respectively.

In the case of a sample in a state of thermal agitation (a person inside a MR scanner at room temperature), the magnetic moments of the hydrogen nuclei are randomly oriented, thus the net magnetic moment is null (Figure 1.2A). On the other hand, when an external static magnetic field is applied (i.e. the static field $B_0$ of the MR scanner), spins immediately align with the field, creating a net magnetic moment (Figure 1.2B). Nevertheless, because of the subatomic nature characterizing their dimensions, protons and neutrons must follow quantum physics laws. This is to say, energy exists in discrete units rather than as a constant electromagnetic radiation. The individual quantities are called quanta. As for hydrogen nuclei, they can only acquire two distinctive energy levels: low and high. As for hydrogen nuclei placed in a magnetic field, the lowest level corresponds to the nuclear magnetic moment parallel to the external magnetic field, whereas the highest may be represented with the nuclear magnetic moment pointing in the
opposite direction of the field (anti-parallel). When the magnetic field is switched on, the nuclear magnetic moments of each single nucleus slowly align with the direction of the magnetic field (Figure 1.2B). Approximately, there is the same number of proton nuclei aligned with the static magnetic field $B_0$ as counter aligned. However, the aligned position is slightly favored, as the nucleus is at a lower energy in this position. This results in a net magnetization vector (NMV) pointing in the direction of the static magnetic field $B_0$ (Figure 1.2B). It is worth mentioning that the energy difference between the high- and low-energy populations is directly proportional to the strength of the static field $B_0$. Therefore, at higher static fields only fewer nuclei posses enough energy to join the high energy level. In terms of net magnetization, this results in greater NMV, thus superior MRI signal.

**Figure 1.2** – Net magnetization vector (NMV).
1.1.5 Magnetic resonance

Resonance, in physics, is a phenomenon that occurs when a system is exposed to an oscillating perturbation with a frequency matching a system preferential frequency. When this occurs, the system begins to oscillate with greater amplitude at this specific frequency, also known as resonant frequency. In terms of magnetic resonance, the concept may be easily understood thinking about a compass. If a compass is placed inside a magnetic field, its needle oscillates about the field direction before stopping. The frequency at which this oscillation occurs is defined resonance frequency. As the compass needle aligns with an external magnetic field, so does the hydrogen nucleus. However, contrary to the needle, the latter also possess an intrinsic angular momentum or spin, which, in combination with an intrinsic electric charge, generates a magnetic moment. When a magnetic moment is placed in a static magnetic field $B_0$, it experiences a torque or a tendency of a force to rotate an object about its axis. This quantity can be exemplified as the entity forcing the compass needle to align along the magnetic field. Considering the hydrogen nucleus, the torque exerted tends to force the nucleus spin to align along the magnetic field. In this particular case, the combined effect of the angular momentum and torque causes the nucleus spin to precess around the direction of the static magnetic field. Specifically, the magnetic moment follows a circular path around the direction of $B_0$. This phenomenon is known as Larmor precession and the precession angular velocity is named the Larmor frequency. The relation between the Larmor frequency and the static magnetic field is given by the Larmor equation as follows:

$$\nu = \gamma B_0$$

where $\nu$ is the frequency in MHz, $\gamma$ is the gyromagnetic ratio in MHz/Tesla for a specific nucleus ($^1\text{H}$ in our case) and $B_0$ is the magnetic field strength in Tesla. The gyromagnetic ratio of a given particle is the ratio of its magnetic moment to its angular momentum. It is a constant of proportionality and it can be exemplified as the precessional frequency of a particle under the effect of a 1 T static magnetic field. In the particular case of the $^1\text{H}$ nucleus, the gyromagnetic ratio is 42.57 MHz/T. This is to say that $^1\text{H}$ will resonate at different frequency when placed inside different magnetic fields. For example, in the typical case of MR scanner fields, the resonance frequency is 63.87, 127.7, 298.04 MHz, respectively at 1.5, 3 and 7 T. An important concept to keep in mind is the fact that different nuclei are characterized by different gyromagnetic ratios, thus different Larmor frequencies. This specificity allows us to image exclusively specific nuclei.

1.1.6 Spin excitation

At nuclear levels, an MR imaging experiment involves reorienting NMV relative to the static field $B_0$ and observing the behavior of NMV. By placing a conductive coil around the sample, it is possible to measure any change induced to NMV. When the net magnetization vector lays in the same direction of the static field and thus it is at equilibrium, no signal is detected by the surrounding coil. According to the
electromagnetic induction principle, the NMV must vary its orientation in order to induce any detectable signal in the conductive coil. Therefore, it is necessary to perturb NMV generating a component of NMV in the transverse plane, which is perpendicular to the principal direction of static field $B_0$. Hereafter, we will refer with $z$ to the direction of $B_0$, whereas the transverse plan will be described by $x$ and $y$. The deflection of NMV from its equilibrium state is done applying an electromagnetic pulse ($B_1$ field) to the sample. Exposure of individual nuclei to this radiofrequency (RF) radiation at the Larmor frequency causes nuclei in the lower energy state to jump into the higher energy state (anti-parallel state). The resonance phenomenon occurs. This results because some of the low-energy nuclei are given enough energy to join the high-energy level. It is worth noting that if energy is delivered at a different frequency to that of the Larmor precession frequency of the $^1$H nuclei, resonance does not appear. When $B_1$ is switched on, NMV gets tilted out of alignment away from the longitudinal axis $z$. The angle that NMV rotates is commonly called flip angle. The amount of this deflection thus the flip angle value is proportional to the duration and amplitude of the RF pulse. This phenomenon is called excitation. The other effect of the magnetic resonance on the hydrogen nuclei is the phase coherence phenomenon. When the RF pulse is released, all the magnetic moments move to the same place on the precessional pathway and begin precessing in unison (in phase) with each other. This results in the NMV transverse component rotating at the Larmor frequency on the transverse plane. On the other hand, magnetic moments that are out of phase are not in the same position at any given time. This condition causes all the nuclear magnetic moments to cancel out with each other, hence a null NMV component on the transverse plane.

1.1.7 MR signal

After $B_1$ is switched off, the NMV begins to precess at a rate equal to the Larmor frequency. In terms of quantum theory, the hydrogen nuclei start loosing the energy previously received from the RF pulse. This phenomenon is known as relaxation. As previously discussed, whenever a conductive coil is placed in a varying magnetic field, a current is induced in the coil. The position of the coil is a very important aspect to ensure a good signal quality. In particular, it must be located in the transverse plane perpendicular to the static field $B_0$ (Figure 1.3). In this way the rotating NMV transverse component is able to generate magnetic field variations inside the coil. The resulting current, thus the voltage, constitute the actual MR signal. As relaxation occurs, the NMV experiences two distinct and independent phenomena at the same time. The magnitude of the NMV gradually increases in the longitudinal plane parallel to the $B_0$ field and the amount of magnetization in the transverse plane gradually decreases. They are known as recovery and decay, respectively. Eventually, the NMV realigns with the static field $B_0$. Once again, because of the electromagnetic induction, only when a component of NMV exists in the transverse plane will a signal be revealed. Therefore, the largest transverse component can be detected following a 90° radiofrequency pulse. On the other hand, if the pulse duration is doubled, thus generating a 180° pulse, no component of NMV can be seen in the transverse plane.
1.1.8 Relaxation times

One of the peculiarities that make MR imaging stand out from the rest of medical imaging modalities, is the broad range of image contrast attainable. By defining specific input parameters in the MR console, it is possible to highlight certain types of tissues and structures. As we learnt before, once the net magnetization is perturbed from its equilibrium state, a particular event known as relaxation initiates. This process of restoring the longitudinal magnetization is generally described using two defined relaxation phenomena. The complete recovery of the longitudinal magnetization (oriented along the $z$ axis) of the NMV is referred to as $T_1$ relaxation or $T_1$ recovery, whereas the decay process of the transverse magnetization ($x$-$y$ component) is known as $T_2$ relaxation or $T_2$ decay.

1.1.8.1 $T_1$ relaxation

Once the excitation pulse $B_1$ is switched off, a proton interaction with surrounding tissues, the lattice, begins. Hydrogen nuclei start transferring energy with neighboring molecules in order to return to the previous thermal equilibrium with the lattice. In other words, this process allows the spin population to equilibrate itself reestablishing the former proportion of parallel and antiparallel spins. Stated more precisely, the random motion of surrounding molecules can be translated as random fluctuations of small
magnetic fields. In general, a large spectrum of molecular motion over a broad frequency band is experienced. Though, it is the magnitude of the fluctuation component at the Larmor frequency the factor enabling an efficient transfer of energy. The process describing the recovery of \( \text{NMV}_z \) to \( \text{NMV}_0 \) is controlled by an exponential function (Figure 1.4) characterized by a time constant known as \( \text{T1 relaxation time} \) (time for the magnetization to recover the 63% of its equilibrium value). In addition, due to the fact that T1 relaxation involves an interaction with the surroundings this phenomenon is also known as spin-lattice relaxation. Different lattice features explain how different tissues can experience distinctive T1 relaxation times. For instance, the cerebrospinal fluid (CSF) is mainly composed by free water, which is characterized by a molecular motion rate greater than the specific Larmor frequency at common magnetic fields. This results the CSF to have a slow T1 relaxation. On the other hand, a restricted molecular motion is experienced in the white matter (WM) and gray matter (GM), with the former showing a more prominent component at the Larmor frequency. This causes the WM to have a faster relaxation rate than the GM. As to the Larmor equation, we have previously seen how an increase in the static magnetic field generates a proportional increase in the Larmor frequency magnitude. In terms of T1 relaxation, this can be translated into a concomitant expansion of relaxation times within tissues. Indeed, a marked increase in the Larmor frequency generates a less efficient exchange of magnetization.

![Figure 1.4 – T1 recovery.](image)

From an analytical point of view, this phenomenon is generally represented using a set of equations known as Bloch equations. They are used to calculate the net magnetization vector...
NMV = (NMV_X, NMV_Y, NMV_Z) as a function of time and relaxation times. The component describing the T1 relaxation process is defined as following:

$$\frac{dNMV_Z(t)}{dt} = -\frac{NMV_Z(t) \cdot NMV_0}{T_1}$$  \hspace{1cm} (1.2)

with NMV_Z representing the longitudinal component of the net magnetization and NMV_0 denoting the net magnetization at equilibrium expressed in the z direction. This is a linear differential equation and its solution is:

$$NMV_Z(t) = NMV_0 - NMV_0 \cdot e^{-t/T_1}$$  \hspace{1cm} (1.3)

where NMV_Z(0) is the longitudinal net magnetization at time t=0. A simplified version may be obtained considering a null magnetization along the longitudinal component at t=0 (NMV tilted at 90° with respect to the z axis):

$$NMV_Z(t) = NMV_0 \cdot (1 - e^{-t/T_1})$$  \hspace{1cm} (1.4)

Typical values of T1 time relaxation are given in Table 1.2.

### Table 1.2 – Average values of T1 relaxation times (msec) in human brain tissues.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>1.5 T¹</th>
<th>3 T¹</th>
<th>4 T¹</th>
<th>7 T¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Matter</td>
<td>756</td>
<td>832</td>
<td>1040</td>
<td>1285</td>
</tr>
<tr>
<td>Grey Matter</td>
<td>1200</td>
<td>1331</td>
<td>1410</td>
<td>1954</td>
</tr>
</tbody>
</table>

From: a(Whittall et al., 1997), b(Wansapura et al., 1999), c(Duewell et al., 1996), d(Wyss et al., 2013).

#### 1.1.8.2 T2 relaxation

As seen previously, starting from a steady state condition (NMV aligned with the static field B₀) a 90° RF pulse will cause the NMV to tilt and lay in the transverse plane. In this precise moment following the excitation pulse, all the spinning nuclei in the sample are in phase rotating exactly at the Larmor frequency. Almost instantly, they start losing coherence as some begin to spin slightly slower/faster than the Larmor frequency. This broadening of spinning frequencies is mainly due to reciprocal interactions between neighboring nuclei. As two hydrogen nuclei come close to each other, they experience a change in the magnetic field, which results into a change in the precession frequency. This dephasing effect results in the transverse component of the magnetic vector to decrease exponentially. This phenomenon is also known as spin-spin relaxation or T2 relaxation. The decaying process is also described with an exponential function (Figure 1.5). The rate of decay is defined by a time constant known as T2 relaxation time (time for the transverse magnetization to lose 63% of its initial value).
T2 relaxation generally proceeds more rapidly than T1 recovery, though different biological tissues have different T2. For example, fluids such as CSF have longer T2 than fat based tissues such as WM. This is mainly due to the slow motion of hydrogen nuclei in the fat which rate is closer to the Larmor frequency. In fat based tissues, molecules are tightly packed together, thus spin-spin interactions are more efficient. This results in a loss of transverse magnetization as magnetic moments rapidly dephase. On the other hand, in water, the molecules are sparsely located therefore spin-spin interactions become less likely to occur. This causes a gradual loss of transverse magnetization with magnetic moments slowly dephasing. As the T1 recovery, also the T2 decay can be formalized using Bloch equations. The component describing the T2 relaxation process is defined as following:

\[
\frac{d \text{NMV}_{X,Y}(t)}{dt} = - \frac{\text{NMV}_{X,Y}(t)}{T2}
\]

with \( \text{NMV}_{X,Y} \) representing the transverse component of the net magnetization. This is a linear differential equation which solve to give:

\[
\text{NMV}_{X,Y}(t) = \text{NMV}_{X,Y}(0) e^{-t/T2}
\]

Figure 1.5 – T2 decay.
where \( \text{NMV}_{X,Y}(0) \) is the transverse magnetization at \( t=0 \). Typical values of T2 time relaxation are given in Table 1.3.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>1.5 T(^1)</th>
<th>3 T(^1)</th>
<th>4 T(^1)</th>
<th>7 T(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Matter</td>
<td>77.4</td>
<td>110</td>
<td>50</td>
<td>38</td>
</tr>
<tr>
<td>Grey Matter</td>
<td>79.9</td>
<td>80</td>
<td>50</td>
<td>49</td>
</tr>
</tbody>
</table>

From: \(^a\)Whittall et al., 1997, \(^b\)Wansapura et al., 1999, \(^c\)Duewell et al., 1996, \(^d\)Wyss et al., 2013.

### 1.2 Image formation

#### 1.2.1 Introduction to pulse sequences

As soon as the RF pulse is switched off, the NMV begins to realign with the static field \( B_0 \). In practice, what the receiving coil will record is an oscillating signal characterized by a decaying exponential envelope. As the transverse magnetization decay, so does the induced voltage in the receiving coil. The frequency of oscillation is exactly the Larmor precession frequency, whereas the amplitude of the decay signal corresponds to the amount of the transverse magnetization. This phenomenon is known as free induction decay (FID), which is nothing but the resonance signal from the precessing spins (Figure 1.3). As we learnt before, the duration of the signal (transverse magnetization) generated by the NMV relaxation is limited by the T2 relaxation phenomenon. However, in reality, when the RF pulse is switched off, the relaxation and therefore the decay of the FID signal are shortened. The reason of this phenomenon is the presence of static field \( B_0 \) inhomogeneities, which accelerate the dephasing mechanism. These inhomogeneities or distortions are the result of production tolerances and environmental magnetic field disturbances. Also the patient susceptibility creates additional inhomogeneities in the \( B_0 \) field. As we previously seen, the Larmor frequency of precessing nuclei is proportional to the \( B_0 \) field magnitude. Therefore, hydrogen nuclei placed in an inhomogeneous \( B_0 \) field experience different precessional frequencies leading to an intensification of the dephasing rate. This leads to a faster drop of the signal and shortens the transverse relaxation time. The relaxation time is known as \( T_2^* \) decay (Figure 1.3). Dephasing caused by magnetic field inhomogeneities is a crucial aspect in MRI. Even though manufactures are trying to make the static magnetic field as homogeneous as possible, a residual amount of distortions is unavoidable. Therefore, to accurately measure proper relaxation times the signal must be restored. One way to regenerate the signal is by using particular MRI pulse sequences. An MRI pulse sequence is a programmed set of changing magnetic fields and RF pulses. The two main families of pulse sequences are the spin echo (SE) and the gradient echo (GRE) pulse sequences.
1.2.1.1 Spin echo (SE) pulse sequence

A possible technique to regenerate the signal is to use an additional RF field (Figure 1.6). Initially, a 90° excitation pulse is delivered causing the NMV to flip onto the transverse plane. As soon as it switched off, an induced voltage is generated in the receiving coil producing the common FID signal (Figure 1.3). The NMV begins to dephase at rate given by the T2* decay. As a way to counteract the loss of coherence of the NMV, a 180° RF pulse is delivered. This rotates the NMV by 180° inverting the dephasing process. This is to say, all the magnetic moments are still in the transverse plane but flipped by 180°. In particular, the precession direction is preserved so that the fast moving spins start to recoup on the sluggish spins. At a precise time, all the magnetic moments realign. The transverse magnetization is then on phase, generating a maximum in the signal. This signal is known as a spin echo. However, as soon as the signal peak is obtained, the dephasing mechanism reoccurs again. In terms of timing parameters, the principals are the repetition time (TR) and the echo time (TE). In a typical MR experiment, a specific delay between consecutive excitation pulses has to be maintained in order to allow the net magnetization to restore its equilibrium state (no x-y component). This delay is generally referred to as TR. It determines how much longitudinal magnetization has been recovered between each pulse. It is expressed in milliseconds. On the other hand, the time between the initial excitation pulse and the spin echo is known as TE.
1.2.1.2 Gradient echo (GRE) pulse sequence

The two main peculiarities that differentiate GRE from SE pulse sequences are a variable flip angle and the use of magnetic gradients to rephase the lost signal due to magnetic field inhomogeneities (Figure 1.7). In terms of signal rephasing, a magnetic gradient with a negative polarity is activated. This causes the magnetic moments and therefore the FID to rapidly dephase in a controlled manner. Afterwards, its polarity is reversed causing the magnetic moments to rephase. A gradient echo is then produced. The major advantage of GRE with respect to SE sequences, it is the possibility to shorten the image acquisition. Since gradients are able to rephase faster than the 180° RF pulse (by simply inverting the polarity), the minimum TE can be reduced, hence the TR can also be shortened. As for the flip angle, a reduced value allows us to decrease the TR duration and therefore the scan time. On the other hand, the main drawback of GRE sequences is the incapacity to compensate for magnetic field inhomogeneities, which are the main cause of imaging artifacts (e.g. magnetic susceptibility artifact).
1.2.2 Basic concepts of image formation

As previously described, when a RF excitation pulse at the Larmor frequency of hydrogen is released, the magnetic resonance phenomenon occurs. Regardless the specific pulse sequence adopted, a transverse component of the NMV occurs. As soon as the coherent magnetization start precessing, the signal or the induced voltage is detected on the receiving coil located in the transverse plane. However, in order to spatially locate the source of the signal, an encoding mechanism is implemented. This is the very first step of the image formation in MRI. An effective three-dimensional spatial localization is performed using magnetic gradients (Figure 1.8). Once the scan plan is defined, a slice is selected. Afterwards, the locality of the signal on the selected slice is encoded. This step is accomplished with two mechanisms: the frequency encoding and the phase encoding (Figure 1.8). Each of these three spatial encoding steps is performed using a set of magnetic gradients. There are three gradient coils within the bore of the magnet that generate a magnetic slope along each of the three directions ($x$, $y$ and $z$). Specifically, a magnetic gradient is a linear variation in the magnetic field magnitude across one specific direction in the imaging volume. The coil used to create the gradient is a dedicated electromagnetic coil or an electric conductor in which a current is passed to generate a magnetic field. In practice, a gradient superimposed to the static magnetic field $B_0$ creates a linear variation in the magnetic strength of the second. This is to say, the left extremity of $B_0$ will experience a decrease while the right one an increase in the strength. The intersection between the three magnetic gradients (along the $x$, $y$, and $z$ direction) is called magnetic isocenter and the magnetic field magnitude in this precise point is equivalent to $B_0$ field strength. According to the Larmor equation, an increase in the static field magnitude equals a proportional increase in the precessional frequency of hydrogen nuclei. Therefore, the presence of a magnetic gradient along one of the three directions will result in a linear variation of the precessional frequency. In other words, knowing the gradient properties, the locality along each direction can be defined according to the related precessional frequency. In this regard, in order to spatially locate the signal in a three-dimensional volume three gradients are required. As previously mentioned, each of them has to perform one of the following task: slice selection, frequency encoding and phase encoding. The first is required to define a slice on the scanning plane, whereas the other two are needed to retrieve the two-dimensional position on the previously selected slice. In a static field $B_0$, all the hydrogen nuclei in a given volume can only resonate at a particular Larmor frequency. On the other hand, when a gradient is switched on, the precession rate of hydrogen nuclei become gradient-dependent. Therefore, a RF pulse with a given frequency band allows a selective excitation of a subpopulation of nuclei. This is to say, a specific slice can be excited by releasing a RF pulse with a band frequency coinciding with the precessional frequency of hydrogen nuclei in that specific slice. Hydrogen nuclei in the rest of the volume do not resonate because of the mismatch between their precessional frequency and the RF pulse frequency. The RF pulse band frequency is known as the transmit bandwidth. In general, the gradient is turned off and it is switched on only when the excitation pulse is released. Its slope along with the bandwidth determines the slice thickness. Normally, a steep slope and a narrow bandwidth are associated with thin slices, while a shallow gradient and a broad bandwidth permit to
increase the slice thickness. Once the slice has been defined and the nuclei in it excited, the signal must be located in the two dimensions. The first dimension to be encoded is the long axis of the anatomy. This process is known as frequency encoding. As for the slice selection gradient, the aim of the frequency encoding gradient is to linearly alter the precessional rate. Because this gradient is switched on during the echo formation, therefore when the signal is recorded, it is commonly referred to as readout gradient. Specifically, it is applied before the maximum of the echo (rephasing) and is turned off after (dephasing). The degree of the slope is inversely proportional to the size of the anatomy to cover along the gradient axis or field of view (FOV). In practice, the greater the slope degree is, the smaller the FOV.

![Image of three-dimensional spatial localization timing diagram.](image)

*Figure 1.8 – Three-dimensional spatial localization timing diagram.*

Eventually, the signal location must be encoded along the remaining direction, the short axis of the anatomy. This process is known as phase encoding. When the phase encoding gradient is switched on the precessional frequency changes in accordance to the linear gradient. The frequency change can also be interpreted in terms of phase differences in the precessional path of hydrogen nuclei. In other words, some nuclei slow down and others speed up, therefore the first will move backward in the precessional path,
while the second will move forward. When the gradient is switched off the accumulated phase difference will be maintained. The gradient slope defines the phase shift along the phase encoding axis. Usually, in a spin echo sequence, this gradient is switched on after the RF excitation pulse is switched off.

1.2.2.1 K space and image formation

Once the slice is selected, the application of the phase and frequency encoding gradients produce a phase and a frequency shift along the respective direction. To easily convey the concept, let us focus on what happens during a simple spin echo sequence. During the excitation pulse, a slice selection gradient is activated causing the protons in the given slice to resonate. The excitation pulse is then switched off and so is the gradient. Afterwards, a specific slope of the phase encoding gradient is applied. This results in a precise phase shift of the precessing nuclei. The gradient is then switched off while the accumulated phase shift of the precessing nuclei is maintained. Later, the 180° refocusing pulse is released and the rephasing mechanism begins. At this point, during the echo formation, the frequency encoding gradient is turned on causing a frequency shift of the precessing nuclei. Concomitantly with the application of this gradient, the signal induced in the receiving coil is recorded and digitized. Finally, the frequency encoding gradient is switched off and the whole process is repeated. This time, the same slice is excited but the slope of the phase encoding gradient is changed. When data of each specific phase-frequency set are collected, the information is stored as numeric values in a matrix of the MR acquisition computer. This matrix is called K space and it is equivalent to the raw data (Figure 1.9). In MR imaging, the raw data are the transverse component of the NMV of the imaged object after excitation, sampled from the signal induced in the receiving coil as a function of time. The frequency encoding information is saved in the horizontal dimension, while the phase information in the vertical. In the classic K space filling, data are stored line by line and for each TR a line is filled. The line to be filled is determined by the polarity and slope of the phase gradient. Specifically, positive polarity slopes are associated with lines in the top half of the K space matrix, whereas negative polarity slopes with lines in the bottom half. In order to cover the whole K space, the phase gradient is changed every TR. The filling of a specific line is determined by the frequency encoding mechanism. When the frequency encoding gradient is switched on, frequencies in the echo are digitized and stored as matrix elements of the K space line selected by the phase encoding gradient. Therefore, the K space represents the spatial frequency information in two dimensions of the imaged object. It is defined by the space covered by the phase and frequency encoding data. The relationship between the K space and the actual image is defined by the Fourier transformation. A point in the K space matrix does not correspond to the same point in the equivalent MR image slice. The outer rows and columns of the K space carry information regarding high spatial frequencies in the image (e.g. borders and contours), while the inner elements of the matrix provide information on the general contrast of the image.

To produce the actual MR image $s(r)$ from the acquired K space matrix $S(k)$, a mathematical algorithm called Fourier transform is implemented:
where \( \vec{r} \) denotes the actual spatial position and \( \vec{k} \) denotes the position in the K space. The raw MR imaging measurements are equivalent to the Fourier transform \( S(\vec{k}) \) of the imaged object. Therefore, the solution of the image reconstruction problem is an inverse Fourier transform:

\[
\begin{align*}
S(\vec{k}) &= \int s(\vec{r}) e^{i2\pi \vec{k} \cdot \vec{r}} d\vec{r} \\
\text{s}(\vec{r}) &= \int S(\vec{k}) e^{-i2\pi \vec{k} \cdot \vec{r}} d\vec{k}
\end{align*}
\]
1.3 Image contrast

As previously mentioned, specific tissue types are characterized by different relaxation properties. Therefore, the longitudinal relaxation time $T_1$ and the transverse relaxation time $T_2$ can be considered as tissue-specific. Thus, it is possible to profit from these differences to define specific image contrasts. This is achieved by directly weighting the contrast towards specific parameters while dimming others. Two standard imaging contrasts in MRI are known as T1-weighted ($T_1$-w) and T2-weighted ($T_2$-w). By exploiting different pulse sequences and by varying imaging parameters, the image contrast can be controlled.

1.3.1 T1- and T2-weighted contrasts

The $T_1$-w is one of the fundamental imaging contrasts in MRI. It demonstrated the difference in the $T_1$ relaxation times between tissues. As learnt before, $T_1$ relaxation relies upon the recovery of the NMV longitudinal magnetization. After the RF pulse excitation, not all tissues realign to the static field $B_0$ at the same pace. Therefore, differences in the $T_1$ recovery are a potential source of imaging contrast. In order to highlight this contrast, while lessening the other two, the TR and the TE must be short. In particular TR should be short enough to avoid a complete longitudinal recovery of different brain tissues (Figure 1.10).

![Figure 1.10 – T1 recovery of different brain tissues.](image)
For instance, fat-based tissue (e.g. WM), characterized by a shorter T1, will recover faster than water (e.g. CSF). Therefore, a short TR will allow a better contrast between these two tissues. Eventually, in order to minimize the T2-w contrast or the T2 relaxation time differences between tissues, a short TE must be used. T1-w images are especially useful for anatomic information, particularly at the interface between the brain and the CSF as well as for the exquisite contrast between WM and GM regions (Figure 1.11).

![Figure 1.11 – T1-weighted image.](image)

Along with T1-w, the T2-w contrast is another important imaging contrasts in MRI. It highlights differences in the T2 relaxation time between tissues. After the RF excitation pulse is switched off, magnetic moments begin to dephase and the NMR transverse magnetization undergoes the classic T2 decay. Differences in this decay are captured in the T2-w contrast. A proper T2-w image is one in which this contrast is maximize as opposed to T1-w contrast which is minimized. Contrary to T1-w imaging, the T2-w requires long TR and TE (Figure 1.12). A long TR permits different tissues to totally recover their longitudinal magnetization reducing T1-w contrast. On the other hand, a long TE allows different tissues to lose the initial transverse magnetization and therefore increase the differences in the T2 decay. T2-w images demonstrate a marked tendency to delineate pathological features characterized by abnormal water content within structures (e.g. inflammation or infection) (Figure 1.13).
Figure 1.12 – T2 decay of different brain tissues.

Figure 1.13 – T2-weighted image.
1.3.2 Magnetization Prepared Rapid Gradient Echo sequence

Due to a considerably reduced acquisition time, excellent image quality and the overall capability to depict focal lesions in the brain, the Magnetization Prepared RAdip Gradient-Echo imaging (MP-RAGE) sequence is typically used for T1-w imaging (Brant-Zawadzki et al., 1992; Fellner et al., 1996; Mugler and Brookeman, 1990). MP-RAGE belongs to the family of fast 3D gradient echo pulse sequences. Technological advances in MRI have permitted rapid switching of gradients, which makes it possible to obtain images using short TR (in the order of 4-10 ms) and TE times (2-4 ms). By using a preparatory 180° RF pulse it is possible to alter the longitudinal magnetization as desired before the sequence itself begins. This allows the operator to achieve an outstanding T1-w image contrast (Brant-Zawadzki et al., 1992; Fellner et al., 1996). As shown in Figure 1.14, the sequence can be divided into three major parts: preparation of the magnetization, data acquisition and a recovery period. The sequence begins with a magnetization preparation (MP) period in conjunction with a rapid gradient-echo (RAGE) sequence (Mugler and Brookeman, 1990).

![Figure 1.14 – MP-RAGE pulse sequence timing diagram (Fellner et al., 1996).](image-url)
After the longitudinal magnetization is encoded with the desired contrast during the MP period, a RAGE sequence is used to sample the prepared magnetization. An incoherent gradient echo (e.g. FLASH) sequence with very short TR, TE and extremely low flip angle is used to acquire the image data. It is a sequence characterized by a variable flip angle excitation pulse and the use of gradient rephasing to produce a gradient echo. The steady state is maintained so that the residual magnetization is left over following previous repetitions. The key feature of this sequence is the dephasing or spoiling of this residual magnetization so that its effect on image contrast is almost absent. After the signal is collected in the form of a gradient echo, a gradient pulse called spoiler is used to destroy any remaining transverse magnetization. The shortness of the repetition time in this sequence requires the use of a spoiler, otherwise, some transverse magnetization would remain at the beginning of the following excitation pulse. Therefore, only transverse magnetization from the previous excitation is enrolled, enabling T1 contrast to dominate while lessening T2 contrast. Following the inversion pulse all 3D phase encoding steps are performed by changing the slice-selection gradient step by step, keeping the 2D phase encoding gradient constant. After the acquisition period, a magnetization recovery follows. The duration of the recovery period is an additional parameter, which determines by the desired contrast properties of the image. Eventually, after the recovery period the whole procedure is repeated changing the phase gradient for the next 2D phase encoding step (Fellner et al., 1996). The complete 3D MP-RAGE image is formed repeating the cycle described above until the desired K space volume is covered.

1.3.3 Fast Spin Echo sequence

The Fast Spin Echo (FSE), or Turbo Spin Echo (TSE), is a well-established technique for the rapid acquisition of T2-w images with true spin-echo contrast features (Atlas et al., 1993; Jolesz and Jones, 1993). Depending on the desired protocol, FSE images may be obtained almost three to 16 times faster than conventional SE images. Advantages aside from the rapid imaging include decreased motion artifacts, decreased magnetic susceptibility effects when imaging metallic objects and skull base (Jones et al., 1992; Thomas et al., 2004). As the name implies, the FSE is a modified version of a conventional SE sequence. The main difference is a much shorter scan time. The reduction in the acquisition time is achieved by performing more than one phase encoding step (therefore more than one K space line is filled) per TR. Using a train of 180° RF rephasing pulses instead of a single one, it is possible to generate an echo train (Figure 1.15). Therefore, for each 180° pulse, a different phase encoding step is implemented and for each echo formation a frequency encoding is applied. The number of rephasing pulses for each TR corresponds to the number of K space lines filled. This number is called the turbo factor. The higher the turbo factor the shorter the scan time. An important aspect to consider is the fact that different echoes are produced at different TE times. As seen before, in the conventional spin echo sequence the K space filling is performed at a specific TE, so that the reconstructed image is defined by a particular contrast. On the other hand, the FSE sequence is characterized by different TE times, which generate different contrasts. Therefore, each K space line is equivalent to a specific weighting on the actual image. The selected TE is only an effective
TE. To achieve the contrast imposed by the effective TE, the phase encoding is performed in a predefined order. Steep gradient slopes, mainly responsible for image borders and spatial resolution, are located away from the effective TE, thus away from the center of the K space. On the contrary, shallow slopes, which account for the average image signal and therefore relevant for the actual contrast, are placed in correspondence of the effective TE, hence in the center of the K space.

Figure 1.15 – FSE pulse sequence timing diagram.
1.4 Image properties

In the first part of this chapter, we have presented the fundamental principles behind the MR acquisition and image formation process. Hereafter, we will introduce two of the main drawbacks affecting conventional MR imaging: intensity non-uniformity artifacts and intensity scale inconsistency. The main effect of both elements is experienced as unpredictable intensity variations. The first type consists of intensity changes within the same tissue class, throughout a single volume. Indeed, signal intensity measured from homogeneous tissue regions is not completely uniform, but it varies smoothly across the image. The second type involves inter-scan signal intensity variability. MR intensities vary between different scans with the same acquisition protocol even for intra-patient studies. Therefore, a defined measured intensity cannot be associated with a specific tissue class.

1.4.1 Intensity non-uniformity in MR data

Intensity non-uniformity (INU) artifacts in MR imaging refers to the presence of smoothly varying (non-anatomic) intensity variations within tissues (Belaroussi et al., 2006; Boyes et al., 2008). It is also denoted as intensity inhomogeneity or more generally as bias field. In extreme cases, this inhomogeneous field, added to the image, makes it impossible to differentiate between GM and WM (Figure 1.16). As a result, this loss of image contrast makes automated processing techniques and even visual inspection particularly troublesome (Likar et al., 2001). The presence of this artifact may affect both qualitative and quantitative image analyses. Indeed, all the analyses conducted on structural MR images, from which quantitative features of anatomic structures are extracted, generally requires a preprocessing step to remove this artifactual component. The presence of the INU throughout the image may be induced by several factors, which can be conventionally grouped as: factors related to the acquisition technique and factors associated with the imaged subject (Arnold et al., 2001; Belaroussi et al., 2006; Collins et al., 2005; Lutti et al., 2010; Van De Moortele et al., 2005). Regarding the first group, numerous are the triggers. Starting from the static magnetic field $B_0$, an imperfect spatial homogeneity is the main cause of slow intensity variations across the imaged volume as well as partial deformations of the imaging object. Gradient fields also play an important role in the artifact generation process. The slope of the gradient-encoding fields is not truly linear and the eddy currents driven by the switching of field gradients affect the spatial appearance of INU artifacts. Along with the aforementioned causes, intensity inhomogeneities are mainly reflected in flip angle variations, i.e. the value of this angle is not homogeneously distributed all over the brain. RF coil inhomogeneity, which depends on its geometry and tuning, is the principal source of this spatial non-uniformity. The inhomogeneity is commonly experienced in the transmitting as well as in the receiving coil, but with different dependence on the RF field.

In terms of the subject-related factors, tissue-dependent properties such as conductivity and dielectric constants also influence the INU field appearance. In particular, these properties become a significant issue with an increase of the static field $B_0$ strength (Bernstein et al., 2006; Moser et al., 2012;
Umutlu et al., 2014). As the static field strength increases, the frequency of the RF field increases with a concomitant decrease of its wavelength. The wavelength decreases even further in biological tissues, since many tissues have high relative dielectric constants, until the point it reaches a length comparable, or even less than, the dimension of the human brain (Vaughan et al., 2001). As a result, wave propagation effects appear inside the head (Figure 1.16). The high dielectric constant of brain tissue results in standing waves inside the head that greatly perturb the effective RF field.

In addition, due to electrical conductivity of biological tissues, the induced voltage at high frequencies produces eddy currents also capable of disturbing the RF field. As reported in literature, when the wavelength of the RF is not much longer than the size of the object, wave effects spatially modulate the image intensity (Bernstein et al., 2006). Despite several correction procedures using phantoms, shimming techniques, and dedicated hardware solutions have been proposed and routinely implemented, a supplementary post-processing correction is often necessary (Belaroussi et al., 2006). Eventually, the presence of INU artifacts can significantly reduce the accuracy of all the post-processing chain following image acquisition. In particular, the image segmentation and registration steps may result the most
affected, hence decreasing the reliability of subsequent quantitative measurements. In the following chapters, we will address this problem with the main goal of removing or at least reducing this artifact.

### 1.4.2 Intensity inconsistency

A major difficulty in conventional MR imaging analysis is the fact that intensities do not have a tissue-specific numeric meaning, not even for the same MRI protocol or for images of the same patient obtained on the same scanner (Jager and Hornegger, 2009; Nyul and Udupa, 1999). After a successful correction of intensity inhomogeneities, despite a considerable gain in reliability, it is still not possible to quantify the observed signal intensities. This implies that MR images cannot be globally displayed at fixed intensity scales; rather an arbitrary scale may have to be defined per case (Figures 1.17 and 1.18). Moreover, for most post-processing applications (e.g. image segmentation, registration and features quantification), this lack of standardization is a major drawback which hampers an accurate and precise image analysis (Bagci et al., 2010; Bezdek et al., 1993; Clarke et al., 1993; Kikinis et al., 1992; Udupa et al., 1997). In comparative and longitudinal studies, it is important that the image intensity for a given tissue remains comparable across all examinations. Although the flip angle is maintained constant across datasets and that signal variation due to INU has been corrected, there is still a global scaling effect on images because of the variation in scanner sensitivity related to how the scanner is calibrated. The scanner sensitivity varies not only between scanners but also within scanner due to variations in coil loading and receiver attenuation setting (Wang et al., 1998).

Several attempts have been made to calibrate MR signal intensity at the time of acquisition using phantoms (Edelstein et al., 1984; Tofts, 1998). Although it is reasonable to perform such standardization for each subject, it is cumbersome. Not only is this approach not applicable to datasets already acquired, but also the execution time is clearly a disadvantage.

On the contrary, post-processing techniques directly applied to the acquired MR image are undoubtedly more appealing. An initial approach towards an automatic processing of MR data is the one proposed in (Wendt, 1994). By means of two image statistics, the median and the standard deviation, the method is able to adjust image contrast and brightness in an automated manner. Although such unsupervised processing may allow the user to display the MR signal with an acceptable degree of uniformity, it is not adequate for quantitative image analyses, since the image intensity still do not posses any tissue-specific connotation.

The first real attempt to overcome the standardization issue using a post-processing approach was made in (Wang et al., 1998). In this study the authors propose a new method based on shape similarity between an image histogram and a reference histogram. This approach is also known as histogram matching.
Figure 1.17 – Intensity distribution inconsistency on T1-w images collected with different MR scanners.

Figure 1.18 – Intensity distribution inconsistency on T2-w images collected with different MR scanners.
By means of two scaling factors, the image histogram to be standardized is stretched along the frequency and the intensity axis. In an iterative approach the process is repeated until minimization of the histogram differences occurs.

Another intensity standardization method, is the one described in (Nyul and Udupa, 1999) and then adopted in (Bagci et al., 2010; Ge et al., 2000; Madabhushi and Udupa, 2005; Nyul et al., 2000) which also implements a histogram matching approach. First, some landmarks (e.g., percentiles, modes) are extracted on a training set of histograms. When a new image is acquired, the detected landmarks of its histogram are matched to the previously computed standard landmarks. Finally, in order to create a continuous intensity mapping, the histogram positions between the landmarks are linearly interpolated.

Despite, preliminary studies performed on these histogram-matching methods indicate that image analysis methods are considerably improved in terms of parameter settings consistency and degree of automation (Ge et al., 2000; Madabhushi and Udupa, 2005), some drawbacks should be acknowledged. It is a technique particularly vulnerable to distortions in the histogram shape due to the fact that the entire intensity standardization relies exclusively on a given histogram profile. In particular, the landmark identification process may be inaccurate in images characterized by high levels of noise and INU artifacts. Indeed, as shown in (Madabhushi and Udupa, 2005), the optimal processing approach for enhancing image standardization can be achieved by preceding it with inhomogeneity correction. Furthermore, the intensity regularization may be erroneous in pathological cases in which the intensity distribution, thus the histogram shape, is considerably altered.

A more sophisticated histogram based method was proposed in (Hellier, 2003). The algorithm is based on a mixture of Gaussian model, which, given an image, estimates a combination of Gaussians distributions that best approximates its histogram. Afterwards, it computes a polynomial correction function that aligns the mean intensities of the different tissue classes represented by the mixture of Gaussians. Unfortunately, this approach is predisposed to the same downsides characterizing other histogram matching techniques. In particular, the presence of image noise and image non-uniformity largely affects the quality of the Gaussian fit.

Later on, Weisenfeld and Warfield proposed a combined intensity standardization and inhomogeneity correction method (Weisenfeld and Warfield, 2004). This method used the Kullback-Leibler divergence as a measure of relative entropy between an image intensity distribution and the template intensity distribution. Relative entropy confers a degree of robustness to the histogram matching, but even this method fundamentally relies on a histogram-matching scheme and ignores relevant spatial information. Without the use of spatial information to reinforce the matching by using image-specific properties, standardizing the histograms does not necessarily guarantee a standardization of the intensities of the different tissue types.

Recently, a new standardization method based on the joint histograms of two image modalities (e.g. T1- and T2-w images) has been presented in (Jager and Hornegger, 2009). A key feature of this approach is the fact that the image intensity distributions are represented in multi-dimensional joint histograms instead of one-dimensional histogram. In order to normalize the probability density function of
a newly acquired image set, a non-rigid image registration is performed between a reference and the joint histogram of the acquired images. From this matching a non-parametric transformation is generated, which describes a mapping between the corresponding intensity spaces. Consequently, this mapping is used to transform the intensity scale of the acquired images to a given standard. As the proposed intensity standardization is based on the probability density functions of the image sets only, it is independent of spatial consistency or prior segmentations of the reference and current images. Moreover, it can be directly applied to different acquisition protocols and body region.

All the standardization techniques described here are indeed largely used in literature. The main concept behind these approaches relies on the assumption that a perfect match between histograms results in a significant gain in the intensity scale similarity of the resulting images. Although this approach may convey satisfactory results when applied to high quality images of healthy subjects, its performance may drastically decrease in pathological populations. Indeed, without considering tissue-specific information, abnormal intensity variations due to pathological processes may be erroneously accounted for (Robitaille et al., 2012). In the next chapters, we will try to overcome this issue, which is currently still unresolved. We will propose a novel technique, which can be applied in both healthy and pathological subjects.

### 1.5 Multimodal MR imaging

As reported in the literature, T1-w images are characterized by better contrast-to-noise ratio in WM regions, whereas T2-w sequences can be used to better discriminate structural differences in fluid-filled regions (Braga et al., 2012). Hence, the combination of both modalities may prove significant in investigations focused on an unknown spectrum of brain pathologies. Therefore, a multi-modal structural imaging approach based on the fusion of T1-w and T2-w images may be helpful for the study of brain tissue abnormalities. Glasser and Van Essen in 2011 (Glasser and Van Essen, 2011) were the first to emphasize the potential effect of the ratio approach. In their study, they pioneered a new multimodal MR methodology for the cortical mapping of brain areas based on myelin content. Specifically, they showed how the contrast related to myelin could be increased by performing the voxel-wise ratio between T1-w and T2-w images. By means of a surface-based registration approach, the ratio image from each subject is mapped to the cortical surface and thus aligned across individuals. Then, the spatial gradient of the myelin map provides a measure of transition in myelin content across the surface. They validated the technique by showing good agreement between myelin map gradients and published atlases of cytoarchitectonically defined cortical areas. One of the main assumptions underlying this technique relies on the notion that the sensitivity profile of the RF receiver coil is the same in both modalities. Thus, the ratio of the T1-w to the registered T2-w image mathematically cancels the intensity inhomogeneities related to the sensitivity profile of the receiver coil. On the other hand, the intensity variations due to transmit field inhomogeneities may be considered minimal when a body coil is used, since such coils produce uniform transmit fields over the head. Eventually, to provide an effective visualization of the resulting myelin maps,
an intensity normalization approach was employed. Myelin maps values, obtained from the ratio of T1-w and T2-w imaging, are indeed dimensionless quantities. To overcome this issue, the authors implemented a histogram equalization technique displaying the actual myelin maps between the 3rd and 96th percentile of the overall intensity distribution. In this way, a better characterization of intensity variations across subjects and datasets was ensured.

It is worth noting that, in order to extend this methodology for the mapping of pathological changes in brain structure, two major limitations should be acknowledged. The first is the presence of spatial inhomogeneities in image intensity, produced by interactions between the subject’s body and the MR scanner. As stated by the same authors, although the method is generalizable across different 3T scanners and pulse sequences, its effectiveness is not granted for other acquisition protocols. In particular, these assumptions may not hold at higher field strengths where local transmit coils are regularly used and transmit inhomogeneity fields are much stronger (Umutlu et al., 2014; Uwano et al., 2014; Van De Moortele et al., 2005). In addition, despite the method is able to convey excellent results on the cortical surface, no guarantee is given in terms of whole brain analyses. Secondly, the lack of image intensity standardization, mainly related to the MR instrumentation and acquisition sequence used, poses an important limitation. Although the proposed intensity normalization approach may allow a sufficient representation of contrast transitions between adjacent areas, no tissue-specific information can be extrapolated from the standardized images.

1.6 Thesis overview

In this thesis, the main goal will be to develop an extension of the T1-w/T2-w approach for the assessment of structural changes in the brain. In chapter 2, we will tackle the problem of the INU artifacts in MR data. In particular, we will address the problem of quantitatively comparing standard INU correction techniques for T1-weighted images. Notably, the same approach can be applied to T2-weighted images as well. We will use simulated data to assess the correction performance by those methods for controlled inhomogeneity magnitudes and image noise levels. Then, in chapter 3, we will introduce an optimized methodology for INU correction in real MR data. Here we will address the debated question of how to define the optimal configuration for a given INU correction method. In chapter 4, we will deal with the development of the T1-w/T2-w technique originally proposed by Glasser and Van Essen (Glasser and Van Essen, 2011). We will tackle the problem of intensity scale inconsistencies across different datasets, and integrate it in a novel analysis workflow for the standardization of T1-w/T2-w intensities in the brain. In chapter 5, we will apply the developed T1-w/T2-w pipeline for the assessment of structural impairments in the diseased brain. We will extensively comment on the work presented in the thesis, addressing general topics in MR imaging that are relevant to this work, in chapter 6.
Chapter 2

QUANTITATIVE EVALUATION OF INTENSITY INHOMOGENEITY CORRECTION METHODS FOR STRUCTURAL MR BRAIN IMAGES

The correction of intensity non-uniformity (INU) in magnetic resonance (MR) images is extremely important to ensure both within-subject and across-subject reliability. Here we tackled the problem of objectively comparing INU correction techniques for T1-weighted images, which are the most commonly used in structural brain imaging. We focused our investigations on the methods integrated in widely used software packages for MR data analysis: FreeSurfer, BrainVoyager, SPM and FSL. We used simulated data to assess the INU fields reconstructed by those methods for controlled inhomogeneity magnitudes and noise levels. For each method, we evaluated a wide range of input parameters and defined an enhanced configuration associated with best reconstruction performance. By comparing enhanced and default configurations, we found that the former often provide much more accurate results. Accordingly, we used enhanced configurations for a more objective comparison between methods. For different levels of INU magnitude and noise, SPM and FSL, which integrate INU correction with brain segmentation, generally outperformed FreeSurfer and BrainVoyager, whose methods are exclusively dedicated to INU correction. Nonetheless, accurate INU field reconstructions can be obtained with FreeSurfer on images with low noise and with BrainVoyager for slow and smooth inhomogeneity profiles. Our study may prove helpful for an accurate selection of the INU correction method to be used based on the characteristics of actual MR data.

Published as:
2.1 Introduction

Magnetic Resonance Imaging (MRI) is a valuable technique for studying the structural properties of the human brain. Due to its non-invasive nature, significant imaging contrast, high spatial resolution, and reasonable acquisition times, it is largely used to investigate alterations in brain structure associated with neurodegenerative and neuropsychiatric disorders (Braga et al., 2012; Chen et al., 2011; Duncan et al., 2013; Frisoni et al., 2010; Tillema and Pirko, 2013). A serious issue in the analysis of MR structural images is, however, the reproducibility of the imaging results, both within and across subjects, that arises from smooth intensity variations across the whole MR image (Belaroussi et al., 2006; Bernstein et al., 2006). These variations can be referred to as intensity non-uniformity (INU), but also intensity inhomogeneity or spatial bias. The magnitude and spatial profile of the INU may be influenced by several factors, among which static field inhomogeneity, reduced radiofrequency (RF) coil uniformity, RF penetration, gradient-driven eddy currents, inhomogeneous reception sensitivity profile, and overall subject anatomy and position (Belaroussi et al., 2006; Mihara et al., 1998; Simmons et al., 1994; Sled and Pike, 1998; Vovk et al., 2007). The acquisition pulse sequence and the field strength play also an important role in determining the INU (Belaroussi et al., 2006; Bernstein et al., 2006; Boyes et al., 2008). In particular, an imperfect spatial homogeneity of the static magnetic field $B_0$ is the main cause of slow intensity variations across the imaged volume when the MR field strength is relatively low. At higher MR field strengths the contribution of $B_0$ will diminish as the other effects, as for example tissue-dependent distortions produced by MR gradients, will start to become much more significant and not behave in a manner that suits the INU correction methods assumptions (Bernstein et al., 2006; Moser et al., 2012; Umutlu et al., 2014).

To correct the INU in MR structural images, a number of prospective calibration methods were proposed (Belaroussi et al., 2006; Vovk et al., 2007). These are generally intended to account for hardware-related factors that hamper MR image quality. For instance, it was suggested that the INU may be compensated by acquiring supplementary images of uniform phantoms (Axel et al., 1987), combining information from different coils (Brey and Narayana, 1988; Murakami et al., 1996), merging data obtained from multiple datasets (Liney et al., 1998), and designing dedicated imaging sequences (Deichmann et al., 2002; Mihara et al., 1998). Nevertheless, it is important to note that those methods can eliminate hardware-related but not subject-induced inhomogeneities. Furthermore, the usefulness of prospective approaches is narrowed by the need of dedicated acquisitions, the limited stability and the sensitivity to input parameters (Belaroussi et al., 2006; Likar et al., 2001; Vovk et al., 2007).

Given the limitations of prospective methods, retrospective INU correction methods, which rely only on image features to remove spatial inhomogeneities, are nowadays more widely used. Notably, they can be applied to structural images with different features, and can theoretically account for both hardware-related and subject-induced INU components (Hou, 2006; Vovk et al., 2007). Many retrospective INU correction methods have been proposed in the last years. In spite of their different implementations, a common characteristic is that they model the INU field as a spatial function describing
slowly changing intensity variations across the volume. This INU field is typically assumed to be multiplicative, in that the intensity of the inhomogeneity is proportional to that of the INU-free MR image at the same location (Ashburner and Friston, 2005; Axel et al., 1987; Belaroussi et al., 2006; Vovk et al., 2007). Also, the presence of an additive hardware-related noise should be taken into account when dealing with actual structural MR images. Notably, performance of an INU correction method can be influenced both by INU characteristics and by noise in various ways, depending on its specific features and implementation.

Since an effective INU correction is critical for investigations of brain structure, previous studies have attempted to compare the performance of several approaches. A number of comparative studies on retrospective INU correction methods were conducted (Arnold et al., 2001; Likar et al., 2001; Velthuizen et al., 1998; Vovk et al., 2006), but none of them recently. This leaves an unanswered question on whether and to what extent newly developed methods outperform older ones. Furthermore, previous comparative studies focused exclusively on default parameters for each INU correction method. Nonetheless, since each method performs better or worse depending on the specific settings used, the selection of optimal parameters is becoming an important topic in MRI (Boyee et al., 2008; Uwano et al., 2014; Weiskopf et al., 2011; Zheng et al., 2009).

In this study we conduct a quantitative assessment of INU correction methods for T1-weighted (T1-w) images, which are the most commonly used images to investigate brain structure. Specifically, we focus on the methods implemented in the most recent versions of Statistical Parametric Mapping (www.fil.ion.ucl.ac.uk/spm), FMRIB Software Library (www.fmrib.ox.ac.uk/fsl), FreeSurfer (www.freesurfer.net) and BrainVoyager (www.brainvoyager.com), respectively. We use simulated data to compare the method results with a ground truth at different INU field magnitudes and image noise levels. Furthermore, we examine a wide range of input parameters for each method, so that we can define their enhanced configuration and compare its performance with those obtained using default input parameters.
2.2 Methods

2.2.1 INU correction: theory and algorithms

In this section, we first introduce the theoretical background for modeling of the INU effects on MR images, and then we describe how each method attempts to remove it from the data.

2.2.1.1 Modelling of MR intensity inhomogeneities

According to the RF field mapping theory (Insko and Bolinger, 1993; Stollberger and Wach, 1996), intensity inhomogeneities can be modeled as multiplicative. The majority of the studies (Arnold et al., 2001; Belaroussi et al., 2006; Dawant et al., 1993; Pham and Prince, 1999; Wells et al., 1996) suggested that the corruption of MR images by intensity inhomogeneity can be formalized as follows:

\[ u(x,y,z) = v(x,y,z) \cdot b(x,y,z) + n(x,y,z) \]  

where \( u(x,y,z) \) is the actual image and \( v(x,y,z) \) is the INU-free noiseless image, \( b(x,y,z) \) is the INU field and \( n(x,y,z) \) is additive spatial noise. While the INU field is slowly varying, the noise has high spatial frequency and its values show a Rician distribution (Andersen, 1996; Belaroussi et al., 2006; Gudbjartsson and Patz, 1995; Henkelman, 1985). Intensity inhomogeneity in MR images can be corrected by estimating the INU field \( b(x,y,z) \). This, indeed, permits to approximate the INU-free image \( v(x,y,z) \) starting from the actual image \( u(x,y,z) \). It is worth noting that the contribution of the noise, due to its high spatial frequency, cannot be eliminated by means of the INU correction (Vovk et al., 2007). To accomplish the INU correction, different methods have been developed in the last years. We will focus hereafter on methods implemented in widely used MR imaging software packages.

2.2.1.2 Methods under investigation

Statistical Parametric Mapping (SPM)

The INU correction in SPM12 is implemented within the unified segmentation module (Ashburner and Friston, 2005). It is indeed integrated with brain segmentation, as this allows the joint optimization of both analysis steps. Specifically, since intensity inhomogeneity is detrimental for the image segmentation process, the INU correction is iteratively performed until convergence of the segmentation results. This is achieved combining a Finite Gaussian Mixture (FGM) model with a deformable template (tissue probability atlas). In this respect, a mixture of Gaussians is used to model the intensity distribution from different tissue types. By default, SPM uses more than one Gaussian for each tissue, since tissue probability maps may be shared across different classes (partial volume effects). On the other hand, built-in (prior) probability maps of different tissues are registered to the subject image. Afterwards, Bayesian statistics rules are used to calculate posterior probabilities that combine the template information with that
Chapter 2

contained in brain tissues. The INU field correction algorithm of SPM models smooth intensity variations by a linear combination of discrete cosine transform (DCT) basis functions. In other words, SPM represents the MR image as a sum of sinusoids of varying magnitude and frequency (basis functions). Due to the low-frequency nature of image inhomogeneities, slowly varying INU fields are isolated in terms of DCT components below a certain cut-off threshold.

An effective segmentation, and therefore an effective INU correction, is based on the minimization of the objective function derived from the FGM model along with the deformable template. The fitting of the model (i.e. minimization of the objective function) is performed following an Iterated Conditional Modes (ICM) approach. In other words, each iteration involves an estimate of different groups of parameters, while holding others at their optimal current solution. Upon convergence, the toolbox provides structural images that are INU-corrected and segmented.

**FMRIB Software Library (FSL)**

As for SPM, the INU field correction method implemented in FSL v5.0 is integrated with the segmentation tool, called FMRIB’s Automated Segmentation Tool (FAST). The FSL method estimates the INU field by fusing information from a FGM model and from a Hidden Markov Random Field (HMRF) model (Zhang et al., 2001). The FGM model is used as in SPM to decompose the image histogram into a mixture of Gaussians. On the other hand, the HMRF model considers the image information to be encoded through contextual constraints of neighboring voxels, so that the presence of noise, strong INU fields, and mainly partial volume effects can be taken into account. The objective function to be minimized is derived from a combination of FGM and HMRF, and is optimized using an ICM approach. Model parameters are estimated by means of an expectation maximization (EM) approach in the framework of image segmentation, iterating between tissue classification and intensity inhomogeneity correction. The expectation step consists of computing the maximum a posteriori estimate (MAP) of the INU field and the tissue labels. In turn, the maximization step is accomplished by computing the maximum likelihood estimate of the model parameters using the INU field and the tissue labels of the expectation step. As mentioned above, the MAP principle is employed to obtain the optimal estimate of the INU field starting from the observed intensity values. By comparing the actual voxel intensities with the predicted ones, a residual field is calculated. Then, this residual field is low-pass filtered to obtain an estimate of the INU field. This procedure can be iterated multiple times. By estimating the EM solution, INU correction and brain segmentation can be performed at the same time.

**FreeSurfer (FS)**

FreeSurfer v5.3 includes a famous INU correction method developed by the Montreal Neurological Institute (MNI), and known as N3 (Sled et al., 1998). This method considers the intensity at each voxel as an independent distributed random variable. The basic assumption is that the INU field has a blurring effect on the MR image, reducing the high frequency components that characterize the image. As a result, the method tries to find the smooth INU field that maximizes the frequency content of the
image intensity distribution. This is achieved through an iterative process that goes through three sequential steps: sharpening of the INU-corrupted intensity distribution, INU field estimate and INU field smoothing. The iterative process terminates when no significant changes in the estimated INU field are detected.

BrainVoyager (BV)

BrainVoyager QX implements an intensity inhomogeneity correction method based on a surface fitting approach. By means of low-order polynomials, the algorithm models low-frequency variations across the image (Dawant et al., 1993). The INU detection is accomplished using intensity information of voxels presumably located in the white matter (WM). After brain extraction, the labelling of WM reference points is achieved in two steps. The first phase is based on voxel intensity criteria. The idea behind that is that surface fitting is most reliable if it is estimated on voxels with higher intensity, and these are likely to be found in the WM. However, this assumption may not hold in presence of heavy inhomogeneity profiles as well as low signal-to-noise ratio. To overcome this hindrance, a heuristic fully automated approach based on intensity information from neighbouring voxels to label as reference points in the WM (Hou et al., 2006). Then, the intensities of reference voxels are fit by low-order polynomials following a least-squares approach. In this manner, it is possible to detect intensity variation across the whole volume. Finally, the INU field is generated using the calculated low-order polynomials, and is regressed out from the actual image. Optionally, the procedure of reference point labelling and INU field estimation may be iterated, aiming at minimizing residual errors. It is worth noting that the use of multiple iterations can introduce additional low-frequency noise in the estimated INU-free image if the residual INU to be estimated has very low magnitude and surface fitting becomes less reliable.

2.2.2 Analysis of INU correction methods

In this section we describe how we built a realistic simulation to include spatial inhomogeneity in MR images, how the four INU correction methods were set up and tested, and how the performance of these methods were assessed.

2.2.2.1 Simulated data

A first simulated INU field was created using images generated from the BrainWeb MRI Simulator (brainweb.bic.mni.mcgill.ca/brainweb). We first extracted a realistic INU field map (denoted as “field_A”) for the T1-w imaging modality, generated using known spatial varying perturbation of the RF pulse flip angle (Kwan et al., 1999). This field has slowly varying and smooth spatial profile, consistent with intensity inhomogeneities that are typically observed with 1.5 T MR scanners (Figure 2.1A,B), and intensity values between 0.9 and 1.1 (corresponding to 20% spatial variation). Then, we generated other two intensity inhomogeneity fields, which are intended to better reproduce inhomogeneities from 3 T (Figure 2.1C,D) and 7 T (Figure 2.1E,F) MR scanners, respectively. As proposed by Vovk and co-workers,
the fields were created by cubic B-spline interpolation between equally spaced nodes at 60 and 40 mm in each direction (Vovk et al., 2004). These fields have an increased complexity in the spatial profile compared to the BrainWeb MRI Simulator field, reflecting intensity inhomogeneities that are typically generated by 3 T and 7 T MR scanners, respectively. Intensity values for the two additional INU fields were initially set to have 20% spatial variation. Thereafter, we generated fields also with 40%, 60% and 80% variation, by rescaling the image with 20% variation to have values ranging between 0.8 and 1.2, 0.7 and 1.3, 0.6 and 1.4, respectively. For the sake of simplicity, we will refer to the three INU fields as BIAS 1.5 T, BIAS 3 T and BIAS 7 T, respectively.

![Simulated INU fields. The simulated INU fields at 40% level are represented in coronal (y=4), axial (z=0), and sagittal (x=-11) sections for the BIAS 1.5 T (A), BIAS 3 T (C) and BIAS 7 T (E) profile. Histograms of the respective INU field are reported on the right side for the BIAS 1.5 T (B), BIAS 3 T (D) and BIAS 7 T (F) profile. It is worth noting that the INU field at 80% level is characterized by the same spatial profile of the one at 40% (range from 0.8 to 1.2), whereas the field values range from 0.6 to 1.4.](image)

Figure 2.1 – Simulated INU fields.

From the BrainWeb simulator we also extracted the phantom volume, which is a simulated MR image representing an anatomical model of a healthy brain. The phantom volume is created by combining ten three-dimensional "fuzzy" tissue membership volumes: gray matter (GM), white matter (WM),
cerebrospinal fluid (CSF), fat, muscle, skin, skull, glial matter, connective tissue, and background. In each tissue memberships volume, the value of each voxel represents the probability of the tissue to be found at that specific voxel. The MRI simulator combines the tissue membership volumes using weights estimated by Bloch equations (Kwan et al., 1999).

These weights are assigned by the simulator depending on the pulse sequence parameters chosen, and can reproduce MR image contrast in a realistic manner (Collins et al., 1998; Kwan et al., 1999). For our study, we used default settings of simulator parameters to generate an INU- and noise-free T1-w image (Figure 2A,D) in order to make our results comparable with previous studies on INU correction (Arnold et al., 2001; Ashburner and Friston, 2005; Sled et al., 1998; Tustison et al., 2010; Vovk et al., 2005, 2006; Ying et al., 2009). The image was obtained using Spoiled Fast Low Angle Shot (SFLASH) pulse sequence, with TR=18 ms, TE=10 ms and 30° flip angle. The image space was $181 \times 217 \times 181$ mm, with voxel sampling of 1mm isotropic.

After obtaining the INU- and noise-free T1-w image from the MRI simulator, we multiplied it with the INU field image to generate an INU-corrupted T1-w image. Finally, we also added Rician-distributed noise to the INU-corrupted image. Noise levels were set at 1%, 3%, 5% and 7% standard deviation compared to the intensity of the brightest tissue in the INU- and noise-free image (Figures 2.2 and 2.3).

---

**Figure 2.2** – Levels of Rician noise used in the simulations. Noise with Rician probability distribution is superimposed on an INU-free T1-weighted image. Different noise levels were used in our simulations. The level (0%, 1%, 3%, 5%, 7%) is quantified as the standard deviation of the noise distribution relative to the brightest tissue of the INU-free image.

---

**Figure 2.3** – Modelling of MR intensity inhomogeneities. The INU- and noise-free T1-w image (A) generated by the BrainWeb simulator is characterized by easily identifiable CSF, GM, and WM peaks in the intensity histogram (B). INU-corrupted T1-w images are generated by multiplying the INU- and noise-free T1-w image by the simulated BIAS 1.5 T (C), BIAS 3 T (G), BIAS 7 T (K) fields respectively, at 40% level (with values ranging from 0.8 to 1.2). This leads to a broadening of CSF, GM, and WM peaks in the respective intensity histograms (D,H,L). Rician noise at 3% relative intensity is added to the INU-corrupted T1-w image to generate a INU- and noise-corrupted image (E,I,M). The histograms corresponding to this image (F,J,N) are smoother and have larger CSF, GM, and WM peaks.
2.2.2.2 Method settings

We examined the performance of the four INU correction methods with a wide range of input parameters. Specifically, we were interested in comparing the results obtained with the parameters giving the most accurate estimate among all combinations of examined parameters (enhanced configuration) and those produced using standard input parameters (default configuration). The INU correction methods were characterized by different input parameters, which will be described in the following sections.

FSL, FS and BV are conceived to work on brain-extracted images. Conversely, SPM does not require any brain extraction, as it uses built-in probability maps that delimit the region to be processed. To generate a standard brain mask to be used in FSL, FS and BV, we calculated the union of thresholded GM, WM and CSF probability images \((p > 0.5)\), obtained from the MRI simulator. We also used the Brain Extraction Tool (BET) implemented in FSL \((Smith, 2002)\), to generate a set of brain masks with different extent, which was controlled by varying the fractional intensity parameter between 0.1 and 0.6. Values above 0.6 were excluded from the analysis due to heavy cortical erosion. In order to have comparable results, the masks were obtained running BET on the INU-free image and then applied to each method. The similarity of BET masks compared to the MRI simulator mask was assessed by quantifying the relative extent in terms of voxel number, as well its Dice similarity index \((Zou et al., 2004)\).

**Statistical Parametric Mapping**

The INU correction method of SPM has two input parameters: the *regularization* and the *bias field smoothing*. By tuning the *regularization*, the method may be more or less sensitive to sharp transitions between image structures. Higher values tend to be more suited in the presence of smooth transitions whereas low regularization values make the method more sensitive to high frequency patterns. The default regularization factor in SPM is equal to \(10^{-4}\). In our analyses on simulated data, we examine all the values implemented in the method: 0, \(10^{-5}\), \(10^{-4}\), \(10^{-3}\), \(10^{-2}\), \(10^{-1}\), 1, 10. The *bias field smoothing* permits to model the smoothness of the INU field. The numerical value to be set is the cut-off of DCT bases expressed in mm. Only DCT bases of periods longer than the cut-off are used to describe intensity inhomogeneities. In the presence of a very smooth INU field, if the estimated INU field is not forced to be smooth, then it will demonstrate higher intensity variation due to different tissue types rather than pure intensity inhomogeneity artifacts. The default cut-off in SPM is equal to 60 mm. For our investigations, we varied the *bias field smoothing* between 30 and 150 mm, at 10 mm intervals.

**FMRIB Software Library**

The INU correction method in FSL allows multiple user-adjustable parameters. Among them, we selected the two parameters that, according to the developers \(Zhang et al., 2001\), have the largest impact on the imaging results: the *bias field smoothing* and the algorithm *iterations*. The *bias field smoothing* parameter controls the level of low-pass filtering applied to the estimated INU field. The numerical value to be set is the Full-Width Half-Maximum (FWHM) in mm, which is supposedly larger in case of larger INU smoothness. FAST assumes a default value of 20 mm. In our study, we varied the FWHM from 5 to 50
mm, at 5 mm intervals. The accuracy of the INU field estimate is also characterized by the number of times the intensity inhomogeneity correction algorithm is iterated. By default, FAST implements 4 iterations. We run the FSL method setting this parameter to 4, 8, 16 and 32 iterations.

**FreeSurfer**

N3, the method included in FS, permits the selection of several parameters. Nonetheless, according to the developers (Sled et al., 1998) and as stated in subsequent studies (Boyes et al., 2008; Zheng et al., 2009), two of them are crucial for the intensity inhomogeneity estimate: the deconvolution kernel and the spline smoothing distance. Accordingly, we focused our investigations on these two parameters. The deconvolution kernel controls the width of the probability distribution of the expected INU field, expressed in terms of FWHM. N3 uses a default value of 0.15. In our study, the deconvolution kernel was varied between 0.05 and 0.5, with intervals of 0.05. The smoothing approach implemented in N3 is based on the approximation of data by a linear combination of smooth basis functions, specifically B-splines. The smoothness is determined by the spline smoothing distance in mm, which refers to the distance between basis functions. The default value in N3 is 200 mm. Accordingly, we varied the spline smoothing distance from 50 mm to 300 mm at 50 mm intervals. We also set the maximum number of iterations to 1000 and the stopping threshold (the coefficient of variation in the ratio between subsequent field estimates) to 0.0001 to support accuracy over speed, as in previous studies (Boyes et al., 2008; Zheng et al., 2009).

**BrainVoyager**

The BV method requires the selection of two input parameters, the polynomials order and the number of algorithm cycles of INU correction, which have a major impact on the intensity inhomogeneity detection (Dawant et al., 1993; Hou et al., 2006). Low order values help to model slowly varying INU profiles, while high orders tend to better describe sharp variations in the intensity inhomogeneity. The default order of polynomials is set to 3 and the number of cycles to 2. We examined the effect of a polynomials order between 1 and 7 and a number of cycles between 2 and 5, following the recommendation of the developers (Dawant et al., 1993; Hou et al., 2006).

### 2.2.2.3 Performance assessment

The performance of each algorithm was quantitatively evaluated on the estimated INU field, in line with previous studies (Arnold et al., 2001; Chua et al., 2009). To account for potential inconsistencies due to arbitrary scaling of the INU estimates, all the INU fields were normalized in intensity (Chua et al., 2009). Normalization was implemented by multiplying the estimated INU field by a scalar value $\omega$, according to the formula by (Chua et al., 2009) as follows:

$$\omega = \frac{\sum_{i=1}^{n} (b_{sim,i}b_{est,i})}{\sum_{i=1}^{n} (b_{sim,i})^2} \quad (2.2)$$
where $b_{\text{sim}}$ and $b_{\text{est}}$ are the simulated and the estimated INU fields, respectively, and $n$ is the number of brain voxels. The correspondence between the simulated and the estimated INU fields was then assessed by the root mean square error (RMSE) between the two images. The RMSE was defined as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (b_{\text{sim},i} - b_{\text{est},i})^2}{n}} \quad (2.3)$$

Being the RMSE a distance measure, the smallest RMSE indicated the best reconstruction performance. To complement the RMSE analysis, we also calculated spatial correlations between the simulated and estimated INU fields. Furthermore, the accuracy of the intensity inhomogeneity correction was assessed by visual inspection of the estimated INU fields and the relative histograms. Finally, we also estimated the effect of the INU correction on the actual images. To this end, we calculated the relative difference between the reconstructed and simulated T1-w images without INU contamination. The image $r(x,y,z)$ representing the relative difference between the reconstructed image $v(x,y,z)$ and the simulated image $v(x,y,z)$ is obtained by the following formula:

$$r(x,y,z) = \frac{v(x,y,z) - v(x,y,z)}{v(x,y,z)} \quad (2.4)$$

As a representative value for a region of interest (ROI), we defined the Mean Absolute Relative Error (MARE) as the average of the absolute relative difference $r$ across ROI voxels. Specifically, we evaluated the MARE value in the GM, WM, CSF, as well as in the whole brain, for different simulated INU fields and inhomogeneity correction methods.
Since INU correction algorithms are often used with standard input parameters, we initially examined the ability for each algorithm to detect intensity inhomogeneities according to the default and the enhanced configurations, respectively. To this end, we used the simulated T1-w image with INU 40% and noise 3% level. This analysis suggested that the default parameters do not always provide an estimate that is comparable to the one obtained using the parameters of the enhanced configuration (Figure 2.4).

For instance, the RMSE obtained with SPM with default settings was substantially larger than the minimal one for the BIAS 1.5 T profile (Figure 2.4A), whereas for the BIAS 3 T and BIAS 7 T profile the difference was less pronounced (Figure 2.4E,I). In the case of BV, the RMSE values in the default configuration were almost double than the ones obtained with enhanced input parameters regardless of the INU field profile (Figure 2.4D,H,I). For FS and FSL, the RMSE obtained with default configuration was similar to, but slightly larger than the one corresponding to the enhanced configuration (Figure 2.4C,G,J and 2.4D,H,I, respectively). Furthermore, the analysis of the complete set of RMSE values obtained with each method suggested that the results obtained with SPM and BV are largely sensitive to the input parameters used, whereas FSL and FS provide relatively stable results. The quantitative results of the RMSE analysis were confirmed by a qualitative comparison of the INU fields produced by the four methods under their default and enhanced configurations. In the BIAS 1.5 T simulation (Figure 2.5), we found that only FSL was able to reconstruct the INU field with relatively good accuracy when using default parameters. Much less accurate results were obtained with the other three methods, with BV showing the less reliable INU field estimate. When we examined the INU field maps obtained using the enhanced configurations, we found that the overall differences across algorithms were largely reduced. Among the four methods, BV was still characterized by a higher RMSE compared to FS, SPM and FSL. The spatial profile of the INU field provided by SPM substantially improved, along with the RMSE. When we selected input parameters based on the enhanced configuration, SPM and FSL converged to fairly similar INU field estimates.

Also in the BIAS 3 T simulation (Figure 2.6), SPM and FSL showed the best performance under default settings, while FS and BV were characterized by a less accurate reconstruction. When the enhanced configurations were selected, FS and FSL considerably improved their performance and showed similar results. Despite a slight improvement in the RMSE value, BV was still characterized by the poorest accuracy in the spatial profile. SPM showed the most accurate results, both in terms of RMSE and INU field profile.

The high INU field complexity in the BIAS 7 T simulation revealed a generally less inaccurate reconstruction of the spatial profile both in the default and enhanced configuration (Figure 2.7). When the enhanced parameter configuration was selected, all the four methods showed reduced RMSE values, with SPM yielding the smallest error.
Figure 2.4 – Dependence of INU field estimate on input parameters. In order to define the intensity inhomogeneity correction for each method and for each INU field (BIAS 1.5 T, BIAS 3 T and BIAS 7 T), we analyzed several parameter configurations. We computed the RMSE between the simulated and the estimated INU field for each configuration. The default (indicated with a cross marker) and the enhanced configuration (indicated with a diamond marker) are shown for SPM (A,E,I), FS (B,F,J), FSL (C,G,K) and BV (D,H,L). In SPM, the regularization and the bias field smoothing (FWHM) parameters were varied. The deconvolution kernel (FWHM) and the spline smoothing distance parameters were varied in FS. In contrast, the number of iterations and the bias field smoothing (FWHM) parameters were varied in FSL. Eventually, the polynomials order and the number of iterations parameters were varied in BV.

Figure 2.5 – Estimated BIAS 1.5 T INU field for default and enhanced parameter configurations. BIAS 1.5 T INU field estimated by FS, BV, SPM and FSL with default (left side) and enhanced parameter configurations (right side), respectively. RMSEs between the simulated and the estimated INU field are also indicated.
A scatterplot analysis conducted on INU values also revealed significant improvements using enhanced parameter configurations (Figure 2.8). Furthermore, the analysis of voxelwise correlations confirmed the results obtained using RMSEs in relation to the performance of the different INU correction methods. When we examined the histogram distributions of the estimated INU fields, we also observed that high RMSE (or low voxelwise correlation) values are not only explained in terms of a poor estimate in the spatial profile of the INU field, but also by an altered reconstruction of its values (Figure 2.9). Particularly, we noticed that FS, BV, and SPM under default configuration displayed INU values outside the range 0.8-1.2, which was the range of values of the simulated INU image. These broadened intensity distributions were still noticeable when considering enhanced configurations for FS and BV, but not for SPM and FSL.
It is also important to consider the effect produced by an imperfect INU field estimate on the actual image intensity. For this reason, we evaluated the difference between the simulated and reconstructed T1-w images, focusing on the results obtained by the methods with enhanced parameter configuration (Table 2.1). For this specific analysis and the subsequent ones, we focused on the enhanced parameter configuration of each method to exclude possible confounds associated with the choice of the default parameters. We obtained larger reconstruction errors for GM and CSF compared to WM, suggesting that the major discrepancy was at the outer edges of the brain, where the algorithms may tend to overcompensate/undercompensate the actual non-uniformity leading to spurious brightening/darkening. This finding was confirmed by close inspection of the relative error images. In both BIAS 1.5 T and BIAS 3 T simulations, FS and BV showed marked intensity variations between INU-
corrected and INU-free T1-w. FSL and SPM were characterized by variations of modest magnitude across the whole volume, with the latter method showing generally smaller errors. Also the BIAS 7 T simulation showed SPM to be the method with the lowest reconstruction errors across the whole brain, in line with the analyses conducted using RMSE (Figures 2.5 and 2.7) and spatial correlations (Figure 2.8). With the complex INU profile of this simulation, BV provided a lower reconstruction error than SPM in WM, but higher in GM, CSF, and overall in the full brain.

Since FS, BV and FSL require a brain mask to be given as input, we evaluated the impact of brain extraction on the INU field reconstruction provided by each of these three methods. This analysis indicated that the definition of an image to be used as spatial mask had a relative impact on the INU correction performance (Figure 2.10). The performance of FS were dependent on the extent of the brain mask, as defined by the fractional intensity value set in the Brain Extraction Toolbox (BET). As expected, the most accurate INU correction with the FS method corresponded with the most accurate brain masking (fractional intensity value equal to 0.4), as indicated by the highest Dice Similarity value.

Figure 2.8 – Scatterplot of INU fields obtained with default and enhanced parameter configurations. The scatterplot of the estimated INU field values for FS, BV, SPM and FSL with default and enhanced parameter configurations are illustrated in red and blue, respectively. The correspondence between the estimated and true INU fields was estimated using spatial correlations. The same analysis was performed for the BIAS 1.5 T (A), BIAS 3 T (B) and BIAS 7 T (C) simulations, respectively.
Chapter 2

Figure 2.9 – Histogram of INU fields obtained with default and enhanced parameter configurations. The histogram of estimated INU fields for FS, BV, SPM and FSL with default and enhanced parameter configurations are represented in red and blue, respectively. The histogram of the simulated INU field is also plotted for comparison in shaded grey. The same analysis was performed for the BIAS 1.5 T (A), BIAS 3 T (B) and BIAS 7 T (C) simulations, respectively.

Table 2.1 – Image reconstruction error for each of the four INU correction methods.

<table>
<thead>
<tr>
<th></th>
<th>WM</th>
<th>GM</th>
<th>CSF</th>
<th>FULL BRAIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIAS 1.5 T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS</td>
<td>1.33</td>
<td>1.58</td>
<td>2.35</td>
<td>1.66</td>
</tr>
<tr>
<td>BV</td>
<td>1.51</td>
<td>2.30</td>
<td>3.12</td>
<td>2.19</td>
</tr>
<tr>
<td>SPM</td>
<td>1.02</td>
<td>1.24</td>
<td>1.44</td>
<td>1.20</td>
</tr>
<tr>
<td>FSL</td>
<td>0.90</td>
<td>1.16</td>
<td>1.49</td>
<td>1.14</td>
</tr>
<tr>
<td>BIAS 3 T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS</td>
<td>2.45</td>
<td>2.92</td>
<td>4.37</td>
<td>3.06</td>
</tr>
<tr>
<td>BV</td>
<td>2.63</td>
<td>3.74</td>
<td>6.85</td>
<td>3.97</td>
</tr>
<tr>
<td>SPM</td>
<td>2.28</td>
<td>2.50</td>
<td>2.81</td>
<td>2.48</td>
</tr>
<tr>
<td>FSL</td>
<td>2.56</td>
<td>2.98</td>
<td>4.11</td>
<td>3.05</td>
</tr>
<tr>
<td>BIAS 7 T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS</td>
<td>3.48</td>
<td>4.56</td>
<td>6.08</td>
<td>4.51</td>
</tr>
<tr>
<td>BV</td>
<td>2.30</td>
<td>3.80</td>
<td>4.74</td>
<td>3.49</td>
</tr>
<tr>
<td>SPM</td>
<td>2.86</td>
<td>3.00</td>
<td>4.33</td>
<td>3.22</td>
</tr>
<tr>
<td>FSL</td>
<td>3.60</td>
<td>4.10</td>
<td>5.41</td>
<td>4.21</td>
</tr>
</tbody>
</table>

The mean absolute relative error (MARE) between the INU corrected T1-w image compared and the simulated T1-w image is shown for FS, BV, SPM and FSL with enhanced parameter configurations. The MARE was calculated for each of the three INU profiles using WM, GM, CSF and the full brain as regions of interest.
A different result was found for BV, which yielded relatively stable RMSE values for intermediate extents of the brain mask (fractional intensity between 0.2 and 0.5). On the other hand, BV was characterized by high RMSE with the most conservative and the most extreme brain masking (fractional intensity equal to 0.1 and 0.6, respectively). Also FSL yielded relatively stable RMSE values, discounting the case of most conservative brain masking. Overall, the RMSE values obtained for FSL were substantially lower than for BV and FS.

![Figure 2.10](image)

*Figure 2.10* – Dependence of the INU correction on brain masking. The impact of the brain extraction on the INU field estimate is shown for FS, BV and FSL only, since SPM does not need the specification of an explicit brain mask. Brain masking was performed using the Brain Extraction Tool (BET), using different fractional intensity values as input parameter. To allow the comparability of the results, the RMSE was calculated only for voxels in the intersection volume of the different masks obtained by BET. The RMSE values for INU estimates by FS, BV and FSL, with different fractional intensity as input for BET, are represented in a bar plot. Average values among the BIAS 1.5 T, BIAS 3 T and BIAS 7 T simulations are reported. Dashed lines represent the RMSE values for INU estimates using a standard mask (obtained from the MRI simulator), which are provided for comparison. Volumetric variations in the brain-extracted volume (dV) with respect to the standard mask are indicated over the bar plots, along with their Dice similarity (DS). The RMSE values are computed averaging the results of the three INU field profiles.

Finally, we evaluated the performance of the four methods for different levels of INU field magnitude and image noise (Figure 2.11). This analysis revealed, as expected, an increased INU field magnitude and/or an increase image noise level generally yielded higher RMSE for all methods. Furthermore, all INU correction methods were more effective with a slowly varying INU field (BIAS 1.5 T) than one with complex profile (BIAS 7 T). SPM showed remarkable stability and accuracy for different noise levels and INU magnitudes, regardless of the INU field profile. FSL had comparable RMSE values to SPM for the BIAS 1.5 T field, irrespective of noise level and INU field complexity, whereas it was relatively less accurate with the BIAS 3 T and BIAS 7 T profiles. FS and BV generally underperformed
the other two methods, and showed specific features related to the sensitivity to the two INU field and noise. Specifically, noise was the primary responsible for a reduction in the performance in both methods. Though, BV proved to be much more sensitive to the INU field magnitude than FS.

Figure 2.11 – Sensitivity of INU correction methods to inhomogeneity magnitude and noise. FS, BV, SPM and FSL were compared in the enhanced parameter configurations for the BIAS 1.5 T (A), BIAS 3 T (B) and BIAS 7 T (C) profile. We calculated RMSEs between the simulated and the estimated INU fields for different INU field magnitudes and noise levels. The image with 40% INU and 3% noise levels corresponds to the one used for previous analyses.
2.4 Discussion

The correction of INU in MR images is extremely important to ensure the reliability of investigations on brain structure. For instance, inadequate correction of the INU leads to decreased stability of automated segmentation algorithms (Clarke et al., 1995; Dawant et al., 1993; Pham and Prince, 1999; Zheng et al., 2009) and this may in turn yield false positive and false negatives in voxel-based morphometry (VBM) studies (Ashburner and Friston, 2000; Good et al., 2001; Imabayashi et al., 2013). Here, we tackled the problem of assessing the performance of INU correction techniques, and in so doing we focused on the methods implemented in four widely used software for MR data analysis: FS, BV, SPM and FSL. Unlike several previous studies (Arnold et al., 2001; Boyes et al., 2008; Gispert et al., 2004; Zheng et al., 2009), we conducted our analyses on simulated MR images as this allowed us to assess INU field estimates against ground truth images. Our choice not to use actual MR images is due to the lack of validation studies and pending disputes regarding the reliability of performance evaluation measures (Chua et al., 2009). It should be considered, however, that simulated images may not account for scanner-specific variables and that representing the brain in terms of few tissue classes may not truly replicate the anatomical variation of a real brain. These potential limitations can be partly levied by adding random noise to simulated MR images, as we did in our study. Indeed, noise may be cautiously seen as “pseudo-anatomy” variations within tissues. We believe that our study has provided valuable information with respect to four important aspects, which will be further discussed in the next sections. First, we found a large variability in the solutions provided by the INU correction algorithms. In particular, default parameter configurations did not always provide results sufficiently close to the results obtained with enhanced parameter configurations. Performance comparisons between different methods can be considered more objective if conducted using their enhanced configurations. Second, all methods under investigation are designed to remove slowly varying intensity inhomogeneities, and are relatively less effective with the more complex INU profiles that characterize high-field MR scanners. Third, the brain mask that needs to be given as input to FS, BV and FSL influences their performance. In contrast, SPM does not require the specification of any input mask. Fourth, approaches that integrate INU correction and brain segmentation, such as SPM and FSL, outperform methods dedicated to INU correction only, such as FS and BV, for different levels of INU magnitude and noise. Nonetheless, relatively accurate INU field reconstructions can be obtained with FS on MR images with low noise and with BV when the inhomogeneity magnitude is limited.

2.4.1 Default and enhanced parameter configurations

A large number of previous studies have compared the performance of several INU correction methods using default parameters (Arnold et al., 2001; Likar et al., 2001; Vovk et al., 2006). While those studies provided valuable information about the features of each method, it should be also considered that an objective comparison between methods could be best accomplished by using optimal input parameters.
To get as close as possible to this condition, we examined a very large set of parameters and focused on the configuration with the relative minimal reconstruction error, which we denoted as enhanced parameter configuration. Notably, we carried out the comparison among INU correction methods on simulated data, since so far no approach to reliably measure INU correction performance exists for actual MR data (Chua et al., 2009). The identification of optimal set of input parameters for actual MR images would be extremely valuable, as it could increase the reliability of structural imaging analyses, and future work is warranted to develop an automated tool capable of performing such identification.

Among all algorithms under investigation, FSL was the only one that provided limited differences between the default and enhanced configurations. This suggests that the FSL method may be preferred in case of uncertainty about the selection of the adequate input parameters for INU correction. Conversely, we observed differences in RMSE values between default and enhanced configurations for SPM and FS, and partially also for BV. This can be explained by a substantial sensitivity of the INU reconstruction with respect to the input parameters used, as revealed by the analysis of RMSE values obtained with different combinations of input values (Figure 2.4). For both SPM and FS, we also noticed that similar INU profiles could be estimated by simultaneously changing more than one input parameter. This can be the case when multiple parameters control the smoothness of the estimated INU. For instance, a similar RMSE could be obtained in SPM by increasing/reducing at the same time both the regularization and the smoothing factor (Figure 2.4A,E,I). By close inspection of the RMSEs obtained with different combinations of input values (Figure 2.4), we also noticed the presence of multiple local minima, which may be indicative of a complex pattern of interactions between the MR image to be corrected and the estimated INU field. This is not surprising, since specific MR image features may be detected as part of the intensity inhomogeneity depending on the algorithm parameters used.

It has been suggested that default input parameters for each INU correction method do not depend only on the method itself, but need to be identified also on the basis of the INU field profile, image quality (i.e. the signal-to-noise ratio) and the magnitude of the INU field (Arnold et al., 2001; Madabhushi and Udupa, 2005). Zheng and coworkers highlighted the fact that most of the INU-correction methods were developed over a decade ago, and for this reason they are optimized to work well with low-field scanners only (Zheng et al., 2009). Likewise, other authors suggested specific input parameters should be properly selected to obtain accurate estimates (Boyes et al., 2008; Chua et al., 2009; Weiskopf et al., 2011). From this standpoint, the main elements to be considered are the strength of the static field and the geometry of the receiver coils (Boyes et al., 2008), as these influence also tissue-induced inhomogeneities. The current efforts put in the development of high-field MR scanners (Moser et al., 2012; Umutlu et al., 2014) suggest that parameter selection will become more relevant in the future, as the structural images will be more affected by the INU (Bernstein et al., 2006; Mihara et al., 2005; Uwano et al., 2014; Van De Moortele et al., 2005).
2.4.2 Dependence of the correction performance to the INU spatial profile

It is a matter of fact that an increase of the static magnetic field has a significant impact on the spatial profile of the intensity inhomogeneity (Vaughan et al., 2001). Although several factors may influence the features of these non-anatomical variations, the magnetic field strength is the most important one (Belaroussi et al., 2006). MR images collected with 1.5 T scanners are characterized by a very smooth profile (Figure 2.1A). On the other hand, localized intensity non-uniformity is observed in images acquired using 3 T and 7 T magnetic fields (Figure 2.1B,C), particularly in correspondence of outer brain structures as well as central regions. Furthermore, the spatial variations of the INU profile at 7 T are substantially larger than at 3 T (Bernstein et al., 2006; Collins et al., 2005; Umutlu et al., 2014). In this study, we considered three different INU profiles (referred to as BIAS 1.5 T, BIAS 3 T and BIAS 7 T), each representative of one of these three magnetic field strengths.

As we expected, all INU correction methods were substantially less effective with the BIAS 7 T profile compared to the BIAS 1.5 T and BIAS 3 T ones. When enhanced input configurations were considered, the SPM method proved to be the least affected by the fast spatial variations of the BIAS 7 T inhomogeneity. We argue that the use of a DCT basis functions may help to select the most important spatial frequencies for the INU reconstruction, and that this may be an important advantage of the SPM method compared to other ones. Notably, the FSL method provided accurate estimates with slowly varying INU fields, with performance largely similar to the SPM method (Figures 2.5 and 2.6). On the other hand, the BIAS 7 T simulation clearly revealed a less accurate INU reconstruction with FSL compared to SPM (Figure 2.7). A possible explanation may be in the different implementation of the bias field smoothing, which is not related to the use of DCT basis functions. As an alternative, reconstruction performance in FSL may be lower than in SPM, because FSL does not model the intensity distribution from different tissue types (e.g., GM, WM, CSF), which may be important to effectively discriminate fast spatial variations in the INU from normal intensity variations between brain tissues. As for the FS method, the input parameters that provided best performance were similar for the three INU profiles (Figure 2.4B,F,J), suggesting a limited capability of adapting to an increased spatial complexity of the INU with higher magnetic field strength. This might be due to the fact that the deconvolution of narrow Gaussian distributions implemented in FS cannot easily discriminate the effects of noise and of high INU frequencies in the MR image. Similarly, BV showed a clear sensitivity to fast spatial variations that characterize the BIAS 7 T profile. The INU correction implemented in BV only relies on intensities of selected WM voxels (Hou et al., 2006). As such, the BV method may be less accurate when the spatial inhomogeneity does not vary slowly, as this condition evidently increases the uncertainty in the INU estimate far from the sampled WM voxels.
2.4.3 Effect of brain extraction on INU correction

Brain extraction is an important pre-processing step in brain imaging analysis. It is typically used before inhomogeneity correction, and may therefore influence its performance. We therefore assessed whether or not brain masking has an impact on the accuracy of INU reconstruction. This analysis was conducted on the methods implemented in FS, BV and FSL, since SPM requires no brain mask to be given as input. In line with our expectations, the brain extraction analysis revealed a substantial sensitivity of the INU field estimate with respect to the mask extent (Figure 2.10). As for the BV method, we found very low RMSE variability using masks generated by BET with intermediate fractional intensity values. On the other hand, the poor accuracy shown with a conservative brain extraction (mask volume +52% compared to that of the standard mask) as well as with a severe masking (mask volume -4.6% compared to that of the standard mask) yielded to the conclusion that BV is particularly sensitive to the brain extraction process. A similar result was observed for the method implemented in FSL. In particular, the high RMSE value obtained with a very conservative masking suggested that brain extraction is fundamental to obtain a reliable INU correction. This is consistent with what suggested by the developers of FSL (Zhang et al., 2001). The FS method showed the highest dependence of the INU reconstruction for intermediate fractional intensity values, as compared to the other BV and FSL. Furthermore, our findings for FS are in agreement with those of Boyes and coworkers, who suggested that the inclusion of non-brain structures as well as parts of the background yields poor INU estimates (Boyes et al., 2008).

2.4.4 Sensitivity to INU magnitude and noise level

While we initially evaluated the methods at a single, realistic level of INU magnitude and noise, we then varied these two parameters to gain insights into the characteristics of each method. The two segmentation-based methods, that are the ones in SPM and FSL, proved generally superior to those in FS and BV (Figure 2.11). This may be explained by the fact that the combination of INU correction and brain segmentation within the same framework yields advantages for both processing procedures (Ashburner and Friston, 2005; Zhang et al., 2001). SPM and FSL can generally achieve good results for different noise levels and INU field magnitudes, with SPM being slightly more accurate for low levels of noise or magnitudes of the INU field. SPM makes use of information in template images for GM, WM and CSF. Essentially, these volumes encode the probability of finding different tissues at each spatial location. Notably, the use of tissue probability maps may have the favorable effect of modeling real intensity contrasts between different tissue types (Uwano et al., 2014). The INU correction approach of FSL, in turn, does not make use of prior information from a template and implements spatial encoding through a HMRF model. This means that also FSL uses contextual intensity information, but only with a spatial extent of few neighboring voxels. This approach may be particularly effective in the presence of intermediate levels of noise and INU field magnitudes. However, at very high levels of noise and INU field magnitude, FSL yielded larger RMSE values compared to SPM (Figure 2.11B,C).
One of the main features that allowed N3, the method integrated in the FS pipeline, to become a method widely used by the neuroimaging community is its limited range of assumptions. It is indeed a histogram-based method that does not require any explicit model of the intensity distribution. Due to the intrinsic INU fields unpredictability, the need of no prior information on intensity properties is a noteworthy aspect. Intensity distributions in pathological cases are not known a priori, thus a model-independent assumption may be beneficial. According to our results (Figure 2.11), N3 provides good estimates in presence of moderate INU field magnitude and noise levels, but is much less effective at higher values (i.e., 80% INU level and 7% noise). In heavily corrupted images, restoring high frequency patterns by deconvolving narrow Gaussian distributions from the MR image becomes trivial. As for BV, the sensitivity to noise and INU magnitude becomes more important. The BV method is based on the idea that intensity variations within WM voxels may be used to extract INU information. Partial volume effects introduced by noise may hamper an accurate estimate of intensity inhomogeneity (Hou et al., 2006). Since this method can only make use of a limited subgroup of samples, the intensity normalization may be less accurate in areas where WM concentration decreases, as for example in the inferior part of the brain and the cerebellum (Figures 2.5, 2.6 and 2.7). By the same token, high noise levels also affect BV performance. Since polynomial fitting is used to estimate the INU, the presence of noise may inevitably lead to a less accurate inhomogeneity correction.

2.5 Conclusion

We have conducted a comprehensive assessment of INU correction methods for structural MR brain images. Specifically, we have generated structural images that accurately mimic those typically collected using 1.5 T, 3 T and 7 T MR scanners, respectively. Accordingly, we have modeled intensity inhomogeneities with spatial profiles characterized by increasing complexity levels. It is worth noting that INU correction methods generally assume the intensity inhomogeneity to slowly vary across voxels. Such an assumption may not hold anymore with high-field scanners, for which the INU field variations are comparable to the dimension of the human brain structures (de Graaf et al., 2012). Our findings confirmed a generally worse INU correction for high-field MR images, suggesting that further work is warranted for the development of inhomogeneity correction methods that are effective also with complex INU profiles. Another important element to consider is that selection of valid input parameters of a given INU correction method may be better conducted by taking into account the characteristics of the MR image. Our findings provide a valuable basis for the selection of the INU correction method to be used. On the other hand, our study does not address the question of which input parameters should be used with actual MR data. Our future work will be focused on the development of a dedicated software tool capable of identifying, for any INU correction method, the parameter configuration that is the most appropriate for a given MR image. Such a tool might have a profound impact on the reliability of structural neuroimaging investigations.
Intensity non-uniformity (INU) in magnetic resonance (MR) imaging is a major issue when conducting analyses of brain structural properties. An inaccurate INU correction may result in qualitative and quantitative misinterpretations. Several INU correction methods exist, whose performance largely depend on the specific parameter settings that need to be chosen by the user. Here we addressed the question of how to select the best input parameters for a specific INU correction algorithm. Our investigation was based on the INU correction algorithm implemented in SPM, but this can be in principle extended to any other algorithm requiring the selection of input parameters. We conducted a comprehensive comparison of indirect metrics for the assessment of INU correction performance, namely the coefficient of variation of white matter (CV_{WM}), the coefficient of variation of gray matter (CV_{GM}), and the coefficient of joint variation between white matter and gray matter (CJV). Using simulated MR data, we observed the CJV to be more accurate than CV_{WM} and CV_{GM}, provided that the noise level in the INU-corrected image was controlled by means of spatial smoothing. Based on the CJV, we developed a data-driven approach for selecting INU correction parameters, which could effectively work on actual MR images. To this end, we implemented an enhanced procedure for the definition of white and gray matter masks, based on which the CJV was calculated. Our approach was validated using actual T1-weighted images collected with 1.5 T, 3 T and 7 T MR scanners. We found that our procedure can reliably assist the selection of valid INU correction algorithm parameters, thereby contributing to an enhanced inhomogeneity correction in MR images.

Published as:
Chapter 3

3.1 Introduction

Magnetic Resonance Imaging (MRI) is a technique that delivers detailed images of the human body by analyzing its interactions with radio waves superimposed on a strong magnetic field. Due to high spatial resolution and imaging contrast, MRI has achieved a widespread use in clinical brain imaging. Indeed, it is regularly utilized for the detection of structural changes driven by trauma (Irimia et al., 2012; Sappenfield and Martz, 2013), neurodegenerative disease (Canu et al., 2011; Tillema and Pirko, 2013) and neuropsychiatric disorders (Buchanan et al., 2004; Pomarol-Clotet et al., 2010).

A major drawback in the quantitative as well as qualitative interpretation of structural MR images arises from the presence of artifactual smooth intensity variations across the whole MR image (Belaroussi et al., 2006; Bernstein et al., 2006). These are commonly referred to as intensity non-uniformity (INU), but also intensity inhomogeneity or spatial bias. According to the radio frequency (RF) field mapping theory, intensity inhomogeneities in MR images can be modeled as multiplicative (Insko and Bolinger, 1993; Stollberger and Wach, 1996). The main factors that can influence the magnitude and spatial profile of the INU include: static field strength, reduced RF coil uniformity, RF penetration, gradient-driven eddy currents, inhomogeneous reception sensitivity profile, and overall subject anatomy and position (Belaroussi et al., 2006; Mihara et al., 1998; Simmons et al., 1994; Vovk et al., 2007). To address this problem, INU correction methods that rely on image features to remove spatial inhomogeneities of different sources have been widely employed by the neuroimaging community (Arnold et al., 2001; Belaroussi et al., 2006; Boyes et al., 2008; Uwano et al., 2014; Vovk et al., 2007; Weiskopf et al., 2011; Zheng et al., 2009).

Since an effective INU correction is critical for investigations of brain structure, previous studies have attempted to compare the performance of several retrospective methods (Arnold et al., 2001; Likar et al., 2001; Velthuizen et al., 1998; Vovk et al., 2006). In the vast majority of studies, INU correction is performed using default parameters. Nonetheless, it is a matter of fact that each method performs better or worse depending on the specific settings used (Boyes et al., 2008; Uwano et al., 2014; Weiskopf et al., 2011; Zheng et al., 2009), and the default configuration may provide in some cases much less accurate results than other ones (Ganzetti et al., 2015). For instance, the definition of optimized parameters is particularly important for the INU correction algorithm implemented in SPM, which is one of the most widely used software for MR data analysis (Ashburner and Friston, 2005). Notably, since the INU correction in SPM is integrated within the brain segmentation tool, an inadequate removal of the INU directly affects the estimate of GM and WM maps (Clarke et al., 1995; Dawant et al., 1993; Pham and Prince, 1999; Zheng et al., 2009). It should be considered that the definition of the best set of parameters for the INU correction algorithm in SPM, as well as for any other alternative INU correction algorithm, is still an unsolved issue.

The optimal set of INU correction parameters can be easily identified on simulated data, for which a direct comparison between true and estimated INU fields is possible. In this case, the correspondence with a ground truth image may be assessed by correlation (Arnold et al., 2001), root mean...
square error (Ganzetti et al., 2015), L2-norm (Chua et al., 2009), and voxel-wise distance (Weiskopf et al., 2011). On the other hand, the use of indirect evaluation metrics, which do not require any reference image, is the only option for actual MR data. Popular indirect measures are based on intensity variability, such as the coefficient of variation of white matter (CV_{WM}), the coefficient of variation of gray matter (CV_{GM}), and the coefficient of joint variation between white matter and gray matter (CJV). A common premise about the spatial intensity distribution in MR images is that the gray scale distribution of white matter (WM) and gray matter (GM) is somehow defined. Hence, an effective INU correction should theoretically restore the original intensity distribution amplitude, which was altered by the inhomogeneity field. The distribution variability within WM and GM tissues can be separately quantified by CV_{WM} and CV_{GM}, respectively. CJV does not only quantify the intensity variability in both WM and GM but also accounts for the overlap between their distributions.

In this study, we evaluate to what extent and how indirect metrics can assist the selection of optimal input parameters for a given INU correction algorithm. We conduct our investigation using the INU correction algorithm implemented in SPM12 (Wellcome Trust Centre for Neuroimaging, University College London), the results of which are particularly sensitive to the selected input parameters (Ganzetti et al., 2015). We focus on T1-weighted images, which are the most commonly used images to investigate brain structure, and the ones typically affected by the INU. We generate simulated MR images with INU fields at different magnitudes and with different image noise levels to define a suitable approach for the detection of algorithm input parameters. Therefore, using the same simulated data, we evaluate the relation between direct and indirect metrics in terms of image quality. After defining an optimized strategy to define INU correction parameters based on an indirect metric, we validate it using actual MR images with different INU spatial profiles and magnitudes.
3.2 Methods

3.2.1 Description of the data-driven approach

Our data-driven approach to define optimal parameters for INU correction (see Figure 3.1) requires a raw (unprocessed) structural MR image as input. After defining the whole set of possible INU correction parameters to be examined (parameter space), INU correction and image segmentation are run for each combination of parameters. For each of these runs, the INU-corrected image is spatially smoothed to mitigate the negative effects of noise. In parallel to this, all the gray matter (GM) and white matter (WM) images produced by the image segmentation are processed to derive optimized subject-specific GM and WM masks. After selecting a metric among CV\textsubscript{WM}, CV\textsubscript{GM} and CJV, the INU correction performance is estimated for each combination of input parameters on the basis of the smoothed and INU-corrected MR image and the subject-specific GM and WM masks. A search for the minimum metric value is conducted, leading to the selection of the set of INU correction parameters putatively yielding the best performance. The software implementing this data-driven approach described above is freely available, and can be found at www.bindgroup.eu/index.php/software.

3.2.1.1 Definition of the INU correction parameter space

The INU correction parameters depend on the specific INU correction algorithm chosen. In this study, we tested our approach with the INU correction method implemented in SPM12 (www.fil.ion.ucl.ac.uk/spm). This is incorporated within the unified segmentation module (Ashburner and Friston, 2005) and integrated within the ‘Segmentation’ toolbox. The INU correction algorithm is based on two parameters: the regularization and the bias field smoothing. By decreasing/increasing the regularization, the method may be more/less sensitive to sharp intensity transitions between image structures, whereas the bias field smoothing permits to model the smoothness of the INU field. For our investigations, we run INU correction on the same image using regularization values \((0, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10)\) and bias field smoothing values (between 30 and 150 mm, sampled at 10 mm intervals) that spanned the whole range suggested by the developers. For a description of the INU correction algorithm and a detailed analysis of its performance, please refer to (Ganzetti et al., 2015).
Chapter 3

3.2.1.2 Selection of the indirect metric

Our approach requires the selection of a metric among CV WM, CV GM and CJV to indirectly estimate INU correction performance. These three metrics measure different properties of the image histogram, and are widely used to evaluate to what extent intensity inhomogeneities affect the MR image. They are defined as follows:

\[
CV_{WM} = \frac{\sigma(WM)}{\mu(WM)}, \quad CV_{GM} = \frac{\sigma(GM)}{\mu(GM)}, \quad CJV = \frac{\sigma(WM)+\sigma(GM)}{|\mu(WM)-\mu(GM)|}
\]  

(3.1)

where \(\sigma\) and \(\mu\) indicate the standard deviation and the mean intensity of a given tissue class, respectively. It is commonly accepted that relatively low values of these metrics correspond to smaller presence of INU field and hence better correction performance (Chua et al., 2009).

3.2.1.3 Definition of optimal smoothing level

A drawback concerning the use of CV WM, CV GM and CJV is that their values are sensitive to image noise (Chua et al., 2009). Accordingly, the presence of noise in actual MR data limits their reliability when evaluating the INU correction effectiveness. To address this problem, we integrated in our approach spatial smoothing on the INU-corrected MR image. We used the smoothing algorithm implemented in SPM12, and we set the Gaussian smoothing kernel to have full-width at half maximum (FWHM) equal to or smaller than 3 mm in order to avoid excessive image blurring and limit partial volume effects.
3.2.1.4 Definition of image-specific GM and WM masks

A key aspect that hampers an effective use of CV_WM, CV_GM and CJV for real MR data is the fact that optimized masks for WM and GM are not accessible, and that those generated from population-specific templates may not be sufficiently accurate to ensure reliability. Hence, we developed a procedure to address also this problem. For each parameter configuration, the WM and GM probability maps produced by the ‘Segmentation’ toolbox were registered to the SPM template in MNI space, using the deformation field generated by the toolbox itself. Afterwards, we binarized the WM and GM probability maps registered to MNI space using a threshold equal to 0.9 to minimize the contaminating effect of partial volume voxels. For each parameter configuration, we calculated the Dice Similarity Index (DSI) between the registered and the SPM template masks for both WM and GM (Zou et al., 2004). The mean DSI (mDSI) for each parameter configuration was computed by averaging the two DSI values for WM and GM, respectively. After estimating the mDSI for each parameter configuration, we selected a relative amount of configurations (called R_T hereinafter) that were characterized by the highest mDSIs. For both WM and GM, the probability maps belonging to the selected configurations were averaged together, and the average probability map was thresholded at 0.9 to generate a representative mask. We examined the mDSI of the representative WM and GM masks, obtained for R_T ranging from 50% to 100% at intervals of 5%. Thus, using the simulated data, we identified the R_T value yielding the maximum mDSI value, and consistently used it in subsequent analyses on actual MR data. As such, representative WM and GM mask obtained with the identified R_T value were considered optimized masks, and employed for the calculation of indirect metrics.

3.2.1.5 Identification of the optimal set of INU correction parameters

Rather than implementing an iterative algorithm for the determination of the optimal set of INU correction parameters, we opted for a search across the whole space of possible combinations. In first instance, this choice can be justified by the limited problem size, but also by the simplicity of implementation. A number of INU correction algorithms, for instance SPM, typically show relatively similar performance between parameters configurations that are close in the parameter space. These algorithms are therefore suited for the implementation of an iterative search algorithm, which tries to identify a gradient that leads to the configuration with minimum metric value. Nonetheless, there are algorithms, as for example the one implemented in BrainVoyager (www.brainvoyager.com), for which parameter configurations that are close in the parameter space may have very different accuracy (Ganzetti et al., 2015). The implementation of a search across the whole space of possible combinations may permit to effectively use our data-driven approach with any INU correction algorithm.
3.2.2 Performance analysis

3.2.2.1 Testing on simulated MR data

Creation of simulated MR images

Simulated MR data were obtained from the BrainWeb MRI Simulator (brainweb.bic.mni.mcgill.ca/brainweb). First of all, we extracted a realistic INU field map for the T1-w imaging modality, simulated using known spatial varying perturbation of the RF pulse flip angle (Kwan et al., 1999). This map has a smooth spatial profile, reflecting intensity inhomogeneities that are typically observed with lower magnetic field systems, e.g. 1.5 and 3 T MR scanners. The MRI simulator provides an INU field with 20% spatial variation (intensity values between 0.9 and 1.1). For our study, we also generated INU fields with 40% and 80% variation by rescaling the INU profile from the simulator to have values ranging between 0.8 and 1.2 and between 0.6 and 1.4, respectively (Figure 3.2A).

In order to generalize our results, we generated an additional intensity inhomogeneity field, characterized by higher dynamics. This profile is intended to mimic better inhomogeneities from higher field scanners. As proposed by Vovk and co-workers, the field was created by cubic B-spline interpolation between equally spaced nodes at 40 voxels in each direction (Vovk et al., 2004). Node values, also defined as multiplication factors, were randomly distributed between the same intervals adopted in the previous field (Figure 3.2B).

![Figure 3.2 – Simulated INU fields. Spatial profiles and histograms of the low-dynamic (A-B) and the high-dynamic INU fields (C-D) at 20% level are represented. Both INU fields are displayed in coronal (y=1), axial (z=0), and sagittal (x=15) sections. It is worth noting that the INU fields at 40% and 80% level are characterized by the same spatial profile of the one at 20%, whereas the field values range from 0.8 to 1.2 and from 0.6 to 1.4, respectively.](image-url)
From the BrainWeb MRI simulator we also extracted the phantom volume, which is a simulated MR image representing an anatomical model of a healthy brain. The phantom volume is created by combining ten three-dimensional "fuzzy" tissue membership volumes: GM, WM, cerebrospinal fluid, fat, muscle, skin, skull, glial matter, connective tissue, and background. In each tissue memberships volume, the value of each voxel represents the probability of the tissue to be found at that specific voxel. The MRI simulator combines the tissue membership volumes using weights estimated by Bloch equations (Kwan et al., 1999). These weights are assigned by the simulator depending on the pulse sequence parameters chosen, and can reproduce MR image contrast in a realistic manner (Collins et al., 1998; Kwan et al., 1999). We used default settings of simulator parameters to generate an INU- and noise-free T1-weighted image in order to make our results comparable with previous studies on INU correction (Arnold et al., 2001; Ashburner and Friston, 2005; Ganzetti et al., 2015; Sled et al., 1998; Tustison et al., 2010; Vovk et al., 2005, 2006; Ying et al., 2009). The image was obtained using Spoiled Fast Low Angle Shot (SFLASH) pulse sequence, with TR=18 ms, TE=10 ms and $\alpha=30^\circ$. The image space was $181 \times 217 \times 181$ mm, with voxel sampling of 1 mm isotropic. After obtaining the INU field and the INU- and noise-free T1-weighted image from the MRI simulator, we multiplied these to generate an INU-corrupted T1-weighted image. Finally, we also added Rician-distributed noise to the INU-corrupted image. Noise levels were set at 1%, 3%, and 5% standard deviation compared to the intensity of the brightest tissue in the unbiased image.

Performance analysis on simulated data

First, we evaluated the CVWM, CVGM and CJV in the identification of the optimized parameter configuration using simulated data with different INU magnitude and noise level. To this end, we adopted WM and GM probability maps provided by the MRI simulator. Before extracting tissue distributions, we thresholded each map at 0.9, in order to control for partial volume effects (Chua et al., 2009). Afterwards, $\sigma$ and $\mu$ were computed for both tissues. Finally, we assessed the performance of CVWM, CVGM and CJV, at different levels of noise and INU magnitudes.

The direct performance was quantitatively evaluated on the estimated INU field, rather than on the INU-corrected images. In this way, we examined the INU correction results without our performance measures being directly affected by the noise added to the MR images. To account for potential inconsistencies due to arbitrary scaling of the INU estimates, all the INU fields were normalized in intensity (Chua et al., 2009). Normalization was implemented by multiplying the estimated INU field by a scalar value $\omega$, according to the formula by (Chua et al., 2009) as follow:

$$\omega = \frac{\sum_{i=1}^{n} (b_{\text{sim},i} b_{\text{est},i})}{\sum_{i=1}^{n} (b_{\text{sim},i})^2}$$  \hspace{1cm} (3.2)

where $b_{\text{sim}}$ and $b_{\text{est}}$ are the simulated and the estimated INU fields, respectively, and $n$ is the number of brain voxels.
The deviation (D) of the simulated from the estimated INU fields was then assessed by computing the median of the brain-voxel values in the image T, defined as:

\[ T = \frac{2|b_{\text{sim}} - b_{\text{est}}|}{b_{\text{sim}} + b_{\text{est}}} \tag{3.3} \]

The smallest D value was associated with the best reconstruction performance (Weiskopf et al., 2011). To assess the reliability of the information extracted from the indirect metrics, we used two primary indices: 1) the D value obtained for the input parameter configuration providing the lowest metrics value; 2) the Spearman's correlation coefficient between the matrix of metrics values obtained for all parameter configurations and the corresponding matrix of absolute distances D (matrix-to-matrix correlation, MMC).

### 3.2.2.2 Testing on actual MR images

**Actual MR images**

To validate the proposed method, we used T1-w images from three publicly available datasets, acquired at different magnetic field strength in healthy volunteers. The first was the IXI database of the Imperial College London (biomedic.doc.ic.ac.uk/brain-development/index.php?n=Main.Datasets), the second was the KIRBY21 database of the Kirby Research Center for Functional Brain Imaging in Baltimore (mri.kennedykrieger.org/databases.html). This dataset contained images collected in 21 subjects during two different sessions (Landman et al., 2011), which were used in this study for a test-retest analysis. The third dataset, contributed by Dr. Bennett Landman from the Vanderbilt University, was downloaded from the NITRC neuroimaging data repository (www.nitrc.org/frs/shownotes.php?release_id=2178). Details on scanning parameters for the different datasets are provided in Table 3.1.

**Table 3.1 – Real data: MR imaging sequence parameters.**

<table>
<thead>
<tr>
<th></th>
<th>IXI</th>
<th>KIRBY21</th>
<th>NITRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanner</td>
<td>Gyroscan Intera, Philips</td>
<td>Achieva, Philips</td>
<td>Achieva, Philips</td>
</tr>
<tr>
<td>Magnetic field (Tesla)</td>
<td>1.5</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Pulse Sequence</td>
<td>MPRAGE</td>
<td>MPRAGE</td>
<td>3D TFE</td>
</tr>
<tr>
<td>Coil</td>
<td>Standard TMJ coil</td>
<td>8-channel phased array head coil</td>
<td>16-channel phased array head coil</td>
</tr>
<tr>
<td>TR (ms)</td>
<td>9.8</td>
<td>6.7</td>
<td>5.5</td>
</tr>
<tr>
<td>TE (ms)</td>
<td>4.6</td>
<td>3.1</td>
<td>2.6</td>
</tr>
<tr>
<td>Flip angle (degrees)</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>In-plane resolution (mm×mm)</td>
<td>0.94×0.94</td>
<td>1×1</td>
<td>0.7×0.7</td>
</tr>
<tr>
<td>Slice thickness (mm)</td>
<td>1.2</td>
<td>1.2</td>
<td>0.7</td>
</tr>
</tbody>
</table>

MR images from different datasets (IXI, KIRBY21, NITRC) were used in this study. The three datasets are characterized by a different static field magnitude. The main imaging parameters of each sequence are reported in the table. TR: repetition time; TE: echo time; α: flip angle; MPRAGE: magnetization-prepared rapid gradient-echo; 3D TFE: Three-dimensional Turbo Field Echo; TMJ: Temporomandibular joint.
Performance analysis on actual data

On actual MR data, we run INU correction and segmentation using the same range of input parameters used with simulated data. First, we computed the segmented WM and GM probability maps for each parameter configuration and generated optimized masks (see the procedure described in 3.2.1.4). Then, we estimated the relative noise as compared to the signal intensity in each image under investigation. We quantified both signal and noise on an axial slice cutting the corpus callosum at both ends: the signal corresponded to the maximum intensity within a polygonal ROI at the anterior end of the corpus callosum, and the noise to the standard deviation of the intensity within a circular ROI of 10 mm radius located outside the brain. Based on the estimated noise level, we defined the necessary level of smoothing based on the results of our simulations and we applied to the actual MR volumes. Finally, the optimized masks and the spatially smoothed MR images were then used to calculate indirect metric values, searching for the input parameter configuration potentially yielding the most accurate results.

We checked the accuracy of the INU correction results by visual inspection of the INU-corrected T1-weighted images, as well as the reconstructed INU profile. Importantly, we verified that MR images with higher dynamics in the INU profile (typically associated with a MR scanner with higher static field) lead to the definition of smaller regularization values, and possibly smoothing values. Also, we used the whole set of images from KIRBY21 dataset to conduct a test-retest analysis, aimed at examining whether the INU correction with the parameters determined using our data-driven approach leads to increased image reproducibility compared to the default configuration. To this end, we used as a quantitative index the DSI between GM (and WM) masks in MNI space, derived from each of the two sessions. We assessed significant increases/decreases in DSI values between sessions at the group-level by means of paired t-tests.

3.3 Results

We started our investigations by using the simulated T1-weighted image with INU 40% relative magnitude and 1% noise level, and examining the variability of CJV, CV<sub>WM</sub> and CV<sub>GM</sub> across different configurations of input algorithm parameters. For each metric, the configuration with the lowest value (associated with the putatively best INU estimate) was identified and its accuracy was quantified by comparing the corresponding INU against the simulated INU (Figure 3.3). This analysis revealed that the CJV generally provides lower absolute distances, and therefore more accurate results than CV<sub>WM</sub> and CV<sub>GM</sub>. For CJV, CV<sub>WM</sub> and CV<sub>GM</sub> calculated on low-dynamic profile MR images, D was 0.6%, 1.4% and 1.1%, respectively. For the high-dynamic profile, it was 0.8%, 1.5% and 1.1%, respectively. Not only did a smaller absolute distance characterize the selected parameter configuration, but also the matrix patterns better resemble the matrix of absolute distances D. For the low-dynamic profile, MMC was 0.96, 0.57 and 0.63 for CJV, CV<sub>WM</sub> and CV<sub>GM</sub>, respectively; for the high-dynamic profile, MMC was 0.99, 0.72 and 0.78 for CJV, CV<sub>WM</sub> and CV<sub>GM</sub>, respectively.
Then, we evaluated the performance of the three metrics for different levels of INU field magnitude and image noise (Figure 3.4). As before, the absolute distance $D$ for the parameter configuration selected from a given metric was computed, as well the MMC between the metric and the distance matrices. An increased INU field magnitude and/or an increase image noise level generally yielded higher $D$ for all the metrics. CJV generally outperformed the others at low noise levels regardless the INU field magnitude and the spatial profile, but was relatively less effective on high-noise MR images. In turn, CVGM showed good stability at higher noise levels. CVWM underperformed the other two metrics for most of the INU and noise levels.

By means of a two-way analysis of variance (ANOVA) we examined the influence of noise and INU magnitude on the absolute distance $D$. For both INU profiles, the effect of noise was highly significant ($F = 367.28, p < 0.001$ for the low dynamic, $F = 50.6, p = 0.001$ for the high dynamic), whereas the INU magnitude showed a much less significant effect ($F = 7.56, p = 0.0229$ for the low-dynamic profile, $F = 3.87, p = 0.0834$ for the high-dynamic one). We did not investigate further the dependence of the metrics on the INU magnitude, and reported from this point on only average performance over INU levels.
Next, we evaluated to what extent and how spatial smoothing can influence an accurate INU reconstruction (Figure 3.5). For a smoothing level set at 1 mm FWHM, a marked improvement of CJV, and no clear changes of CVWM and CVGM values, were found. Notably, a smoothing larger than 2 mm of FWHM led to less accurate INU reconstructions. This was evident in CVWM and CVGM, and less pronounced in CJV measures. Based on these results, we selected the CJV to identify input configurations with low INU estimation errors, and conducted further analyses on CJV only.

Figure 3.4 – Sensitivity of different metrics to inhomogeneity magnitude and noise. For each INU field magnitudes and noise level, we calculated the voxel-wise distance $D$ between the simulated and the estimated INU fields, as well as the matrix-to-matrix correlation (MMC). We report in this figure the results for the parameter configuration identified by each metric. CJV, CVWM and CVGM were compared for the low-dynamic (A) and high-dynamic (B) INU profiles.

We addressed the issue of defining subject-based masks to enhance the use of CJV in actual MR images. GM and WM probability maps corresponding to each of the parameter configurations under investigation were estimated, and a subset of them was used to generate average WM and GM masks. Our analysis on both low and high dynamic INU profiles revealed that, on average, including 85% of the masks with the largest correspondence with the SPM template mask in individual space ($R_T$ equal to 85%) is likely to be a reliable approach to ensure an effective use of the CJV (Figure 3.6).
Figure 3.5 – Relation between spatial smoothing and image noise. We assessed the relation between image noise and amount of smoothing applied after INU correction. The voxel-wise distance $D$ between the simulated and the estimated INU fields was computed. We report in this figure the results for the parameter configuration identified by each metric. CJV, CV$_{WM}$ and CV$_{GM}$ were compared for the low-dynamic (A) and high-dynamic (B) INU profiles.

Figure 3.6 – Mean Dice Similarity Index (mDSI) threshold analysis. By defining a subset of segmented WM and GM masks corresponding to each parameter configuration and averaging them together, we generated an improved version of the same masks. This was separately done for low-dynamic (A) and high-dynamic INU profiles, using a relative amount of configurations $R_T$ ranging from 50% to 100%. Each of these values represents the relative amount of included masks with respect to the total number of parameter configurations. The bar plots represent the mDSI, which quantifies the correspondence of each segmented mask with respect to the SPM template mask. The mDSI shown in figure was calculated for the two INU field profiles, averaging together the results over the whole set of simulated data (12 simulated images: 3 INU field magnitudes $\times$ 4 noise levels).
The need of a procedure for the definition of reliable WM and GM masks was confirmed by a complementary analysis conducted on the CJV, using the template (not subject-specific) masks derived from SPM (Figure 3.7). When comparing D values obtained using the SPM template masks and the average-based individual masks, the performance obtained using the former was found to be much inferior. On the other hand, by implementing our data-driven procedure, it was possible to achieve performance similar to the ones derived from the BrainWeb simulator masks used in the first part of the study (for comparison, see Figure 3.5).

![Figure 3.7 – CJV results obtained using optimized and standard masks. We assessed the impact of the masks on the CJV results, expressed in terms of the voxel-wise distance D between the simulated and the estimated INU fields. Specifically, we compared the results obtained using the MNI template masks (A-B) and the optimized masks (C-D). We examined the performance for different noise level and smoothing, both for low- (A-C) and high-dynamic (B-D) INU profiles.](image)

To validate the usefulness of the data-driven approach for the input parameters definition, three MR images, collected with 1.5 T, 3 T and 7 T scanners respectively (Figure 3.8), were used. The noise level was 1.38%, 1.25%, 0.72% for the three images, respectively. As such, a smoothing level equal to 1 mm FWHM was used to estimate the CJV. Then, WM and GM masks were generated with a relative amount of configurations $R_T$ equal to 85%. The analysis of CJV values obtained using different input parameters for the INU correction algorithm revealed different solutions for the three MR images under investigation. The 1.5 T dataset was characterized by a smoothing parameter of 30 mm FWHM and a regularization parameter of 0.1, consistent with the low-dynamic profile. This was supported by a visual inspection of the raw data, which also showed a negligible INU magnitude. A parameter matrix mainly weighted to higher regularization values characterized the 7 T image, which had a highly dynamic spatial profile. In this case, the identified regularization parameter was equal to 0.001. The INU for the 3 T image had intermediate magnitude compared to those of the 1.5 T and 7 T images, as well as low dynamic profile. The analysis of the CJV suggested the regularization parameter to be best set to 0.01, with a smoothing parameter of 30 mm FWHM. When we extended this analysis of all the 42 MR images
collected at 3 T and belonging to the KIRBY21 dataset, our data-driven approach was found to yield the same parameter configuration (smoothing level: 30 mm FWHM; regularization parameter: 0.01). By using the KIRBY21 dataset, we also tested whether our approach yielded increased INU correction reliability. Notably, a significant increase was observed in the test-retest DSI analysis for both GM and WM masks (Figure 3.9) when using the optimized configuration as compared to the default one. Specifically, the average DSI across subjects increased from 0.862 to 0.8675 for GM (p=0.0022) and from 0.9406 to 0.9432 for WM (p=0.0015).

**Figure 3.8** – INU correction on actual MR images. We examined the effectiveness of the CJV analysis on actual MR data, after smoothing and mask optimizations, for three representative images collected using a 1.5 T, a 3 T and a 7T MR scanner (A-B-C), respectively. The diamond marker highlights the input parameter configuration selected on the basis of the CJV results. INU-corrupted image, estimated INU field and INU-corrected image are shown for each dataset (D-E-F).

**Figure 3.9** – Test-Retest analysis of INU correction performance. We used the full set of KIRBY21 images to perform a test-retest reliability analysis. As an indirect measure of INU correction performance, we employed the DSI between GM (and WM) segmented volumes. We compared the DSI values obtained from the optimized and default parameter configurations using paired t-tests. The bar plots show mean and standard error for GM and WM masks, and for optimized (OPT) and default (DEF) configurations, respectively. The probabilities estimated using the t-tests are indicated in the figure as well.
3.4 Discussion

INU correction is a fundamental processing step for structural MR images. It is a matter of fact that the performance of any INU correction method depends on the input setting used (Boyes et al., 2008; Uwano et al., 2014; Weiskopf et al., 2011; Zheng et al., 2009), and a less effective INU correction can substantially hamper the reliability of MR imaging results (Ashburner and Friston, 2000; Good et al., 2001; Pham and Prince, 1999; Zheng et al., 2009). Using simulated MR data, we previously showed that the INU correction using specific parameter configurations may be much more accurate than those obtained using a default one (Ganzetti et al., 2015). However, to the best of our knowledge no reliable method exists to define parameters that most likely yield the best correction for actual MR data. Here we examined the characteristics of different metrics, defined among them the most accurate one, and used it to develop a data-driven approach to address this problem. We conducted our investigation using the INU correction algorithm implemented in SPM12, which is one of the most widely used software for MR data analysis. Notably, this algorithm would largely benefit from an optimization approach, as it is largely sensitive to the selection of the input parameters, namely the regularization and the smoothing factor (Ganzetti et al., 2015).

In actual MR data, a common approach to assess INU correction is the one based on indirect measures relying on intensity variability. Among them, the CJV, the \( CV_{WM} \) and the \( CV_{GM} \) metrics are the most commonly used ones in the literature. Specifically, the CV expresses the normalized standard deviation in a single tissue class, whereas the CJV takes into account the intensity distributions in both classes. In line of principle, the smallest CV and CJV correspond to smaller intensity inhomogeneity residual, thus better performance (Belaroussi et al., 2006; Likar et al., 2001). On the other hand, the CJV not only evaluates the parallel reduction of GM and WM distributions, but also the degree of overlap between the two. Indeed, an effective INU correction produces a consistent increment of contrast in the image, reflected by a clear separation of WM and GM distribution peaks, and thus a decrease of CJV. In addition, a strong INU correction may remove smooth intensity variations characterizing the actual anatomical contrast. In this scenario, while \( CV_{GM} \) and \( CV_{WM} \) decrease, the CJV increases because the WM and GM distributions peaks get closer. Our simulation results on both voxel-wise distance (D) and matrix-to-matrix correlation (MMC) revealed a larger accuracy of CJV compared to the other two, regardless of the spatial profile of the INU (Figure 3.3). CJV combines information about image intensities in both GM and WM. In this manner, it allows the joint assessment of intensity variability within each tissue class as well as in the image contrast between the two structures. In turn, \( CV_{WM} \) and \( CV_{GM} \) are estimates derived from image values only in WM and GM, respectively. When the INU correction tends to overestimate the actual inhomogeneities present in the MR image, the contrast diminishes and the CV may erroneously detect an image improvement simply due to a reduced standard deviation in the intensity distribution. This effect may explain - at least in part - the results obtained on simulated data, for which \( CV_{WM} \) and \( CV_{GM} \) tended to indicate low regularization values and low smoothing factors as yielding better
INU correction (Figure 3.3). Specifically, lower values of regularization allow the INU correction algorithm to follow sharp intensity variations, up to the point that factual anatomical variations may be cancelled. Our findings suggest that considering the overlap between the intensity distributions of distinct tissue classes is very important for the detection of INU correction performance, and that the CJV may be potentially more suitable than CV_{WM} and CV_{GM} for an accurate inhomogeneity correction.

It should be considered that actual MR images may be characterized by various noise levels and INU magnitudes. These may depend on the subject as well as on the acquisition hardware and sequence used. Our findings suggested noise to substantially influence the performance of CJV, CV_{WM} and CV_{GM} (Figure 3.4). Specifically, all the three metrics provided accurate results for low levels of noise (0-1%), with the CJV overperforming CV_{WM} and CV_{GM} both in terms of D and MMC. On the other hand, the CJV was the most sensitive to noise, and underperformed CV_{WM} and CV_{GM} with very noisy MR data (5% noise level). A possible explanation may be the spreading effect spatial noise has in the intensity distribution of both WM and GM. This may hamper a reliable measure of the actual statistical properties of each tissue distribution, thus leading to an improper parameter selection. The low MMC values obtained for CJV at high noise levels seem to confirm this possibility (Figure 3.4). In line with previous studies (Chua et al., 2009), a moderate amount of spatial smoothing (i.e. 1mm FWHM) led to a considerable increase of the CJV accuracy (Figure 3.5). The same solution did not prove to be as effective when using CV_{WM} and CV_{GM} instead of CJV.

After establishing that CJV in combination with spatial smoothing can yield a reliable estimation of INU correction parameters, we addressed the problem of how this metric could be effectively applied on actual MR data. Importantly, GM and WM masks are needed to measure the CJV, and different options exist as for deriving these masks from the actual MR images. One aspect to consider is that the INU correction algorithm of SPM is integrated with brain segmentation, such that GM and WM probability maps are automatically generated. This means that, in line of principle, it would be possible to estimate GM and WM maps for each input parameter configuration, and use them for the CJV calculation. Such a solution, however, does not permit an unbiased comparison across configurations, as the masks would be different case by case. Rather than using the SPM template masks registered to individual space, we implemented an approach that exploits the similarity between those template masks and the ones estimated from the SPM segmentation algorithm, which are subject-specific. By using the mean Dice Similarity Index (mDSI), we searched through the entire configurations space and selected a set of probability maps that had mDSI superior to a certain threshold. For both WM and GM, all the maps satisfying the mDSI criteria were then averaged together, and then employed to create the actual masks. The latter ones were then used across all configurations for the CJV assessment. Our analysis on simulated data revealed that this approach can lead to the definition of masks that are very close to the ground truth masks and are much more precise than the SPM template masks registered to individual space (see Figure 3.5A-B and Figure 3.7). It is our opinion that the approach we implemented limits the possibility of deceptive CJV evaluations due to partial volume effects, which are typically present in the voxels including both WM and GM tissues.
To show the potential usefulness of the developed data-driven approach to estimate INU correction parameters, we used also actual MR images collected with 1.5 T, 3 T and 7 T scanners respectively. One of the main features that influences INU properties is indeed the strength of the static field (Boyes et al., 2008; Uwano et al., 2014). With increasing magnetic field, not only does the INU field magnitude rise, but also the INU spatial dynamic is more variable as a result of tissue-induced inhomogeneities (Bernstein et al., 2006; Mihara et al., 2005; Moser et al., 2012; Umutlu et al., 2014; Uwano et al., 2014; Van De Moortele et al., 2005). The CJV results for MR images at different magnetic fields suggested this metric to be sensitive to INU properties, since the minimum CJV value across the whole set of input parameters was different across MR images (Figure 3.8). For instance, a relatively low regularization parameter was identified as being more accurate for the 1.5 T image, consistent with a low frequency INU pattern compared to the underlying anatomical structures. The intensity had a consistent intensity drop at the center of the 3 T image. This might be related to a RF wavelength shortening as well as the coil sensitivity (Bernstein et al., 2006). Although this intensity inhomogeneity was still low frequency compared with anatomical brain structures, the CJV analysis suggested a larger level of regularization and the same FWHM level of the 1.5 T image. This was putatively due to a larger INU field magnitude. The 7 T image was characterized by a substantially different intensity inhomogeneity profile compared to the 1.5 T and 3 T images. In this case, the CJV values were more weighted towards higher regularization parameters that allowed the INU correction to better follow sharp inhomogeneity variations across the MR image.

When we extended our analysis on actual MR images to the whole KIRBY21 dataset, which was collected with a 3 T MR scanner, we could appreciate a very high stability of the configuration of input parameters selected by our data-driven method. This may indicate that the selected input configuration, rather than being subject-specific, more likely depends on the MR hardware and acquisition sequence used. It remains however to be verified if this finding for images collected at 3 T generalizes also to higher field strengths, for which tissue-induced inhomogeneities are more prominent. This may indeed lead to an increased inter-subject variability in the selected parameter configuration. Importantly, we also observed that the segmentation results for the same subject scanned in two separate sessions were more similar when using optimized than standard configurations (Figure 3.9). It is commonly accepted that intensity inhomogeneity primarily affects the accuracy of image segmentations (Belaroussi et al., 2006; Zheng et al., 2009). Accordingly, this finding might be taken as indirect evidence of an increased INU correction performance. Since we conducted the rest-retest analysis on a single dataset, we suggest that future studies are warranted to evaluate whether the increased INU correction performance is confirmed with other datasets, possibly collected with different scanners.
3.5 Conclusion

To the best of our knowledge, this is the first study that addressed the problem of selecting the most appropriate input algorithm parameters for INU correction of structural MR images. Our analyses were based on the INU correction algorithm implemented in SPM, but the same approach can be in principle extended to any other INU correction algorithm requiring the selection of input parameters. In short, we conducted a comprehensive comparison of indirect metrics for the assessment of the INU correction results. We identified the CJV as the most accurate one, as long as the noise level in the INU-corrected image was controlled by means of spatial smoothing. Based on the CJV, we developed a data-driven approach aiding the selection of the parameters to be used for an accurate inhomogeneity correction in actual MR images. Our findings suggest that it is possible to tailor the parameter configuration of the INU correction algorithm based on the characteristics of the MR image to be processed, leading to a substantial improvement compared to the default parameter configuration. Since substantial progress is being made on the development of high-field MR scanners (Moser et al., 2012; Umutlu et al., 2014), the problem of INU correction is becoming increasingly important (Bernstein et al., 2006; Mihara et al., 2005; Uwano et al., 2014; Van De Moortele et al., 2005). The data-driven approach described here may contribute to address this problem by optimizing the performance of any given INU correction algorithm.
Despite recent advancements in MR imaging, non-invasive mapping of myelin in the brain still remains an open issue. Here we attempted to provide a potential solution. Specifically, we developed a processing workflow based on T1-weighted (T1-w) and T2-weighted (T2-w) MR data to generate an optimized myelin enhanced contrast image. The workflow allows whole brain mapping using the T1-w/T2-w technique, which was originally introduced as a non-invasive method for assessing cortical myelin content. The hallmark of our approach is a retrospective calibration algorithm, applied to bias-corrected T1-w and T2-w images, that relies on image intensities outside the brain. This permits standardizing the intensity histogram of the ratio image, thereby allowing for across-subject statistical analyses. Quantitative comparisons of image histograms within and across different datasets confirmed the effectiveness of our normalization procedure. Not only did the calibrated T1-w/T2-w images exhibit a comparable intensity range, but also the shape of the intensity histograms was largely corresponding. We also assessed the reliability and specificity of the ratio image compared to other MR-based techniques, such as magnetization transfer ratio, fractional anisotropy and fluid-attenuated inversion recovery. With respect to these other techniques, T1-w/T2-w had consistently high values, as well as low inter-subject variability, in brain structures where myelin is most abundant. Overall, our results suggested that the T1-w/T2-w technique may be a valid tool supporting the non-invasive mapping of myelin in the brain. Therefore, it might find important applications in the study of brain development, aging and disease.

Published as:
Chapter 4

4.1 Introduction

Myelin, the dielectric sheath surrounding neuronal axons, is an essential component for efficient brain functioning. Its main role is to facilitate long-range neuronal communication processes supporting higher-order cognitive, sensory, and motor functions. An accurate assessment of myelin in vivo is extremely important for a comprehensive understanding of human neurodevelopment and neurodegeneration (Barkovich, 2005; Deoni et al., 2011; Glasser and Van Essen, 2011; Kizildag et al., 2005; Laule et al., 2006; Laule et al., 2007; Paus et al., 2001; Staudt et al., 1994; Steenweg et al., 2010; van Buchem et al., 2001; Welker and Patton, 2012). Histopathological techniques are the gold standard for the quantitative assessment of myelin, but they can be used only post mortem (Gareau et al., 2000; Laule et al., 2006). Furthermore, histopathological investigations are typically conducted for only a limited number of regions, rather than for the whole brain. To address this problem, non-invasive imaging tools based on magnetic resonance imaging (MRI) were proposed for myelin mapping (Barkovich, 2000; Paus et al., 2001): conventional T1-weighted (T1-w) and T2-weighted (T2-w) imaging, magnetization transfer imaging (MTI), diffusion tensor imaging (DTI), fluid-attenuated inversion recovery (FLAIR), multi-component T2-relaxation imaging (MCRI), and multi-component Driven Equilibrium Single Pulse Observation of T1 and T2 (mcDESPOT).

Early MR studies used T1 and T2 relaxation times (Crooks et al., 1987; Tofts and du Boulay, 1990), which are strictly connected to changes in the interactions between water molecules and tissue macromolecules (Miot-Noirault et al., 1997), to assess the spatial distribution of myelin in the brain. The level of brightness characterizing white matter (WM) in T1-w MRI is associated with the spatial distribution of myelin-bound cholesterol such that the degree of myelin-related contrast can be inferred from T1-w images (Dobbing and Sands, 1973; Koenig, 1991). Conversely, T2 relaxation relates to proton transfers, molecular exchange and diffusion of water. Hydrophobic properties of the lipidic bilayer in myelin restrict molecular motion of protons (Barkovich, 2000; Miot-Noirault et al., 1997) and hypointensity on T2-w images reflects relatively larger myelin content. It is worth noting that T1-w and T2-w images typically provide only qualitative information on myelin distribution in the brain. Therefore, different MR techniques are preferred for clinical studies involving the direct comparison of myelin in patients and healthy controls.

MTI is the most commonly used technique to detect subtle changes in the biochemical architecture and composition of tissues (Barkovich, 2005; Grossman et al., 1994; Rademacher et al., 1999; van Buchem et al., 2001). The fundamental concept behind this modality is the exchange of magnetization between mobile protons (water) and immobile protons bound to macromolecules (non-aqueous tissue). This effect is usually measured as a magnetization transfer ratio (MTR). Despite its high sensitivity towards tissue changes and damage, MTR cannot be considered an absolute marker of myelination. In fact, a low MTR may result either from a change in myelin content or from structural changes following inflammation (Gareau et al., 2000; Laule et al., 2007).
Diffusion tensor imaging (DTI) is a technique sensitive to diffusion processes of water molecules in biological tissue (Beaulieu, 2002). The kinematics of water molecules can be expressed in terms of fractional anisotropy (FA), which serves as a marker of WM development, axonal damage and myelin pathology. However, different studies have provided evidence that myelin is not the sole element of anisotropic water diffusion in axonal fibers (Laule et al., 2007; Madler et al., 2008). Hence, FA should be considered an indicator of fiber tract density, and only indirectly of myelin content.

Additionally, a limited number of studies have speculated about the potential of FLAIR imaging as a suitable marker of myelin maturation (Ashikaga et al., 1999; Kizildag et al., 2005; Murakami et al., 1999). FLAIR is a particular inversion-recovery sequence that can be used in brain imaging to suppress or heavily reduce the signal originated from the cerebrospinal fluid. In this regard, the detection of deep white matter lesions juxtaposed to the ventricles has shown to be extremely important in the recognition of pathological processes such as multiple sclerosis (Miller et al., 1998).

In recent years, other techniques have also been introduced, such as MCRI (Beaulieu et al., 1998; Gareau et al., 2000; Laule et al., 2006; MacKay et al., 1994; Madler et al., 2008; Oh et al., 2006; Vidarsson et al., 2005; Whittall et al., 1997) and mcDESPOT (Deoni et al., 2011; Deoni et al., 2008). These are based on the principle that spin relaxation in a particular inhomogeneous environment may not be assumed as mono-exponential. Accordingly, they employ multiple MR pulse acquisition sequences in order to define the biophysical properties of the tissue under investigation (Laule et al., 2007). This permits to separate the signal belonging to water trapped between the myelin bilayers (myelin water) (MacKay et al., 1994) from the total MR signal, resulting in a myelin water fraction (MWF) measure. MWF is currently considered a reliable marker of myelin (Gareau et al., 2000; Laule et al., 2006). A critical challenge using both MCRI and mcDESPOT is however a perceptibly long scan time (between 10 and 25 minutes) (Deoni et al., 2011; Gareau et al., 2000; Kitzler et al., 2012; Kolind et al., 2012; Madler et al., 2008; Oh et al., 2006; Whittall et al., 1997), which may limit their applicability in clinical studies.

Recently, there has been a resurgence of interest on T1-w and T2-w imaging for myelin mapping. Glasser and Van Essen (2011) proposed to combine T1-w and T2-w images to obtain a myelin-enhanced contrast image (Glasser et al., 2013a; Glasser et al., 2013b). Compared to quantitative methods, which generally require longer acquisitions, fast scanning times make it potentially well-suited for clinical investigations. It is worth noting, however, that the T1-w/T2-w technique as described by Glasser and Van Essen (2011) is a relative measure potentially characterized by intensity scale inconsistencies across datasets, which may be present even for MR images collected with the same scanner on different days. To address this issue, which may hamper within- and between-group statistical comparisons, the use of a calibration approach is strictly necessary. Glasser and Van Essen (2011) introduced an internal calibration based on the image histogram. Importantly, this approach may be unsuited for studies in which myelin changes are expected as a result of a brain disease. Indeed, internal calibration attenuates global differences between patients and controls, to the point that altered myelin levels might not be detected. Also, the shape of the T1-w/T2-w image histogram may be different in patients with respect to controls,
so that local changes in myelin levels in patients may be erroneously observed due to histogram
equalization between patient and control groups.

Here we aimed to further develop the T1-w/T2-w technique, by tackling the problem of intensity
scale inconsistencies across different datasets. We developed an analysis workflow for the calibration of T1-
w/T2-w intensities in the brain using information of T1-w and T2-w intensities extracted from tissue
outside the brain, thereby avoiding the problems related to the use of an internal calibration. To evaluate
the effectiveness of our normalization procedure, we compared T1-w/T2-w images obtained from
different MR scanners, with different sequences and acquisition parameters. Next, we examined the
consistency of T1-w/T2-w across healthy individuals against other MR imaging modalities, such as MTR,
FA and FLAIR. Our results suggest that T1-w/T2-w ratio method can be a reliable and relatively fast tool
for non-invasive myelin imaging.

4.2 Methods

In this section we describe the workflow for the calibration of T1-w/T2-w images, allowing for the
mapping of myelin in the human brain using T1-w and T2-w MR imaging data. Next, we show the
reliability and sensitivity of T1-w/T2-w ratio method as compared to alternative techniques.

4.2.1 Description of the method

4.2.1.1 Theoretical background

Our method is an extension of the method originally proposed by Glasser and Van Essen (2011). They showed that, by calculating the ratio between T1-w and T2-w images of the same subject, it is
possible to increase the contrast related to myelin content (Figure 4.1).

The myelin-enhanced contrast image obtained through this approach is however not
automatically bias-free because the ratio does not attenuate (or cancel) the image bias resulting from
different sensitivity profiles of the receiver coils for the two images (Belaroussi et al., 2006). Furthermore,
the intensity scale of the T1-w/T2-w image is dependent on the specific instrumentation and scanning
parameters used for the T1-w and T2-w images. In general terms, the T1-w/T2-w image can be modeled
as follows:

\[
\frac{T1-w}{T2-w} \approx \frac{a_1 \cdot s_1 \cdot x}{a_2 \cdot s_2 \cdot \left(\frac{1}{x}\right)} = \frac{a_1 \cdot s_1}{a_2 \cdot s_2} \cdot x^2 = \beta \cdot x^2
\]

(4.1)

where the myelin content is represented by \(x\), the sensitivity profiles are denoted by \(s_1\) and \(s_2\) for the T1-w
and T2-w images respectively, and \(a_1\) and \(a_2\) are scaling factors. Accordingly, T1-w/T2-w intensity
depends on the combination of \(s_1, s_2, a_1, a_2, \beta\) in Eq. 4.1). The aim of an offline normalization procedure is
to reach the ideal configuration in which both the differences in the sensitivity profiles of T1-w and T2-w
sequences become negligible (i.e., \( s_1 \rightarrow 1 \), \( s_2 \rightarrow 1 \)), and the values \( \alpha_1 \) and \( \alpha_2 \) are standardized, so that the T1-w/T2-w intensity scaling is comparable across different subjects.

![Figure 4.1 – Myelin enhanced contrast image for a representative subject. The ratio of T1-w (A) to T2-w (B) signal intensity is calculated to obtain the T1-w/T2-w ratio image (C). This is done to improve the mapping by increasing the contrast between different myelinated structures. Since conventional MRI images have arbitrary intensity scales, the three images are showed with a colormap assigned on the basis of the 10th and 90th percentile values. The subject used for this figure is Subject 30 of the KIRBY21 database.](image)

### 4.2.1.2 Method implementation

**Mask creation**

Intensity standardization may be achieved with an internal scaling of intensity values, as previously proposed by Glasser and Van Essen (2011). By implementing this procedure, erroneous representation may occur in the presence of altered myelin levels. In this case, internal scaling may indeed hide substantial differences, preventing valid comparisons between controls and patients. This is the reason why we implemented an external calibration approach. The standardization of the T1-w/T2-w image was achieved through several processing steps (Figure 4.2), for which we used SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK). As a first step, two subject-specific masks were created by warping predefined masks in the stereotaxic space of the Montreal Neurological Institute (MNI) to the individual space, using T1-w images in the two spaces to calculate the necessary spatial transformation (Ashburner and Friston, 1997; Ashburner and Friston, 1999). To ensure the effectiveness of this step, the masks should contain voxels outside the brain and should span image regions with relatively high regional homogeneity. Furthermore, one of them should contain relatively low values on the T1-w image and high values on the T2-w image, and the other mask should have reversed characteristics. We implemented this specification by selecting two masks covering the eyeballs and the temporal muscles, respectively (Figure 4.3). These were defined directly in MNI space by segmenting and thresholding the ICBM152 template images (www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009).
Chapter 4

Figure 4.2 – Calibration of the T1-w/T2-w image: analysis workflow. Workflow of T1-w/T2-w image data processing, including the warping of standard masks from MNI to subject space. The bias correction is a fundamental stage for both T1-w and T2-w raw images. Then each of the bias-free image undergoes the normalization process in order to accomplish a proper scaling. Finally, the T1-w/T2-w image is calculated as the ratio of calibrated T1-w and T2-w images.

Figure 4.3 – Masks used for the calibration procedure. The calibration algorithm is based on values extracted from two anatomical masks (eye and temporal muscle) warped to the subject space. The eye mask (A) is located within the vitreous humour of the eyeball and encloses the gel that fills the space between the retina and the crystalline lens. The temporal muscle mask (B), which is one of the main muscles involved in the mastication process, is set on the bulk of the temporalis muscle that covers the temporal bone. The subject represented in the figure is Subject 30 of the KIRBY21 database, as in Figure 4.1.
Bias correction

In parallel to the creation of subject-specific masks, the original T2-w image was coregistered to the T1-w image through a rigid-body transformation (Collignon et al., 1995). Then, the T1-w and T2-w images were jointly subjected to bias correction to ensure that the sensitivity profile \( s_1 \) and \( s_2 \) in Eq. 4.1 was spatially equalized. Unlike receive field (B1-) inhomogeneities, the T1-w/T2-w ratio cannot completely correct for transmit field (B1+) inhomogeneities in intensity and contrast (Glasser et al., 2013a; Glasser and Van Essen, 2011). Accordingly, instead of removing common spatial inhomogeneity by combining T1-w and T2-w images (Glasser et al., 2013a), we opted for using the intensity inhomogeneity correction tool implemented in SPM8 (Ashburner and Friston, 2005; Weiskopf et al., 2011) on the two images separately. The input parameters for the intensity inhomogeneity correction algorithm, namely the smoothing and the regularization parameters, were set at their default value (equal to 60 mm and \( 10^{-4} \), respectively).

**Figure 4.4** – T1-w and T2-w intensities for the eyeball and temporal muscle masks. We analyzed T1-w and T2-w intensities within the eyeball and temporal muscle masks for Subject 30 of the KIRBY21 database. The eyeball mask values for the T1-w image (A) are always lower than the ones in the temporal muscle mask (B), whereas eyeball mask values in T2-w image (C) are larger than the ones bounded by the temporal muscle mask (D). Since the voxel intensities in the two masks generally showed distributions deviating from a Gaussian trend, we selected reference values as the distribution peak (i.e., the numerical mode) rather than the statistical mean or median.
Intensity standardization

After bias correction, the T1-w and T2-w images were further processed to standardize their intensity by using a linear scaling procedure. Specifically, the distribution peaks (modes) of intensities in the two masks (Figure 4.4) were extracted from either the unbiased T1-w or T2-w images of the single subject, indicated as $X_S$ and $Y_S$, and were then compared with the corresponding values from the ICBM152 template image of the same modality, indicated as $X_R$ and $Y_R$. The modes for the ICBM152 template corresponded to $X_R=58.6$ and $Y_R=28.2$ for the T1-w image, or $X_R=21.1$ and $Y_R=99.9$ for the T2-w image. The linear scaling of either the T1-w or the T2-w image was accomplished using the following formula:

$$I_C = \frac{X_R-Y_R}{X_S-Y_S} \cdot I + \frac{X_S Y_R - X_R Y_S}{X_S-Y_S}$$  \hspace{1cm} (4.2)

where $I$ and $I_C$ are the images before and after calibration, respectively. After calibrating T1-w and T2-w images with the formula described above, their ratio was calculated to produce the calibrated T1-w/T2-w image (see Eq. 4.1).

4.2.2 Method validation

4.2.2.1 Subjects and data acquisition

We used three different publicly available datasets for the method validation. Two of them were extracted from the IXI database of the Imperial College London (biomedic.doc.ic.ac.uk/brain-development/index.php?n=Main.Datasets), whereas the third was from the KIRBY21 database of the Kirby Research Center for Functional Brain Imaging in Baltimore (mri.kennedykrieger.org/databases.html). For the first two datasets we extracted T1-w and T2-w images collected in 21 healthy subjects with a 1.5T MR scanner (Gyroscan Intera, Philips Healthcare) and in 21 healthy subjects with a 3T MR scanner (Intera, Philips Healthcare), respectively. The third dataset contained T1-w, T2-w, MT, FA and FLAIR images collected in 21 healthy subjects with another 3T MR scanner (Achieva, Philips Healthcare). It is worth noting that more than 600 subjects are available in the IXI database, but we selected only 21 of them for each scanner to ensure statistical comparability of the results with those from the KIRBY21 database. The selection was made such that subjects with a comparable age range across databases could be used in our analyses. We identified the optimal age-matching group in the IXI database after assessing each possible group of 21 subjects, as generated by permutation. We calculated the Mann-Whitney U-test on the ages of each possible IXI group using the ages of the KIRBY21 group as reference. Finally, we determined the IXI group that provided the highest probability to the test. Details on subject demographics and scanning parameters for the different image modalities are provided in Tables 4.1 and 4.2, respectively.
Table 4.1 – Demographic data.

<table>
<thead>
<tr>
<th></th>
<th>IXI database (1.5 T)</th>
<th>IXI database (3 T)</th>
<th>KIRBY21 database (3 T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of subjects</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Number of female subjects</td>
<td>12</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Age (min – max)</td>
<td>21-59</td>
<td>21-68</td>
<td>22-61</td>
</tr>
<tr>
<td>Age (mean±sd)</td>
<td>31.7±8.4</td>
<td>32.5±12.1</td>
<td>31.7±9.4</td>
</tr>
</tbody>
</table>

A total of 63 healthy subjects were included in this study. Each dataset consisted of 21 subjects within a comparable age range.

Table 4.2 – MR imaging sequence parameters.

<table>
<thead>
<tr>
<th></th>
<th>IXI database (1.5 T)</th>
<th>IXI database (3 T)</th>
<th>KIRBY21 database (3 T)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1-w</td>
<td>T2-w</td>
<td>T1-w</td>
</tr>
<tr>
<td>TR (ms)</td>
<td>9.8</td>
<td>8178</td>
<td>9.6</td>
</tr>
<tr>
<td>TE (ms)</td>
<td>4.6</td>
<td>100</td>
<td>4.6</td>
</tr>
<tr>
<td>Inversion time (ms)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Resolution X (mm)</td>
<td>1.2</td>
<td>0.94</td>
<td>1.2</td>
</tr>
<tr>
<td>Resolution Y (mm)</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>Resolution Z (mm)</td>
<td>0.94</td>
<td>1.25</td>
<td>0.94</td>
</tr>
<tr>
<td>Flip angle (deg)</td>
<td>8</td>
<td>90</td>
<td>8</td>
</tr>
</tbody>
</table>

MR images collected with different sequences (T1-w, T2-w, MTI, DTI and FLAIR) were used in this study. The main imaging parameters of each sequence are reported in the table. TR: repetition time; TE: echo time.

4.2.2.2 Similarity of image histograms across subjects

We computed the T1-w/T2-w images, both before and after calibration, for each single dataset included in the study, and we compared them to assess the effects of the calibration procedure. We quantified the similarity of the intensity histograms for T1-w/T2-w images from the same MR scanner, as well as from different scanners. Specifically, we divided the whole range of image values into 500 bins and we normalized each histogram by dividing it by the sum over all its elements to account for the different number of brain voxels across individuals. We then estimated mean and standard deviation across histograms of different datasets in a bin-by-bin fashion, to quantify the consistency of the T1-w/T2-w values across subjects.

In addition, we conducted a quantitative analysis on the white matter, where myelin is mostly present. We used the SPM8 segmentation toolbox (Ashburner and Friston, 2005) on T1-w and T2-w images to create a white matter probability map that was thresholded at p>0.5 to obtain a binary white matter mask. Hence, we estimated the numerical mode of the T1-w/T2-w values distribution in the mask, as representative for the whole brain structure. We applied this procedure to each dataset, and we analyzed the resulting values by descriptive and inferential statistics to evaluate a potential increase in
across-subject reproducibility. We first checked that the values were normally distributed by means of a Lilliefors test \((p<0.05)\). Then we assessed whether the differences between databases were reduced by the calibration procedure using t-tests and a single-factor analysis of variance (ANOVA) on values before and after the calibration.

### 4.2.2.3 Comparison of T1-w/T2-w with MTR, FA and FLAIR images

As an additional assessment, we also compared the T1-w/T2-w image with MTR, FA, FLAIR images of the same subjects, using the multi-modal imaging dataset of the KIRKY21 database. FLAIR image was only coregistered to the T1-w/T2-w image whereas MTR and FA values were calculated from the coregistered magnetization transfer images and DTI data, respectively.

MTR, being sensitive to the macromolecular composition of tissue, is classically used for the assessment of alterations in the myelin content (Schmierer et al., 2004). MTR data are characterized by a pulsed sequence using a spoiled 3D gradient echo. For the data in the KIRKY21 database, magnetization transfer (MT) preparation was achieved using a five-lobed, sinc-gauss shaped radiofrequency irradiation (\(B_1=10.5 \text{ mT, duration 24 ms, and offset frequency}=1500 \text{ Hz})\). Also, a reference scan was obtained with the same parameters in the absence of MT preparation. We calculated the MTR image as:

\[
\text{MTR} = 1 - \frac{M_S}{M_0}
\]

where \(M_S\) is the image with MT preparation (in the presence of the radiofrequency irradiation) and \(M_0\) is the reference image without MT preparation. Next, we used SPM8 to register the MTR image to T1-w and T2-w images of the same subject by a rigid-body transformation (Collignon et al., 1995).

DTI is sensitive to the diffusion of water through white matter bundles, and is commonly used to produce a map of FA values across the brain. Notably, since myelin surrounds and protects white matter fibers, the presence of high FA values can be considered an indirect index of large myelin content. The DTI data in the KIRKY21 database were acquired with a multi-slice, single-shot, echo-planar imaging (EPI), spin-echo sequence with fat suppression by spectral presaturation with inversion recovery and with anterior-posterior phase encoding direction. Diffusion weighting was applied along 32 directions with a \(b\)-value of 700 \(s/mm^2\) (Landman et al., 2011). We used the FSL 5.0 software (Oxford Centre for Functional MRI of the Brain, University of Oxford) for the calculation of the FA image. First, we performed a prealignment (similar to motion correction in fMRI data) to correct for head movement during the session and to reduce the effects of gradient coil eddy currents (Horsfield, 1999). We also used the alignment parameters to correct the B-matrix, so that information on diffusion weighting directions was correctly preserved (Leemans and Jones, 2009). Then, the diffusion tensor was calculated using a simple least squares fit of the tensor model to the diffusion data. From this, the FA image was calculated as follows (Basser et al., 1994; Pierpaoli and Basser, 1996):
\[
FA = \sqrt{\frac{3}{2} \left( \frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2} \right)}
\] (4.4)

where \( \bar{\lambda} \) is the mean of the three eigenvalues \( \lambda_1, \lambda_2, \lambda_3 \). After calculating FA across brain voxels, we corrected spatial mismatch between the FA map and the DTI geometric reference image in the KIRBY21 database using the SPM8 normalization tool. Next, we used again SPM8 to coregister the FA image to the T1-w and T2-w images of the same subject.

After T1-w/T2-w, MTR, FA and FLAIR images were generated and were spatially aligned to each other, we transformed them to MNI space using the SPM8 normalization tool. This permitted us to perform across-subject statistical analyses. Specifically, we assessed the across-subject reproducibility of the different image modalities on specific regions of interest (ROIs), which were selected on the basis of previous myelin studies (Barkovich, 1988, 2000; Kizildag et al., 2005; Leppert et al., 2009; Welker and Patton, 2012; Whittall et al., 1997). A first group was composed by ROIs in the white matter and with putatively high myelin content: anterior corona radiata (ACR), superior corona radiata (SCR), pontine crossing tract (PCT), anterior limb of internal capsule (ALIC), genu of corpus callosum (GCC), splenium of corpus callosum (SCC). These ROIs were defined using the stereotaxic white matter atlas of the Laboratory of Brain Anatomical MRI, John Hopkins University School of Medicine, Baltimore, MD, USA (cmrm.med.jhmi.edu). A second group of ROIs included the putamen (PUT), caudate nucleus (CAU) and thalamus (THA), which are all structures with relatively low myelin content. These control ROIs were defined using the ICBM Deep Nuclei probabilistic atlas provided by the International Consortium for Brain Mapping (www.loni.ucla.edu/ICBM).

Since we sought to compare different kinds of images that are putatively characterized by different image intensities and contrasts, we evaluated the image intensity in a single ROI against the average intensity in the whole brain, by using a two-tailed paired t-test. Specifically, we used the following formula:

\[
t_{ROI} = \frac{\sqrt{n-1} \cdot \text{mean}(\Delta_{ROI})}{\text{sd}(\Delta_{ROI})}
\] (4.5)

where \( \Delta_{ROI} = [I_{ROI} - I_{BRAIN}] \) is the vector with the differences between ROI intensity and full-brain mean intensity across subjects, and \( n \) the number of subjects. The resulting t-score reflects how much the ROI intensity differs from the mean value calculated across the brain, taking between-subject variability into account. Determining t-scores for different image modalities (T1-w/T2-w, MTR, FA, FLAIR) allowed us to assess their reliability across individuals, as well as consistency across different ROIs. Additionally, we generated a t-score map from the T1-w/T2-w data by applying the same formula in Eq. 4.5 to each voxel rather than to a single ROI. This t-score map was thresholded at \( p<0.05 \), FDR-corrected for multiple comparisons (Genovese et al., 2002), highlighting brain regions with significantly larger T1-w/T2-w values than the average across the brain.
Chapter 4

4.3 Results

As an initial analysis, we calculated T1-w/T2-w images for each dataset included in the study, using the simple ratio of unprocessed T1-w and T2-w images (as shown in Figure 4.1). We evaluated the variability in image histograms across datasets when no calibration procedure was applied (Figure 4.5). As expected, we observed that the range of intensities was largely inconsistent across the three datasets, and was especially different between the KIRBY21 dataset (Figure 4.5C) and the IXI datasets (Figure 4.5A,B). The inter-subject variability, expressed by the standard error calculated bin-by-bin across histograms, was also uneven among the three datasets. These results suggested that, although the T1-w/T2-w image can permit to map myelin distribution in an individual brain, an intensity calibration is necessary to enable meaningful comparisons across datasets.

Figure 4.5 – Histograms of T1-w/T2-w image intensities before calibration. Mean T1-w/T2-w histograms (with standard deviation in colored shade) are presented for the IXI 1.5 T (A), IXI 3 T (B) and KIRBY21 3 T (C) datasets. The three datasets have inconsistent T1-w/T2-w intensity values. (A) IXI 1.5 T dataset shows a relatively large inter-subject reproducibility with the smallest extent of standard deviation. (B) IXI 3 T dataset displays a similar trend with an increased standard deviation in correspondence to the GM peak. (C) KIRBY21 3 T dataset exhibits the greatest inter-subject variability especially in the right tail of the histogram.

Before using our workflow to standardize the T1-w/T2-w, we first evaluated how the bias affected T1-w and T2-w images separately, and to what extent the bias correction procedure improved the similarity of images belonging to different datasets. Visual inspection of the data suggested that T1-w images were more affected than T2-w images by the spatial bias, and in particular the latter in 3 T datasets had larger magnitude than that in the 1.5 T dataset (Figure 4.6). Importantly, we found that the image histograms of the T1-w images were variable across datasets, and the bias correction procedure strongly reduced this variability (Figure 4.6A,C,E). In turn, no major change in the image histogram was produced for the T2-w images (Figure 4.6B,D,F). Overall, this analysis suggests that bias correction step, independently implemented on T1-w and T2-w, can potentially improve the reproducibility of T1-w/T2-w histograms.
Chapter 4

Figure 4.6 – Effect of bias correction on T1-w and T2-w images. A modality-dependent bias correction procedure was performed for both T1-w and T2-w. To illustrate the relative results, we analyzed the T1-w (A,C,E) and T2-w histogram (B,D,F) of three representative subjects from the IXI 1.5 T (Subject 002), IXI 3 T (Subject 093) and KIRBY21 (Subject 30) databases, respectively. Specifically, we compared the histograms before (blue line) and after bias-correction (red line). Before correction, the T1-w image of IXI 1.5 T dataset (A) was less biased than the IXI 3 T (C) and the KIRBY21 3 T (E), with an average correlation of $r=0.73$. After bias correction, this correlation increased to $r=0.89$. Conversely, in the T2-w image minor changes were observed, with an average correlation between histograms being $r=0.96$ and $r=0.97$ before and after bias correction, respectively.
Next, we applied the linear calibration algorithm to the bias-corrected T1-w and T2-w images (see Figure 4.2), and we calculated again T1-w/T2-w images for the three datasets. Notably, the calibrated T1-w/T2-w image histograms (Figure 4.7) exhibited comparable intensity scales and reduced inter-subject variability within each dataset.

Figure 4.7 – Histograms of intensities in calibrated T1-w/T2-w images. Mean T1-w/T2-w histograms (with standard deviation in colored shade) are presented for the IXI 1.5 T (A), IXI 3 T (B) and KIRBY21 3 T (C) datasets. In agreement with the scaling algorithm, the calibrated T1-w/T2-w images exhibit comparable intensity scales with a reduced inter-subject variability for each dataset. Note the normalized frequency on the vertical scale obtained as the ratio of each subject-specific histogram to the total area beneath the curve.

A quantitative analysis conducted on white matter voxels revealed that T1-w/T2-w image values were significantly more aligned after calibration. Specifically, a clear decrease of inter-subject variability for all three datasets confirmed the effectiveness of our approach (Table 4.3). After checking that the values were normally distributed (Lilliefors test, p<0.05), we also performed a single-factor ANOVA in order to assess the correspondence of the T1-w/T2-w means in the three datasets. The differences were significant before (F(2,62)=568.48, p<0.001), but not after calibration (F(2,62)=1.54, p=0.2236), further suggesting that the calibration procedure improved the reproducibility of the T1-w/T2-w values across datasets. To assess the spatial distribution of the T1-w/T2-w values, we also calculated the average T1-w/T2-w image for each of the three datasets. We set a common colormap scale to highlight potential differences among intensities in the three resulting images. Even so, we observed a very consistent spatial pattern among datasets, with no outlying features (Figure 4.8).

Table 4.3 – T1-w/T2-w reliability assessment.

<table>
<thead>
<tr>
<th></th>
<th>IXI database (1.5 T)</th>
<th>IXI database (3 T)</th>
<th>KIRBY21 dataset (3 T)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>before calibration</td>
<td>after calibration</td>
<td>before calibration</td>
</tr>
<tr>
<td>mean</td>
<td>3.44</td>
<td>2.11</td>
<td>4.52</td>
</tr>
<tr>
<td>sd</td>
<td>0.31</td>
<td>0.15</td>
<td>0.82</td>
</tr>
<tr>
<td>t-score</td>
<td>50.0</td>
<td>64.9</td>
<td>25.3</td>
</tr>
</tbody>
</table>

We quantitatively analyzed T1-w/T2-w values in the WM, in order to assess the potentially increased reproducibility across subject and scanners. Specifically, we compared the numerical mode of the T1-w/T2-w values in the WM before and after retrospective calibration, by means of descriptive (mean and standard deviation) and inferential statistics (t-score).
As a last validation step, we compared the calibrated T1-w/T2-w images with other images, namely FA, MTR and FLAIR images, obtained from the same subjects (Figure 4.9). By means of one-sample t-tests, we specifically tested the across-subject reproducibility and sensitivity of the four image modalities in detecting myelin-related signals. This statistical analysis conducted on different ROIs revealed that T1-w/T2-w had large reproducibility (indicated by large t-scores), which was mostly consistent across all selected white matter structures (Figure 4.10).
The greatest t-score value was measured in the anterior limb of internal capsule, which is consistent with the largest myelin concentration revealed by other studies (Whittall et al., 1997). As expected, the gray matter (GM) ROIs had lower values than white matter structures, both in terms of mean values (Figure 4.11) and t-scores (Figure 4.10). Overall, the t-scores obtained for MTR were inferior to those of T1-w/T2-w, but these two modalities showed a good similarity both in terms of white matter and deep GM structures. Also, we observed high FA values, comparable on average to those of T1-w/T2-w, but much more uneven across brain regions. In structures with multiple fiber crossing, e.g. the anterior corona radiata, FA values were lower than those of T1-w/T2-w and MTR. Conversely, regions with the greater anisotropy, such as the genu and splenium of the corpus callosum, exhibited larger t-scores than T1-w/T2-w. With the exception of the superior corona radiata, FLAIR results had negative t-scores. This is consistent with the specific FLAIR image property, for which more myelinated areas have darker contrast than less myelinated ones. On the other hand, FLAIR images were characterized by low absolute t-score values, indicating a relatively low reliability and sensitivity for myelin mapping.

Figure 4.9 – Single-subject images: comparison between image modalities. T1-w/T2-w (A), FA (B), MTR (C) and FLAIR (D) images for Subject 30 of the KIRBY database are shown in an axial section. Since conventional MRI images have arbitrary intensity scales, the four modalities are scaled according to the 1th and 99th percentiles. Note that higher intensity values in T1-w/T2-w, FA and MTR characterize structures with a greater degree of myelination, whereas an inverted intensity scale defines the FLAIR technique.
Chapter 4

106

Figure 4.10 – Inter-subject reproducibility: ROI analysis. The across-subject reproducibility of the T1-w/T2-w images as compared to MTR, FA and FLAIR were evaluated on specific ROIs. The analysis was conducted on six WM structures and three subcortical GM deep nuclei with putatively high and low myelin content, respectively.

Figure 4.11 – T1-w/T2-w intensities in GM and WM ROIs. Average and standard deviation of T1-w/T2-w values are shown for the nine selected ROIs, three in the GM and six in the WM. As expected, T1-w/T2-w values were lower in the GM than in the WM. The ROIs are labeled as follows: caudate nucleus (CAU), putamen (PUT), thalamus (THA), superior corona radiata (SCR), pontine crossing tract (PCT), anterior limb of internal capsule (ALIC), splenium of corpus callosum (SCC), anterior corona radiata (ACR) and genu of corpus callosum (GCC).
To corroborate our T1-w/T2-w results on selected ROIs, we repeated the same analysis based on t-scores at the single voxel level. The resulting t-score map (Figure 4.12) showed the six WM structures used in the ROI analysis, but not the GM ones, to have significantly larger T1-w/T2-w values than the average in the brain. We also observed additional structures to have significant t-score values, among the posterior thalamic radiation, inferior longitudinal fasciculus, corticospinal tract, middle cerebellar peduncle and red nucleus.

Figure 4.12 – Inter-subject reproducibility: whole-brain analysis. The across-subject reproducibility of the T1-w/T2-w images were evaluated on a voxel-wise basis. The maps illustrate regions with significant t-scores (p<0.05 FDR-corrected) over coronal, sagittal or axial sections of a standard MNI template. The following sections, referring to the MNI coordinate system, are shown: x=3 (A), x=17 (B), y=-20 (C), y=-2 (D), z=5 (E), z=-39 (F). Selected structures with significant t-score are indicated using arrows: anterior corona radiata (ACR), superior corona radiata (SCR), corona radiata (CR), anterior limb internal capsule (ALIC), posterior limb internal capsule (PLIC), internal capsule (IC), external capsule (EC), posterior thalamic radiation (PTR), genu of corpus callosum (GCC), splenium of corpus callosum (SCC), inferior longitudinal fasciculus (ILF), corticospinal tract (CST), pontine crossing tract (PCT), middle cerebellar peduncle (MCP) and red nucleus (RN).
4.4 Discussion

In this study, we have optimized the T1-w/T2-w methodology for non-invasive myelin mapping such that inference can be drawn at group level. Our retrospective calibration procedure yielded consistent ranges of T1-w/T2-w intensities across datasets, and this may enable potential comparisons and meta-analyses across different studies and individuals. Moreover, our statistical analyses suggested that T1-w/T2-w may be a more sensitive tool for myelin imaging than MTR, FA and FLAIR, and may therefore have future clinical applications.

4.4.1 Methodological considerations on the T1-w/T2-w approach

The T1-w/T2-w approach was originally proposed by Glasser and Van Essen (2011), who showed how the contrast related to myelin content can be increased by performing the simple ratio between T1-w and T2-w images (Glasser et al., 2013a). An important caveat of this approach is that the sensitivity profile of T1-w and T2-w images should be similar to yield a reliable T1-w/T2-w image. According to Belaroussi et al. (2006), this is an unlikely scenario, and this is also confirmed by our analyses (see Figure 4.6). The different image bias between the T1-w and T2-w images of the same subject might primarily depend on the fact that pulse sequences, and in particular the repetition time (TR) and the number of echoes, significantly influence the spatial uniformity of image intensities (Belaroussi et al., 2006). To address the issue of different image sensitivity in T1-w and T2-w images, we have included a bias-correction step in our analysis workflow (Figure 4.2). This substantially attenuates the slowly changing and smooth spatial variation in signal intensity that depends on the scanning hardware, the imaging parameters and the subject themselves (Belaroussi et al., 2006; Vovk et al., 2007), thereby leading to a more reliable intensity calibration.

From a methodological point of view, the image normalization is probably the most important step in our processing workflow for the T1-w/T2-w image. Frequently, a qualitative comparison between different images is achieved with an internal scaling of intensity values. This procedure, normally known as histogram equalization, consists of rescaling the image on the basis of the whole brain intensity distribution only. In this case, a color palette can be used for a visual evaluation of the image (Glasser and Van Essen, 2011), but no quantitative analysis across different images can be conducted. In general, a prospective approach permitting quantitative analyses on data produced by a single MR scanner is the use of a phantom-based calibration (Tofts, 1998). On the other hand, a retrospective approach would nevertheless be needed to perform quantitative multi-scanner comparisons. On grounds of these considerations, we implemented a retrospective calibration using image values from outside the brain. This involved the definition of reference T1-w and T2-w intensity values in the eye and temporal muscle masks to obtain a calibration curve. By using a linear scaling, we aimed to translate the intensity scale of a single image into a set of standardized values. The comparison of image histograms within and across the three datasets confirmed the effectiveness of our retrospective calibration. The two IXI datasets had the same scanning
parameters, but they were collected with a 1.5 T and 3 T MR scanners, respectively (Table 4.2). This may be the reason why their T1-w/T2-w images spanned a different range of values (Figure 4.5). Also, the KIRBY21 dataset deviated consistently from the other two, showing an altered pattern mainly on the right tail of the histogram, likely because of the different pulse sequence parameters (Table 4.2). As a matter of fact, variations of repetition time (TR) and echo time (TE) may yield different histogram distributions. In addition to these differences between datasets, large differences within datasets were also evident before calibration. These differences may be due to instrumentation factors, such as temperature and humidity, or by interactions with the subject’s tissues. After calibration, the T1-w/T2-w histograms had comparable intensity scale and a similar standard deviation across datasets (Figure 4.7), suggesting that differences in intra- and between-dataset reproducibility were substantially reduced. Furthermore, the consistency of representative T1-w/T2-w values for the WM across subjects confirmed this finding in a quantitative manner (Table 4.3). The observed effectiveness of the calibration procedure to standardize T1-w/T2-w values across subjects opened up the way to numerical analyses focused on the reliability of the T1-w/T2-w approach with respect to other myelin-related imaging techniques (Figure 4.10).

4.4.2 Myelin-related information in T1-w/T2-w images

Previous studies documented that myelin is distributed unevenly between WM and GM structures (Barkovich, 2005; Paus et al., 2001). Thus, we clustered these structures in two groups to assess the specificity of the T1-w/T2-w technique. The analysis that we conducted on selected ROIs showed high T1-w/T2-w scores in those WM structures where myelin is most abundant (Barkovich, 1988; Kizildag et al., 2005; Leppert et al., 2009; Welker and Patton, 2012). In large accordance with our T1-w/T2-w results, previous studies reported a high degree of myelination for projection fibers, e.g. the internal capsule, corona radiata, and commissural fiber tracts including the genu and the splenium of the corpus callosum (Barkovich, 2000; Deoni et al., 2011; Rademacher et al., 1999; Steenweg et al., 2010). Furthermore, the T1-w/T2-w value in the anterior limb of internal capsule was the highest among all investigated WM structures, which is concordant with previous reports on the spatial distribution of myelin in the brain (Whittall et al., 1997). As for the GM structures, the thalamus exhibited higher T1-w/T2-w scores than did the putamen and the caudate nucleus, corroborating results reported in previous studies (Madler et al., 2008; Whittall et al., 1997). Overall, the results of our ROI analysis for the T1-w/T2-w were also consistent with previous T2-multicomponent relaxation and MTR studies. For example, by using T2-multicomponent relaxation, Vidarsson et al. (2005) found the greatest values of myelin water fraction in the internal capsule, genu and splenium of corpus callosum. Markedly reduced myelin content was also found in the putamen (Vidarsson et al., 2005). Smith et al. (2006) reported high MTR values in correspondence of densely packed WM regions, such as the callosal fibers and the internal capsule as compared to less densely packed structures. On the other hand, they found lower, but not negligible values in GM structures, such as putamen and caudate nucleus, in accordance with our findings (Smith et al., 2006). Since MTR is one of the most widely used techniques to study myelination (Schmierer et al., 2004),
the correspondence that we observed in terms of t-scores between T1-w/T2-w and MTR in our study (Figure 4.10) may be considered as an indirect evidence for the potential effectiveness of T1-w/T2-w for quantitative myelin mapping.

A substantial difference in the ROI analysis results was found between T1-w/T2-w and FLAIR, with overall lower values for the latter modality. The FLAIR technique was previously employed for qualitative analyses on pathological processes related to myelination (Ashikaga et al., 1999; Murakami et al., 1999), but to the best of our knowledge it has not been employed in quantitative studies. Specifically, our comparative analysis showed that FLAIR images had relatively low t-scores in both WM and GM structures. Accordingly, myelin assessment may not be considered the key hallmark of this technique.

Another important finding in our ROI analysis was that T1-w/T2-w and FA values were generally high, but T1-w/T2-w scores were substantially more uniform than FA across WM structures. In first instance, this might be interpreted as evidence that FA is a sensitive technique to detect quantitative differences between regions. Nonetheless, closer inspection of FA t-scores across ROIs indicates that the FA variability may be partly due to the crossing fibers problem (Madler et al., 2008; Wedeen et al., 2008), which specifically affects DTI-derived measures. In line with previous studies (Barkovich, 2005; Provenzale et al., 2007), FA was indeed found to be high in structures with a highly organized fiber placement, such as the corpus callosum and the internal capsule, whereas it was relatively lower in regions where fibers with different orientation cross, as for example in a significant portion of the anterior corona radiata (see Figure 4.10) (Assaf and Pasternak, 2008; Wedeen et al., 2008). This is in agreement with the proposal that, although myelin sheets contribute to anisotropy, other factors such as axonal membrane might substantially contribute to large FA values (Beaulieu, 2002; Huang et al., 2006).

4.4.3 Potential limitations of the method

Our analyses suggested that our T1-w/T2-w workflow may be potentially useful for the myelin mapping in the human brain. Nonetheless some potential limitations of our study should be acknowledged. A first limitation is that only a limited number of datasets were used in this study. Images collected with very different pulse sequences may generate inconsistent results in terms of image contrast. Accordingly, the effectiveness of this approach for meta-analyses should be assessed in future studies, by examining a broader range of datasets. Secondly, our calibration procedure strongly depends on the accuracy of the calibration masks, which is in turn influenced by the effectiveness of the spatial warping from the MNI space to the subject space. To address this issue, we have extracted calibration values using the numerical modes of the mask intensity distributions. This is likely to mitigate the problem of the mask definition. It is also worth noting that our T1-w/T2-w calibration relies on the assumption that across-subject variability in the tissue selected through the masks is negligible compared to the potential differences that can be observed across the brain of different subjects. In this regard, our analysis on healthy subjects yielded largely similar T1-w/T2-w image histograms, thereby suggesting that such an assumption may generally hold. Furthermore, we utilized external calibration using eye and temporal
muscle masks in alternative to internal calibration because the latter type of scaling could hide quantitative differences between healthy groups and those with altered myelin. However, diseases that cause altered myelin levels also might affect the external calibration points, e.g. temporal muscle wasting/composition change. In this case, differences between healthy and pathological groups would be underestimated or overestimated using an external calibration approach. Another aspect to be considered is that the T1-w/T2-w image in diseased individuals may be altered not only due to demyelination, but also to edema, inflammation, iron accumulation or atrophy. This needs to be further investigated by using information from histological samples. Finally, we could not compare the T1-w/T2-w technique with all existing MR techniques for myelin mapping. Future studies should be conducted, for instance, to quantitatively compare T1-w/T2-w and mcDESPOT modalities. Additionally, a comparison with histological samples would be warranted to evaluate T1-w/T2-w as a sensitive marker of myelin content.

4.5 Conclusion

In this study, we implemented a new analysis workflow for the standardization of T1-w/T2-w images, thereby enabling the use of the T1-w/T2-w technique for a non-invasive mapping of myelin at group level. Our statistical analyses on selected ROIs suggested that T1-w/T2-w may permit extracting reliable information on myelin distribution, with potentially larger sensitivity than other techniques such as MTR, FA and FLAIR. Future work is warranted to examine the potential utility of the T1-w/T2-w technique for myelination studies on development and aging, as well as for comparative investigations between healthy individuals and patients with neurological and psychiatric disease.
Chapter 5

MAPPING PATHOLOGICAL CHANGES IN BRAIN STRUCTURE BY COMBINING T1- AND T2-WEIGHTED MR IMAGING DATA

A workflow based on the ratio between standardized T1-weighted (T1-w) and T2-weighted (T2-w) MR images has been proposed as a new tool to study brain structure. This approach was previously used to map structural properties in the healthy brain. Here we evaluate whether the T1-w/T2-w approach can support the assessment of structural impairments in the diseased brain. We use schizophrenia data to demonstrate the potential clinical utility of the technique. We analyzed T1-w and T2-w images of 36 schizophrenic patients and 35 age-matched controls. These were collected for the fBIRN collaborative project, which had IRB approval and followed HIPAA guidelines. We computed T1-w/T2-w images for each individual and compared intensities in schizophrenic and control groups on a voxel-wise basis, as well as in regions of interest (ROIs). Our results revealed that the T1-w/T2-w image permits to discriminate brain regions showing group-level differences between patients and controls with greater accuracy than conventional T1-w and T2-w images. Both the ROIs and the voxel-wise analysis showed globally reduced gray matter (GM) and white matter (WM) values in patients compared to controls. Significantly reduced values were found in regions such as insula, primary auditory cortex, hippocampus, inferior longitudinal fasciculus and inferior fronto-occipital fasciculus. Our findings were consistent with previous meta-analyses in schizophrenia corroborating the hypothesis of a potential “disconnection syndrome” in conjunction with structural alterations in local GM regions. Overall, our study suggested that the T1-w/T2-w technique permits to reliably map structural differences between the brains of patients and healthy individuals.

Published as:
Chapter 5

5.1 Introduction

The non-invasive detection of altered brain anomalies in neurological and psychiatric patients largely relies on magnetic resonance imaging (MRI) data. In particular, images from conventional T1-weighted (T1-w) and T2-weighted (T2-w) MRI, being characterized by high contrast and spatial resolution, have been widely used to investigate structural properties of the brain (Raz et al., 2010; Steenweg et al., 2010; Welker and Patton, 2012; Zhou et al., 2013). Since being introduced into clinics several decades ago, the interpretation of T1-w and T2-w images is presently a routine task for the majority of clinicians. Importantly, the two images differ with respect to their intrinsic sensitivities towards biophysical properties of brain tissues. The T1-w image is thus more weighted towards the predisposition of a tissue to absorb energy into its lattice whereas the T2-w image is more influenced by spin-spin interaction processes (Westbrook et al., 2011). This implies that, in the case of brain pathologies, structural alterations resulting from edema, inflammation, tumor infiltration, iron accumulation, or atrophy, will manifest themselves differently in the two kinds of images. However, T1-w and T2-w images also have some common limitations, which primarily relate to: 1) the presence of spatial inhomogeneities in image intensity, produced by interactions between the subject’s body and the MR scanner, and 2) the lack of image intensity standardization, mainly related to the MR instrumentation and acquisition sequence used. These factors limit the use of conventional structural MR images for the study of pathological changes in brain structure.

Nowadays, clinicians most often tend to infer brain tissue integrity through visual comparison of T1-w and T2-w images, when both are available. Also, a number of clinical studies have combined the information from the two kinds of images to study, for instance, multiple sclerosis (van Walderveen et al., 2001) or mesial temporal lobe epilepsy (Braga et al., 2012). The parallel analysis of T1-w and T2-w images represents a first attempt to implement a multi-modal structural imaging approach for the study of brain tissue abnormalities. However, direct fusion of the two kinds of images, so as to produce a multi-modal T1-w/T2-w image (Ganzetti et al., 2014; Glasser and Van Essen, 2011) could prove to be an interesting alternative. In particular, T1-w sequences are characterized by better contrast-to-noise ratio in WM regions, whereas T2-w sequences can be used to better discriminate structural differences in fluid-filled regions (Braga et al., 2012). Therefore, the advantages of each of these MRI techniques should be taken into account when designing clinical investigations focused on an unknown spectrum of brain pathologies. Furthermore, the comparison of different imaging modalities is advisable when no clear hypothesis can be made about which brain structures is likely to be altered.

In this study, we apply the T1-w/T2-w analysis workflow (Ganzetti et al., 2014) to structural MRI data collected in individuals with schizophrenia and age-matched healthy controls, then compare the resulting images between groups to verify whether the technique can detect disease-related structural impairment. Since T1-w and T2-w images are most often collected in clinical scanning protocols and do not require long acquisition times, the T1-w/T2-w technique may become a widely-used tool for mapping potential pathological changes resulting from brain disease.
5.2 Methods

5.2.1 Structural MR data

To validate the T1-w/T2-w approach for the detection of structural abnormalities in patients, we used MR data obtained from the fBIRN Phase II dataset (fbirnbdr.birncommunity.org:8080/BDR). This was contributed by a consortium of brain imaging centers in the US to explore the full potential of a multi-site neuroimaging approach to elucidate clinical and cognitive abnormalities in schizophrenia. The dataset includes structural and functional MR imaging (fMRI) data, as well as behavioral data, demographic information, and clinical assessments, collected in large samples of schizophrenia patients and age-matched healthy controls. Patients with schizophrenia or schizoaffective disorder were diagnosed based on the clinical interview of the Diagnostic and Statistical Manual of Mental Disorders, fourth version revised (DSM-IV-TR). Controls subjects were excluded if they had a current or past history of a major neurological, psychiatric, medical illness, previous head injury, substance or alcohol dependence or an IQ below 75. MR imaging data included structural T1-w and T2-w scans, two cognitive fMRI tasks (auditory oddball and serial item recognition tasks), two calibration fMRI scans (breath hold and sensory motor tasks), and a resting state fMRI scan. Informed consent was obtained from the participants in this study, which had itself received ethical approval from the relevant institutional review board (IRB). In the current study, we retrospectively evaluated T1-w and T2-w images from 36 patients (10 female, 40.4±11.8 years old) and 35 controls (14 female, 39.1±12.9 years old), coming from sites coded as 0006, 0010 and 0018. MR data from other sites were not included as they either did not contain both T1-w and T2-w images for the same subject or the scan did not cover the whole brain. MR data extracted from the sites 0006 and 0018 were collected with a 3 T Siemens Trio, whereas the data extracted from the site 0010 were collected with a 1.5 T Siemens Trio scanner. MR images were acquired in both patients and controls with the same scanner, using the same scanning parameters. T1-w scans were acquired in the sagittal plane using a magnetization-prepared rapid gradient echo (MPRAGE) sequence. The scanning parameters were: 256×256 matrix, 24 cm FOV, 160-170 slices, slice thickness 1.2-1.5 mm. T2-w scans were acquired using a Turbo Spin Echo (TSE) sequence in the oblique axial plan tilted to be parallel to the AC-PC plane. The scanning parameters were: 256×192 matrix, 22 cm FOV, 27 slices, slice thickness 4 mm with 1 mm gap.

5.2.2 Imaging analysis

T1-w and T2-w images were preprocessed and combined using a dedicated workflow as described in (Ganzetti et al., 2014). This includes bias correction and intensity calibration on each of the two images, and the subsequent calculation of the ratio between preprocessed T1-w and T2-w images (Figure 5.1).
Chapter 5

The entire processing of the T1-w/T2-w image was conducted using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK). Initially, the original T2-w image was coregistered to the T1-w image through a rigid-body transformation (Collignon et al., 1995). Then, the T1-w and T2-w images were separately subjected to bias correction using the ‘New Segmentation’ tool implemented in SPM8 (Ashburner and Friston, 2005; Weiskopf et al., 2011). The input parameters for the bias correction algorithm, namely the smoothing and the regularization parameters, were set at their default value. After correcting for intensity non-uniformity, the T1-w and T2-w images were further processed to standardize their intensity by using a linear scaling procedure systematically described in Ganzetti et al. (Ganzetti et al., 2014). Eventually, after calibrating T1-w and T2-w images their ratio was calculated to generate the calibrated T1-w/T2-w image. To conduct statistical comparisons between groups, we spatially transformed the T1-w/T2-w image in subject space to MNI space and carefully checked the effectiveness of this registration (Figure 5.2).

**Figure 5.1** – Processing workflow for the generation of the T1-w/T2-w image. The T1-w and T2-w images are subjected to bias correction to remove the slow intensity variations related not only to the MR hardware but also its interaction with the subject’s cranial tissue. Subsequently, the bias-corrected images are calibrated using intensities from masks outside the brain, thereby avoiding problems related to the use of an internal calibration. Finally, the multi-modal T1-w/T2-w image is calculated as the ratio of the standardized T1-w and T2-w images.

**Figure 5.2** – GM and WM maps for the fBIRN dataset images. We tested the registration to MNI space of images from healthy controls and schizophrenia patients. To this end, the grey and white matter masks in subject space were transformed to MNI space by applying the spatial normalization estimated from T1-w and T2-w images. The MNI-space masks belonging to the control and schizophrenia groups were then averaged across subjects to create group-specific probability maps (with values between 0 and 1). The probability maps associated with the MNI template are shown for comparison.
5.2.3 Statistical analysis

First of all, we assessed the T1-w/T2-w technique against other processing strategies and modalities. To this end, we conducted an analysis using small regions of interest (ROIs). Using the MarsBaR toolbox (marsbar.sourceforge.net), we created spherical ROIs of 6mm radius, centered over the coordinates reported in a recent meta-analysis (Ellison-Wright and Bullmore, 2010) summarizing structural alterations in schizophrenia as revealed by a conspicuous number of medical imaging studies (see Table 5.1 for details).

Ta 5.1 – Regions of interest: MNI coordinates.

<table>
<thead>
<tr>
<th>ROI name</th>
<th>MNI coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Left Insula</td>
<td>-38</td>
</tr>
<tr>
<td>Right Insula</td>
<td>44</td>
</tr>
<tr>
<td>Left thalamus: medial dorsal nucleus</td>
<td>-3</td>
</tr>
<tr>
<td>Left frontal lobe: medial frontal gyrus (BA10)</td>
<td>-1</td>
</tr>
<tr>
<td>Left anterior cingulate (BA32)</td>
<td>-1</td>
</tr>
<tr>
<td>Left deep frontal lobe</td>
<td>-12</td>
</tr>
<tr>
<td>Right frontal lobe: medial frontal gyrus (BA9)</td>
<td>3</td>
</tr>
<tr>
<td>Left posterior cingulate (BA23)</td>
<td>-7</td>
</tr>
<tr>
<td>Right globus pallidus</td>
<td>18</td>
</tr>
<tr>
<td>Left caudate head</td>
<td>-5</td>
</tr>
</tbody>
</table>

Spherical ROIs of 6 mm radius were defined according to the meta-analysis of Ellison-Wright and Bullmore (2010). Coordinates of the sphere center are in MNI stereotaxic space.

We also included in this analysis the masks of GM, WM and cerebrospinal fluid (CSF), provided by the International Consortium for Brain Mapping (www.loni.usc.edu/ICBM). For each ROI, we used an unpaired t-test to test the difference between the T1-w/T2-w values in the schizophrenia and control groups. Specifically, we calculated the average intensity across the voxels of a given ROI for n controls and m schizophrenic patients, to define the vectors C and SZ respectively. The t-score associated with the unpaired t-test between the two vectors was obtained using the formula:

$$t_{ROI} = \frac{\bar{C} - \bar{SZ}}{\sqrt{\frac{s_C^2}{n} + \frac{s_{SZ}^2}{m}}}$$

(5.1)

where $\bar{C}$ and $\bar{SZ}$ are the sample means and $s_C$ and $s_{SZ}$ the sample standard deviations for controls and patients, respectively. Statistical significance was assessed by using permutation-based non-parametric testing (Nichols and Holmes, 2002). For this analysis, we examined the results of the unpaired t-test on: 1) the ratio between the unprocessed T1-w and T2-w images; 2) the ratio between unbiased, but not
Chapter 5

...calibrated T1-w and T2-w images; 3) the ratio between calibrated, but not unbiased T1-w and T2-w images; 4) unbiased and calibrated T1-w images; 5) unbiased and calibrated T2-w images. Using t-values as an indicator of the capability of discriminating between schizophrenic and control groups, we evaluated the effects of bias correction and image calibration individually, as well as the enhancement of contrast introduced by the ratio between T1-w and T2-w images.

In a second step, we focused on T1-w/T2-w signal differences between patients and controls. We calculated image histograms and compared them for the two groups, in order to examine whether schizophrenia patients had globally increased/decreased values with respect age-matched healthy controls. Subsequently, we analyzed T1-w/T2-w values more in detail using an additional ROI analysis. The ROIs were based on masks for specific GM and WM regions. The GM regions included: frontal lobe, parietal lobe, occipital lobe, temporal lobe, insula, cerebellum, putamen, caudate, thalamus. These were defined using the ICBM probabilistic atlas provided by the International Consortium for Brain Mapping (www.loni.usc.edu/ICBM). The WM regions included: anterior thalamic radiations (ATR), superior longitudinal fasciculus (SLF), inferior longitudinal fasciculus (ILF), inferior fronto-occipital fasciculus (IFOF), pontine crossing tract (PCT), corticospinal tract (CST), forceps major, forceps minor, middle cerebellar peduncle (MCP). These ROIs were defined using the DTI-based WM atlas of the Laboratory of Brain Anatomical MRI, John Hopkins University School of Medicine, Baltimore, MD, USA (cmrm.med.jhmi.edu).

Eventually, we carried out a whole-brain analysis for a more exploratory investigation of regional differences in T1-w/T2-w intensity. In particular, we generated a t-score map from the T1-w/T2-w data by applying the formula in Eq. 5.1 to each voxel rather than to a ROI. We used defined a statistical threshold using permutation-based non-parametric inference (Nichols and Holmes, 2002), applying threshold-free cluster enhancement (TFCE) to account for multiple comparisons (Smith and Nichols, 2009). The resulting probability map (corrected for multiple comparisons) was thresholded at p=0.005, highlighting brain regions with significantly larger differences between the two groups.
5.3 Results

First of all, we used selected ROIs that are considered key regions in the schizophrenia pathology to validate T1-w/T2-w approach and accurately examine its effectiveness for the mapping of differences between patients and healthy individuals (Figure 5.3). We tested 10 ROIs in total, of which 8 were significant (p<0.05) when using the T1-w/T2-w images. Overall, the results of the unpaired t-tests comparing the T1-w/T2-w values of the schizophrenia and control groups demonstrated that the bias correction, and more considerably the external calibration, yielded an increased T1-w/T2-w signal differences between groups. The T1-w/T2-w workflow, which combines bias correction and image intensity calibration, indeed yielded higher t-scores than images obtained by using only one or none of these processing steps.

![Figure 5.3](image)

Notably, the t-scores obtained using only the calibration procedure were much closer to T1-w/T2-w t-scores than those obtained with only the bias correction. We observed a similar trend when using GM and WM as regions of interest. A remarkable result was that the difference between patients and controls in the CSF, where no structural differences were expected, was minimal for the T1-w/T2-w approach. We also assessed the importance of taking a multi-modal approach combining T1-w and T2-w images. Accordingly, we compared the T1-w/T2-w results with those obtained using the unbiased and
calibrated T1-w and T2-w images independently. In this case as well, we obtained higher t-scores for the T1-w/T2-w in the schizophrenia-specific regions, as well as in GM and WM. In turn, the t-score in the CSF was lower for the T1-w/T2-w image than for separate T1-w and T2-w images, suggesting that the ratio minimizes spatial inhomogeneities that may be unrelated to the structural properties of the brain and are present both in T1-w and T2-w images.

After this initial validation step, we focused on the calibrated T1-w/T2-w images, and examined the global distribution of their intensities in schizophrenic patients and age-matched healthy volunteers, respectively. The patient group revealed globally reduced values, along with increased inter-subject variability (Figure 5.4). More specifically, the reduction appeared more marked in regions with higher T1-w/T2-w intensity, which can be putatively attributed to the WM.

![Figure 5.4 – Histograms of T1-w/T2-w intensities in patients and controls. Mean T1-w/T2-w histograms (with standard deviation shaded) are shown for patients and controls, separately. Two main peaks are visible in the histograms of both groups. The intensity of the first peak (associated with the GM) is similar in patients and controls, whereas the second peak (associated with the WM) has a lower intensity in patients than in controls. Furthermore, the schizophrenic patients are characterized by a higher standard deviation than healthy subjects.](image)

To confirm this result, we analyzed average T1-w/T2-w intensities within GM, WM and CSF (Figure 5.5). Accordingly, we found significant T1-w/T2-w reductions in schizophrenia for GM and WM, with the latter yielding the largest difference compared to controls.
Figure 5.5 – T1-w/T2-w intensities in patients and controls for CSF, GM and WM. Mean T1-w/T2-w values (with standard deviation) are shown for patients and controls. Both GM and WM have a lower intensity in patients than in controls, with the second yielding the largest difference. Furthermore, no statistical difference was observed for the CSF (* = p<0.05 ; ** = p<0.01, *** = p<0.005 uncorrected).

We then focused our investigation on specific brain regions within the GM and WM for a local assessment of image intensities (Figure 5.6). No regions with T1-w/T2-w increase were found in schizophrenic patients. Among the GM regions, the frontal lobe, temporal lobe, and insula showed significant reduction (p<0.05). The parietal lobe, occipital lobe and cerebellum did not show a significant difference. As for the subcortical nuclei, only the thalamus showed a significant reduction (p<0.05), whereas no differences were found in putamen and caudate nucleus. Among the WM regions included in the analysis, the ATR, SLF, ILF, IFOF, forceps major, forceps minor and MCP, showed a significant reduction (p<0.05) in T1-w/T2-w values, whereas no significant difference was detected in PCT and CST.

To corroborate and extend our T1-w/T2-w results beyond the selected ROIs, we repeated the same analysis at the single-voxel level, thereby generating an unthresholded map of differences between groups (Figure 5.7). Schizophrenic patients exhibited hypointensity in areas thought to play a central role in the pathology. In terms of GM structures, we found T1-w/T2-w reductions in correspondence of the left hippocampal formation (HF), right inferior temporal gyrus (ITG), left superior temporal gyrus (STG), left visual cortex (VC) V2, right ventromedial prefrontal cortex (vmPFC), left primary auditory cortex (PAC), bilateral insular cortex (IC). As for WM tracts, decreased values were reported in the MCP, ILF, IFOF, uncinate fasciculus (UF), posterior thalamic radiations (PTR), splenium of corpus callosum (SCC), and callosal body (CB). The differences between schizophrenia and control groups were particularly significant (p<0.005, TFCE corrected) in the left temporal regions, and specifically in PAC, HF and IC for the GM, and in ILF and IFOF for the WM (Figure 5.8).
**Figure 5.6** – T1-w/T2-w intensities in patients and controls: GM and WM regions. Mean T1-w/T2-w values (with standard deviation) are shown for patients and controls. A global reduction of GM (a) and WM (b) intensity was reported. In terms of GM structures, the temporal lobe showed the largest significant reduction. Among the subcortical nuclei, caudate and thalamus exhibited decreased values. No statistical difference was revealed in the parietal lobe and putamen. As for WM tracts, the left hemisphere along with the frontal region showed the largest reduction (* = p<0.05; ** = p<0.01, *** = p<0.005 uncorrected).
Figure 5.7 — T1-w/T2-w differences between schizophrenia patients and healthy controls: exploratory analysis. T1-w/T2-w differences in schizophrenic subjects compared with age-matched healthy controls were assessed on a voxel-wise basis using an unpaired t-test. Regions with significantly greater T1-w/T2-w values in controls are shown in yellow-white. No regions were found with significantly greater T1-w/T2-w values in patients. The maps were represented using coronal (a), sagittal (b) and axial (c) sections. Fundamental structures with high t-scores are indicated using arrows: inferior longitudinal fasciculus (ILF), middle cerebellar peduncle (MCP), inferior fronto-occipital fasciculus (IFOF), insular cortex (IC), hippocampal formation (HF), uncinate fasciculus (UF), superior temporal gyrus (STG), inferior temporal gyrus (ITG), visual cortex V2 (VC V2), ventromedial prefrontal cortex (vmPFC), posterior thalamic radiations (PTR), primary auditory cortex (PAC), splenium of corpus callosum (SCC), callosal body (CB).
Figure 5.8 – T1-w/T2-w differences between schizophrenia patients and healthy controls: thresholded map. T1-w/T2-w differences in schizophrenic subjects compared with age-matched healthy controls were assessed on a voxel-wise basis using an unpaired t-test. Regions with significantly reduced T1-w/T2-w values in patients (p<0.005, TFCE corrected) are shown in red. No regions were found with significantly greater T1-w/T2-w values in patients. The maps were overlaid on a standard MNI template, and represented using coronal (a), sagittal (b) and axial (c) sections. Selected structures with significant t-scores are indicated using arrows: primary auditory cortex (PAC), hippocampal formation (HF), insular cortex (IC), inferior longitudinal fasciculus (ILF), inferior fronto-occipital fasciculus (IFOF).

5.4 Discussion

The T1-w/T2-w approach was initially proposed by Glasser and Van Essen (Glasser and Van Essen, 2011) for a qualitative mapping of the myelin present in the GM of an individual subject. It was subsequently further developed by Ganzetti et al (Ganzetti et al., 2014) for quantitative assessment of myelin content in the whole brain. In this current study, we evaluated the potential utility of the T1-w/T2-w approach by conducting an analysis on T1-w and T2-w data collected in schizophrenic patients and age-matched controls. A considerable number of studies have focused on the etiology behind structural alterations in schizophrenic patients, though no consensus has yet been fully reached. As for the GM, detailed examinations of tissue reductions in the schizophrenic brain have been commonly achieved by voxel-based morphometry (VBM) (Ellison-Wright and Bullmore, 2010), while diffusion tensor imaging (DTI) has obtained particular attention for the mapping of reduced anisotropic patterns within WM tracts possibly linked to myelination changes (Ellison-Wright and Bullmore, 2009; Kubicki et al., 2007). Nonetheless, cortical atrophy (Bonilha et al., 2008; McCarley et al., 1999; Shenton et al., 2001), neuroinflammation (Jeste et al., 2011), degradation of the myelin sheath (Kubicki et al., 2005; Kyriakopoulos et al., 2008) and accumulation of extra-cellular water (Kochunov et al., 2014), have most frequently been...
related to the disease. All these factors can, in principle, be detected as T1-w/T2w intensity differences between patients and healthy individuals.

We began our investigations by selecting a set of brain regions that are most robustly found to be impaired in schizophrenia patients, based on the recent meta-analysis of Ellison-Wright and Bullmore (Ellison-Wright and Bullmore, 2010). The results of our ROI analysis showed that T1-w/T2-w image had increased T1-w/T2-w signal differences between groups, compared to images obtained with either bias correction or linear calibration. Furthermore, our analysis also suggested that the ratio between the bias-corrected and calibrated T1-w and T2-w images yielded larger detection accuracy with respect to the single modalities taken separately. Not only was the ratio between T1-w and T2-w critical for detecting true positives, but also allowed for an attenuation of any putatively detrimental component shared between the two modalities. As shown by the CSF results, subsequent pipeline stages generate smaller difference when comparing schizophrenic patients and healthy controls, whereas in those regions characterized by a high probability of finding structural alterations, the difference becomes significant. Furthermore, the results of the analysis conducted using WM and GM as regions of interest (Figure 5.5) suggested a global reduction of T1-w/T2-w intensity, with the first one showing relatively more reduced values in schizophrenia. This is in line with several functional and structural studies suggesting that psychotic and cognitive symptoms are linked to abnormalities of cortical tissues, such as in limbic and frontal-temporal networks (Assaf et al., 2006; Buchanan et al., 2004; Ellison-Wright and Bullmore, 2010; Goghari et al., 2010; Honea et al., 2005; Lawrie and Abukmeil, 1998). Likewise, the presence of differences in WM implies fiber tract impairment as a major key feature in the pathology (Davis et al., 2003; Pomarol-Clotet et al., 2010). Thus, functional dysfunctions in brain networks may result from structural changes both in WM and in GM (Bassett and Bullmore, 2009; Hugdahl et al., 2009). These results are in line with current studies suggesting that global disruptions in WM connectivity combined with local GM alterations may be the foundation for the clinical and cognitive outcome of the pathology (Kubicki et al., 2007; Kubicki et al., 2005; Kyriakopoulos et al., 2008).

In terms of local structural impairments in GM (Figures 5.6 and 5.7), we found agreement between our findings and those of voxel-based morphometric studies, in particular for insula, thalamus, hippocampal formation (HF), temporal and frontal lobe (Chan et al., 2011; Ellison-Wright and Bullmore, 2010; Ellison-Wright et al., 2008; Glahn et al., 2008; Honea et al., 2005). The results of the ROI and voxel-wise analyses were in large agreement, with the voxel-wise analysis revealing additional areas such as HF and PAC (Figure 5.8). Our analyses revealed prominent alterations in schizophrenic patients to be localized in the left hemisphere, in correspondence of the PAC (Gupta et al., 2014; Honea et al., 2005; Rajarethinam et al., 2000; Sun et al., 2009; Takahashi et al., 2006) (Figure 5.8). Weakened activity in this brain region was also reported by Gaser and coworkers (Gaser et al., 2004), who suggested that abnormalities in this area may underlie hallucination-specific disorders (Kompus et al., 2013). Regarding subcortical structures, no significant difference was reported in the putamen and caudate (Figure 5.6). However, we found a significant T1-w/T2-w reduction in the right globus pallidus (Figure 5.3). Several structural studies showed alterations in the basal ganglia, among which the globus pallidus, to underlie
problems in higher cognitive functions such as attention and working memory, which are typically impaired in schizophrenia (Brandt and Bonelli, 2008). Not only are our findings consistent with previous neuroimaging studies, but also with the main behavioral features of the pathology. For instance, we found significant differences in the frontal cortex, in particular the vmPFC, which has been implied in social cognition and especially in emotion regulation (Amodio and Frith, 2006; Grossmann, 2013). Also, alterations found in the insula are consistent with problems related to mental representation of emotions (Park et al., 2008).

As for the WM results, our analyses identified two major locations of T1-w/T2-w reduction in schizophrenic brains: one in the frontal areas and the other in the left temporal region (Figures 5.6 and 5.7). Similar results have also been reported in several meta-analyses, showing consistent WM aberrations in prefrontal areas pathways (Ellison-Wright and Bullmore, 2009; Kubicki et al., 2007; Kyriakopoulos et al., 2008). Both our ROI- and voxel-based results highlighted significantly reduced values in the left temporal lobe compared to the contralateral (Figure 5.8), with the inferior longitudinal fasciculus (ILF) and the temporal radiations of IFOF as the most affected in patients. Reductions of WM in patients were also reported in the cerebellum; in particular we observed a significant reduction in the middle cerebellar peduncle (MCP) (Figure 5.6B). These findings are consistent with previous structural and functional neuroimaging studies linking cerebellar deficits to neurological signs in schizophrenia, via impaired coordination of mental processes (Duan et al., 2015; Picard et al., 2008). Smaller cerebellar volumes for both hemispheres (Hirjak et al., 2015) along with WM volume reductions in the cerebellum (Kubicki et al., 2007; Singh et al., 2014) are often reported. Furthermore, the observed impairments in the ATR (Fig. 6b), which contains fibers interconnecting thalamus and prefrontal cortex, are consistent with the results of Mamah et al (Mamah et al., 2010), who documented that ATR integrity is inversely correlated with cognitive dysfunctions. In conclusion, our study mostly confirmed and extended the “disconnection syndrome” theory in the schizophrenic brain (Duan et al., 2015; Kubicki et al., 2007; Liang et al., 2006).

More generally, our current results from schizophrenia data suggest that the T1-w/T2-w technique can be more reliably used to map differences compared to healthy controls than methods based either on T1-w and T2-w images. A number of potential limitations of our study should be however considered. A first one is that we used only a limited amount of datasets for the method validation. Since images collected with very different pulse sequences may generate inconsistent results in terms of image contrast, we suggest the T1-w/T2-w pipeline should be used on a broader range of datasets to evaluate its potential use for large-scale neuroimaging analyses. Further investigations are warranted to confirm our preliminary findings and to explore the utility of the method on different brain pathologies. In line of principle, better spatial accuracy can be achieved with images acquired at high spatial resolution. Our results suggested that image resolution is not critical for the use of the T1-w/T2-w technique. The availability of T1-w and T2-w images at high spatial resolution is critical for the use of surface-based mapping methods. This would provide more accurate spatial registration across subjects as well as would permit evaluating cortical folding patterns and cortical thickness together with T1-w/T2-w image
intensity. The integration of surface-based methods in the pipeline for the T1-w/T2-w analysis may be an important methodological advancement.

5.5 Conclusion

In summary, our findings suggest that the T1-w/T2-w technique can be reliably used to map differences in brain structure between patients and healthy individuals with a greater accuracy than by using methods based on T1-w or T2-w images alone, supporting the definition of more reliable disease biomarkers. The use of other MR-based techniques (e.g. proton spectroscopy, magnetization transfer techniques and relaxation times) and positron emission tomography (PET) can certainly help to clarify the pathological processes associated with the detected T1-w/T2-w signal differences between patients and healthy individuals.
We developed a processing workflow based on T1-w and T2-w MR imaging to generate an optimized multimodal image. By means of a voxel-wise ratio between T1-w and T2-w images, the workflow allows whole brain structural mapping. Here we developed the T1-w/T2-w technique to primarily address issues of image standardization across different datasets and then evaluated for its potential utility in the quantitative assessment of structural alterations. The overall workflow includes three major pre-processing and data analysis steps: intensity non-uniformity (INU) correction, intensity standardization and T1-w/T2-w image generation. Technical limitations, future developments as well as potential applications will be discussed hereafter.

6.1 MR inhomogeneity correction

One of the main objectives of this work is the quantitative assessment of intensity inhomogeneity in structural MR imaging. After a quantitative evaluation of the major INU correction methods, we presented a method-specific approach for the optimization of input parameters. The correction of intensity inhomogeneities in MR imaging is considered as one of the most crucial processing step in MR data analysis. Indeed, an accurate INU field assessment ensures reliable examinations of structural properties and therefore guarantees interpretations consistent with biological processes underlying a certain pathological status. It is worth noting that almost all INU correction methods currently available have been developed over a decade ago, and are therefore tailored to MR data acquired at low to medium magnetic field strength (e.g. 1.5 T and 3 T). When the MR field strength is relatively low, non-uniformity in the spatial profile of the static magnetic field is considered the major source of intensity inhomogeneity. This results in smooth and slowly varying intensity changes across the imaged volume. On the contrary, when the field strength increases, the inhomogeneity dynamic becomes particularly complex. At 7 T, tissue-related effects produced by the interaction of the radiofrequency (RF) excitation pulse and brain structures can be considered as the main cause of non-uniformity. Furthermore, the spatial pattern of inhomogeneities does not respect standard INU correction methods assumptions of smooth intensity variations across the imaged volume (Bernstein et al., 2006; Moser et al., 2012; Umutlu et al., 2014). The
relatively low correction accuracy we obtained for the 7 T INU field suggested that improved INU correction methods are required. In particular, emphasis should be laid upon the development of methods that are effective also with complex INU profiles that are observed at high magnetic fields. Regarding the acquisition process, despite a conspicuous amount of studies have been conducted, further efforts may be concentrated on optimizing RF field mapping. Indeed, as suggested by our results, at high fields is not always possible to extract an accurate estimate of the INU field from the corrupted data. Thus, acquiring extra information during the scanning process (i.e. the RF map) may turn out to be an effective approach (Glasser et al., 2013b). As for the post-processing, it is commonly accepted that INU correction methods based on intensity distribution models, such as the ones implemented in SPM and FSL, are more efficient than the others. Methods developed over a decade ago, such as N3, claimed to be independent of scanner parameters (e.g. pulse sequence, acquisition coils) and to eliminate the dependence of the INU field estimate on the anatomy. The fact that they do not require supplementary information regarding the image to be corrected has always been considered the main feature of these types of method. Likewise, using no prior knowledge, coming from instance from brain atlases or reference points, may also be a drawback.

After assessing the performance of each specific correction technique, we attempted to provide a solution to the much debated problem of how to define the optimal input configuration for a given INU correction method. It is commonly accepted that the performance of any method may greatly depend on the specific input parameters employed (Boyes et al., 2008; Uwano et al., 2014; Weiskopf et al., 2011; Zheng et al., 2009). As described in chapter 2, the majority of INU correction algorithms available to the neuroimaging community require input parameters to be set. By defining distinct sets of parameter configurations, quite different INU field estimates can be obtained. However, due to the intrinsic complexity and unpredictability characterizing the INU, the use of a unique configuration may not be possible for all MR images. To this end, we have proposed a data-driven approach to identify the most appropriate configuration of input parameters for a given INU correction method. The peculiarity of our optimization approach relies on the identification of the optimal set of parameters after spanning over the whole configuration space. We tested it on the algorithm implemented in SPM, due to its widespread use in the community. Our results on simulated and actual MR data demonstrate that this method may assist the user to achieve more accurate INU corrections (see chapter 3).

Finally, it should be mentioned that our comparative analysis was only performed on T1-w images. In this regard, future studies are warranted to assess the performance of INU correction algorithms on MR data collected with different sequences, and in particular T2-w images. Our preliminary analysis on T2-w images (Figure 4.6, chapter 4) revealed a smaller effect of intensity inhomogeneity artifacts compared to T1-w images. However, such results should be confirmed and validated on simulated data. Moreover, the proposed data-driven optimization approach needs to be also evaluated on T2-w images. In principle, the ability to detect the optimal set of parameters within the whole parameters configuration space should be preserved. Theoretically, as long as a discernible difference between the WM and GM intensity distributions exists, the use of our optimized procedure is
still viable. It should however be considered that, in the case of imaging modalities characterized by a poor contrast between WM and GM or poor spatial resolution of the MR images, the accuracy of approach may decrease substantially.

6.2 MR intensity standardization

In chapter 4, we have introduced a new automatic method to correct for intensity variations across different MR acquisitions. Contrary to other approaches based on the histogram-matching concept (Hellier, 2003; Jager and Hornegger, 2009; Nyul and Udupa, 1999; Wang et al., 1998), we proposed a landmark-based correction method. Specifically, the intensity correction is derived from anatomical information and does not rely exclusively on the intensity distribution of the image. The assumption that image intensity of GM and WM are consistent across healthy subject cannot hold in brain disease (Ormerod et al., 1986). To address this potential issue, our proposed standardization relies on image values extracted from structures outside the brain to avoid potential inaccuracies due to WM and/or GM alterations in brain disease.

Our intensity standardization method is efficient, easy to implement, and fully automated. It can be incorporated in a Picture Archiving and Communication System (Huang, 1998), so that images are directly transformed or complemented with the appropriate intensity transformation when they are visualized in a workstation. Given its simplicity and the possibility to be directly applied to existing datasets, we foresee potential clinical applications. Nonetheless, some limitations should be acknowledged. A first point is that our intensity normalization technique relies on a linear transformation matching the intensity of the input image to that of a standard scale. However, significant variations in the MR scanning protocol, specifically the MR acquisition parameters (echo time and repetition time) seem not to affect linearly the image intensity (Hellier, 2003; Robitaille et al., 2012). This is to say, the intensity correction between two different MR pulse sequences may be nonlinear, thus a higher order transformation might be required (Robitaille et al., 2012). It is also worth noting that, as long as different subjects are acquired using the same protocol and scanner, our approach is able to deliver accurate results. On the other hand, in the case of datasets characterized by a heterogeneous set of acquisition protocols, the calibration may require further adjustments to fully account for intensity inconsistencies. Depending on the objective of the study, alternative techniques may be taken into account in order to achieve a nonlinear transformation of brain intensities. In the particular case of investigation focusing on localized tissue alterations (e.g. brain lesions), histogram-matching approaches may replace or integrate the one we implemented. In the presence of focal tissue alterations, the histogram is expected to be substantially unaltered. This allows potentially precise matching between the brain intensity histogram of the input image and that of a standard scale. By the same token, histogram-matching would perform poorly in the case of severe tissue pathology (Madabhushi and Udupa, 2006). It has indeed shown that relying solely on histogram landmarks extracted from the brain tissues distribution may yield erroneous interpretations (Ghassemi et
al., 2015). In this particular case, a promising solution might be conceived as a fusion of histogram matching with our external calibration approach. First, intensity values of non-brain structures (e.g. cerebrospinal fluid, skull, skin, muscle, fat) can be extracted from the input image and a reference image. A nonlinear function mapping the input distribution on the standard distribution is then generated (Hellier, 2003; Weisenfeld and Warfield, 2004). Eventually, the estimated intensity standardization function is globally applied to the whole image. In this way, a nonlinear normalization relying on intensity values extracted from anatomical structures outside the head can be accomplished.

### 6.3 T1-w/T2-w approach

In the preceding sections we have provided a critical description of the major issues underlying the generation of the T1-w/T2-w image. Hereafter, we will focus on the T1-w/T2-w technique as a whole. We will discuss specific areas of possible improvement, and we also illustrate a number of potential future applications.

**Technical considerations**

First of all, some considerations should be made regarding the image acquisition process. As MR imaging continues to evolve, emphasis should always be placed on employing sequences that exhibit the best combination of speed of acquisition, image quality and sensitivity to disease. A balance must always be reached between number of sequences performed, comprehensive evaluation and total study directions. Depending on the objective of the study and in particular the pathology investigated, a dedicate set of pulse sequences may be preferred to another. However, the time advantage of fast sequences such as MPRAGE and FSE may be used simply to obtain images much more rapidly or to acquire images of much higher contrast or spatial resolution compared with conventional SE images in a comparable period of time. The MPRAGE sequence yields both high image quality and GM/WM contrast (Fellner et al., 1996). This is mainly due to the stronger T1-w contrast provided by the initial 180° inversion pulse for magnetization preparation, in a way similar to inversion recovery technique (Fellner et al., 1996; Mugler and Brookeman, 1990). As for T2-w images, an important feature is the bright appearance of fat on T2-w FSE sequences, although this may be eliminated easily by using routine fat-suppression techniques. In the particular case, the true T2-w contrast needs to be preserved, conventional SE sequences are desirable (Jones et al., 1992; Thomas et al., 2004).

As illustrated in chapter 5, the proposed T1-w/T2-w workflow offers the potential for quantifying brain structural alterations between different patients populations. However, the effectiveness of such approach is limited by image reproducibility errors, which can translate into inaccuracies hampering the interpretation of morphometric data used in structural investigations. Therefore, it is crucial to identify and correct for the various sources of variance that can affect image reliability, within- and across-sites. In this work we focused on two major sources of image intensity variability: intensity inhomogeneity, mainly
due to the RF field $B_1$ (Belaroussi et al., 2006; Vovk et al., 2007), and intensity inconsistencies arising from global scaling effects, such as variations in coil loading and receiver attenuation setting (Wang et al., 1998). A key characteristic of our workflow is the separate correction of INU artifacts and intensity inconsistencies on T1-w and T2-w images. As reported in literature and then replicated in our results, the INU profiles are not the same in the two images, therefore modality-specific fields need to be estimated independently. This is obviously crucial for any segmentation and registration procedure, which utilizes intensity information in order to classify voxel into different tissue classes. Hence, the mutual contribution of an optimized INU correction along with an improved contrast-to-noise ratio may potentially result in more accurate tissue classifications.

However, as reported in (Glasser et al., 2013b), other factors such as scanner-specific non-linearity in the imaging gradients and readout distortions may further increase reproducibility errors in multi-site studies. The first is a critical distortion artifact affecting MR scanners characterized by significant gradient non-linearity (Glasser et al., 2013b). It results in reduced geometric and image intensity accuracy. The goal of a distortion correction is to transform the original distorted image into a corrected version by repositioning each voxel to its true location. In addition, each voxel intensity needs to be scaled to account for voxel-size distortions. The displacements along each of the three dimensions can be calculated from a model of the magnetic field generated by each of the gradient coils. This model is usually provided by the manufacturer and used to generate the required image transformation (Jovicich et al., 2006).

The second type of artifact causing inequality in the spatial distortion profile of T1w and T2w images is introduced during the readout phase (van der Kouwe et al., 2008). These distortions differ between the two modalities due to different readout dwell times, thus different bandwidths. It is a minor artifact in comparison to gradient non-linearity, but it is greatest in regions with high $B_0$ inhomogeneity due to magnetic susceptibility differences (orbitofrontal cortex and inferior temporal cortex especially). Readout distortions can be corrected using a field map as the distortion field and scaling it according to the readout dwell time (Glasser et al., 2013b). It is worth noting that an appropriate correction of such spatial distortions undoubtedly ensures a better cross-modal registration, thus yields a more precise structural characterization.

**Potential applications**

It is important to consider the potential benefits of acquiring MR images at high resolution. As a matter of fact, high-resolution imaging provides improved characterization of numerous abnormal tissues, and can potentially advance current clinical techniques. Among routinely used MR modalities, conventional T1-w and T2-w data have the major advantage to achieve a finer level of detail while keeping clinically acceptable acquisition times. The ratio approach also provides superior image quality and enhanced diagnostic information compared to the T1-w and T2-w modalities taken separately.

A potential application of this technique may be in morphometric studies. In voxel-based morphometry (VBM) of structural MR images, signal intensities are not directly measured. Instead, they are used as input for image segmentation and subsequent morphometric analysis. Therefore, structural
alterations are not directly detected by an increase or decrease of T1-w/T2-w intensity but as brain areas with an increased or decreased probability that specific voxels belong to a predefined tissue class. As stated in (Glasser and Van Essen, 2011), the ratio method improves areal localization by increasing the contrast to noise ratio between heavily and lightly myelinated areas. In an equivalent manner, the WM/GM contrast is also amplified enabling a more accurate evaluation of boundary voxels, thus better tissue delineation. Furthermore, the improved image contrast may be beneficial for surface-based registration methods. By combining both T1-w and T2-w images may be possible to overcome many registration hurdles (e.g. classification errors resulting in topological defects) by enhancing the laminar structure of the cortical gray matter (Glasser et al., 2013b). Specifically, a finer level of contrast may be achieved at the interface of cortical GM and other structures such as WM and pial surface. Essentially, this method may be extended to investigate cortical structure, as originally proposed by (Glasser and Van Essen, 2011). Cerebral cortex properties, such as its laminar organization and topographic organizations, can only be evaluated in the context of two-dimensional structure of the cortical surface. By extracting the surface and than registering it across subjects, further information in folding patterns and cortical thickness could be provided together with T1-w/T2-w intensities. When such measures are combined with automated procedures for normalizing individual subjects into a standardized space, it becomes possible to create statistical maps of T1-w/T2-w differences between individuals and/or groups.

The T1-w/T2-w technique may have a number of potential applications also in clinical settings. The method is fully automated, and integration with existing MRI analysis software seems promising. The combined information provided by T1-w and T2-w images is expected to allow more accurate assessment over a full range of brain pathology. Indeed, pathological features are commonly reflected by changes in MR image intensity. Both the scanning technique and the nature of the pathology influence the acquired MR signal, so that disease-related intensity variations can differ on T1-w and T2-w images. The presence of abnormality and differences in component concentrations may affect T1-w and T2-w images in different ways. When T1-w and T2-w variations are in opposite directions, the contrast enhancement offered by the T1-w/T2-w technique may lead to an improved detection of pathologic features. For instance, lesions characterized by axonal loss, severe gliosis, edema and inflammation are in principle detectable on T1-w/T2-w images. In these cases, T1-w hypointensities are normally accompanied by consistent T2-w hyperintensities, thus resulting in high T1-w/T2-w intensity (Bakshi et al., 2001; Neema et al., 2007; Tillema and Pirko, 2013; van Walderveen et al., 1998; Zhou et al., 2010). On the other hand, pathologic features such as excessive iron depositions and melanin-containing lesions are commonly reported as hyperintensities on T1-w and hypointensities on T2-w, hence an overall low T1-w/T2-w intensity may be revealed (Bakshi et al., 2001; Ginat and Meyers, 2012; Neema et al., 2007; Vymazal et al., 1999; Zhou et al., 2010).

We envision that our method has potential for clinical use. Since T1-w/T2-w images have a standardized intensity scale, their analysis may have diagnostic value. Due to spatial registration and intensity normalization of the T1-w/T2-w image, signal changes in patients can be quantified and directly compared to the values of healthy controls. It can be easily applied across different scanners, which is
essential for multicenter investigations of brain disorders. Since this method relies on conventional MR images such as T1-w and T2-w images, it can be retrospectively applied to existing datasets acquired with standard imaging protocols. This is a distinct advantage over other MRI approaches. If integrated in dedicated software of MRI post-processing, the T1-w/T2-w analysis may assist the clinician in the automated assessment of structural alterations. It may support clinicians in the understanding of the pathogenesis of several brain diseases, but more importantly we foresee important applications in individual patient care. The potential ability to standardize across images acquired in the same patient can be a valuable asset in the evaluation of structural changes over time. This would allow screening, monitoring disease progression and treatment effects at all stages of the pathology.
Bibliography


