MRI QUANTIFICATION OF LONGITUDINAL MORPHOLOGICAL CHANGES IN NEURODEGENERATION: APPLICATION TO MULTIPLE SCLEROSIS AND ALZHEIMER'S DISEASE

Nicolas Guizard

Department of Biomedical Engineering

McGill University, Montreal, Qc, Canada

April 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

In memory of Albert Guizard (1925-2005)

Contents

LIST OF ILLUSTRATIONS

Figures

Tables

ACKNOWLEDGMENTS

The research included in this dissertation could not have been performed if not for the assistance, patience, and support of many individuals. I would like to extend my gratitude first and foremost to my thesis advisor D. Louis Collins for mentoring me over the course of my undergraduate and graduate studies. His insight led to the original proposal to work on spatio-temporal analysis of neurodegenerative diseases. He has helped me through extremely difficult times over the course of the analysis and the writing of the dissertation and for that I sincerely thank him for his confidence in me.

I would additionally like to thank Drs Daniel Garcia, Vladimir Fonov and Douglas L. Arnold for their support in both the research and especially the revision process that has led to the articles in this thesis. With her knowledge and understanding of the written word, Dr. Marta Kersten has allowed me to fully express the concepts behind this research.

I would also like to extend my appreciation to Dr. Pierrick Coupé who has served as a voice of quiet wisdom in matters ranging from the most basic aspects of the experiments, to the paths that my career has eventually taken.

This research would not have been possible without the assistance of the IPL lab which built the foundations for the data analysis. Also, I would like to thank NeuroRx for sharing their data and guiding me through the different clinical evaluations. I also express my gratitude to the MS society of Canada for funding my research.

On a more personal note, I would like to thank my patient and understanding partner, Ilana, for supporting my self-inflicted impoverishment. Over the years, she has fed me and even clothed me from time to time, but also corrected so many of my English mistakes.

Finally I would like to extend my deepest gratitude to my family, Marion, Damien, Chantal and Alain, without whose love, support and understanding I could never have completed this doctoral degree.

COMMON ABBREVIATIONS

AD	Alzheimer's disease
BBB	Blood Brain Barrier
CEL	Contrast Enhancing Lesion
CNS	Central Nervous System
CSF	Cerebral Spinal Fluid
DBM	Deformation-Based Morphometry
FLAIR	FLuid Attenuated Inverse Recovery
Gad	Gadolinium
GM	Grey Matter
MCI	Mild Cognitive Impairment
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
MS	Multiple Sclerosis
N3	Non-parametric Non-uNiformity correction
NAGM	Normal Appearing Grey Matter
NAWM	Normal Appearing White Matter
NLM	Non-Local Means
PDW	Proton-Density Weighted
RDF	Random Decision Forest
RRMS	Relapsing Remitting Multiple Sclerosis
SPMS	Secondary Progressive Multiple Sclerosis
T1W	T1-Weighted
T2LL	T2 lesion load
T2W	T2-Weighted
TBM	Tensor-Based Morphometry
TE	Echo Time
TI	Inversion Time
TR	Repetition Time
VBM	Voxel-Based Morphometry
WM	White Matter

ABSTRACT

In Canada, Alzheimer's disease (AD) and multiple sclerosis (MS) affect respectively 2.1 and 0.3% people. AD affects memory, day-to-day cognitive ability and represents 60 to 80 percent of dementia cases in the elderly while MS is the most common neurological disability in young adults. These two common neurodegenerative diseases have no cure, but approaches to help with early diagnosis, treat their symptoms, and delay their progression are currently under heavy international research investigation. While the neurodegenerative processes that occur during these diseases are quite different, they both can result in widespread and sometimes subtle anatomical changes over time. Neurodegenerative changes, such as loss of neurons or loss of other building blocks of the brain (e.g., myelin), can – over time – result in cerebral atrophy, defined as the macroscopic loss of brain parenchymal volume. Reliably quantifying these changes will provide new insights into disease progression with immense potential implications for disease prognosis, treatment strategy and monitoring therapeutic effects. Magnetic resonance imaging (MRI) enables the macroscopic visualization of brain structure and anatomy, and allows the non-invasive assessment of neurodegenerative phenomena (e.g., atrophy), as well as their temporal evolution. Although automatic image analysis methods for the measurement of anatomic changes in the brain have been used extensively for the estimation of morphological differences, they have been in large part limited by technological approaches that are inherently cross-sectional (single time-point).

Recently, the emergence of numerous databases with longitudinal MR images provides the necessary data to develop, test and investigate the longitudinal structural changes of brain anatomy at a much larger scale. The main objective of this thesis was to develop robust, accurate and fully automatic measures of focal longitudinal changes in the brain on MRI. Such methodology will enable a more detailed understanding of the initiation and spread of neurodegeneration, with the overall goal to comprehend the dynamic spatio-temporal distribution of brain atrophy in neurodegenerative diseases. As registration-based methods to

detect brain atrophy are confounded by image intensity changes due to focal lesions, the first aim was to automatically detect the presence of focal lesions. Second, we proposed a new filling strategy to remove these lesion-induced intensity irregularities that could potentially affect subsequent analysis. For the third aim, we proposed and assessed a new method to remove potential bias in pair-wise non-linear registration used to estimate brain atrophy. Finally, we developed a spatio-temporal image registration framework, which accounts for potential biases in longitudinal image analysis. The methods developed in this thesis improve longitudinal morphological measurements and can find application in monitoring and diagnosing neurodegenerative diseases.

ABRÉGÉ

Au Canada, la maladie d'Alzheimer (MA) et la sclérose en plaques (SP) affectent respectivement 480,600 et 100,000 patients. AD affecte la mémoire, la capacité cognitive quotidienne et représente 60 à 80 pourcent des cas de démence alors que la SP est le handicap neurologique le plus répandue chez les jeunes adultes. Ces deux maladies neurodégénératives sont incurables, mais des traitements pour leur symptôme, leur diagnostic précoce et pour retarder leur apparition sont actuellement évalué internationalement. Les processus neurodégénératifs résultants de ces maladies peuvent entraîner des changements anatomiques spatio-temporaux globaux et focaux. Quantifier de manière fiable ces changements permettra d'offrir de nouvelles perspectives sur la progression de la maladie et aura d'immenses implications pour le pronostic, les stratégies de traitement et le suivi des effets thérapeutiques sur la maladie. L'imagerie par résonance magnétique (IRM) permet la visualisation de macrostructure anatomique du cerveau, permet de saisir les phénomènes neurodégénératifs tels que l'atrophie, et d'enquêter de manière non invasive l'évolution temporale. Bien que des méthodes ont été largement utilisées pour l'estimation robuste de l'atrophie, elles se sont limitées aux analyses de type transversal.

L'émergence récente de nombreuses bases de données longitudinales fournit les données nécessaires pour développer, tester et étudier les changements structurels longitudinaux de l'anatomie cérébrale. L'objectif principal de cette thèse était de développer une mesure robuste, précise et entièrement automatique des changements longitudinaux focaux en IRM, permettant une compréhension détaillée de l'origine, de la propagation et de la distribution spatio-temporelle de l'atrophie cérébrale des maladies neurodégénératives. Le premier objectif était de détecter automatiquement la présence de pathologies telles que des lésions. Deuxièmement, les lésions peuvent avoir un impact sur les analyses d'image. Nous avons donc proposé une nouvelle stratégie de remplissage des lésions pour éliminer les irrégularités d'intensité qui peuvent affecter les analyses ultérieures. Pour le troisième objectif, nous avons proposé et évalué une

nouvelle méthode permettant d'éliminer le biais potentiels du recalage non-linéaire. Finalement, nous avons développé une approche de recalage spatio-temporelle d'images avec pour but de supprimer la variabilité longitudinale. Les méthodes développées dans cette thèse offrent un grand potentiel clinique en termes de diagnostic, traitement et suivi des maladies neurodégénératives.

ORIGINAL CONTRIBUTIONS

The original scientific contributions claimed in this thesis include:

I. **Chapter 3 The development, implementation and validation of a new segmentation algorithm for MS lesions**:

We developed a novel rotation-invariant distance metric for multi-atlas non-local means (NLM) image segmentation. We developed an approach based on only two MRI contrasts (T2W and FLAIR) that out-performed state-of-the-art automatic lesion segmentation methods. We evaluated our approach in both voxel- and lesion-wise frameworks, which showed the ability of the method to detect small lesions with high sensitivity and specificity.

II. **Chapter 4 Development, implementation and validation of an automatic inpainting method to fill focal image intensity irregularities such as lesions**:

We proposed a novel approach using non-a priori non-local information to replace signal intensity changes due to MS lesions with intensities representative of surrounding healthy tissue in order to remove focal inhomogeneities that might adversely affect image registration quality. We demonstrated a higher fidelity of the reconstructed images from simulated lesion data and an increase in power to detect brain atrophy.

III. **Chapter 5 Evaluation of non-linear registration symmetry of different popular registration algorithms and development of a method to force registration symmetry**:

We evaluated bias in non-linear registration algorithms and proposed a novel solution to force inverse consistency ("symmetry") in the estimation of the registration transformation. We demonstrated that the symmetrization constraint does not impair the accuracy of the registration and improves the ability to detect morphological changes.

IV. **Chapter 6 Design, implementation and validation of a new method for robust longitudinal MRI data registration**:

We proposed a new longitudinal registration framework that corrects for longitudinal variability mainly due to relate to image acquisition factors such as intensity non-uniformity and distortion.

CONTRIBUTION OF AUTHORS

Manuscript #1 (chapter 3): Rotation-invariant multi-contrast non-local means for MS lesion segmentation

Authors: Nicolas Guizard, Pierrick Coupé, Vladimir S. Fonov, José Manjon Hererra, Douglas L. Arnold, D. Louis Collins

Nicolas Guizard had the idea of using the large library of manually segmented MS lesions available at the MNI to perform multi-atlas segmentation using the rotation-invariant non-local mean (NLM) patch-based segmentation technique. Nicolas Guizard also designed, implemented, and performed the validation experiments, interpreted the results, and wrote the manuscript describing the method.

Pierrick Coupé and José Manjon Hererra had the original idea of using rotation invariant distance measures within the NLM for image denoising, and helped interpreting the results.

Vladimir S. Fonov helped with the implementation of the lesion-wise validation framework. Douglas L. Arnold provided the clinical MS data used in the multi-template library.

D. Louis Collins supervised the project and revised the manuscript.

Manuscript #2 (chapter 4): Non-local means inpainting of MS lesions in longitudinal image processing

Authors: Nicolas Guizard, Kunio Nakamura, Pierrick Coupé, Douglas L. Arnold, D. Louis Collins

Nicolas Guizard had the original idea of using the non-local inpainting to fill pathological tissue with surrounding normal appearing tissues. He designed, implemented, and performed the

validation experiments, interpreted the results, and wrote the manuscript of the method. Kunio Nakamura had the idea of using simulated MS lesions on healthy subjects for the validation framework. Pierrick Coupé provided the original NLM implementation for image denoising. Douglas L. Arnold provided the clinical MS data. D. Louis Collins supervised the project and revised the manuscript.

Manuscript #3: Impact of Non-Linear Registration Symmetry in Longitudinal MRI Studies

Authors: Nicolas Guizard, Vladimir S. Fonov, Bérengère Aubert-Broche, D. Louis Collins

Nicolas Guizard implemented and performed the comparison of some of the most popular non-linear registration algorithms. He led the experiments, interpreted the results, proposed a solution to improve registration symmetry and wrote the manuscript of the method. Vladimir Fonov and Bérengère Aubert-Broche helped with the validation framework and the writing of the article. D. Louis Collins supervised the project and revised the manuscript.

Manuscript #4: Spatio-temporal regularization for longitudinal registration to an unbiased 3D individual template

Authors: Nicolas Guizard, Vladimir S. Fonov, Daniel García-Lorenzo, Kunio Nakumura, Bérengère Aubert-Broche, D. Louis Collins

Nicolas Guizard had the idea of using Taylor series to decompose the longitudinal non-linear deformation of brain images and perform regularization on this decomposition. He designed, implemented, and performed the validation experiments, interpreted the results, and wrote the manuscript describing the method. Daniel García-Lorenzo implemented the initial longitudinal pre-processing pipeline. Vladimir Fonov and Bérengère Aubert-Broche helped with the implementation of the longitudinal pipeline framework. Vladimir Fonov also provided support in the mathematical modeling of the longitudinal deformation fields. Kunio Nakumura

provided other external method results and helped with the power analysis. D. Louis Collins supervised the project and revised the manuscript.

CHAPTER 1 INTRODUCTION

Neurodegeneration is a phenomenon by which neurons deteriorate and ultimately die. This process underlies major human neurological diseases such as Alzheimer's disease (AD), Parkinson's disease, multiple sclerosis (MS), Huntington's disease, and amyotrophic lateral sclerosis (ALS, also known as Lou Gehrig's disease), but remains poorly understood. With the widespread use of magnetic resonance imaging (MRI), morphological changes, and in particular brain atrophy, have emerged as clinically relevant biomarkers of neurodegeneration. Such biomarkers have been used to monitor disease progression and evaluate treatment since they can be measured in-vivo in a reproducible manner at both local and global levels.

The main objectives of this dissertation were to develop and validate the necessary image processing algorithms to quantify brain volume changes on local (voxel-wise) metrics in order to characterize atrophy due to neurodegeneration over the course of the disease. In order for these longitudinal metrics to be robust and stable, it was important to ensure that they were not being influenced by the underlying pathology and that they were relatively immune to the potential acquisition biases affecting the medical images.

In Chapter 2, the different background materials required to understand the objectives of this thesis are introduced. The clinical and pathological aspects of neurodegeneration in the context of MS and AD are described, as well as how MRI data are reconstructed and analysed.

The body of this thesis is composed of 4 articles presented in Chapters 3 to 6. Chapter 3 and Chapter 4 focus on methods to limit the influence of focal intensity irregularities such as lesions on image processing steps. As such, in Chapter 3, we propose a new method to automatically detect MS lesions. This method uses a patch-based approach to learn from a large library of manual segmentations by looking at the local image characteristics. The results of our method

show promising results compared to other state-of-the-art approaches. The manuscript has been published in *NeuroImage: Clinical*.

Chapter 4 presents a new automatic method to minimize the impact of lesions on longitudinal image analysis. The proposed method uses previously generated lesion segmentations and fills the regions with anatomically realistic normal appearing tissue intensities. The results exhibit better power to detect to detect longitudinal changes, related to natural course of disease and/or treatment than if no filling is carried out. The manuscript has been published in *Frontiers of Neuroscience*.

Following the detection and compensation of pathological biases such as lesions as described in Chapters 3 and 4, a better estimation of morphological changes using non-linear registration can be performed. However, in itself, non-linear registration for longitudinal analysis can be subject to various other sources of bias, namely asymmetry in the estimation of the registration transformation. Thus, Chapter 5 contains an evaluation of the impact of non-linear registration symmetry in the context of longitudinal MR images analysis. Non-linear registration asymmetry can generate significant bias depending on the direction of the registration (from volume *A* to *B*, or from *B* to *A*) and can favour atrophy or growth for example. In the manuscript, we concluded that symmetry is an important step in pairwise non-linear registration, and the proposed method to force symmetry has shown to be easy to implement and to reduce unwanted biases. The manuscript was presented at the *Spatio-Temporal Image Analysis (STIA) workshop,* 2010.

As part of an extension from pairwise to multi time-point longitudinal analysis, in Chapter 6 we propose a method to estimate longitudinal changes by means of the creation of an individual 3D template using spatio-temporal regularization. We validate our approach on a large population of AD patients. We found that spatio-temporal regularization enables a significant decrease in the number of subjects required to detect group differences (i.e, an increase in statistical power) in the setting of a clinical trial. The manuscript has been published in *Plos-One*.

Finally, we discuss the novelty, strengths and limitations, and the relevance of our findings, as well as suggestions for future work before concluding in Chapter 7.

CHAPTER 2 BACKGROUND

The following chapter is organized in four sections to provide the necessary background to understand the purpose of the thesis. The first section describes briefly the neurodegeneration phenomenon with the example of MS and AD. The second is on the imaging of neurodegenerative disease, after introducing the principles of MRI. In this pathological context, the third section provides the fundamental image processing techniques used in MRI. This chapter concludes with a review of atrophy measurement approaches and their potential as a surrogate metric of neurodegeneration.

2.1. Neurodegeneration: MS and AD as disease examples

Neurodegeneration is a broad term comprising the progressive loss of structure or function of neurons, as well as the death of neurons. Many neurodegenerative diseases including amyotrophic lateral sclerosis (ALS), Parkinson's (PD), Huntington's (HD), AD and MS occur as a result of different forms of neurodegenerative processes. This work focuses primarily on MS and AD, which are fundamentally different diseases, with different symptoms, causes and pathogenesis. MS is a chronic inflammatory disease, which is believed to be initially driven by an auto-immune response characterized by demyelination of the white and grey matter (WM and GM). In contrast, AD is a neurodegenerative disease associated with the aggregation of misfolded proteins. However, recent studies indicate some similarity in the neurodegenerative mechanisms of the two diseases. In the following sections, the basics of neuroanatomy and neurodegeneration are described, followed by the clinical and pathological aspects of both diseases.

2.1.1. Neuroanatomy

The human brain, as illustrated in Figure 2.1, has similar features across most vertebrate brains. It is divided into a forebrain, hindbrain and midbrain, surrounded by cerebrospinal fluid (CSF). The forebrain can be broken down structurally into the cerebral cortex, subcortical structures (thalamus, hypothalamus, hippocampus, basal ganglia and amygdala) and the corpus callosum. The hindbrain is composed of the cerebellum, brainstem and pons.

The human cerebral cortex can be segmented into four lobes: frontal, parietal, occipital and temporal. Gyri are the cortical ridges of these lobes and are bordered by grooves (sulci), which create the characteristic folded appearance of mammalian brains. Most of the GM is in the cortical layers and is distinguished from WM by its numerous neuron cell bodies (Figure 2.2). WM is usually deeper in the hemispheres and holds mainly glial cells (Figure 2.2) and myelinated neuronal axons which connect the cell bodies. Axons are integral part of a neuronal cell. Many of the axons are parts of neurons that originate in the cortex. The biggest WM fiber tract bundle is the corpus callosum which connects both hemispheres. Deep grey matter structures are embedded within the WM such as the basal ganglia and thalamus. The brainstem connects the brain and the spinal cord while the cerebellum is at the base of the brain.

Most of the lobes are demarcated following skull anatomy but they also have different functional roles. The division in lobes is convenient for reference and each of them have different main cognitive and motor functions. The frontal lobe functions control attention, abstract thinking, behavior, problem solving tasks, physical reactions and personality. The occipital lobe's main functions are visual reception, visual-spatial processing, movement and color recognition. The temporal lobe controls auditory and visual memories, language, and some hearing and speech functions. And the parietal lobe, also known as the somatosensory cortex, is essential to process and integrate sensory information.

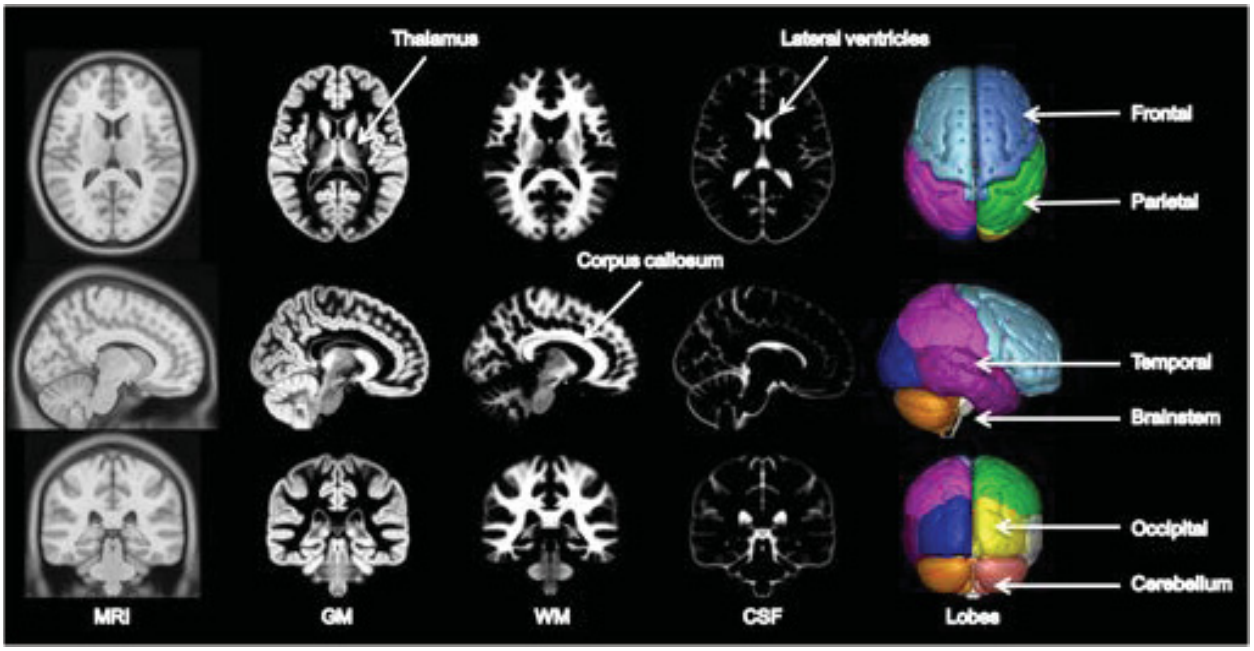


Figure 2.1: Human brain neuroanatomy. Magnetic resonance image (MRI), grey matter (GM), white matter (WM), cerebral spinal fluid (CSF) and brain lobes are illustrated with annotation and from top to bottom the axial, sagittal and coronal views.

2.1.2. Neurodegeneration

Neurodegeneration occurs when the central nervous system (CNS), and in particular the neurons (Figure 2.2), begin to deteriorate. The degeneration of neuronal cells alters their functionality and eventually leads to their death. In neurodegenerative diseases, as neurons deteriorate, patients may at first experience mild symptoms, but as the number of affected neurons increases, symptoms progressively worsen and can lead to death. Aging affects many cellular processes and is the greatest risk factor for neurodegeneration (Trapp, Peterson et al. 1998, Niccoli and Partridge 2012). This process is still poorly understood but recent progress in research has found many pathological parallels between neurodegenerative diseases (e.g., AD, MS, ALS, PD, and HD) (Durrenberger, Fernando et al. 2014). However, the causes and the neurodegenerative mechanisms are not yet fully understood and in the following sections we summarize the current understanding of these processes.

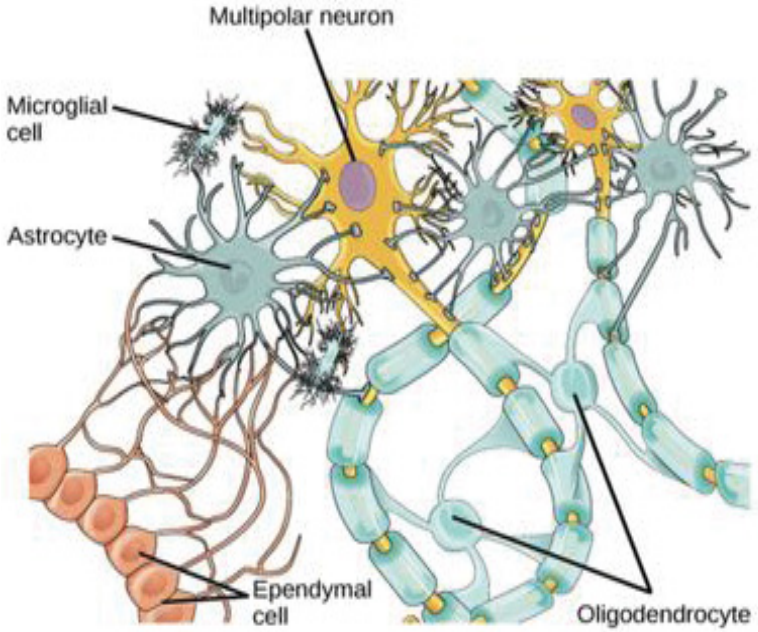


Figure 2.2: Neurons and glia cells in the human nervous system. The neurons and their surrounding blood vessels interact with different categories of glial cells: oligodendrocytes create the myelin sheath around axons to speed up neuronal transmission. Astrocytes play a supportive role for neurons and the blood brain barrier but are also involved in providing nutrient, repairing, scaring and maintenance of extracellular ion balance. Microglia act as the first and main form of active immune defense in the CNS. Ependymal cells produce cerebrospinal fluid that cushions the neurons.

Most neurodegenerative diseases show neuropathological changes mainly in the form of focal loss of neurons (cell death) with reactive changes of glia cells (or "gliosis"). Improving the understanding of the mechanisms underlying neurodegeneration is a major challenge in experimental neuroimmunology and the development of treatments. The potential mechanisms of neurodegeneration can be summarized in three different biological mechanisms (more details can be found in Ramanan et al. (2013)):

- Intra-cellular mechanism: Three major morphological types of cell death have been found: apoptotic, autophagic and necrotic. Apoptosis is generally understood as the process of programmed cell death (Lockshin and Williams 1964), and is characterized by a specific morphologic sequence of changes in the dying cell (i.e. membrane and nuclear fragmentation). Autophagy is the mechanism regulating the degradation of unnecessary and dysfunctional cellular compartments. Controlled autophagy can clear

aggregated or dysfunctional proteins which could contribute to neurodegeneration (Bredesen, Rao et al. 2006). Necrosis is the result of unregulated digestion of cell components due to external factors such as infection, toxins, or trauma (Proskuryakov, Konoplyannikov et al. 2003).

- Extracellular environment: The extracellular matrix is the extracellular environment made of molecules secreted by cells that provide structural and biochemical support to the surrounding cells. The disturbance of this environment can contribute to neurodegeneration. For instance, cellular adhesion, which involves the binding of cells between each other or to extracellular tissue, are important to maintain the synaptic contacts, blood brain barrier (BBB) integrity, neurotransmission efficiency and the intracellular signalling (Horwitz 2012).

- Systemic environment: Complex biological system deregulation in inflammatory, vascular, or endocrine domains has also been considered as a potential precursor of neurodegeneration. In general, inflammation is a protective response to various cell and tissue injuries. The effect of uncontrolled immune responses initiates excessive cell and tissue damage that result in destruction of normal tissue and chronic inflammation (Lee and Yang 2012). Inflammation and immune deregulation being causal to neurodegeneration or secondary to apoptosis is still an ongoing debate and might be specific to the pathology.

Genetic and environmental are the most studied factors of neurodegenerative diseases include genetic and environmental factors. Some disorders have shown familial occurrence suggesting deterministic genetic factors such as HD which follows an autosomal dominant pattern (Myers 2004). Genetics may also play a role in roughly 10% of AD cases (Lautenschlager, Cupples et al. 1996). Recently in MS, epigenetic modifications have raised the question about potential risk genes instead of deterministic genes (Huynh and Casaccia 2013). Indeed, toxic environment factors might play a crucial role in the initiation of the neurodegenerative disease. The prevalence of geographical and temporal clusters of patients

supports this hypothesis as it has been shown in the case of exposure to synthetic opioid analgesics in a severe variant of PD (Przedborski and Vila 2001) for example. Similar neurological conditions have shown to correlate with toxic environmental exposure risk factors for specific socio-geographical and professional contexts. However, most patients suffering from neurodegenerative disorders do not belong to specific population clusters. These findings suggest that neurodegenerative diseases are possibly the result of a combination of both factors, environmental and genetic, and these factors might potentially initiate some neurodegenerative mechanisms.

2.1.3. Multiple Sclerosis

MS is an inflammatory demyelinating disease that affects the patient's CNS. First described in 1868 by the French neurologist Jean-Martin Charcot (Charcot 1868), MS was initially called "sclérose en plaques" in reference to the sclerosis (hardening) of the CNS tissues and the appearance, which is described as 'plaques' for their sectional plate shape. In young adults in Northern Europe, North America, and similar temperate latitudes in Australasia, it is the most common cause of neurological disability (Weinshenker 1996). The average age of onset is around 30 and women are 2-3 times more susceptible to be affected by MS than men.

The symptoms, prognosis and disease progression vary considerably between patients. Treatment can help to recover from and prevent new attacks, however, the cause of the disease is still unknown and there is no known cure. Despite the different treatments, MS patients have a life expectancy 5 to 10 years lower than the population average (Compston and Coles 2008). In Canada, 55,000 to 75,000 people are affected by this disease with an annual therapy cost of up to $40,000cnd per patient (estimated to be more than $1 billion cnd per year (1998)).

In the following, we summarize the clinical and pathologic characteristics of MS to better understand this complex and multifaceted disease.

MS Symptoms and diagnosis

MS is characterized by unpredictable acute attacks on the CNS (or relapses), where the patients can present with a variety of neurological symptoms. Following these relapses, the patients can thereafter partially or completely recover. Neither the attacks, nor the progression of the disease can be predicted by any surrogate marker.

The location of the lesions in the CNS determines the related symptoms and can vary considerably within and between patients. Autonomic (digestion, bladder, sexual arousal), visual, motor (spasticity, gait ataxia, and weakness in arms and legs) and sensory nervous system symptoms (numbness, pain, tingling sensation, and coordination problems) are the most common symptoms (Noseworthy, Lucchinetti et al. 2000, Compston and Coles 2008). Other MS symptoms have been widely reported such as cognitive (slower speed processing and concentration) and memory impairments (Chiaravalloti and DeLuca 2008) as well as fatigue and depression (Ziemssen 2009).

Relapses usually persist for days (Naldi, Collimedaglia et al. 2011),), but the patients tend to recover partially or completely, more so at the early stage of the disease. Following these relapses, there is a period of remittance during which the disease is silent for a relatively long period of time. The frequency of the attacks is rarely less than twice per year (Scalfari, Neuhaus et al. 2010).

The similarity between MS symptoms and other pathologies (ie. Acute Disseminated Encephalomyelitis, Systemic Lupus Erythematosus…) can make the diagnosis difficult, and a at least 2 clinical relapses over time have been required in the past to confirm the MS diagnosis (Poser, Paty et al. 1983). More recently, the revised McDonald diagnosis criteria use MRI and permit determination of a spatio-temporal dissemination of lesions, and thus allows for an earlier diagnosis and thus earlier treatment of MS (Polman, Reingold et al. 2011).

MS disease course and treatments

In most cases, MS leads to eventual chronic disability (Noseworthy, Lucchinetti et al. 2000). MS patients are classified according to clinical disease course in different subtypes (Lublin and Reingold 1996) that include: relapsing remitting (RRMS) and secondary progressive (SPMS), as well as two less common forms of MS being primary progressive (PPMS) and benign.

The different subtypes of MS are described as follows (Hurwitz 2009):

- RRMS: Well-defined relapses with partial to full recovery followed by a remittance period without progression of neurological symptoms.
- SPMS: Continuous progression of the disease following the RRMS phase. The progression might occasionally relapse with short remissions or plateaus.
- PPMS: Continuous progression of the disease from onset, with occasional improvement or plateauing.
- Benign MS: Continuous mildest form of the disease.

The most common type of MS is RRMS with about 85% of the patients at onset, followed by the SPMS pattern (Confavreux and Vukusic 2006). Figure 2.3 describes the typical evolution of disability for the RRMS patients before converting to SPMS, where during the preclinical phase, the underlying tissue morphology (i.e., brain volume) may have already started to show alterations. After an initial neuroinflammatory event, if a second relapse occurs or imaging evidence of spatio-temporally disseminated lesions is detected, the patient is diagnosed with MS (Polman, Reingold et al. 2011). RRMS patients usually convert to SPMS when their disability worsens with almost no relapses (Confavreux and Vukusic 2008).

There is currently no known cure or preventive measures for MS. However, there are treatments that can affect the symptoms and disease-modifying therapies (DMT) that might reduce relapses and slow disease progression. For RRMS patients, interferon-β or glatiramer

acetate are the most commonly used subcutaneous or intramuscular DMTs but new immunomodulating oral DMTs have been approved since 2010 (Teriflunomide (O'Connor, Wolinsky et al. 2011), Dimethyl Fumarate (Gold, Kappos et al. 2012), Fampridine (Goodman, Brown et al.)) and are expected to gain in popularity (Miller 2011). However, the DMT's efficacy, with respect to disability progression and relapse rate, is limited as some patients do not respond to these therapies at all (Rudick, Lee et al. 2004). These DMTs were initially investigated for RRMS patients to prevent attacks, but they have shown some effects on SPMS patients. In the case of severe acute attacks, to reduce the inflammation and the severity of the relapse, high doses of corticosteroids are often administered.

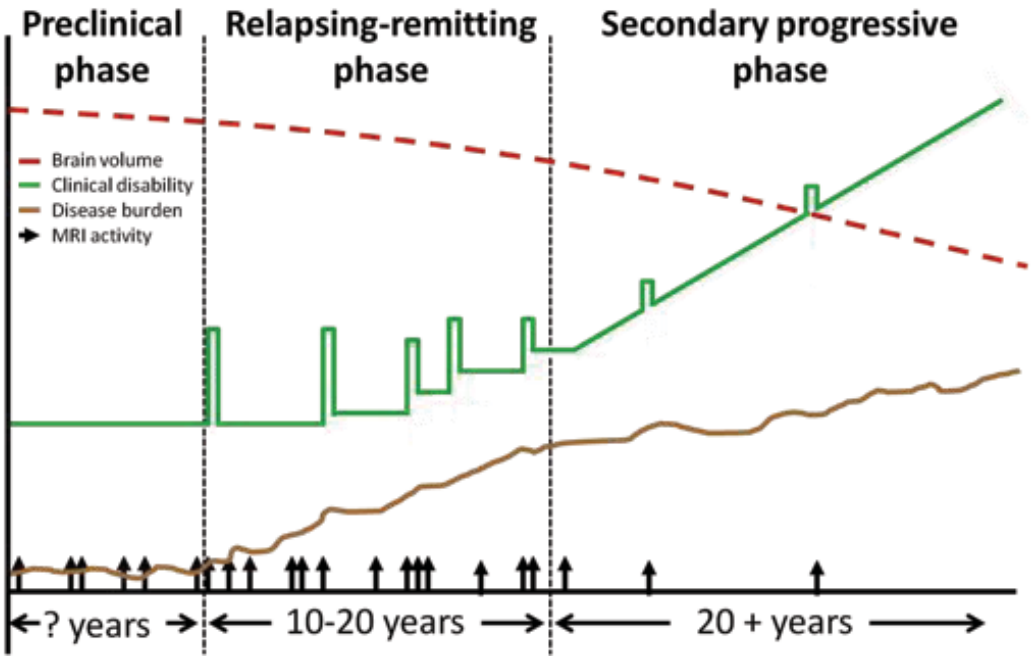


Figure 2.3: Typical clinical and MR image course of MS. MR image activity (vertical arrows) indicates an inflammatory process as measured on brain MR imaging by a gadolinium-enhancing or a new T2 hyperintense brain lesion. An increase in disease burden (brown line) can be measured by the total volume of MR-imaged lesions and is an indication of permanent tissue damage. A loss in brain volume (dashed red line) is also thought to occur early in the disease and gradually progresses over time. Importantly, MR image activity is typically more frequent than clinical relapses as shown by the clinical disability (green line). Adapted from (Fox, Bethoux et al. 2006).

Neuropathology in MS

MS is considered an autoimmune disorder but recent hypotheses suggest that MS could be primarily a neurodegenerative disease (Chaudhuri 2013). It is believed that MS is due to nervous system dysfunction, secondary to focal inflammatory demyelination as a result of autoimmune responses acting upon myelin and/or the oligodendrocytes that produce myelin (Evangelou, Esiri et al. 2000) (Figure 2.2). Post-mortem investigations suggest a focal model of MS WM lesions (Bitsch, Kuhlmann et al. 2001) and has been widely studied on MRI based on their appearance (Section 2.2.2). However, more global or diffuse damage is occurring in the normal appearing brain tissues (NABT) (Prineas and Connell 1978, Rodriguez and Scheithauer 1994).

Focal pathology in MS: Lesions are the most obvious pathology in MS and are mainly visible in WM but can be found also in GM (Figure 2.7). MS lesions are more frequently located in the peri-ventricular or sub-cortical region of the brain. They vary in size, location and volume but are usually elongated along small vessels. The lesions are highly heterogeneous and include different underlying processes: focal breakdown of the BBB, inflammation, destruction of the myelin sheath (demyelination), astrocytic gliosis, partial preservation of axons and remyelination. The temporal inflammation activity is usually used to classify lesions in the following manner (Trapp, Peterson et al. 1998):

- Acute active lesions are often referred to as “new lesions” and present abundant macrophages that contain early and late hypercellular degradation products.
- Chronic active lesions are demyelinated and present an increasing concentration of macrophages from their center to their edge.
- Chronic inactive lesions are the most common subtype of lesion in the brain. They are hypocellular, demyelinated and often have gliosis and large extracellular spaces.
- Remyelinating lesions are characterized by thin irregularly formed myelin sheaths.

***Global pathology in MS*:** The previously described focal involvement may be combined with a more global model, where a more diffuse process occurs in WM and GM, which could explain the focal and global loss of tissue due to the decline of neuronal density in the normal appearing white matter (NAWM) and the normal appearing grey matter (NAGM) (Arnold, Matthews et al. 1990, Kidd, Thorpe et al. 1993, Ceccarelli, Rocca et al. 2008). Potential re-organization of cortical networks (cortical plasticity), and CNS redundancy, may hide clinical symptoms of the pathological changes caused by axonal loss (Kidd, Thorpe et al. 1993, Evangelou, Esiri et al. 2000). In terms of volumetric changes, since axons contribute up to 46% of the WM volume while myelin represents only 24%, (Miller, Barkhof et al. 2002), thus cerebral atrophy should reflect their irreversible loss. However, these volumetric changes might be confounded by the swelling effect of inflammation and gliosis.

2.1.4. Alzheimer's disease

Alzheimer's disease (AD) is a neurodegenerative disease of the brain, named after the German physician Dr. Alois Alzheimer and described as a progressive dementia and the presence of neuropathological alterations in the brain.

In the elderly population, AD is the most common cause of dementia, with a world-wide prevalence of 44 million in 2014 (2014). In Canada, 480,600 people are affected by this disease with an annual cost estimated to range from $9,000cnd to $37,000cnd for severe disease cases (2012).

AD is a progressive disease where the earliest symptom is usually memory loss, followed by functional and cognitive decline until the patient is fully dependant on care givers (De Leon and Braak 1999). There is currently no cure, nor any disease-modifying therapy available for AD. Mental function and behavioral symptoms can sometimes be maintained with treatment, but the

focus of current pharmaceutical trials are on the development of new treatments to delay and/or prevent its onset or progression (Salomone, Caraci et al. 2012).

In the next section, we describe the clinical and pathologic characteristics of AD in order to emphasize the need to quantify neurodegeneration to better understand and assess the changes due to the disease.

AD Symptoms and diagnosis

The most common and often the first symptom is short-term memory loss. However, other numerous cognitive impairments can affect AD patients:

- Familiar daily tasks such as cooking and cleaning might become difficult.
- Language problems, such as word-finding problems or reduced vocabulary in speech or writing.
- Confusion with location or passage of time.
- Impaired judgment (e.g. health treatment, proper clothes for outside temperature).
- Difficulty of performing abstract tasks (e.g. planning public transport itinerary)
- Having visual or space difficulties, such as not understanding distance in driving, getting lost or misplacing items.
- Change in mood, behavior, personality

The diagnosis of AD is difficult because of the large spectrum of symptoms but consensus criteria have been proposed to guide the classification of patients as definite, probable or possible AD (McKhann, Drachman et al. 1984), theses criteria have been improved by using biochemical and neuroimaging biomarkers (McKhann, Knopman et al. 2011). These consensus criteria have

proven to be accurate up to 80% when confirmed by neuropathology gold standards (Ranginwala, Hynan et al. 2008).

However, significant challenges remain to diagnose AD including pre-symptomatic diagnosis, differential diagnosis and the evaluation and prediction of disease progression Sperling, Aisen et al. 2011).

AD disease course and treatments

AD is believed to start many years before the first clinical symptoms, when underlying pathology phenomena are starting. Figure 2.4 describes the clinical stages and their relationship with different biomarkers. One common belief is that AD starts with the deposition of beta-amyloid plaques (Aβ) and neurofibrillary tangles (Tau) in the pre-symptomatic phase. Then, the disease progresses and affects brain structures with atrophy and loss of memory and function.

Before developing AD, many patients go through a mild cognitive impairment (MCI) phase. This phase involves problems with general cognitive function (their mental abilities such as thinking, knowing and remembering). Although MCI significantly increases someone's risk of developing dementia, more than 50% of MCI patients will not convert to dementia (Mitchell and Shiri-Feshki 2009). From a meta-analysis of 41 robust cohort studies, it was shown that MCI patients convert to AD with a yearly rate of about 4.8-8.5%.

There is currently no cure for AD but some symptomatic treatments are available to improve the cognitive and behavioral symptoms, however they appear to delay progression by a few months to a few years. Disease-modifying drugs to alter AD progression and improve patient's quality of life are under study and people with a risk of developing AD will greatly benefit from any disease-modifying drugs or preventive treatments. In fact, as the number of AD patients is expected to roughly quadruple by 2050 worldwide, delaying the onset of the disease by 5 years would reduce the number of AD cases by half (Brookmeyer, Johnson et al. 2007).

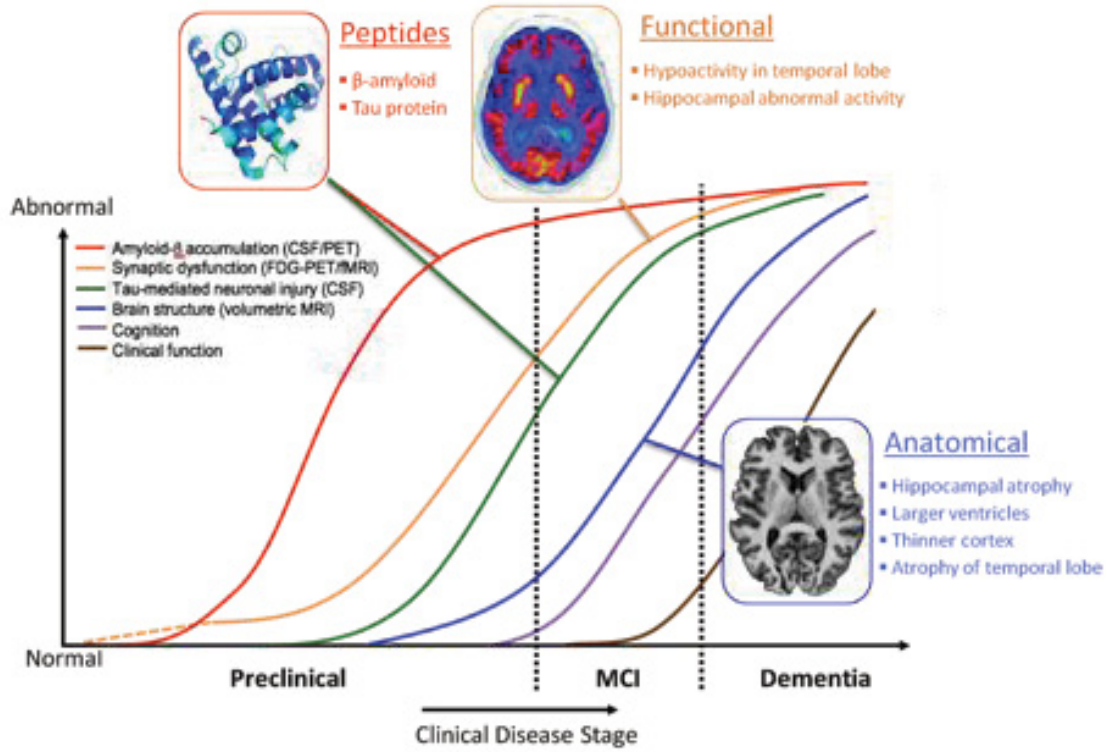


Figure 2.4: Dynamic biomarkers of the Alzheimer's pathological cascade. Aβ, Tau-mediated neuronal injury and dysfunction, memory, brain structure and clinical function evolution in AD are represented as a function of patient's clinical disease stage. Adapted from (Sperling, Aisen et al. 2011).

Neuropathology of AD

The principal neuropathological characteristics of AD are the abundance of extra-cellular amyloid plaques and intra-cellular neurofibrillary tangles, which are clearly visible on post-mortem microscopy (Crimins, Pooler et al. 2013). It is believed that the accumulation of the amyloid deposits creates a cascade of neuropathological processes leading to neuronal loss (Figure 2.4). This phenomenon is called "the amyloid cascade hypothesis" (Jack, Knopman et al. 2010).

Amyloid plaques are dense, insoluble deposits of beta-amyloid (Aβ) peptides and cellular material that aggregate around the neurons. These Aβ peptides are produced from the successive scission of larger amyloid precursor proteins (APP). $A\beta_{40}$ and $A\beta_{42}$ can be found in amyloid

plaques but $A\beta_{42}$ is believed to be more neurotoxic and have a greater tendency to aggregate (Mohandas, Rajmohan et al. 2009).

Neurofibrillary tangles are insoluble aggregates of the microtubule-associated hyperphosphorylated tau protein. The hyperphosphorilation of tau occurs during the primary phase of neurodegeneration of individual neurons and is associated with the accumulation of tau and microtubules-associated proteins inside the cells. This disturbance of the microtubules, which are an important part of intra-neuronal communication, leads to cell death (Mohandas, Rajmohan et al. 2009).

Neuronal loss results in brain atrophy when tissue volume is not compensated for by gliosis and inflammation. Gross brain atrophy is visible with lateral ventricles enlargement while the cerebrum, hippocampi and cortex shrink (Figure 2.7). These changes occur in both healthy ageing and AD populations, however, it accelerates drastically during the progression of AD in particular in the temporal lobes.

To conclude, as is the case for both MS and AD, neurodegeneration represents irreversible loss of neurons, which will inevitably lead to brain atrophy. MRI provides a non-invasive in-vivo tool to investigate the gross effects of the underlying microscopic pathology and will be described in the next section.

2.2. Imaging neurodegeneration with MRI

The discovery of X-rays by Wilhelm Roentgen at the end of the 19th century revolutionized our ability to investigate otherwise hidden anatomical structures in a non-invasive way. Since then, other non-invasive imaging techniques have been developed to explore the human body such as X-ray computed tomography, positron emission tomography and MRI. In the following section, we describe MRI as it has proven to be a useful, robust and non-invasive imaging tool to

explore in-vivo soft tissues such as the brain, and is particularly useful for quantifying the effects of neurodegeneration.

2.2.1. MRI principles and modalities

MRI can produce high-resolution anatomical images of soft tissue by exploiting the nuclear magnetic resonance (NMR) phenomenon. This work on the NMR phenomenon resulted in a double 1952 Nobel Prize in Physics for Felix Bloch and Edward Mills Purcell. The NMR phenomenon is the physical principal that describes the absorption and re-emission of electromagnetic radiation of an object when placed in a magnetic field. The development of analytic chemistry and biochemistry using NMR as well as electromagnetic and electronic technology enabled its introduction into medical use. The transfer of NMR to medical imaging came thanks to two major contributors: Paul Lauterbur for the spatial localization using magnetic field gradients (Lauterbur 1973) and Peter Mansfield for the mathematical model to reconstruct 3D images (Mansfield and Granell 1973). Their work was awarded a Nobel Prize in 2003 and MRI is now widely used in a variety of applications ranging from anatomical imaging to quantitative mapping, functional mapping, connectivity analysis to name a few.

Principles

MRI relies on the NMR principle of nuclear resonance of certain atoms when they are placed in a magnetic field. When nuclei holding an odd number of protons (or neutrons) are introduced into a magnetic field (B_0), a small fraction of the nuclei align themselves with the direction of the field and precess around it at the Larmor frequency, which is dictated by their gyromagnetic ratio. The precession of charged particles, such as nuclei, produces a small net magnetic field or moment. Due to the abundance of water (H_2O) in the human body, hydrogen protons (H) are by far the most studied in conventional MRI. Their Larmor Frequency is proportional to B_0 and to the proton gyromagnetic ratio, which is equal to 42.575 MHz/T for protons. The resulting net

magnetic moment (M_0) generated by the precession of the hydrogen nuclei is aligned with B_0 and has small amplitude (in comparison to B0) making the signal difficult to measure. However, by applying a radio-frequency (RF) magnetic pulse (B_{rf}) perpendicular to B_0 with a frequency equal to the Larmor frequency, M_0 can be rotated to the transverse plane and parallel to B_{rf}. Once the B_{rf} pulse is removed, the nuclei will realign with B_0 and the magnitude of the signal will recover based on the proton density (PD), the longitudinal (T1) and the spin-spin or transverse (T2) relaxation times (Figure 2.5). A receiver perpendicular to B_0 can detect the signal emitted by the relaxing nuclei. In order to reconstruct a 3D image of this process, magnetic field gradients are used to spatially encode this signal. By applying linear gradients along different axes, the frequency and phase of the processing nuclei are modulated based on their different spatial locations. A 2D Fourier transform applied to this data results in an image in the spatial domain.

For a given tissue type, the relaxation time constant T1 describes the exponential return to equilibrium of the longitudinal magnetization Similarly, the transverse equilibrium magnetization decays with a time constant of T2 (Figure 2.5). By manipulating imaging parameters such as the repetition time (TR) of the RF pulse, the flip angle and the time at which the image is acquired (echo time TE), different contrasts between tissue types can be achieved. In the following section, we describe a few of the conventional MRI contrasts.

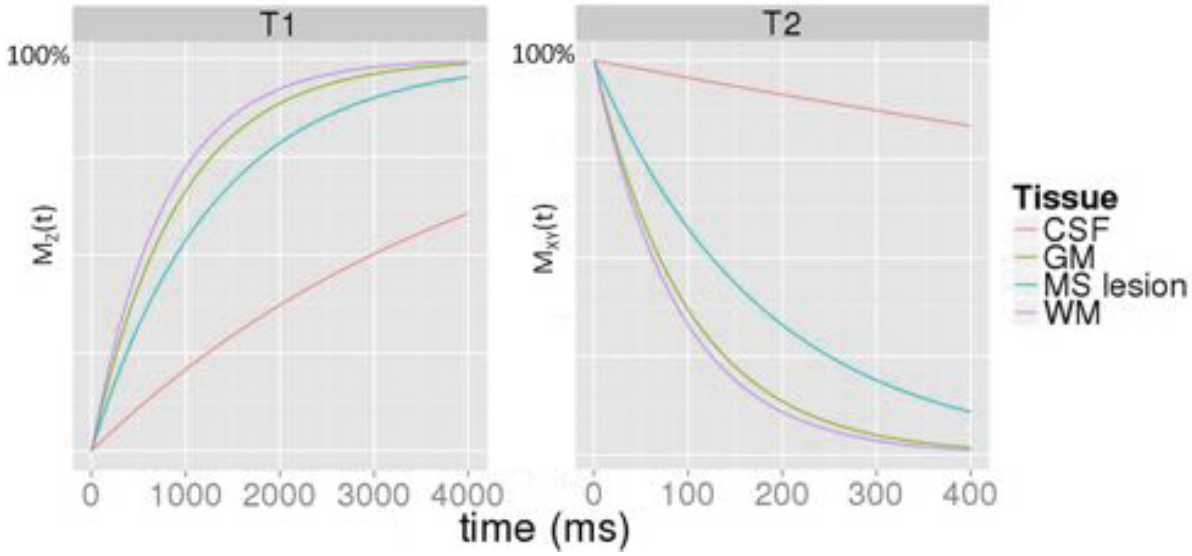


Figure 2.5: Time course of T1 relaxation and T2 decay after the RF pulse (B_{rf}). The signal (magnetization on z axis (M_z) and on the xy plane (M_{xy})) of different tissue types (CSF, GM, WM and MS lesion) is shown as a function of the time (t), T1 and T2 relaxation behaviors that can be described by exponential curves and vary considerably between the tissue types.

T1-weighted imaging

T1-weighted (T1W) images (Figure 2.6) exhibit contrast between the long T1 relaxation of fluid (i.e. dark CSF), tissues with water content (i.e.: mid-grey grey-mater) and short T1 for fat-based tissues (i.e.: bright WM). These different T1 relaxation times are related to the gyromagnetic ratio and the exchange of energy between the spins and the surrounding nuclei lattice, called "spin-lattice relaxation". The mobility of the lattice structure will increase the dissipation of energy and thus will require a longer relaxation time (i.e., for liquids such as CSF). T1W images are usually acquired with a short TE (10 to 20ms) and TR (400 to 600 ms) in order to limit the effects of T2 relaxation. Because of the high tissue contrast provided by T1W imaging, it is often used as the "anatomical" reference in neuroimaging.

Concurrently with imaging, the injection of a contrast agent can be used to enhance the visibility of certain tissues or abnormalities. For example, gadopentetate dimeglumine ("Gadolinium" or "Gad") is a contrast agent commonly used to enhance blood vessels and highlight breakdown of the BBB (Figure 2.6). This is because Gad is a large molecule that

causes T1 shortening and normally does not cross the BBB, such that vasculature and areas where the BBB is damaged are highlighted on T1w images.

T2-weighted imaging

On T2-weighted (T2W) images (Figure 2.6), fatty tissues appear darker and tissues with high water content appear brighter. T2W images are based on the dephasing of the spins after the application of the transverse B_{rf} pulse. The "spin-spin" interactions are greater in solids or semi-solid tissues (i.e. white and grey matter) compared to liquids (CSF) leading to a difference in contrast. T2W images can be obtained with a long TE (> 75 ms) and TR (> 1500 ms) in order to limit the effects of T1 relaxation.

Proton density weighted (PDW) images (Figure 2.6) are often simultaneously acquired with T2W images, using a short echo time in order to minimize the effects of T1 and T2.

Inversion recovery imaging

Inversion recovery (IR) is an imaging technique that allows the suppression of specific T1 signals from a T1W or T2W image and therefore can help to better visualize tissues of interest. To achieve this attenuation, a 180° pulse is first used to flip the longitudinal magnetization. Then, during the recovery of this magnetization, a 90° RF pulse is applied at a specific time (TI) equal to the relaxation time of the component to be suppressed. At TI, the transverse magnetization of that component will be nulled and thus effectively absent from the signal to be imaged.

A typical example is Fluid Attenuated Inversion Recovery (FLAIR) (Figure 2.6), which is a T2W sequence often used to detect WM pathologies. In this case, the bright CSF is nulled such that abnormalities such as lesions, that are also hyper-intense, are more conspicuous.

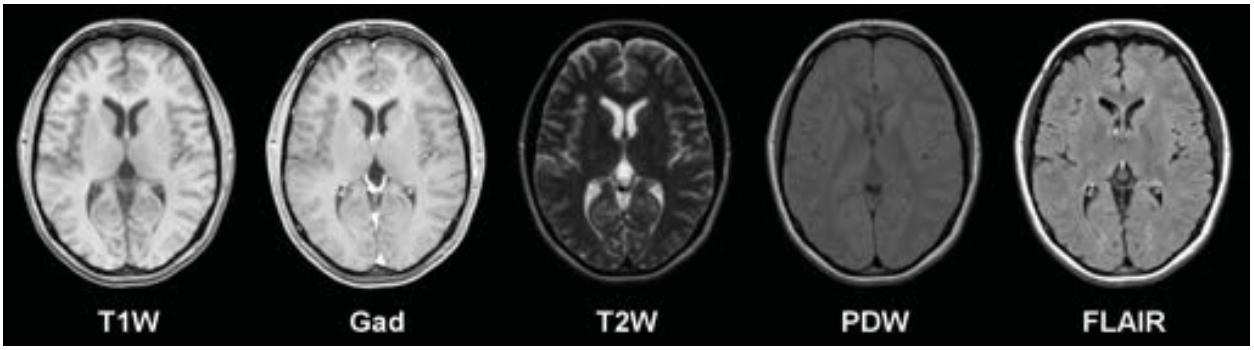


Figure 2.6: T1W, Gad, T2W, PDW and FLAIR MRI images of the same axial (or transverse) slice of an healthy young subject. These images demonstrate the polyvalence of MRI and show few typical image modalities that can be obtained.

2.2.2. Imaging neurodegenerative disease

Imaging MS

Conventional MRI acquisition protocols for MS often include the following contrasts: T1W, T2W, PDW, FLAIR and contrast-enhanced T1W. This is due to the fact that these qualitative MRI modalities are sensitive to the inflammatory and demyelinating changes directly associated with relapses, which lead to "MS lesions" and atrophy (Figure 2.7). As such, MRI is often used to monitor, identify and quantify MS progression (Fazekas, Barkhof et al. 1999).

In particular, T2W images are usually used for the diagnosis of MS where patients typically show peri-ventricular or subcortical hyperintense regions in the WM. These hyperintense regions indicate either inflammation or scar tissue and are commonly called "T2-lesions". T2W FLAIR images enable a better discrimination of T2-lesions from CSF. The formation of new lesions can be observed on contrast-enhanced T1W images using Gad. These contrast-enhanced lesions (CEL) typically last about 2 months and correlate with the BBB breakdown as well as inflammation. Thus the inclusion of enhanced T1W imaging enables the differentiation between acute, active lesions and chronic, inactive lesions (Lassmann 2008). New CELs are almost always associated with the presence of T2-lesions and they shrink in size and their intensity

decreases on T2W after a few months (Meier and Guttmann 2003, Meier, Weiner et al. 2007). When lesions cease to enhance, about 40% of them remain hypointense on T1W images ("black holes") (Sahraian, Radue et al. 2010) and are associated with irreversible tissue damage (Brück, Bitsch et al. 1997, van Walderveen, Kamphorst et al. 1998, Bitsch, Kuhlmann et al. 2001).

More recently, a measurement of cerebral atrophy on MRI images has been shown to be a good indicator of MS status and long-term progression as well as correlated with clinical measures (Clark, James et al. 1992, Huber, Bornstein et al. 1992, Barkhof, Elton et al. 1998, Pelletier, Suchet et al. 2001, Owens 2003, Confavreux and Vukusic 2006). The average whole brain atrophy rate in MS has been shown to be 0.5-1% per year versus 0.1-0.3% per year for healthy subjects and appears to be relentless (Coffey, Wilkinson et al. 1992). The different forms of MS (e.g. RRMS, SPMS) have all shown some extent of atrophy (Simon, Jacobs et al. 1999, Turner, Ramli et al. 2001, Bakshi, Dandamudi et al. 2005, De Stefano, Giorgio et al. 2010). Regional atrophy of the corpus callosum (Huber, Bornstein et al. 1992, Barkhof, Elton et al. 1998, Pelletier, Suchet et al. 2001) and ventricular enlargement (Clark, James et al. 1992, Dastidar, Heinonen et al. 1999) have shown moderate to strong correlations with clinical scores of disability (i.e.: expanded disability status score (EDSS)). All in all, measures of atrophy have shown to be strongly correlated with irreversible neurological dysfunction, which implies that atrophy might be a good representative surrogate marker for the progression of MS (Jagust and Noseworthy 2000).

The relationship between lesions and atrophy is complex and not fully understood. Some studies have found that MS T2-lesion load correlates with late brain atrophy (Chard, Brex et al. 2003). In addition, a weak correlation was found during early disease between T2-lesion load and cortical atrophy (Charil, Dagher et al. 2007, Amato, Portaccio et al. 2010), as well as brain (Sailer, Losseff et al. 2001) or spinal cord atrophy (Evangelou, Esiri et al. 2000). CELs (Meier, Weiner et al. 2004, Richert, Howard et al. 2006) and black hole lesion load (Chard, Griffin et al. 2002, Bermel, Sharma et al. 2003) have also shown correlations with brain atrophy. These results are consistent with the hypothesis that focal damage is related to diffuse atrophy.

Other less conventional MR modalities, such as MR spectroscopy, have been used to measure the integrity of the myelin sheath and axonal loss in the non-lesional NAWM and the NAGM in patients with MS (Casanova, Martínez-Bisbal et al. 2003, Vrenken, Rombouts et al. 2006). These results were consistent with previous histopathological findings (Ferguson, Matyszak et al. 1997, Trapp, Peterson et al. 1998, Bjartmar and Trapp 2003) highlighting the existence of diffuse axonal demyelisation in the NABT. Similar results were found when measuring the axonal integrity in the NAWM with diffusion weighted imaging (DWI) where possible chronic axonal changes were found in NAWM (Kezele, Arnold et al. 2008, Fox, Cronin et al. 2010).

Imaging AD

MCI patient longitudinal outcome is heterogeneous. In fact, despite having a higher risk to convert to AD, a significant number of MCI patients remain stable (Petersen, Roberts et al. 2009) or might even become clinically normal (Manly, Tang et al. 2008). Brain atrophy measurements have shown to strongly correlate with cognitive and neuropsychological scores and can help to predict the transition from MCI to AD (Jack, Shiung et al. 2004). Furthermore, as shown on Figure 2.4, AD brain atrophy rates accelerate faster than normal ageing and starts before the onset of the disease. Because of its high sensitivity in tracking AD patient progression from cognitively normal to cognitively impaired, structural brain atrophy measures from MRI are currently included as biomarkers in the diagnostic of AD (Dubois, Feldman et al. 2007). Structural brain atrophy has proven to be more accurate than standard neuropsychological evaluations (Fox, Cousens et al. 2000) and could potentially be used as a surrogate biomarker in clinical trials (Frisoni, Fox et al. 2010).

The brain atrophy pattern follows the earliest site of Tau deposition (Scahill, Schott et al. 2002) and encompasses the regions where structural measurements correlate highly with cognitive decline: hippocampi (Jack, Shiung et al. 2004), enthorinal cortex (Cardenas, Chao et al. 2011), temporal lobes (Scheltens, Leys et al. 1992), cortical thinning (Dickerson, Bakkour et al. 2009), as well as lateral ventricle enlargement (Ridha, Barnes et al. 2006). T1W images are the

standard acquisition protocol used to assess brain atrophy both manually and automatically and a few of these mentioned atrophy patterns can be appreciated on Figure 2.7. Despite their small size and convoluted shape, hippocampi have been extensively assessed and validated with manual and automatic segmentation on T1W images. These structures have also provided promising results to predict accurately the conversion of MCI subjects to AD (Yuan, Gu et al. 2009).

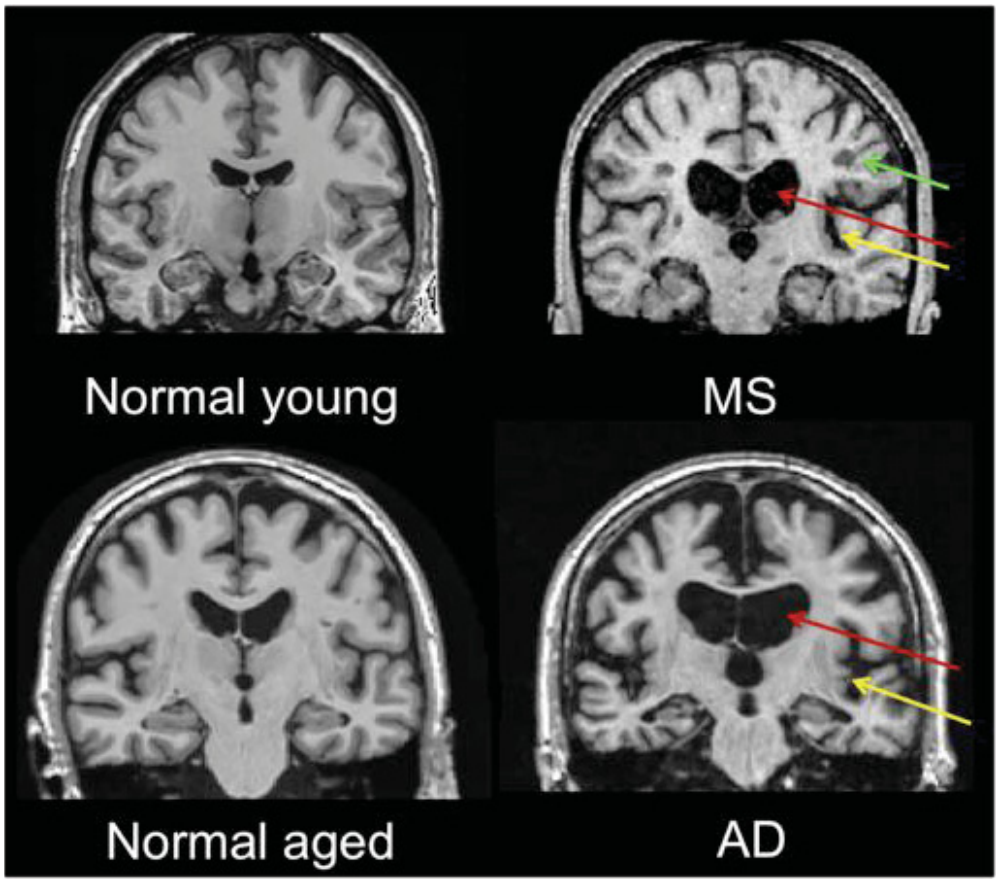


Figure 2.7: Comparison of a normal young, normal aged, MS and AD patient brains. Lateral ventricles enlargements are represented in red and the enlarged sulci in yellow. MS lesions are shown in green.

Other image modality techniques are not used clinically but have proven to be good markers for early change and could help the diagnosis of AD. DWI (Kantarci, Petersen et al. 2005) and magnetization transfer ratio (MTR) (Ridha, Symms et al. 2007) are able to capture the micro-

structural changes (axonal, dendritic, myelin), while MRS can capture the chemical environment due to the pathology (Miller, Moats et al. 1993). More recently, perfusion MRI with arterial spin labeling (Luckhaus, Jänner et al. 2010) and the default mode network from resting-state fMRI (Damoiseaux 2012) have shown promising results in terms of detecting early changes, but still need further clinical validation.

Despite all the progress in MRI acquisition techniques, the unpredictable course and short-term changes caused by neurodegeneration make clinical trials of MS and AD difficult. Thus, specific image processing is required to improve, normalize and compare MRI images in the context of longitudinal and neurodegenerative diseases.

2.3. MRI processing for neurodegenerative diseases

MRI is now the standard imaging technique used to support the diagnosis of neurodegenerative diseases and is commonly used to assess disease progression by a neurologist. Qualitatively, the increasing variety of MRI protocols has improved sensitivity and specificity in determining pathology. However, visual and quantitative assessment of images taken under a variety of protocols can be time consuming for a neurologist and difficult in large-scale MRI studies. In order to process and analyze a significant volume of images with high quality quantitative measures, computer assisted image processing has become an integral part of medical imaging. Over the last 2 decades, biomedical image processing has combined the expertise of interdisciplinary fields from applied mathematics, computer science, engineering, statistics, physics, medicine and biology to enable the development of MRI processing.

In the following section, we provide an overview of the different pre-processing steps required by most of image processing tools but first, we give an introduction to MRI brain atlases which are commonly use as a spatial reference system in medical imaging.

2.3.1. MRI brain atlas

Brain atlases or templates are referential brain images or histological series of brain slices associated with a corresponding set of anatomical labels that provide a standardized coordinate frame. In 1988, Talairach and Tournoux proposed a brain atlas, based on an individual post-mortem brain, that associated anatomical landmarks to a spatial coordinate system (Talairach and Tournoux 1988). In order to obtain a more representative atlas of normal population anatomical MR images, the Montreal Neurological Institute (MNI) and the International Consortium for Brain Mapping (ICBM), created an anatomical average of 152 MR images of young normal healthy controls (Mazziotta, Toga et al. 2001) after linear and non-linear co-registration (Collins, Neelin et al. 1994, Collins and Evans 1997) (Section 2.3.3) to the Talairach and Tournoux atlas. This template, known as the ICBM152, remains the most commonly used today, and was used in this thesis as a common reference space. Each normal control (NC) MRI that was used to create the template, was automatically segmented into different anatomical regions (Collins and Evans 1997) and these regions were used to create statistical probability anatomy maps (Figure 2.8). Furthermore, as part of this thesis, we created AD and MS population templates to identify more specifically the anatomy of each of these populations.

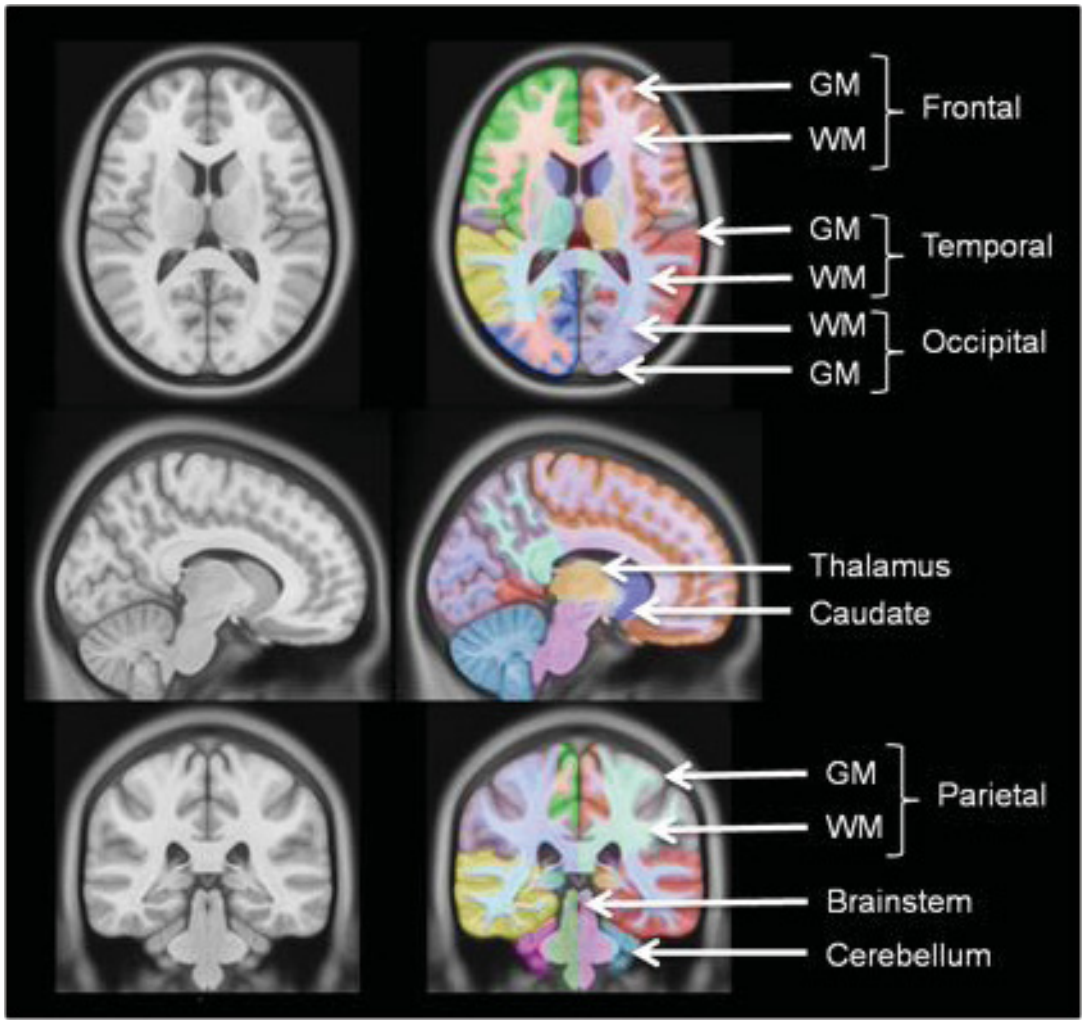


Figure 2.8: Non-linear ICBM152 T1W template. The maximum probabilistic anatomical atlas of the brain lobes are overlaid on top of the template T1W image.

2.3.2. MRI pre-processing

The corruption of MRI data can alter the final outcome of applied image processing techniques. Artifacts, a common problem in MRI, result from the complexity of the imaging technique and arise from different sources of errors (Bellon, Haacke et al. 1986). Scanner hardware limitations and physical constraints, such as field inhomogeneity can also contribute to the problem. Additionally, different source of spurious noise, related to the sequence and the

patient, can degrade the images. These artefacts and constraints can degrade the images but they can also lead to more subtle effects (i.e. intensity contrasts) which can lead to misinterpretation of the images and have an impact on the following processing methods: segmentation, tissue classification, and registration. For these reasons, MRI pre-processing techniques mentioned in this section can be used to remove or minimize the effects of artifacts.

Intensity non-uniformity correction

A multiplicative model with additive noise is commonly used to describe an image corrupted by a non-uniformity field:

$$x(i) = g(i).\hat{x}(i) + n(i) \qquad \text{Equation 2.1}$$

Where $x(i)$ is the measured intensity and $\hat{x}(i)$ is the original intensity of the image at the location $i = (x, y, z)$, $g(i)$ is the non-uniformity effect, and $n(i)$ is the noise.

Intensity non-uniformity is partially independent of the anatomy of the scanned subject and is mainly caused by local variation of the flip angle, eddy currents and standing waves in tissue. An example of the non-uniformity bias field is provided in Figure 2.10.E. The combination of the MRI sequence and the shape of the object being scanned can trigger these local non-uniformity variations. Despite careful correction of the inhomogeneity of the static magnetic field B_0, residual variations are possible and result in intensity non-uniformity as well as spatial distortion (Bridcut, Redpath et al. 2001). RF coil geometry (Collins, Li et al. 1997), gradient fields (Langlois, Desvignes et al. 1999) and the shape of the imaged object (Sled and Pike 1998) can also have an impact on the inhomogeneity.

Intensity inhomogeneity manifests by smooth intensity variation across the image and can significantly alter the performance of image segmentation and registration which are sensitive to the spurious variation of image intensity. Numerous different approaches have been proposed to

minimize the adverse effects of inhomogeneity. The estimation of the inhomogeneity effect can be done prospectively or retrospectively and then simply divided from the image signal (*x(i)*). Prior information that are obtained from measures on an MRI phantom (Narayana, Brey et al. 1988) or mathematical model (McVeigh, Bronskill et al. 1986) can also be used to predict the inhomogeneity map but neither of these methods cannot estimate the inhomogeneity induced by the object. The most common approaches are retrospective, estimating the inhomogeneity directly from the image itself. These can be classified as follows:

- Spatial filtering: As inhomogeneity are slowly varying in the image domain, using low-pass or homomorphic filtering in the frequency domain can help extract the true signal. The simplicity of the low-pass filtering methods make them popular, however, the overlapping intensity spectrum from the inhomogeneity and the imaged object and the possibility of introducing ringing artifacts limits the efficiency of these methods (Brinkmann, Manduca et al. 1998).

- Surface fitting: Since inhomogeneity fields are smooth, they can be approximated by a smooth function using splines (Dawant, Zijdenbos et al. 1993) or polynomial basis functions (Tincher, Meyer et al. 1993) but only for 2D slices. In both of these approaches, single or multi-pass fitting can be used, with control points being obtained manually or from automatic tissue segmentation.

Statistical methods: Statistical methods developed to segment images can also be used to model inhomogeneity. Here, inhomogeneity is assumed to follow a distribution (i.e. Gaussian) (Wells, Grimson et al. 1996) or can be modelled as a random process (i.e. Markov random field) (Held, Rota Kops et al. 1997). However, other non-parametric approaches, such as the non-parametric non-uniform intensity normalization (N3) technique proposed by Sled et al. (Sled, Zijdenbos et al. 1998) consider the problem as a deconvolution and therefore do not require segmentation. N3 simplifies the inhomogeneity distribution as a maximization problem by iteratively estimating a smooth Gaussian multiplicative field which maximize the frequency content of the tissue intensities. In this thesis, we used the N3 algorithm as our work focuses on

images acquired on a 1.5T MRI magnetic field (B_0), for which the N3 was developed. For higher MRI fields, N4 was proposed to improve the original N3 B-spline smoothing estimation strategy (Tustison, Avants et al. 2010).

Denoising

In MRI, the Rician distribution of noise can be caused by many factors, including the receiving coil resistance, induction loss in the imaged object, the B_0 strength, the object size, the voxel size, receiver bandwidth, the number of averages and most importantly the thermal noise in the amplifier (Gudbjartsson and Patz 1995). Noise removal, therefore needs to be robust to the different MRI image protocols and the object being scanned, while preserving relevant information. Many approaches have been proposed to estimate the original image ($\hat{x}(i)$) contaminated with noise ($x(i)$): Gaussian filters (Ashburner and Friston 2000), anisotropic diffusion filters (Gerig, Kubler et al. 1992), wavelet-based methods (Donoho and Johnstone 1994), and bilateral filtering (Tomasi and Manduchi 1998). More recently, Buades et al. (Buades, Coll et al. 2005) proposed the non-local means (NLM) approach to denoise images. This approach takes advantage of image redundancy to average similar local realizations of the noisy image (Figure 2.9). The noise of the image is reduced by averaging the voxels that would have a similar intensity in the noise-free image. The denoising of voxel $x_{(i)}$ is achieved, by comparing the patch $P(x_{(i)})$ centered on i with all the patches $P(x_{(j)})$ centered on j of the images in the neighbourhood Ω such that:

$$\hat{x}(i) = \frac{\sum_{j\in\Omega} w(i,j)x(j)}{\sum_{j\in\Omega} w(i,j)} \quad where\ w(i,j) = e^{-\frac{\|P(x(i))-P(x(j))\|_2^2}{h^2}} \qquad \text{Equation 2.2}$$

The term $w(i,j)$, a weight based on the similarity of patches $P(x_{(i)})$ and $P(x_{(j)})$, is designed to attribute a smaller weight to the greater L2-norm ($\|.\|_2$) (or sum of square differences (SSD)) distance measures. The h^2 term is a smoothing parameter that is proportional to the noise variance. This denoising technique shows outstanding results in MRI denoising (Coupé, Yger et

al. 2008, Manjón, Thacker et al. 2009). An example of NLM denoising results on MRI is provided in Figure 2.10.B.

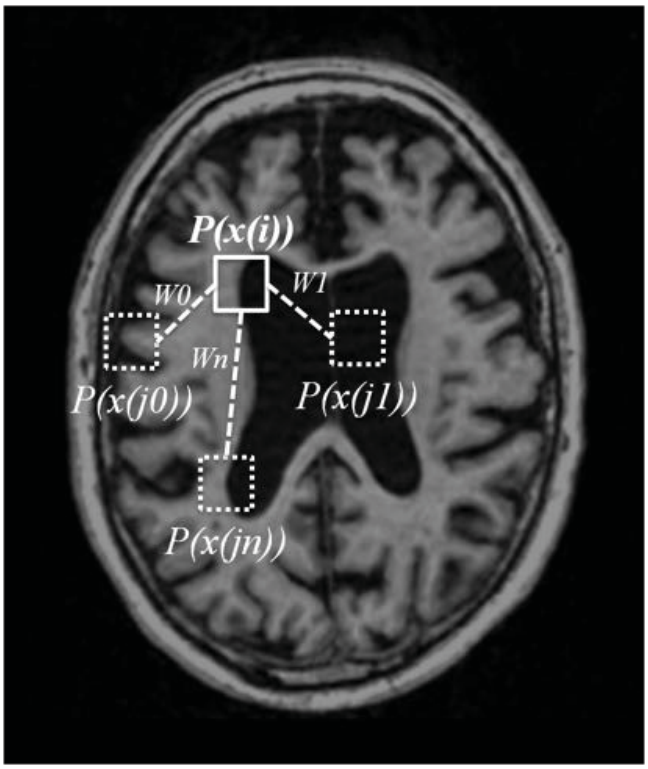


Figure 2.9: NLM estimator principle. When evaluating patch *P(x(I))* (white square), similar patches such as patch *P(x(jn))* (white dotted squares), give a larger weight (*Wn*) while smaller weights (W0 and W1) are given to more different patches (*P(x(j0)* and *P(x(j1)*).

Intensity normalization

Another drawback of MRI is the lack of quantifiable standard image intensities, as is possible with CT. Indeed, even in an image acquisition of the same body part on the same subject with the same MRI protocol/sequence, the image intensities do not hold a fixed quantitative meaning making direct quantative voxel comparison impossible. This lack of normalization requires the adjustment of the intensity window setting in order to display the images with the optimal contrast. This intensity variability can also limit automatic segmentation (Zhuge and Udupa 2009) and registration techniques (Bağcı, Udupa et al. 2010). To normalize the intensity between

protocols, subjects and image templates, different MRI calibration methods have been proposed. Similar to inhomogeneity correction, methods that do not require extra acquisitions (i.e. MRI phantoms) are usually preferred.

The most common intensity normalization approaches generally aim to match the intensity histograms of two images, a reference (or model/template) and the considered image. The basic idea of these methods is to find a transformation of the image histogram so that it matches the reference image histogram by minimizing the bin count difference. To make the methods more robust and more specific to the region of interest, thresholds or masks can be used to remove the background. This transformation of the image histogram can be linear (Wang, Lai et al. 1998) where a multiplicative scale factor is applied to modify the gain ($g(i)$), piece-wise linear (Nyúl and Udupa 1999, Nyúl, Udupa et al. 2000) or non-linear (Jäger, Deuerling-Zheng et al. 2006). Other methods, which use region or tissue class specific normalization where the source and the reference histograms are modeled (i.e. Gaussian mixture models (Hellier 2003)) or where tissue model parameters are defined manually (Wells, Grimson et al. 1996), have also been proposed.

The effectiveness of these methods was mainly assessed on "normal" MRI images without pathological lesions or atrophies where histogram matching can be challenging. As demonstrated by Shah et al. (Shah, Xiao et al. 2011) for pathological intensity normalization in MS, piece-wise linear normalization such as in Nyul et al. (Nyúl and Udupa 1999) can offer a good compromise between computational burden and better image processing.

Linear and piece-wise linear normalization of histograms is both easier to customize to various anatomical regions and more robust. Furthermore, these methods do not rely on specific statistical properties of tissue classes that can vary with pathologies. Another advantage is their lower computational complexity. For these reasons, linear and piece-wise linear normalization are the most common methods for normalization in the presence of pathology and therefore were used in this thesis. An example of the piece-wise linear normalization can be appreciated in Figure 2.10.B where image Figure 2.10.A is normalized to the ICBM152 T1W template (Figure 2.10.C).

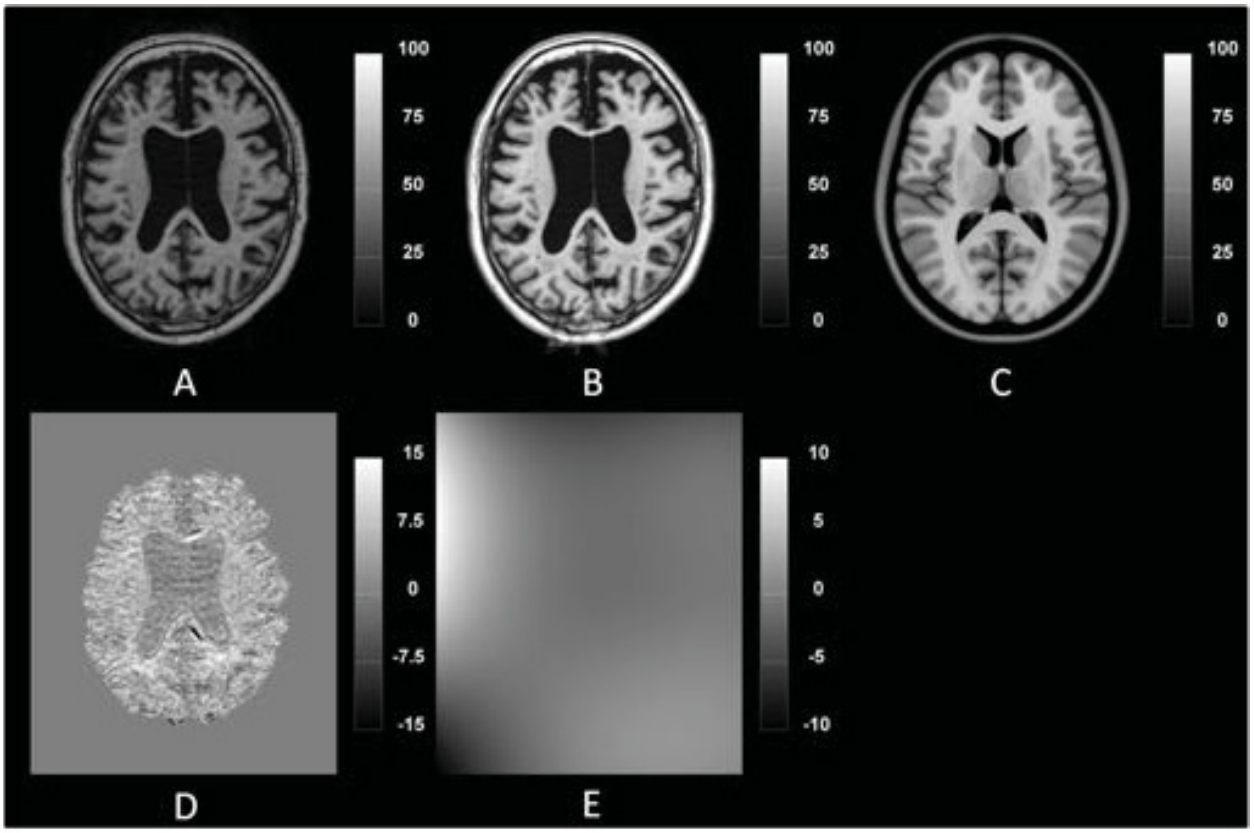


Figure 2.10: T1W image pre-processing. A) original T1W image, B) pre-processed T1W image (denoising, non-uniformity correction and intensity normalization to image C), C) ICBM152 T1W template D) Intensity difference image after denoising, normalization and non-uniformity correction (or A minus B), E) N3 bias field (Note the different ranges on the scale bars).

2.3.3. MRI registration

After the pre-processing steps, images should present little to no intensity non-uniformity and their intensity distribution should be normalized and have minimal noise. Following these pre-processing steps, image registration is a crucial task for medical imaging with numerous applications, including: i) longitudinal individual comparison of anatomical changes to determine treatment effect, ii) statistical analysis of population anatomy when comparing to a reference template, iii) non-linear registration-based segmentation, and iv) template creation. Registration is still an active field of research and numerous methods have been proposed. For

the purpose of this thesis we focus on intensity based registration techniques. A detailed review of image registration can be found in Hajnal et al. (Hajnal and Hill 2014).

The aim of image registration is to find the spatial transformation, $T(i)$, between two images, a moving image, *I(i),* and a reference image, *R(i),* which maximizes their similarity, such that:

$$R(i) = I\big(T(i)\big) \Leftrightarrow T: I(i) \rightarrow R(i) \qquad \text{Equation 2.3}$$

This problem can be rewritten under the form of an objective function where the optimization consists of finding its minimum or maximum depending on the similarity metric used:

$$M\left(R, I\big(T(i)\big)\right) + W(T(i)) \qquad \text{Equation 2.4}$$

In the following sections, we will describe the different components of this problem: transformation model (*T*), optimization, similarity metric (*M*), regularization (*W)* and interpolation).

Transformation model

Depending on the final application of the registration, different transformation models (*T*) can be chosen, such as rigid-body, affine and non-linear.

Rigid transformations are global in nature; they cannot model local anatomical differences (Figure 2.11.B) and they can be described as a translation and rotation. While affine transformations combine translation, rotation and scaling and allow for shearing and skewing, thus preserving points, straight lines and planes. A 3D affine transformation can therefore be defined using 12 parameters. Rigid or affine transformations are often used as an initial step for non-linear registration and for aligning multiple image modalities of a subject.

Non-linear transformations are more appropriate to estimate the local differences that might occur between different time-points of an MR image of an AD patient, for example. Indeed, non-linear transformations can locally warp the source image to the reference image through local deformations as illustrated in Figure 2.11.C. The choice of the transformation is important as a trade-off between computation cost and anatomical correspondence. We can classify the non-linear transformation models into two main categories: physical models and interpolation models. Physical models are used to enforce the topology property using theories derived from continuum mechanics of solid materials (i.e. elastic body) and fluids (i.e. viscous fluid). Interpolation theory methods use basis function expansion to model the transformation. Many mathematical basis functions are used to interpolate the deformations, for example: radial, b-spline, thin-plate splines and wavelets. Besides these two categories, non-linear transformation models can also be described by the constraints applied to the deformation when specific knowledge is known (i.e. geometric transformations deriving from biomechanical models of specific organs).

We provide a summary of non-linear transformation models, inspired by the works of Holden et al. (2008) and Sotiras et al. (2013), in Table 2.1.

Table 2.1: Summary of non-linear transformation methods, with their principles and authors. For Further details, the reader can refer to the reference.

Non-linear transformation models		Principles	Authors
Physical models	Elastic body	Navier-Cauchy Partial Differential Equation (PDE) Riemannian elasticity energy	(Davatzikos 1997) (Christensen and Johnson 2001) (Leow, Huang et al. 2005)
	Viscous fluid	Navier-Stokes equation (velocity field)	(Bro-Nielsen and Gramkow 1996, Christensen, Rabbitt et al. 1996)
	Optical flow	Maxwell's demons(Diffusion model) Diffeormorphic flow (Large Deformation Diffeomorphic Metric Mapping (LDDMM))	(Thirion 1998) (Miller, Trouve et al. 2002, Vercauteren, Pennec et al. 2007)
Interpolation theory	Radial basis functions	Thin plate splines (TPS)	(Bookstein 1989, Younes 2006)
	Free Form Deformations (FFD)	Cubic B-spline	(Rueckert, Aljabar et al. 2006, Shi, Zhuang et al. 2012)
	Locally Affine Models	Block-matching or poly-affine	(Collins and Evans 1997)

The transformation model is closely related to regularization, which can be applied to the transformations in order to obtain specific properties of the transformation (symmetry, topology preservation, diffeomorphism).

The majority of non-linear registration algorithms are asymmetric and provide different results depending on direction of the registration ($I \rightarrow R$ or $R \rightarrow I$). For longitudinal non-linear registration, this symmetry, or inverse consistency is crucial to prevent bias from a specific deformation direction. Christensen et al. (Christensen and Johnson 2001) suggested inverse consistency by combining both the forward and the backward transformation and penalizing the inconsistency during the registration process. Other types of approaches use symmetric objective functions to estimate the transformation (Leow, Huang et al. 2005), while others estimate the forward and the backward transformations by optimizing a standard objective function. In these cases, both transformations map the image to a common space (often called half-way

space)(Noblet, Heinrich et al. 2008). The final transformation (i.e. from the moving to the reference image) can be obtained by combining the forward (to halfway space) and the inverse of the backward transformation (from halfway space). Further details on the inverse consistency technique are provided in Chapter 5.

The inverse consistent transformation can preserve topology but other methods require topological constraints to prevent collapsing and crossing of deformation fields. Topology preservation or homeophormism can be achieved by stretching and compressing without tearing. Figure 2.11.C provides an example of a homeomorphic transformation map. A one-to-one mapping between images is not always possible and can, without sufficient regularization, create a discontinuous mapping. To preserve the topology and provide an invertible and continuous one-to-one mapping, two conditions are mandatory (Musse, Heitz et al. 2001): (1) transformation is bijective and (2) the Jacobian determinant (*J*) of the transformation is positive (Equation 2.5). Indeed, *J* represents the amount of local volume change with its expansion (>1) or contraction (<1) and negative values describes that the deformation collapse onto itself.

$$J\big(T(i)\big) = det\begin{bmatrix} \frac{\partial T(i)_1}{\partial i_1} & \cdots & \frac{\partial T(i)_1}{\partial i_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial T(i)_m}{\partial i_1} & \cdots & \frac{\partial T(i)_m}{\partial i_n} \end{bmatrix} \qquad \text{Equation 2.5}$$

These conditions can be integrated into the objective function or the optimization scheme (Noblet, Heinrich et al. 2006, Durrenberger, Fernando et al. 2014). To conclude on transformation models, diffeomorphic transformations preserve the topology, by definition, as they provide an invertible and differentiable transformation mapping (Vercauteren, Pennec et al. 2009).

In this thesis, we used the in-house ANIMAL (Automatic Nonlinear Image Matching and Anatomical Labeling) non-linear registration approach (Collins and Evans 1997), which is based on a small deformation assumption that is estimated from local block-matching. Such a

constraint is justified when looking at the small deformations expected in AD and MS. ANIMAL regularizes the estimated deformations with Gaussian kernels that approximate a linear-elastic model. These types of non-physical modeling techniques present a major advantage over explicit physically based approaches, as they are computationally efficient. Indeed, physical modeling, such as viscous fluid and elastic body, is problematic as typically the observed physical phenomena are very complex and finding the underlying model's parameters requires a multitude of assumptions and/or very expensive computation. Furthermore, the approximated linear-elastic model such as the one used in ANIMAL has shown to correspond well to real tissue deformation when compared to elastic model (Rogelj and Kovacic 2004). Another advantage of using small deformations is that the arithmetic manipulation of displacement fields is well defined and allows for symmetric constraints as well as group-wise and/or longitudinal registration (Joshi, Davis et al. 2004).

Optimization

In image registration, the optimization aims to estimate the transformation parameters that maximize the similarity of the moving and the reference image according to an objective function. The difficulty to optimize the registration process correlates with the complexity of the transformation model. Linear registration can require optimizing up to 12 parameters for affine registration and could in theory use a global optimization scheme on the whole image.

This task is much more complex for non-linear registration where the number of parameters is much higher. To solve this problem, different sophisticated optimization algorithms have been proposed and choosing an appropriate algorithm depends highly on the objective function. Indeed, if the objective function is differentiable and the variables are real values, continuous optimization such as gradient descent, Gauss-Newton and Levenberg Marquardt can be used. On the other hand, discrete methods such as Graph-based and linear programming need to be chosen when the parameters take a finite set of values (ie. grid-based deformation (Broit 1981)). A summary of a few of the most popular optimization methods, inspired by the work of (Sotiras, Davatzikos et al. 2013), is provided in Table 2.2.

Table 2.2: Summary of optimization methods, with their principles and authors that used them for registration. For Further details, the reader can refer to the reference.

Optimization method		Principles	Authors
Continuous method	Gradient descent	Follows the direction which decreases the energy of the objective function (negative gradient)	(Viola and Wells 1997)
	Gauss-Newton	Minimize the sum of squared differences	(Vercauteren, Pennec et al. 2007, Ashburner and Friston 2011)
	Levenberg-Marquardt	Similar to Gauss-Newton but faster to converge	(Sawhney and Kumar 1997, Thevenaz and Unser 1998)
Discrete methods	Graph-based	Max-flow min-cut principle	(So, Tang et al. 2011)
	Linear-programming	Linear programming (i.e. Simplex)	(Collins and Evans 1997)

The continuous and discrete methods are limited regarding what objective functions and structures they can optimize. Furthermore, the type of deformation of an image determines the complexity of the registration problem.

Using a hierarchical approach with a coarse-to-fine strategy can prevent the convergence to an incorrect solution, termed a “local minimum” or “local maximum”. In hierarchical strategies first larger scale deformations are estimated, these are then followed by finer scales (Vanderbrug and Rosenfeld 1977). Starting from a lower resolution, this strategy, which usually follows a Gaussian pyramid, reduces the search domain and thus decreases the computational burden. In such strategies, regularization is often performed through convolutions with a Gaussian kernel at each level of the pyramid, such as in the ANIMAL algorithm (Collins and Evans 1997).

Similarity metrics

The similarity metrics (*M*) measure the correspondence between the images to be registered. Different features can be used to match the images such as geometric descriptors (i.e. SIFT (Urschler, Bauer et al. 2006)) but for the purpose of this thesis we only describe intensity-based similarity metrics. Because intensity-based similarity are performed at the voxel level, they can better estimate dense deformation fields and capture the underlying physical image differences.

The simplest and most intuitive similarity measure is the sum of square difference (SSD) or L^2-norm. This metric is very sensitive and can be used only if the images have similar intensities but, as mentioned above, image intensities of MRI vary.

$$SSD = \frac{1}{x \times y \times z} \sum_{i=1}^{x \times y \times z} \|I(T(i) - R(i))\|_2^2 \qquad \text{Equation 2.6}$$

Cross-correlation (CC), which assumes a linear relationship of their voxel intensities when aligned, is widely used for mono-modal image registration (Leese, Novak et al. 1971):

$$CC = \frac{\sum_i \left(I(T(i)) - \overline{I(T(\imath))}\right)(R(i) - \overline{R(\imath)})}{\sqrt{\sum_i \left(I(T(i)) - \overline{I(T(\imath))}\right)^2} \sqrt{(R(i) - \overline{R(\imath)})^2}} \qquad \text{Equation 2.7}$$

where $\overline{I(.)}$ and $\overline{R(.)}$ are the mean intensity of image *I(.)* and *R(.)*.

In the case of multi-contrast image registration (i.e. T2W registered to T1W images), a linear correlation between both images cannot be assumed. The information theory field provides measures that can capture the "mutual information" (*MI*) of two images (Collignon, Maes et al. 1995, Viola and Wells 1997), defined as:

$$MI(I,R) = H(I) + H(R) - H(I,R)$$ Equation 2.8

where $H(.)$ is Shannon's entropy (Shannon 1948) and represents the randomness (or dispersion) of the images I, R and of the joint probabilities distribution $p(I,R)$, such that:

$$H(I) = -\sum_i p(I(T(i))) \log\left(p(I(T(i)))\right)$$

and

$$H(R) = -\sum_j p(R(j)) \log\left(p(R(j))\right)$$ Equation 2.9

$$H(I,R) = -\sum_{i,j} p(I(i),R(j)) \log\left(p\big(I(i)\big),p(R(j))\right)$$ Equation 2.10

Thus maximizing MI is equivalent to minimizing the joint distribution entropy of I and R. The joint probability distribution of the images (I and R) can be estimated from their joint histogram and this joint histogram shows smaller dispersions when both image are aligned.

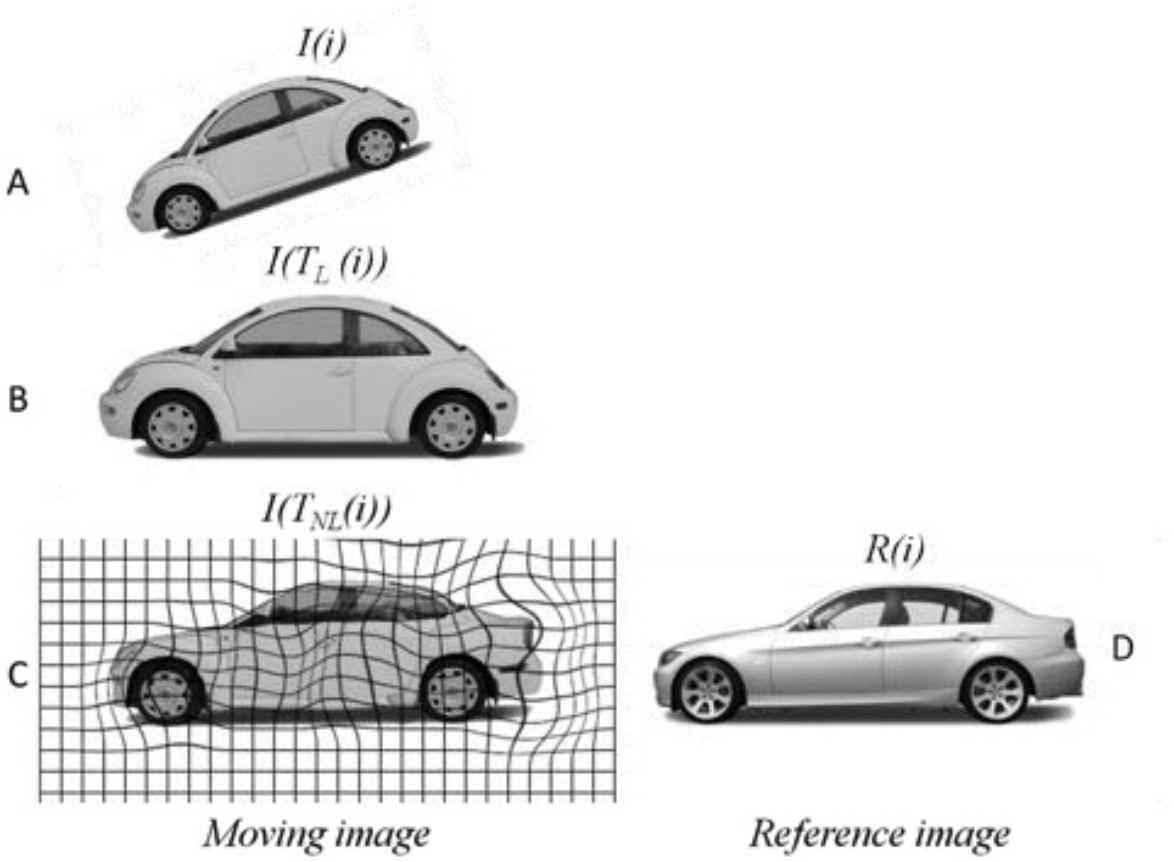


Figure 2.11: Image registration transformations. The image shows the reference image, *R(i)*, in D and the source, *I(i),* image in A. The linear registration $I(T_L(i))$ is shown in B with rotation, translation and scaling while in C the non- linear registration $I(T_{NL}(i))$ with its deformation map overlay on top of the image is displayed. Figure inspired by (Avants, Grossman et al. 2006)

2.3.4. MRI segmentation

Image segmentation consists of dividing an image into homogeneous regions (intensity, texture, color...) by assigning a label to "similar" features. Whereas MRI segmentation allows for the extraction of anatomical structures (cerebrum, lateral ventricles...), tissue classification is used for assigning tissue classes (WM, GM, CSF...) to voxels. Tissue pathology such as MS lesion can be seen as a classification and/or a segmentation problem. In this thesis, MRI images

are used to segment anatomical regions of the brain but also MS lesions. In the following section, we provide a brief summary of different MRI segmentation approaches.

Manual segmentation

Manual segmentation is still recognized as the gold standard for MRI segmentation, but it is time consuming and despite precise protocols and expert training, manual segmentation results in inter- and intra-rater variability. It requires a detailed segmentation protocol for the structure to be identified and a good understanding of brain anatomy from the person segmenting. In the case of structures with high variability (i.e. shape, size, contrast…), such as MS lesions, manual segmentation (on the same subject by different experts) had an inter-rater reliability agreement of only 25% (García-Lorenzo, Francis et al. 2013). For these reasons manual segmentation is difficult to apply to large cohort studies. Conversely, automated or semi-automated segmentation techniques hold the potential of being more robust and reducing variability when the resulting segmentations are of sufficient quality.

Anatomical segmentation methods

There are numerous automated segmentation techniques that enable identification of anatomical structures in MRI images. An exhaustive review of the literature is available in Cabezas et al. (2011). Example techniques include deformable models or region growing (Ghanei, Soltanian-Zadeh et al. 1998, Shen, Moffat et al. 2002, Chupin, Mukuna-Bantumbakulu et al. 2007), appearance-based models (Duchesne, Pruessner et al. 2002, Hu and Collins 2007), and atlas/template-warping techniques (Collins and Evans 1997, Fischl, Salat et al. 2002, Rohlfing, Brandt et al. 2004, Zhou and Rajapakse 2005, Heckemann, Hajnal et al. 2006, Hammers, Heckemann et al. 2007, Barnes, Foster et al. 2008, Gousias, Rueckert et al. 2008, Aljabar, Heckemann et al. 2009).

The in-house template-warping technique proposed by (Collins and Evans 1997) named ANIMAL performs non-linear registration using *minctracc* (cf. Section 2.3.3) of the image to be

segmented to the ICBM152 template. Then, the inverse of the non-linear deformation is used to interpolate the anatomical atlas onto the subject image space (Figure 2.14).

An individual template many not be sufficient to capture the individual or pathological variability of the anatomy. Thus, to overcome this limitation, multi-atlas segmentation approaches have been proposed. In this case, a library of MR images with their respective manual expert-based segmentations is used. Then linear and/or non-linear registration is used to register the subject image to the library. Next, the fusion of the labels is done using a majority-voting rule. To decrease the anatomical variability of the library, a pre-selection of the most similar templates is possible (Aljabar, Heckemann et al. 2009). Multi-atlas segmentation has demonstrated better accuracy to segment anatomical structures, yielding less error than single template approaches (Rohlfing, Brandt et al. 2004, Heckemann, Hajnal et al. 2006, Hammers, Heckemann et al. 2007, Gousias, Rueckert et al. 2008, Aljabar, Heckemann et al. 2009, Collins and Pruessner 2010). Multi-atlas label fusion segmentations have improved over individual exemplar segmentation approach results. However, these methods are still limited by the atlas biases, mis-registration, and interpolation error.

Recently, a solution to these limitations of structural segmentation has been proposed in terms of patch-based methods that use the NLM operator (Buades, Coll et al. 2005) (described in section 2.3.1). Indeed, the non-local approach exploits image redundancy within a library of images and their respective segmentation to obtain a larger number of samples by looking for similar patches in a given search area. In these approaches, the intensity-based distances between the patch under study and corresponding patches in the template library are used to perform a weighted label fusion based on the NLM estimator (Figure 2.12). Furthermore, the weighted average of the most similar patches provides a more robust labelling than traditional majority voting where the same weight is given to all the samples. Patch-based segmentation methods have gained wide popularity and have shown impressive results, despite their simplicity. They have been applied to segment a multitude of anatomical structures including hippocampus (Coupé, Manjón et al. 2011, Tong, Wolz et al. 2013), brain (Eskildsen, Coupé et al. 2012), lateral ventricles (Fonov, Coupé et al. 2012), deep nuclei (Xiao, Fonov et al. 2014), intracranial cavity

(Manjón, Eskildsen et al. 2014) and other structures of the brain (Rousseau, Habas et al. 2011, Zhang, Guo et al. 2012).

In this thesis, we applied NLM segmentation, as well as template-warping approaches to segment different structures. Example segmentations are provided in Figure 2.13 for brain (Eskildsen, Coupé et al. 2012), lateral ventricles (Fonov, Coupé et al. 2012) and hippocampi (Coupé, Manjón et al. 2011).

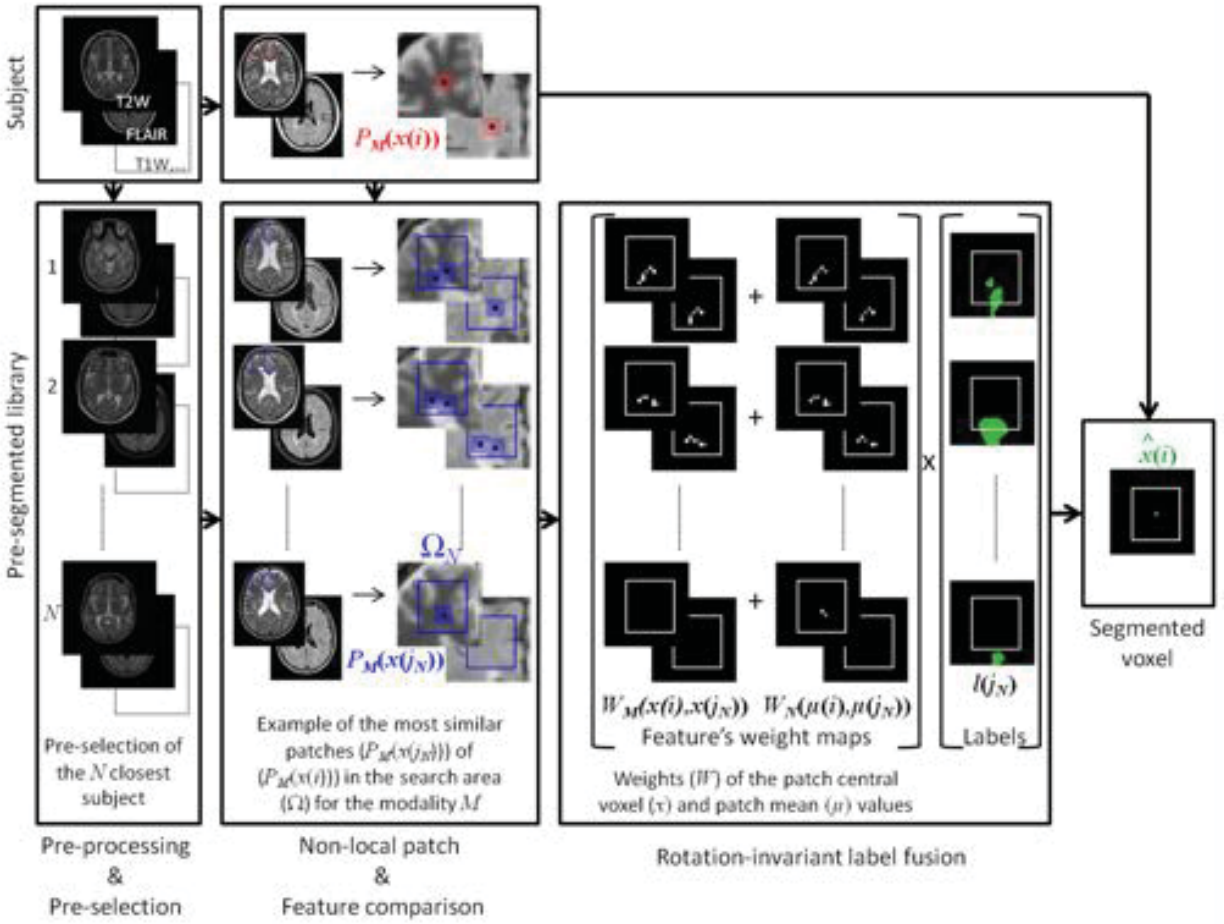


Figure 2.12: Multi-atlas NLM segmentation overview. The patch $P(x)_i$ of the image modality M to be segmented (in red) is compared with the patch in the search area Ω of the pre-selected N training subjects. The most similar patches are represented in blue and their corresponding weights are represented on the weight maps. Then a weighted average of the library of manual segmentation $l(j_n)$ is estimated to obtain the final label (green). In this exemple, the weighted average is estimated from the central ($x(i)$) and the mean value ($\mu(i)$) of the patch as described in Chapter 3. Graphical abstract from (Guizard, Coupé et al. 2015).

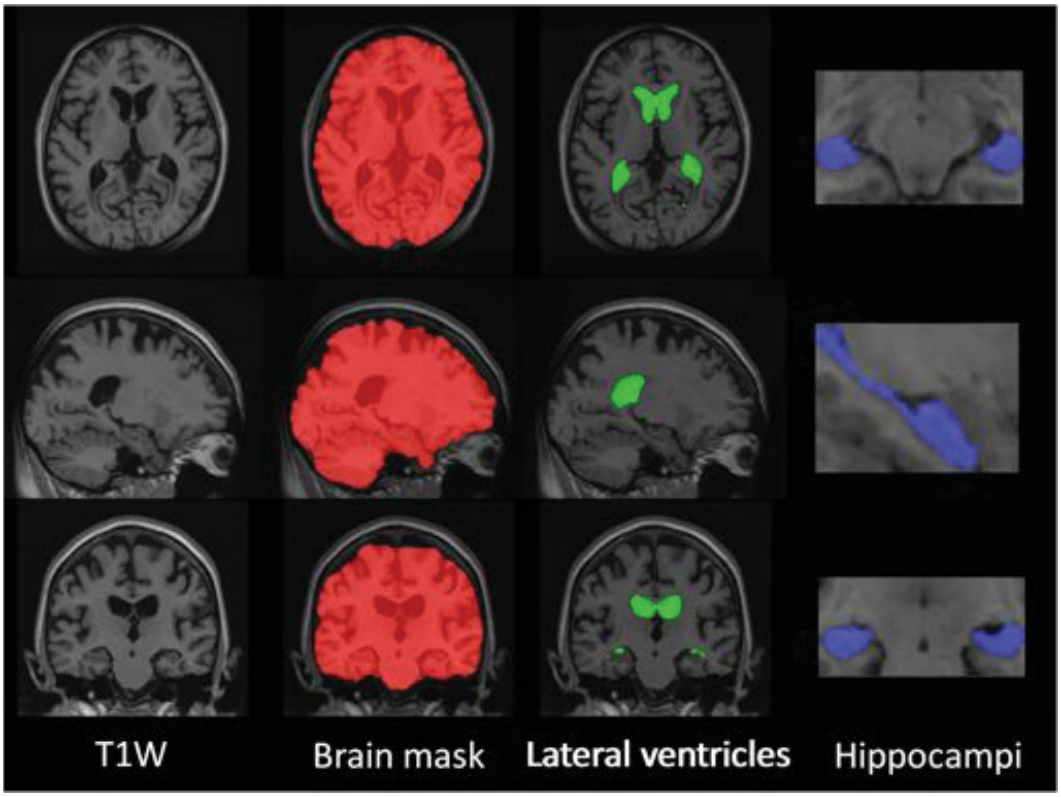


Figure 2.13: NLM segmentation examples. From left to right, T1W image, the brain (red), lateral ventricles (green) and hippocampi (blue) segmentation..

Tissue classification and lesion segmentation methods

Tissue classification can be considered as a form of image segmentation where the task is to classify the voxel into tissue types rather than structures; for example, WM, GM and CSF (Figure 2.14). MRI tissue classification represents an important problem in medical image analysis. In cases of disease or pathology, one may want to identify lesion or tumour. Tissue classification has numerous applications related to diagnosis, surgical planning, image-guided interventions, monitoring therapy, and clinical drug trials. Tissue class volumes can be used as a disease progression biomarker (Bendfeldt, Klöppel et al. 2012, Weygandt, Hummel et al. 2015). Tissue classification is also used to improve image-processing techniques (i.e. normalization, registration…) and for further processing (e.g., cortical thickness, partial volume effects…). For the purpose of this thesis, we only describe classification techniques that can identify MS lesions.

From two recent literature reviews of MS lesion segmentation methods (Lladó, Oliver et al. 2012, García-Lorenzo, Francis et al. 2013), two main categories of automatic classifiers emerge from the literature: unsupervised and supervised.

Supervised methods learn features from previously segmented datasets (intensity, local gradients, etc.). The features in new images to be segmented are compared with those in the training sets in order to estimate the lesions. These methods can use different techniques to select the features, including: artificial neural networks (ANN) (i.e. INSECT (Intensity Normalized Stereotaxic Environment for Classification of Tissues) from (Zijdenbos, Dawant et al. 1994)), k-nearest neighbours (K-NN) (Vinitski, Gonzalez et al. 1999), decision trees (Kamber, Shinghal et al. 1995), random decision forests (RDF) (Geremia, Clatz et al. 2011), and Bayesian frameworks (Harmouche, Collins et al. 2006). Patch-based techniques have also been used for tissue classification (Cordier, Menze et al. 2013). We propose an adaptation of the NLM patched-based segmentation technique to the MS lesion classification in Chapter 3.

Unsupervised methods do not require manual segmentation of the lesions a priori, but estimate tissue classes or clusters of similar voxels based on voxel features such as intensity, with or without the help of anatomical and MRI knowledge. Unsupervised techniques were initially developed to classify healthy brain tissues from MRI intensities into three classes (CSF, WM and GM) using a fuzzy C-means and Gaussian mixture models and expectation maximization (EM) (Wells, Grimson et al. 1996). To detect MS lesions, some groups adapted the Gaussian models by adding an extra class for MS lesions (Kikinis, Guttmann et al. 1999, Souplet, Lebrun et al. 2008) while others defined lesions as outliers of the mixture model (Van Leemput, Maes et al. 2001). Other approaches define lesions as outliers when comparing the spatial and intensity information of the images to be segmented and the library of healthy subjects (Tomas-Fernandez and Warfield 2011). To correct for noise and image artefacts, graph-cut techniques combine spatial information from local neighborhoods (García-Lorenzo, Prima et al. 2008, Khayati, Vafadust et al. 2008). Unsupervised techniques suffer from image intensity non-uniformity when present. Lesions, which present different intensity characteristics with

respect to their stage (acute, chronic...), make it difficult to apply a single global model as it cannot capture this variability.

Image and lesion variability are common limitations of both supervised and unsupervised techniques. However, by using image pre-processing, along with the appropriate training set and image features, supervised techniques can potentially capture MS lesions and image intensity variability, as we will describe in Chapter 3.

All these tissue classification and segmentation techniques are usually performed in a cross-sectional manner and have shown good results. However, these techniques are usually limited in the context of longitudinal analyses because they do not leverage the constraints available when considered longitudinal data as an ensemble. More specifically in MS, lesion segmentation is often considered as an end-point where lesion count or the total lesion volume, are derived from this segmentation. . Furthermore, in the context of longitudinal morphological analysis, lesions represent potential confounding and can introduce pathological variability. Thus, accurate lesion segmentation will help automating longitudinal processing and improve the atrophy measure sensitivity.

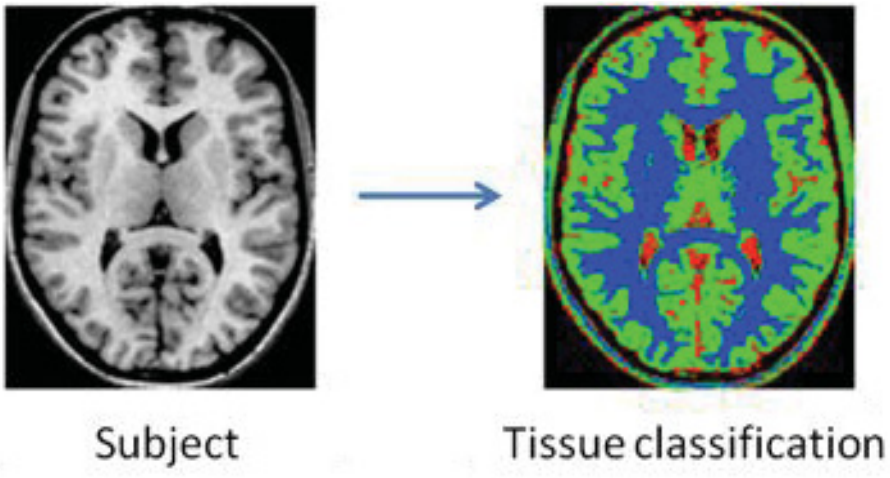


Figure 2.14: Tissue classification of a normal control subject. WM is in blue, the GM in is green and the CSF is represented in red. Note that in this example the skull, dura-matter and skin are also mis-segmented as either GM, WM or CSF.

2.4. MRI atrophy measurements as a surrogate of neurodegeneration

Most of the different tools mentioned above have been used to assess morphological changes in brain images due to disease progression. The recent progress in MR imaging techniques and image processing has helped to increase our knowledge about the pathophysiology of MS and AD. Brain atrophy has been shown to be a relevant feature for both pathologies and has been proposed as a surrogate marker for disease burden in MS (Bakshi, Dandamudi et al. 2005, Zivadinov and Leist 2005, Bermel and Bakshi 2006, Simon 2006) and in AD (Thompson, Hayashi et al. 2004, Schott, Price et al. 2005, Barnes, Godbolt et al. 2007, Frisoni, Fox et al. 2010). Brain atrophy measures are promising surrogates for the pathological progress of the disease because:

1) Atrophy measures represent the net effect of tissue destruction.

2) Atrophy quantification does not require a specialized MRI protocol.

3) Atrophy can be measured in an automatic and robust manner.

4) Atrophy can be detected over small periods of time depending on the disease.

5) Atrophy metrics should have a significantly smaller variance as compared to clinical disability scales.

Many studies have assessed longitudinal brain atrophy and there are different approaches possible for measuring these changes. The recent availability of large datasets of longitudinal MRI data has given rise to the need for longitudinal image processing techniques. Indeed, different meta-analyses of the literature have shown strong evidence of heterogeneity in the measure of brain atrophy rates. One of these meta-analysis from Barnes et al. (Barnes, Bartlett et al. 2009) investigated the hippocampal atrophy rate on 595 AD patients. In their random-effect

meta-analyses, they reported an overall mean atrophy rate of 4.66% per year (95% CI 3.92, 5.40) in AD and an estimated "between-study" standard deviation of 0.77%. They also reported significant differences between-study groups revealing the strong evidence of study heterogeneity (Figure 2.15). This variability found between studies could be explained by many factors such as (Jovicich, Czanner et al. 2009, Frisoni, Fox et al. 2010, Vrenken, Jenkinson et al. 2013):

- Pathological changes: treatment effect, severe atrophy or lesion.
- Intrinsic image acquisition factors: spatial resolution, incomplete head coverage, positioning, tissue contrast, noise and artifacts.
- Methodological limitations.

Clinical studies usually control for scanner, treatment and patient pathological factors. But the inherent methodological errors can be subtle and must be taken into considerations. In the following sections, we thus describe the different approaches and their limitations to assess brain atrophy globally and (Section 2.4.1) and focally (Section 2.4.2).

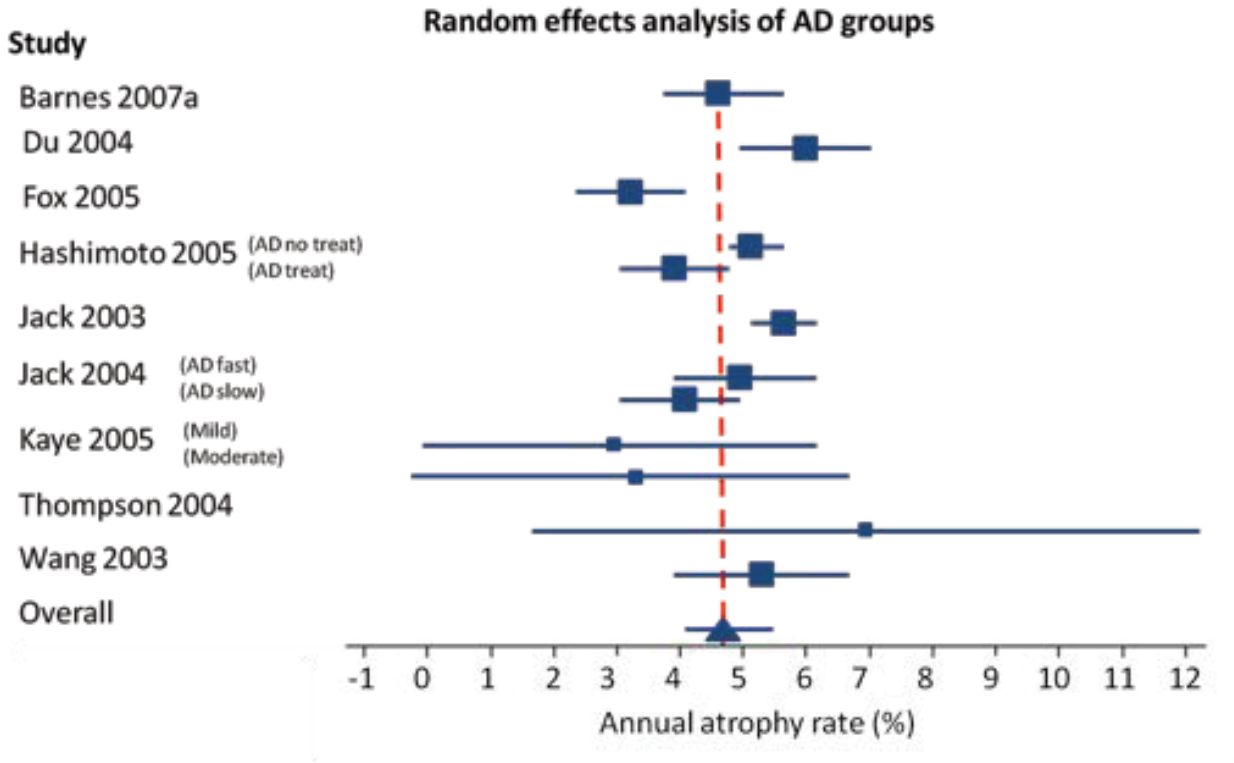


Figure 2.15: Meta-analysis of hippocampal atrophy rates of AD patients From nine studies (Jack, Slomkowski et al. 2003, Wang, Swank et al. 2003, Du, Schuff et al. 2004, Jack, Shiung et al. 2004, Thompson, Hayashi et al. 2004, Fox, Black et al. 2005, Hashimoto, Kazui et al. 2005, Kaye, Moore et al. 2005, Barnes, Godbolt et al. 2007). The square size is proportional to the inverse of (the study variance + the between-study variance estimate) while the solid lines represent the 95% confidence intervals. Adapted from (Barnes, Bartlett et al. 2009).

2.4.1. Global brain atrophy

Global atrophy is the overall atrophy of brain tissue, or of a particular tissue compartment (WM or GM). It is measured globally, and carries no specific information of location of lost tissue. We can classify global brain atrophy metrics into two main categories: cross-sectional and pair-wise atrophy measure.

Cross-sectional global atrophy measure

Until recently, most longitudinal analyses have been performed globally, using cross-sectional (i.e. from structural segmentation), or pair-wise measures. Cross-sectional global atrophy measurements can be achieved with any segmentation technique as mentioned in section 2.3.4, in order to estimate an absolute volume of the total brain parenchyma (Dalton, Chard et al. 2004), the corpus-callosum cross sectional area (Pelletier, Suchet et al. 2001), the cerebellum volume (Liu, Edwards et al. 1999), the ventricle volume (Lukas, Minneboo et al. 2010), or the volume of individual brain lobes (Benedict, Zivadinov et al. 2005).

Non-linear registration can be used to obtain a Jacobian determinant of the transformation at each voxel (Equation 2.5). This determinant can be integrated over a pre-segmented brain region to obtain an approximation to the volume different between a source and target image in that region. This technique has been applied to many brain structures (ie. whole brain (Freeborough and Fox 1998) or hippocampus (Crum, Scahill et al. 2001)) and we showed that it can improve the estimation of longitudinal cortical grey matter atrophy in MS (Nakamura, Guizard et al. 2014).

Another approach to measure global brain atrophy uses brain segmentation to compute a ratio of the brain tissue volume and the intra-cranial cavity to obtain a normalized head-size metric such as the brain parenchymal fraction (BPF) (Rudick, Fisher et al. 1999). A similar fully automated approach was developed at the MNI to measure the brain to intra-cranial capacity ratio (BICCR) (Collins, Montagnat et al. 2001). BPF has shown to be smaller in MS than in healthy normal controls (Rudick, Fisher et al. 1999), and to decrease faster than in healthy normal controls (Chard, Brex et al. 2003). Similar results have been found in AD (Ramirez, McNeely et al. 2014).

Pair-wise global atrophy measures

The first set of methods uses longitudinal data where volume changes are estimated at the edge of the brain between two successive images of registered data (registered using affine

transformations). Because of their accuracy and robustness, the boundary shift integral (BSI) (Freeborough and Fox 1997) and SIENA (Structural Image Evaluation, using normalization of atrophy) (Smith, Destefano et al. 2000) in the FMRIB Software Library (FSL) have been widely used for global brain atrophy measures in research and clinical trials:

BSI: Given the boundary of the segmented structure (ie. brain (Freeborough, Fox et al. 1997)), BSI estimates the amount of displacement of this boundary between two time-point images. A region around the boundaries is defined using morphological operators. In this region, the volume change is estimated by integrating the normalized intensity difference of both images.

SIENA: SIENA starts with brain and skull segmentation (Smith 2002) to perform skull-based registration in a half-way space. Then, the brain:non-brain boundary is estimated from tissue classification (Zhang, Brady et al. 2001) before computing the perpendicular displacement between the brain boundaries of the two time-points. Finally, the surface displacement is averaged to obtain a global estimate of percentage of brain volume change (PBVC).

Global atrophy measures require a priori knowledge of the anatomical structures being assessed and might exclude potential focal or regional relevant regions from the analysis.

2.4.2. Focal brain atrophy

With the advent of more sophisticated techniques, it is possible to investigate focal longitudinal changes, which will invariably provide more insight into specific atrophy patterns. In addition to manual, automated or semi-automated outlining methods, studies have explored the whole brain in a voxel-based manner to measure focal atrophy at the individual level but also across whole populations. Here, we describe possible approaches to measure voxel-wise tissue changes or atrophy using: voxel-base morphometry (VBM), deformation- and tensor-based morphometry (DBM).

***VBM*:** Initially proposed by Ashburner et al. (Ashburner and Friston 2000), VBM is believed to reflect the tissue changes at a voxel level. These changes are estimated by first aligning the images into stereotaxic space (cross-sectional or longitudinal), followed by GM segmentation and voxel-wise statistical comparison of the smoothed GM images. In *optimized VBM,* the Jacobian determinant of the non-linear registration to a template space can be used to weight the GM density (Good 2001). The statistical comparison employs the general linear model (GLM), which allows a variety of different statistical tests at the voxel-wise level such as group comparisons and correlations with covariates of interest. VBM has been widely used to identify focal brain atrophy in normal aging brain (Kalpouzos, Chételat et al. 2009) and neurodegenerative diseases such as Alzheimer's disease (Baron, Chételat et al. 2001). The VBM approach enabled the identification of atrophy in GM structures such as the cortex and thalamus (Audoin, Davies et al. 2006, Ceccarelli, Rocca et al. 2008) in MS.

DBM: This morphometric method has been developed to estimate macro-anatomical or shape differences between two MRI images using non-linear registration. In this thesis, we use the term DBM to describe methods of studying the deformation fields or scalar field derived from the deformation fields (ie. Jacobian determinant). The term tensor-based morphometry (TBM) is sometimes used in the literature to describe DBM analyses based on Jacobian determinants. In DBM, vector-values or derived scalar properties of the deformation field are used to compare and correlate morphological changes directly instead of using intensity as an indirect handle on shape change such as within the techniques described for VBM. Tao et al. (Tao, Datta et al. 2009) used DBM in MS to evaluate the deep GM structure atrophy and they demonstrated a correlation between the thalamus volume and clinical scores. Hua et al. (Hua, Leow et al. 2008) performed statistical analysis of the Jacobian determinant maps between normal AD, MCI and NC to reveal correlations with clinical measurement and genes associated with AD. However, one limitation of DBM analysis is that multiple solutions exist for non-linear registration. This is due to the fact that there are different strategies used to estimate the deformation (Section 2.3.3), which need to appropriately handle the potentially opposing requirements for deformation and regularization for the global and local fitting.

One main limitation of morphometric analyses is that they attempt to find a one-to-one anatomical correspondence between subjects, a correspondence that might not exist. These is particularly true in cross-sectional research were VBM and DBM are mainly used. However, DBM advantages are its ability to detect subtle anatomical differences that could potentially provide precise longitudinal morphological measures where the one-to-one correspondence exists over time. Furthermore, DBM does not require tissue segmentation or a-priori information about specific anatomical structural changes.

The variability mentioned above in medical image acquisitions and methods has limited the wide use of longitudinal morphological studies. These potentially confounding factors need to be controlled to increase the sensitivity of atrophy measures. Furthermore, due to the vast variety of registration algorithms, no widely accepted standard for DBM studies exists, which has prevented its incorporation into major neuroimaging software packages. To overcome these limitations, non-linear registration and their resulting measures for longitudinal analysis should satisfy the requirements mentioned by Fox et al. (2011):

1. Symmetry: The order of the non-linear registration should produce the same absolute change regardless of the registration direction ($I \rightarrow R$ or $R \rightarrow I$).

2. Transitivity: In the case of multiple images (here A, B and C), the sum of the changes from $A \rightarrow B$ and $B \rightarrow C$ should be equal to the direct measure $A \rightarrow C$.

3. State of the art comparison: The measure should be compared on publically available data and compared to other approaches.

4. "Scan-rescan" reproducibility: Brain changes should tend to zero when estimated on scans acquired over a short period of time. Therefore, subjects scanned the same day should have zero brain changes and any deviations will provide insight into potential biases and robustness of the method.

5. Power analysis: Statistical power analysis such as sample size and effect size measures should be performed and normal ageing should be taken into account to assess the performance of the method to differentiate the groups or decrease the longitudinal variability.

Efforts have been made to improve longitudinal image analysis, but longitudinal non-linear registration still needs improvement in order to be used in research and clinical settings. This improvement will at term increase the sensibility of longitudinal measurements and increase the sensitivity to detect morphological changes. This will lead to a better power to evaluate disease progression and/or treatment effect, which could reduce the number of participants required in clinical trials and thus reduce the cost of potential new treatments.

2.5. Objectives

To summarize, neurodegenerative diseases such as AD and MS affect a large proportion of our population and their occurrence is expected to increase, especially in the case of AD. Given our current knowledge of these diseases, there is a need to understand and track the pathological processes in order to ultimately reliably assess potential treatments. MRI offers the possibility of investigating brain changes in a non-invasive manner. However, efficient, reliable and stable longitudinal measures from MRI are challenging to obtain and numerous factors can affect these measures, ranging from pathological to acquisition-related variability.

Therefore, the objectives of this thesis are to develop the necessary image processing approaches to attempt to circumvent the potential longitudinal variability in MRI studies while being effectively applicable in clinical settings.

In the following (Chapter 3), we begin to address the issue of pathological confounds by first proposing a method to detect MS lesions in a sensitive manner using an example-based approach, which exploits the use of large available libraries of manual lesion segmentations. In order to limit the impact of lesions on longitudinal measurements, we propose, in Chapter 4, a

method to synthesise normal tissue to be used as a replacement in regions of affected tissue. In terms of longitudinal analysis, Chapter 5 assesses the bias and the symmetry of non-linear registration approaches. Finally, in Chapter 6, we propose a spatio-temporal non-linear registration framework to reduce the longitudinal variability primarily caused by acquisition differences.

CHAPTER 3 AUTOMATED FOCAL PATHOLOGY DETECTION

Preface

This chapter presents a method for automatic segmentation of lesions from mulit-contrast MRI data. As discussed in the previous chapter, detection and quantification of brain morphological change has resulted in heterogeneous findings, perhaps due to the potential pathological confounds such as lesions. MS lesions are a clear example of sporadic and scattered visible pathological changes occurring in MR images in patients with MS. Despite poor correlation between MR Lesion volume and clinical disability, manual MS WM lesions have been used to follow MS disease progression and it is still perceived as the gold standard. Therefore, providing a reliable and fully automatic longitudinal measure of brain atrophy is important to detect the presence of potential pathological confounding but also as a pathological marker on its own.

The following manuscript proposed a new approach to detect MS lesion and was published in *NeuroImage: Clinical (Guizard, Coupé et al. 2015)*. The method exploits, in an optimal fashion, a large library of MRI images with corresponding pre-segmented lesion labels using label fusion segmentation principles. The method has shown highly competitive results on both public and local private datasets. The flexibility of this supervised approach makes it potentially applicable to other focal pathology (ie. AD T2W hypointensity lesions or FLAIR hyperintensities).

Rotation-invariant multi-contrast non-local means for MS lesion segmentation

Nicolas Guizard, Pierrick Coupé, Vladimir S. Fonov, Jose V. Manjón, Douglas L. Arnold, D. Louis Collins

3.1. Abstract

Multiple sclerosis (MS) lesion segmentation is crucial for evaluating disease burden, determining disease progression and measuring the impact of new clinical treatments. MS lesions can vary in size, location and intensity, making automatic segmentation challenging. In this paper, we propose a new supervised method to segment MS lesions from 3D magnetic resonance (MR) images using non-local means (NLM). The method uses a multi-channel and rotation-invariant distance measure to account for the diversity of MS lesions. The proposed segmentation method, rotation-invariant multi-contrast non-local means segmentation (RMNMS), captures the MS lesion spatial distribution and can accurately and robustly identify lesions regardless of their orientation, shape or size.

An internal validation on a large clinical magnetic resonance imaging (MRI) dataset of MS patients demonstrated a good similarity measure result (Dice similarity=60.1% and sensitivity=75.4%), a strong correlation between expert and automatic lesion load volumes (R^2=0.91), and a strong ability to detect lesions of different sizes and in varying spatial locations (lesion detection rate=79.8%). On the independent MS Grand Challenge (MSGC) dataset validation, our method provided competitive results with state-of-the-art supervised and unsupervised methods. Qualitative visual and quantitative voxel- and lesion-wise evaluations demonstrated the accuracy of RMNMS method.

3.2. Introduction

Multiple Sclerosis (MS) is a chronic, inflammatory demyelinating disease, which mainly affects the white matter of the central nervous system (CNS) but may also affect the cortex. The disease presents itself with a wide range of clinical manifestations, usually beginning with a relapsing remitting (RRMS) phase. RRMS is characterized by attacks of worsening neurologic function (relapses) that are followed by partial of full recovery (remissions). Relapses are directly related to an underlying inflammation of the CNS, which affects the myelin of the axons and consequently leads to focal "MS lesions". Because magnetic resonance imaging (MRI) is sensitive to inflammatory and demyelinating changes, it is often used to monitor, identify and quantify MS lesions (Fazekas, Barkhof et al. 1999) that are hyperintense on T2-weighted (T2W) magnetic resonance (MR) images and may become hypointense on T1-weighted (T1W) images. Lesion counts are often used to assess the disease burden and track disease progression as new lesions are related to current disease activity. Both counts are used to assess the efficacy of new therapies (Polman, Reingold et al. 2011). For the purpose of this article, we focus on lesions commonly called "T2-lesions" (those that are hyperintense on T2W images) and do not consider other sub-types of lesions (i.e. gadolinium enhancing "active lesions", "black holes" and cortical lesions). MS lesions in MR images are extremely difficult to identify because of inter-subject anatomical variability, lesion location, size, texture and shape. Manual segmentation of MS lesions is still recognized as the gold standard in MS, but it is time consuming and subjects to intra- and inter-expert variability. As an alternative, a multitude of automatic techniques to detect and segment MS lesion have been proposed. However, recent reviews of the literature (Lladó, Oliver et al. 2012, García-Lorenzo, Francis et al. 2013) concluded that automatic MS lesion segmentation is still an unsolved topic. Although promising progress has been made in this field open problems and limitations persist. For example, many techniques are not robust across imaging centers or differing MRI protocols.

Two main categories of classifiers emerge from the literature: unsupervised and supervised. Unsupervised methods do not require manual segmentation of the lesions, but estimate tissue

classes or clusters of similar voxels with or without the help of anatomical and MRI knowledge. Many unsupervised techniques were initially developed to classify healthy brain tissues based on MRI intensities into three classes (cerebral spinal fluid (CSF), white matter (WM) and grey matter (GM)). This was done by using fuzzy C-mean and Gaussian mixture models with expectation maximization (EM) (Wells, Grimson et al. 1996). To detect MS lesions, some groups adapted the Gaussian models by adding an extra class for MS lesions (Kikinis, Guttmann et al. 1999, Souplet, Lebrun et al. 2008) and/or added topological constrains (i.e. the publically available approach LesionTOADS (Shiee, Bazin et al. 2010)). Others defined lesions as outliers of the mixture model (Van Leemput, Maes et al. 2001, Schmidt, Gaser et al. 2012, Cabezas, Oliver et al. 2014) or as outliers when comparing the spatial and intensity information of the images to be segmented and library of healthy subjects (Tomas-Fernandez and Warfield 2011, Tomas-Fernandez and Warfield 2015). To correct for noise and image artifacts, graph-cut techniques have been used to combine spatial information from the local neighborhoods with the intensity model (García-Lorenzo, Prima et al. 2008, Khayati, Vafadust et al. 2008). Unsupervised techniques suffer from both intensity non-uniformity, in the whole image and in the lesion since this variability in intensities cannot be captured with a single global model. Furthermore, the properties of each image need to be specifically defined which can be difficult when artifacts have properties similar to lesions.

The supervised methods use machine-learning techniques to extract relevant features (e.g. intensity or local gradient) and train automatic classifiers from manual or automatic lesion segmentation datasets. Then, the features of new images to be segmented are compared with the training sets to estimate the lesions. These methods can use different machine-learning approaches including: artificial neural networks (ANN) (Zijdenbos, Dawant et al. 1994), k-nearest neighbors (K-NN) (Vinitski, Gonzalez et al. 1999), decision trees (Kamber, Shinghal et al. 1995), random decision forests (RDF) (Geremia, Clatz et al. 2011), Bayesian frameworks (Harmouche, Collins et al. 2006) and logistic regression (Sweeney, Shinohara et al. 2013). The common limitation of both supervised and unsupervised techniques is their sensitivity to image and lesion variability. However, using the appropriate training set and image features, supervised

techniques can potentially identify the MS lesion and compensate for variability in image intensities.

Indeed, it has been shown that many supervised library-based (or multi-atlas) techniques outperform unsupervised model-based segmentation methods (Lao, Shen et al. 2008). For example, patch-based methods using non-local means (NLM) for structural segmentation have gained in popularity and shown promising results despite their simplicity (Coupé, Manjón et al. 2011). Patch-based approaches have been applied to segment a multitude of anatomical structures including the hippocampus (Coupé, Manjón et al. 2011), brain (Eskildsen, Coupé et al. 2012), lateral ventricles (Fonov, Coupé et al. 2012), deep nuclei (Xiao, Fonov et al. 2014), intracranial cavity (Manjón, Eskildsen et al. 2014), brain tissues (Cordier, Menze et al. 2013) and other structures of the brain (Rousseau, Habas et al. 2011). Although NLM has proven to be useful in segmenting anatomically well-defined structures (e.g. hippocampus and lateral ventricles) they have not yet been applied intensively to MS lesion segmentation.

Given the success of patch-based approaches, we present a library-based NLM approach where voxels with similar surrounding neighborhoods (or patches) are used to estimate the presence of lesions. Contrary to the original patch-based segmentation method (Coupé, Manjón et al. 2011), we offer two main contributions in order to efficiently address the problem of MS lesion segmentation: i) a rotationally-invariant similarity metric for patch comparison which better captures lesion shape variability and ii) a multi-contrast framework that takes advantage of information derived from T2w and FLAIR images. Indeed, in the context of MS lesion segmentation the dimension, shape, orientation and position of lesions vary greatly (Figure 3.3). The sum of squared differences (or L2-norm), which originally used as patch-based distance measure (Buades, Coll et al. 2005), is sensitive to the orientation of the patch. While this is good for structure segmentation, this could potentially miss lesion labels (Figure 3.1). Similar to the work by Manjón et al. (2012), we replace the L2-norm distance measure with a rotation-invariant distance (RI) where only the intensity of a central voxel and the mean intensity of the patches are considered. Furthermore, the existing NLM segmentation algorithms used a single contrast library (e.g. only T1W images), however, a single image contrast does not hold enough

information to separate lesions from healthy tissues. Most lesion segmentation algorithms use T2W and FLAIR MR images, as most MS lesions appear hyperintense on these modalities indicating inflammation or scar tissue. On T1W images, lesions appear hypointense but present a larger intensity variability which might reflect the different sub-types of lesion such as the so-called "black hole" associated with irreversible tissue damage (van Walderveen, Kamphorst et al. 1998). Inspired by the work of Coupé et al. (2013), which introduced multi-contrast NLM for image super-resolution and Xiao et al. (2014) for dual-channel NLM segmentation of deep brain structures, we propose an adaptation of the NLM segmentation algorithm to take advantage of multi-contrast images for MS lesion segmentation. This library-based approach captures the potential global and local variability of the anatomy as well as the intensity variability in lesions.

To our knowledge, despite the increasing popularity of patch-based techniques, only a few recent methods have been developed to segment MS lesions. Weiss et al. (2013) presented a supervised segmentation technique using sparse coding with patches from a library of healthy subjects to reconstruct MS patient images. The reconstruction estimates an error map, which detects outliers believed to describe MS lesions. Their method shows promising preliminary results but was not assessed on a large clinical cohort and might not be specific enough to distinguish MS lesions from artifacts when detecting image outliers. The approach by Roy et al. (2014) also uses sparse techniques. In their method, they estimate a weighted average of the most similar patches of a kd-tree using a nearest neighbor search on a library of pre-segmented multi-contrast (T1W and FLAIR) images. While sparse strategies present the advantage of decreasing the dimensionality of the library, kd-tree removes the 3D spatial knowledge of the training images which may hold additional pertinent information that can help identifying MS lesions. Indeed, despite careful pre-processing, the local MS lesion properties in MR images (i.e. intensity, contrast, noise, etc.) depend on the local anatomical and/or spatial location of the lesion (Meier and Guttmann 2003). Thus, by using a 3D volume library, this local information can help capture the spatial layout of different MS lesions. Finally, neither Weiss et al. (2013) or Roy et al.'s (2014) patch-based approaches for lesion segmentation use rotationally invariant features.

As will be demonstrated in Section 3.3, this aspect is crucial in the context of MS lesion detection.

We assessed our rotation-invariant multi-contrast non-local means segmentation (RMNMS) approach on 108 RRMS patients from a multi-site clinical study. Our method obtained a Dice similarity measure of 60.1±16.4%, a sensitivity 75.4±15.7% and a precision 55.0±20.1% in cross-validation. Using the parameters established for our initial evaluation, we compared RMNMS to several different state-of-the-art techniques using the datasets from the MS lesion Grand Challenge (MSGC, MICCAI 2008 (Styner, Lee et al. 2008)), on which we obtained very competitive results holding the first rank at the time of the submission.

3.3. Methods

In the following section we first describe the developed algorithm (3.3.1), then the datasets (3.2.2), and lastly our evaluation techniques (3.3.3).

3.3.1. The algorithm

Our algorithm adapts the NLM estimator (a.) to account for multi-modal images (b.) and rotation-invariant distance measure of the patches (c.).

a. The non-local mean approach

NLM estimator

The NLM estimator, which takes advantage of image redundancy, was initially proposed by Buades et al. (2005) for image denoising. The idea of the NLM is to reduce the noise of the image by averaging the voxels that would have a similar intensity in the noise-free image. To achieve the denoising of voxel *x(i)*, the patch *P*(*x(i)*) centered on *i* is compared with all the patches *P*(*x(j)*) centered on *j* of the images in the neighbourhood Ω such that:

$$\hat{x}(i) = \frac{\sum_{j \in \Omega} w(i,j)\,x(j)}{\sum_{j \in \Omega} w(i,j)} \quad where \quad w(i,j) = e^{-\frac{\|P(x(i)) - P(x(j))\|_2^2}{h^2}} \qquad \text{Equation 3.1}$$

Where the term *w*(*i,j*) is a weight based on the similarity of between the patches *P*(*x(i)*) and *P*(*x(j)*), and is designed to attribute a smaller weight to the greater L2-norm ($\|.\|_2$) distance measures. The term h^2 is a smoothing parameter proportional to the noise variance.

NLM segmentation

The NLM approach has been used for structural segmentation (Coupé, Manjón et al. 2011) by employing a library of atlases with co-registered anatomical images and manually segmented structures to segment those particular structures on new subjects. For NLM segmentation, the weights are estimated between intensities of the subject patch, *P*(*x(i)*), and patches from a subject *S* from library of *N* pre-segmented subjects, *P*($x(j_s)$), such that:

$$\hat{x}(i) = \frac{\sum_{s=N}\sum_{j\in\Omega} w(i,j_s)\,l(j_s)}{\sum_{s=N}\sum_{j\in\Omega} w(i,j_s)} \qquad w(i,j_s) = e^{-\frac{\|P(x(i))-P(x(j_s))}{h^2}}$$ Equation 3.2

where in this case the h^2 parameter is set based on the patch minimum distance of the search area (Coupé, Manjón et al. 2011). Thus, if similar patches are found in the library the minimum distance h^2 will be low and the weight function will decay quickly such that it is not influenced by other patches.

b. Multi-contrast NLM segmentation (MNLM)

Multi-contrast NLM estimator

In MS, multi-contrast images for manual and automatic segmentation have shown to improve the identification of MS lesions. Inspired by previous work on multi-contrast NLM (MNLM) for denoising (Manjón, Thacker et al. 2009) and for super-resolution (Coupé, Manjón et al. 2013), we apply the NLM weighting function to allow for various contrasts (M) such that:

$$w(i,j_s) = e^{-\left(\sum_M \frac{\|P_M(x(i))-P_M(x(j_s))\|_2^2}{{h_M}^2}\right)}$$ Equation 3.3

Here M represents the different MR contrasts commonly used in MS lesion segmentation: T1W, T2W, PD, or FLAIR for example. It is important to note that the smoothing parameter h_M^2 is estimated for each considered contrast (i.e., the per contrast minimum distance). Similarly, the L2-norm distance is estimated between patches of the same contrast.

Multi-contrast training subject pre-selection

Subjects with similar lesion load and spatial distribution may be more similar with respect to their brain intensity characteristics. Therefore, focusing the weight estimation on the most similar subjects should potentially hold more similar patches, and also presents the advantage of reducing computation. In the context of label fusion segmentation methods, Aljabar et al. (2009) proposed a pre-selection for single contrast images of the most similar structures present in the training library. In our multi-contrast method, we seek the most similar training subjects by measuring the multi-contrast L2-norm (ML2-norm) distance of the subject being segmented and the training subjects across their brain mask region, defined as:

$$\text{ML2-norm} = \sum_{M} \left\| I_M(x(i)) - L_M(x(j_s)) \right\|_2 \qquad \text{Equation 3.4}$$

The N subjects with smallest ML2 distances are selected as they represent the most similar training subjects and thus provide the most similar set of features.

c. Rotation-invariant multi-contrast non-local mean segmentation (RMNMS)

Previous NLM segmentation implementations have shown convincing results in segmenting anatomically well defined structures (e.g., hippocampus and lateral ventricles (Coupé, Manjón et al. 2011)). Anatomically, these structures present a relatively small variability of shape, contrast and spatial location making the orientation of the structure to be segmented an important constraint when looking for similar patches in the library. However, this strong advantage for structural segmentation could be a drawback in the context of MS lesion segmentation where no structural, orientation nor spatial location can be assumed. Indeed, Kincses et al. (2011) show that MS lesions can be found almost anywhere in the brain. However, there is spatial predilection for lesions to occupy the peri-ventricular area, the cortico-spinal tract and the optic radiations the lesion distribution probability map as can be shown in Figure 3.3. The lesions themselves do not appear to have a constraint on their size, shape or number (as shown in Figure 3.3 and Figure 3.10 on three MS cases).

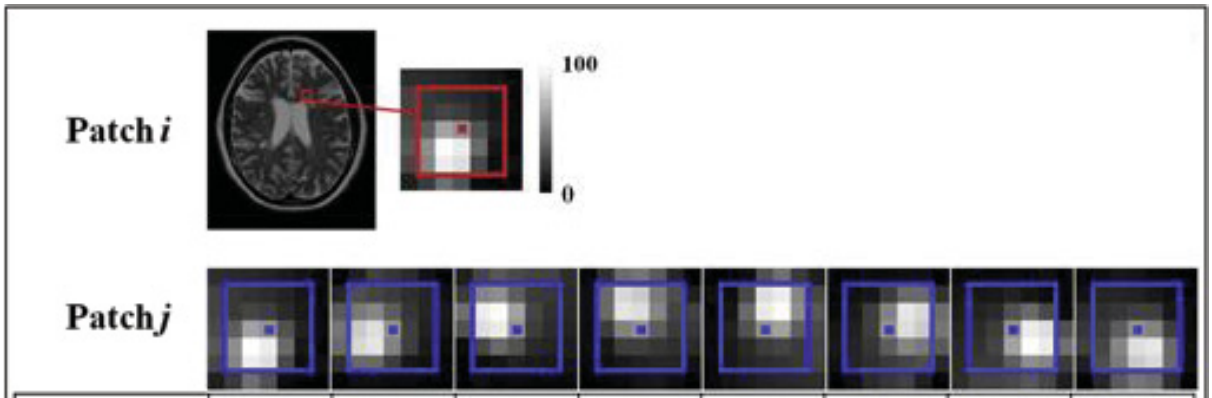
In order to increase the ability of the NLM segmentation approach to detect MS lesions, we propose a rotation-invariant distance (RI) metric instead of the L2-norm metric that is used in the multi-contrast NLM framework. Similar to the work of Manjón et al. (2012) on sparseness and self-similarity for MRI denoising, we replaced the L2-norm with a RI distance measure. Therefore, only the intensity of a central voxel (x) and the mean of surrounding patch (μ) are considered:

$$w(i, j_S) = e^{-\left(\sum_M \frac{(x_M(i) - x_M(j_S))^2 + \alpha(\mu_M(i) - \mu_M(j_S))^2}{h_M{}^2} \right)} \qquad \text{Equation 3.5}$$

In our experiments, we found that the intensity difference of the central voxels $(x_M(i) - x_M(j_S))^2$, was roughly the same as the intensity difference of the patch average $(\mu_M(i) - \mu_M(j_S))^2$, thus we chose α=1, whereas Manjón et al. (2012) used α =3. The need for a difference in alpha might be due to our pre-processing and in particular the denoising step, which tends to smooth the neighbouring intensity values surrounding the central voxel. The image denoising step used in our pre-processing (Coupé, Yger et al. 2008) is indeed a crucial step as it removes the variability of the central voxel with respect to its neighborhood and therefore allowing a better identification of similar RI features.

In order to be fully invariant to rotation, patches should be spherical, however, we found that cubic patches significantly reduce computational burden while preserving the distance accuracy. A graphical example is provided in Figure 3.1, where the distance of a cubic patch containing a lesion and the identical patch subject to different rotations is measured. Indeed, RI provides identical distance measures for different rotations of the same patch, only varying due to sampling and/or interpolation error, while L2-norm varies greatly and favors a larger distance between patches.

Another advantage of the RI distance measure is a reduced computational cost owing to considering only the central voxel and the mean of the patch, rather than all voxels in the patch. To further reduce computational cost we used multithread processing.



Rotation	0	π/4	π/2	3π/4	π	-3π/4	-π/2	-π/4
L2-norm	0.00	23.86	41.09	44.05	48.30	43.97	41.09	24.81
RI	0.00	0.69	0.00	0.69	0.00	0.69	0.00	0.69

Figure 3.1: Comparison of L2-norm and RI distance. Different rotations are applied to the extracted patch *i* (red) to obtain the patches *j* (blue). The L2-norm and the RI distance metric are then computed between these two patches.

3.3.2. Datasets

As mentioned by García-Lorenzo et al. (2013), simulated data sets, e.g. BrainWeb, enable a good proof of concept validation for image processing methods by providing ground truth images and lesion masks. However, BrainWeb images present multiple limitations: synthetic images are much easier to segment, only one phantom anatomy exists, and BrainWeb lacks contrasts such as FLAIR. Therefore, in this article we focus on two clinical datasets, an RRMS multi-center clinical dataset and the MICCAI 2008 MS Grand Challenge dataset (Styner, Lee et al. 2008).

Clinical MS dataset

One clinical multi-center dataset of 108 RRMS patients [age=42.6±10.7, 72 females] was used to assess the proposed segmentation algorithm. The dataset contains T1W [TE=9-11 ms, TR=30-40 ms, flip angle=30°, in-plane resolution=0.977x0.977 mm^2, slice thickness=3 mm], T2W [TE=66-100 ms, TR=3550-6610 ms, flip angle=90°, in-plane resolution=0.977x0.977 mm^2, slice thickness=3 mm], PD [TE=10-18 ms, TR=1867-3750 ms, flip angle=90°, in-plane resolution=0.977x0.977 mm^2, slice thickness=3 mm] and FLAIR [TE=59-94 ms, TR=7977-9630 ms, TI=1993-2500 ms, flip angle=90°, in-plane resolution=0.977x0.977 mm^2, slice thickness=3 mm] for all subjects. The MRI data were acquired at 32 sites on 1.5T scanners from different manufacturers: GE (n=19), Philips (n=3) and SIEMENS (n=10). We do not have access to demographic information for this dataset.

This dataset also contains gold standard MS lesion segmentation labels that were first automatically segmented by a multi-spectral Bayesian classifier (Francis 2004) with the T1W, T2W and PD images and manually assessed and corrected by expert raters who underwent extensive training on similar MS patient MRI data. In a previous study (Caramanos, Francis et al. 2012), seven raters with similar expertise, corrected the automatically generated lesion labels and were evaluated on a set of 10 MS patients with similar MRI protocols to those used in this study. Thanks to the initial automatic segmentation, this evaluation revealed an excellent inter-rater reliability relative to their trainer's reference segmentations (DSC=93.5±1.5%) and intra-class correlations (ICC=99.0±0.5%).

This RRMS cohort presents a large range of lesion loads (0.5-48.8 ml) and lesion counts (1-156 lesions) which are depicted in Figure 3.2. In this gold standard delineation protocol, only lesions with at least three connected voxels (or a lesion volume of 0.009 ml) in the 3D volume are kept and considered in our experiments. The MRI data and the expert lesion masks were used to form the template library of our proposed algorithm, which was tested in a leave-one-out fashion.

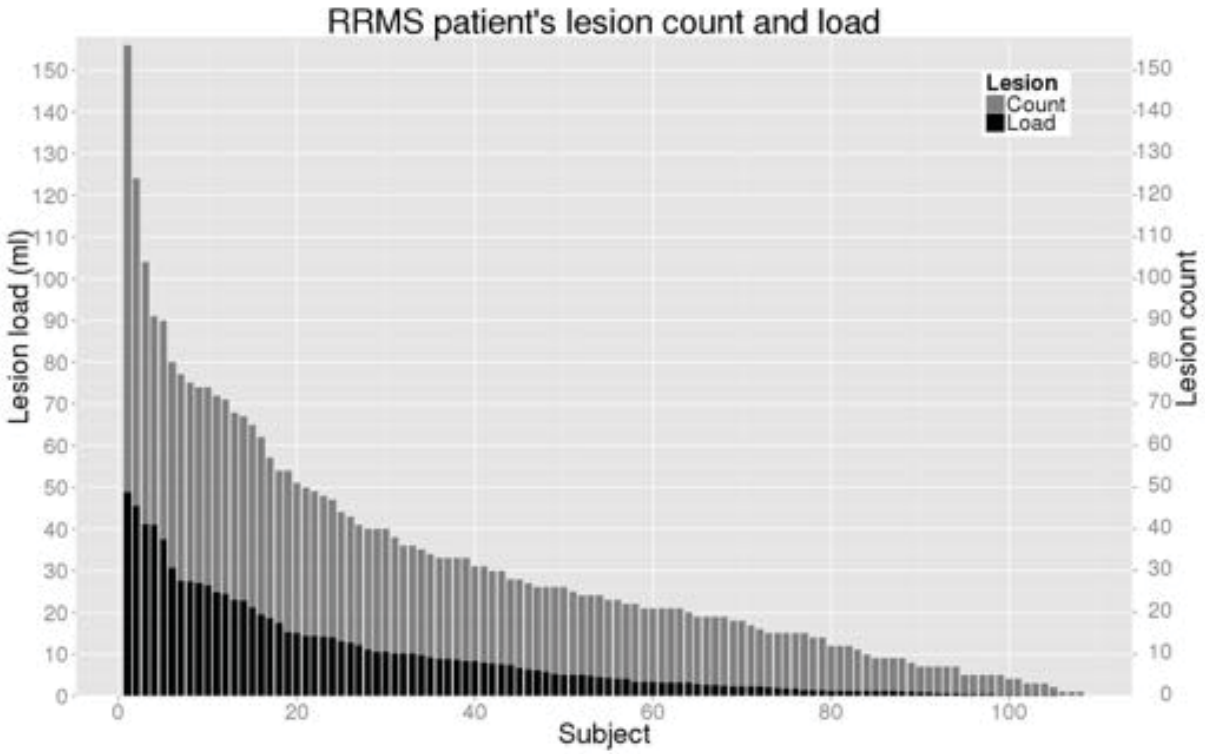


Figure 3.2: Lesion count and load for each RRMS subject of the clinical cohort. Only lesions with more than three connected voxels (or a lesion volume > 0.009 ml) are considered. The lesion count represents the number of non-connected lesions in grey. The lesion load represents the total volume of lesion (ml) in black. We can note that the lesion load volume is coarsely proportional to the number of lesions.

MS Grand Challenge dataset of MICCAI 2008 (MSGC)

Our proposed RMNMS algorithm was further validated using the clinical data provided by the MS lesion segmentation challenge introduced at MICCAI 2008 (Styner, Lee et al. 2008). From the MS challenge website[1], 20 training MR datasets with ground truth manual lesion segmentations and 23 testing cases were provided from the Boston Children's Hospital (CHB) and the University of North Carolina (UNC). While lesions masks for the 23 testing cases are not

[1] http://www.nitrc.org/projects/msseg/

available for download, an automated system is available to evaluate the output of a given segmentation algorithm.

We downloaded the co-registered T1W, T2W, FLAIR images for all 43 datasets as well as the ground truth lesion mask images for the 20 training datasets. All images were interpolated at 0.5 mm^3 isotropic resolution. We used the MSGC training set as a library to segment the MSGC T2W and FLAIR images.

Pre-processing and training library

All the images from both MS datasets (clinical RRMS and MSGC) were processed using the same pipeline, which does:

a) NLM image denoising (Coupé, Yger et al. 2008).
b) Intensity non-uniformity correction (N3) (Sled, Zijdenbos et al. 1998).
c) Linear intensity normalization of the image histogram to our in-house MS templates that were created with an unbiased template creation algorithm (Fonov, Evans et al. 2011) from the 108 T1W images of the RRMS patients (Figure 3.3).
d) Linear registration of each T1W image to our MS template which is in the MNI152 template space (Collins, Neelin et al. 1994).
e) Rigid registration of the T2W and FLAIR to the T1W image, followed by resampling onto a 1x1x1mm grid in the MNI space. Note that for the purpose of the validation describe here, we used the T1W as the reference image for registration, but other modalities (T2W, FLAIR…) could be chosen if a T1W image is not present or required.
f) Brain extraction (Eskildsen, Coupé et al. 2012).

Contrary to some recent patch segmentation approaches (Bai, Shi et al. 2013), we did not apply non-linear registration to segment and create the library. This is due to the fact that non-

linear registration and interpolation of MS lesions could alter the anatomical and intensity characteristics of MS lesions.

After pre-processing, all of the images and their respective manual segmentation lesion maps are spatially aligned and their intensity distributions are normalized. The denoising step of the pipeline is crucial for the RI distance measure as the central voxel value of a patch is given as much weight as its surrounding patch average. The MS library, used for the segmentation, was built using the output images from the pre-processing stages d, e and f. To double the size of the library and increase the spatial distribution of MS lesions, left-right-mirrored copies of each dataset were added to the library (Figure 3.3).

3.3.3. Evaluation metrics and experiments

In the following Section, we describe the general evaluation strategy (2.3.1) and the different metrics (2.3.2) used to assess the proposed segmentation method.

General evaluation strategy

On the clinical RRMS dataset, we performed a leave-one-out cross-validation of the proposed RMNMS method. This leave-one-out cross-validation is achieved by first removing the subject and its respective left-right-mirrored images from the training library, then the multi-modal pre-selection of the N closest subjects are selected and finally the segmentation is performed.

We first evaluated the performance of RMNMS with respect to the search area radius of the patches in the library and the number of pre-selected training subjects (as described in 3.3.3).

We also assessed RMNMS using i) different contrast combinations (T1W+FLAIR, T2W+T1W+FLAIR and T2W+FLAIR contrasts), ii) the original L2-norm distance measure

version of the NLM segmentation algorithm using T2W+FLAIR contrasts ("T2W+FLAIR NLM"), iii) without the left-right mirror addition to the training library, and iv) the LesionTOADS[2] approach proposed by Shiee et al. (2010). LesionTOADS is an iterative atlas based segmentation technique that uses a topological and a statistical atlas within the fuzzy C-means algorithm. As it was originally developed to segment healthy brain tissues (Bazin and Pham 2008), the algorithm was adapted to use multi-contrast (T1W+FLAIR) and an extra lesion class within the WM class. LesionTOADS was chosen as it is publically available and obtained one the best results during the 2008 MSGC (Styner, Lee et al. 2008). Note, that the algorithm was used with its default parameters.

Furthermore, we explored the effect of patients' total lesion-load and lesion-by-lesion detection measures on the RMNMS method.

Two experiments were done using the MSGC dataset. First, our algorithm was validated on the training set using leave-one-out cross-validation. Second, our segmentation results on the testing MSGC dataset were submitted online[3] and compared with other published techniques including i) LesionTOADS, ii) Souplet el al. (Souplet, Lebrun et al. 2008), winner of the MSGC at MICCAI 2008, iii) a recent supervised technique by Geremia et al. (2011) and iv) Tomas-Fernandez et al. (2011), who hold the current best score on the MSGC website before our submission.

For the online MSGC evaluation, we provided the lesion mask in native space after interpolation. The organizers normalized different metric results between 0 and 100, where 100 is a perfect score and 90 is the typical score of an independent rater as described by Styner et al. (2008). The different metrics (volume difference "VolD", surface distance "SurfD", true positive rate "TPR" and false positive rate "FPR" (Table 3.1)) were measured by comparing the automatic segmentation to the manual segmentation of two experts ("CHB" and "UNC").

---

[2] https://www.nitrc.org/projects/toads-cruise/
[3] http://www.ia.unc.edu/MSseg/

Evaluation metric

The quantitative evaluation of our method is carried out using different metrics, summarized in Table 3.1 as suggested by Styner et al. (2008), and García et al. (2013).

A high precision (PPV) and sensitivity (TPR) indicate that the automatically segmented lesions correspond well to the manually labeled lesion voxels. A low fall-out (FPR) indicates that the procedure does not identify voxels as lesion when they are not. We measure the absolute volume difference (VolD) of the manual versus the automatic segmentation (0% indicates a perfect lesion volume agreement) and the symmetric surface distance (SurfD) estimates the Euclidean distance between the surfaces of both segmentations at each voxel of their contours (0 mm indicates a perfect match of the surfaces). To estimate the SurfD values, we first estimate the distance transform from the binary segmentation using a 3D-Euclidean metric (Borgefors 1988) where the surface has a value of 0. Then, we look at the value of the binary segmentation and the corresponding transform distance value to estimate the distance to the surface. Usually, the true positive (*TP*), false positive (*FP*) and false negative (*FN*) rates are voxel-based; however, this measure can also be performed in a lesion-wise manner.

Indeed, in some studies, detecting small lesions is more important than properly identifying their borders. In these comparisons, we use LPPV and LTPR, which are lesion-wise version of the PPV, and TPR metrics where lesion wise *TP* (*LTP*), *FP* (*LFP*) and *FN* (*LFN*) are measured at each distinct lesion (Table 3.1). In this case, instead of applying the metrics on a voxel-by-voxel basis, we measure the ability of the method to detect the presence of a lesion. Following the expert manual segmentation protocol, only lesions with at least 3 voxels (or 0.009 ml in the native image space) and an overlap of at least 1 voxel (or 0.003 ml) were considered (Karimaghaloo, Shah et al. 2012).

Finally, we assess the behaviour of our method with regard to the patient's lesion load, size and location.

Table 3.1: Nomenclature of the evaluation metrics. The evaluation metrics are listed in the table below and estimated using the following abbreviations: true positive (*TP*), lesion-wise rue positive (*LTP*), false positive (*FP*), lesion-wise false positive (*LFP*), false negative (*FN*), lesion-wise false negative (*LFN*), automatic lesion volume (V_a), manual lesion volume ($V_{m,}$), d_a^{am} and d_a^{ma} are the Euclidean distances between the automatic (*a*) and the manual (*m*) lesion surface voxels, and n_a and n_m are the number of surface voxels for each segmentation.

Name	**Abbr.**	**Equation**	**Units**
Dice similarity measure	DSC	$\frac{2\times TP}{FP+FN+2\times TP}$	%
true positive rate or sensitivity	TPR	$\frac{TP}{TP+FN}$	%
	LTPR	$\frac{LTP}{LTP+LFN}$	%
positive predictive value or precision	PPV	$\frac{TP}{TP+FP}$	%
	LPPV	$\frac{LTP}{LTP+LFP}$	%
volume difference	VolD	$\frac{\lvert V_a - V_m \rvert}{V_m}$	%
false positive rate or fall-out	FPR	$\frac{FP}{FP+TP}$	%
symmetric surface distance	SurfD	$\frac{1}{(n_a+n_m)}\left(\sum_{i=1}^{n_a}\lvert d_a^{am}\rvert+\sum_{j=1}^{n_m}\lvert d_m^{ma}\rvert\right)$	mm

3.4. Results

The T1W, T2W, and FLAIR RRMS average templates created for the pre-processing stages of our method are presented in Figure 3.3 along with the spatial distribution of T2W lesions. While one can appreciate the anatomical definition of the different contrast templates, we can

still visually identify the hypo-intense intensity distribution around the lateral ventricles on the T2W and the corresponding hyper-intense intensity on the FLAIR. As expected, the T2W lesion spatial probability distribution is higher is the peri-ventricular region. However, the presence of lesions is diffuse, and our library of MS patients holds enough spatial variation to capture the spatial distribution of lesions.

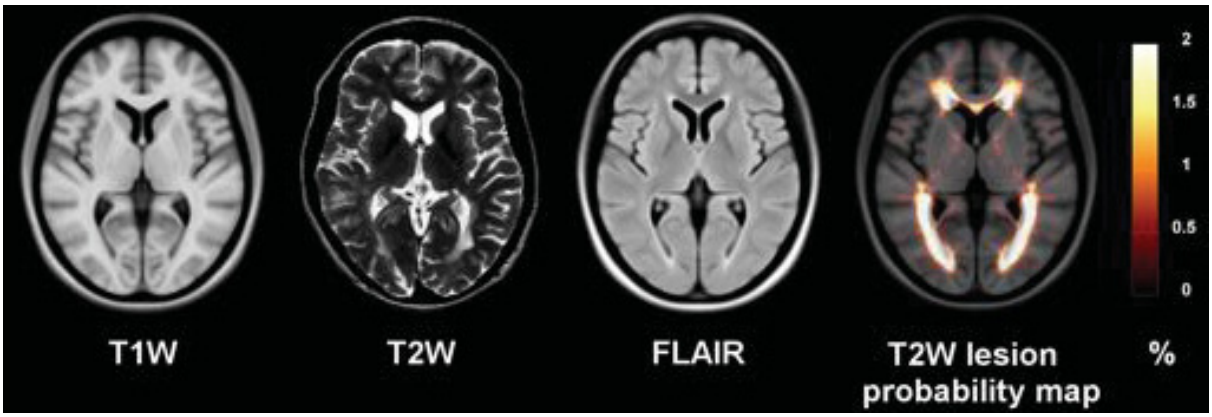


Figure 3.3: RRMS templates (T1W, T2W, FLAIR) and T2W lesion probability map overlay on the T1W RRMS template.

3.4.1. Evaluation on the clinical RRMS dataset

Impact of the search area radius

The results for different search area radii on the different metrics (DSC, VolD, TPR, PPV, LTPR and LPPV) are presented in Figure 3.4 for RMNMS using T2W+FLAIR. In our experiments, we found that using a patch size of 3x3x3 provides the best compromise between accuracy and computational burden. With a 3x3x3 patch (i.e., radius of 1 voxels) and a pre-

selection of 50 subjects, the DSC, TPR, PPV, LTPR and LPPV results plateau at their best results with a search area radius of 5 voxels (i.e., 11x11x11 search area). The volume difference (VolD) results are not as clear but the best results are achieved for any search area radius bigger than 2 voxels. Increasing the search area increases the chance of capturing a patch that is more similar to the considered patch, thus it is not surprising that better results are achieved with bigger search areas. However, increasing the search area needs to balance against computational cost where for instance, increasing the search area from 5 to 6 voxels increases the computational time by 15%. We found a search area radius of 5 voxels to be a good compromise (median results: DSC=60.1±16.4%, TPR=75.4±15.7%, PPV=55.0±20.1%, VolD=33.5±68.9%, LTPR=79.8±14.6% and LPPV=85.7±24.2%) and was chosen for the rest of the evaluation.

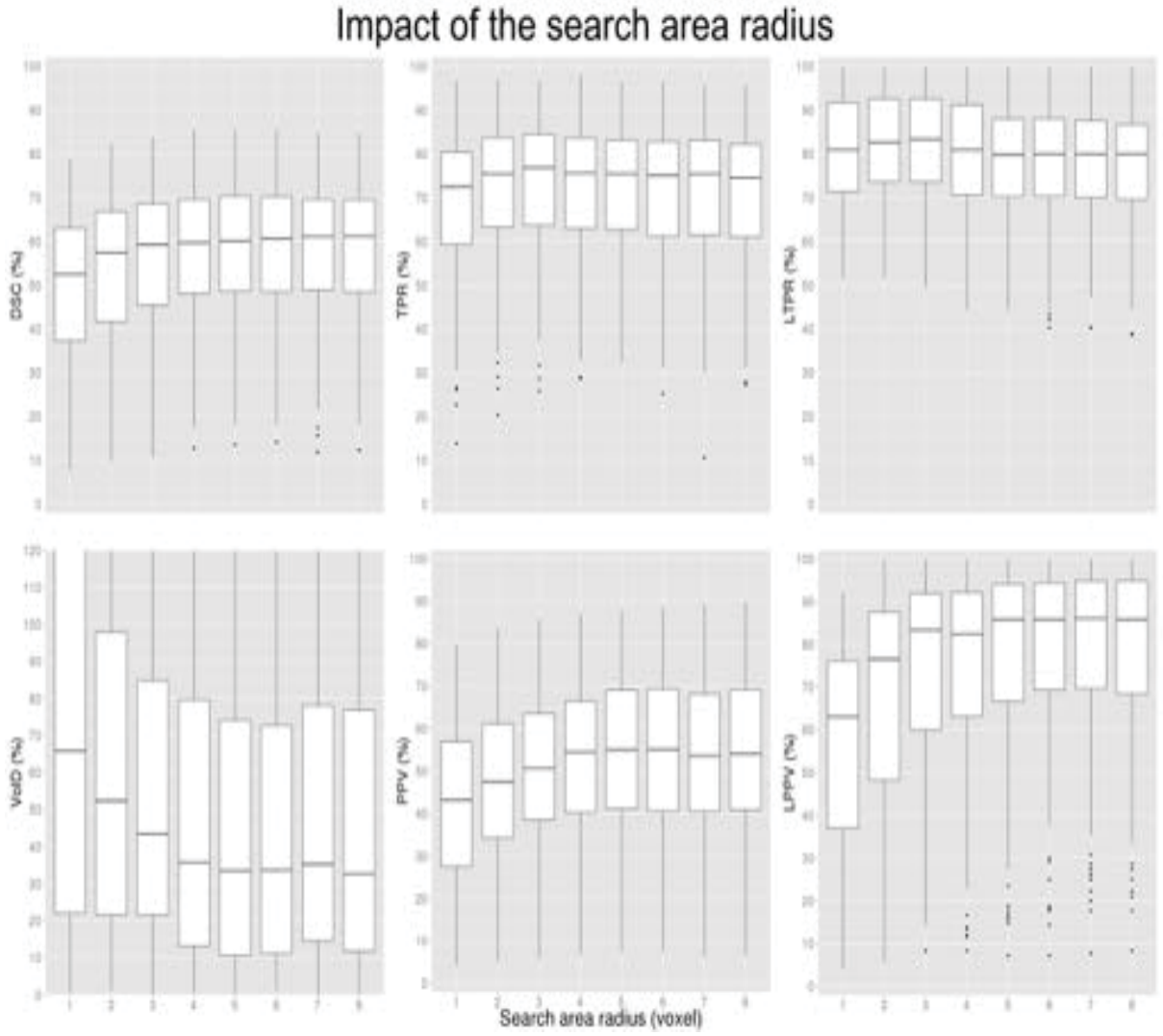


Figure 3.4: Impact of the search area radius (1-8 voxels) on DSC, VolD, TPR, PPV, LTPR and LPPV distributions. The boxes represent the lower and upper quartile with the median as the central black line. The whiskers extend to the most extreme data point. The experiment was conducted with a patch size radius of 1 voxel, and a pre-selection of 50 subjects.

Impact of the number of pre-selected training subjects

Pre-selecting more subjects from the template library can increase segmentation accuracy. In Figure 3.5, the results for the RMNMS method using T2W+FLAIR with different numbers of pre-selected training subjects on the different metrics (DSC, TPR, VolD, PPV, FPR and VolD) are presented. The experiment was performed with a patch radius of 2 (voxels), and a search area

radius of 5 voxels while varying the number of pre-selected training subjects from 10 to 80. As expected, increasing the number of subjects in the library improves the quality of the segmentation. Using 50 subjects provides a good comprise between segmentation results and computational cost (median results with 50 pre-selected subjects are the same as in section 3.4.1.a, as we used the same parameters) and was chosen for the rest of the evaluation.

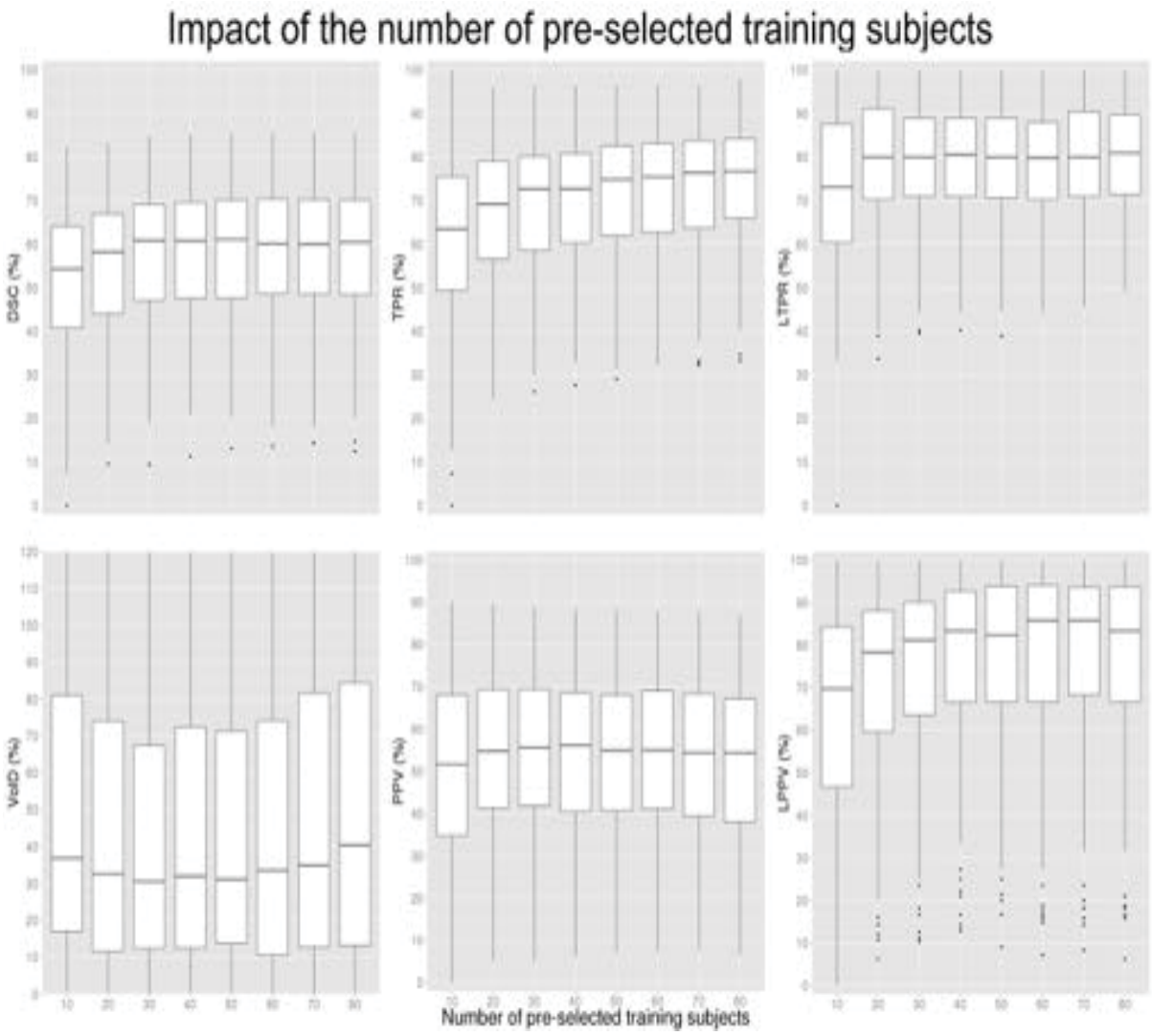


Figure 3.5: Impact of the number of pre-selected training subjects on DSC, VolD, TPR, PPV, LTPR and LPPV distributions. The experiment was conducted with a patch radius of 1 voxel, and a search area radius of 5 voxels on the 108 RRMS subjects.

Impact of the methods and modalities

Here, we compare RMNMS using T1W+FLAIR, T2W+T1W+FLAIR, T2W+FLAIR with and without the mirrored library images as well as the previous MNLM technique using T2W+FLAIR images, and LesionTOADS using T1W+FLAIR images. RMNMS with T2W+FLAIR was selected as the baseline for comparison and the similarity metric results are summarized in Figure 3.6.

The main result made evident by Figure 3.6 is that RMNMS T2W+FLAIR provides a higher LTPR (79.8±14.6%) than T2W+FLAIR MNLM (67.3±18.6%). Furthermore, T2W+FLAIR RMNMS consistently obtains the highest results (DSC, VolD, PPV, TPR, LTPR and LPPV) when compared to the different modalities used with RMNMS but also when compared with the unsupervised LesionTOADS approach.

Using T2W+FLAIR images provides overall better segmentation results than the other modality combination and the addition of the left-right mirrored images to the training set improves consistently the segmentation results of T2+FLAIR RMNMS.

The computational time for RMNMS using 16 threads on an Intel Core i7-950 processor at 3.06 GHz was around 40 min per subject. Our method with these settings is about three times faster (p-value <0.01) than similar MNLM patch-based methods with the same parameter settings and the computation time for the methods using the entire training set are provided in Table 3.2.

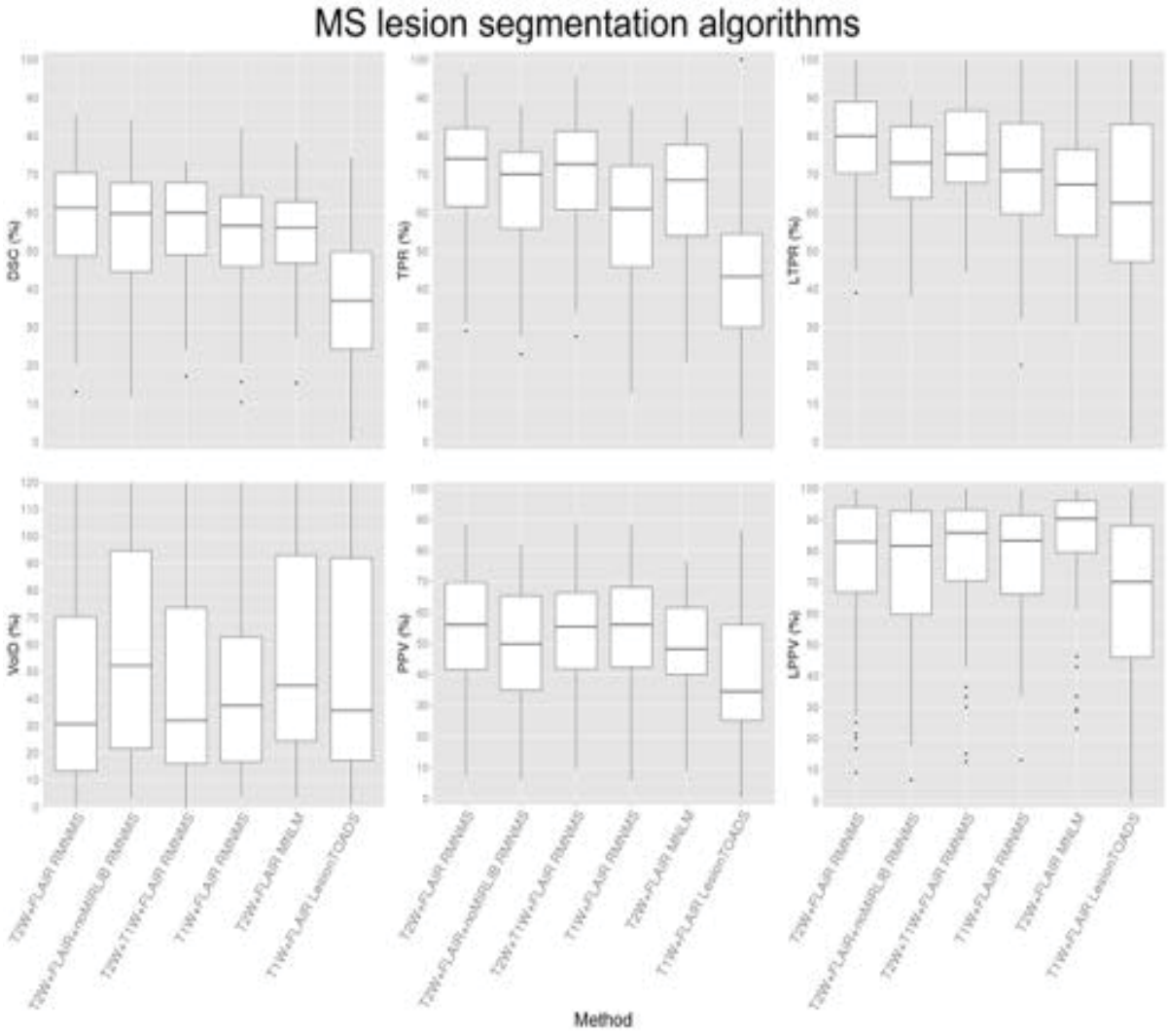


Figure 3.6: DSC, VolD, TPR, PPV, LTPR and LPPV distributions for different NLM MS lesion segmentation techniques (MNLM and RMNMS), different image modalities (T2W+FLAIR, T1W+FLAIR and T2W+T1W+FLAIR) as well as T2W+FLAIR RMNMS with (T2W+FLAIR RMNMS), without the left-right mirrored of each dataset (T2W+FLAIR+noMIRLIB RMNMS) and T1W+FLAIR LesionTOADS. The experiment was conducted on the 108 RRMS subjects, and for the NLM approaches a patch radius of 1 voxel, and a search area radius of 5 voxels were chosen.

Table 3.2: Computational time results on the RRMS clinical dataset. The proposed method RMNMS, with T2W+FLAIR images, is compared to the original NLM segmentation approach with multi-contrast (MNLM) and a T1W+FLAIR and T2W+T1W+FLAIR version of RMNMS (T1W+FLAIR and T2W+T1W+FLAIR RMNMS). The best measures are shown in bold and the significant difference when comparing with T2W+FLAIR RMNMS is shown in red. The experiment was conducted with a patch radius of 1 voxel, a search area radius of 5 voxels, a pre-selection of 50 subjects for all the methods.

Method	**Computation time (min)**		
	Mean	**Std**	***p*-value**
T2W+FLAIR MNLM	111.88	±11.76	<0.01
T1W+FLAIR RMNMS	42.15	±4.73	0.23
T2W+T1W+FLAIR RMNMS	72.15	±5.13	<0.01
T2W+FLAIR RMNMS	41.81	±4.52	-

Impact of lesion load and sizes

The segmentation results for patients with different lesion loads are shown in Figure 3.7. Subjects with larger lesion loads have better results with lower variability. However, we found that the mean TPR of the method is less affected by the lesion load than the other metrics (i.e., DSC and PPV). Note that DSC is sensitive to object size and smaller DSC is expected for smaller lesions. The linear regression of the manual lesion volume and RMNMS lesion volume shows good correlation with a R^2 of 0.91, a slope of 1.01 and an intercept of 1.5 ml.

Figure 3.8 shows the ability for the RMNMS segmentation to capture the presence of a lesion for different lesion size groups (<0.05, 0.05-0.10 and bigger than 0.10 ml). Sixty percent of all manually segmented lesions are smaller than 0.05ml and not surprisingly, it is easier to capture the presence of bigger lesions as demonstrated by the LTPR and LPPV (median results: LTPR=

100.0±16.2% and LPPV=100±17.6%). For the lesions smaller than 0.05ml, the results are not as good (median results: LTPR= 62.5±20.9% and LPPV=71.7±26.2%).

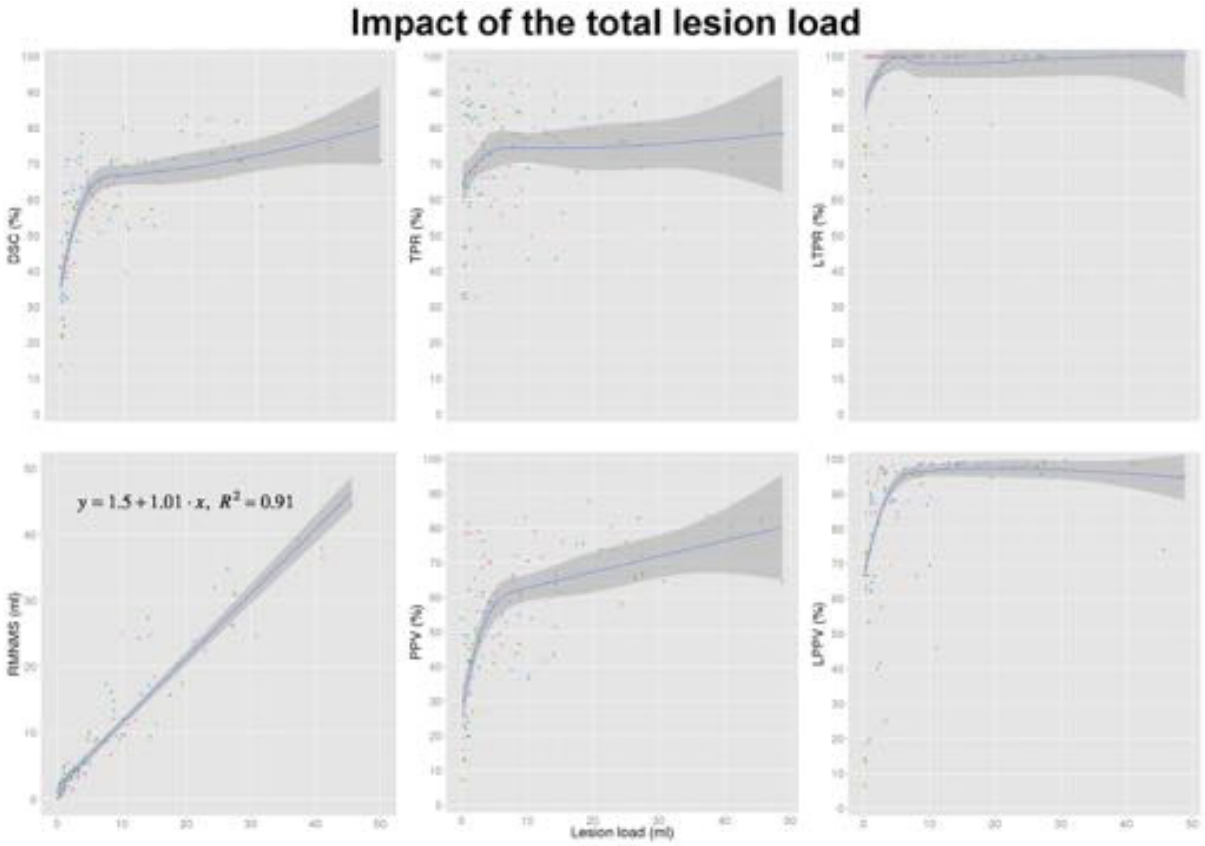


Figure 3.7: Impact of the lesion load on DSC, manual lesion load linear correlation with RMNMS, TPR, PPV, LTPR and LPPV. The experiment was conducted with a patch radius of 2 voxels, a search area radius of 5 voxels and a pre-selection of 50 training subjects on the 108 RRMS subjects (represented by colored dots on the graph). The blue line represents a non-parametric fitting using a nearest neighbour approach with a locally weighted regression for DSC, TPR and PPV and a linear fitting for the linear regression of the manual lesion load and RMNMS lesion volume. The darker grey shading represents the 95% confidence and for the linear correlation, the slope, the intercept and the residual error (R^2) are provided on the graph.

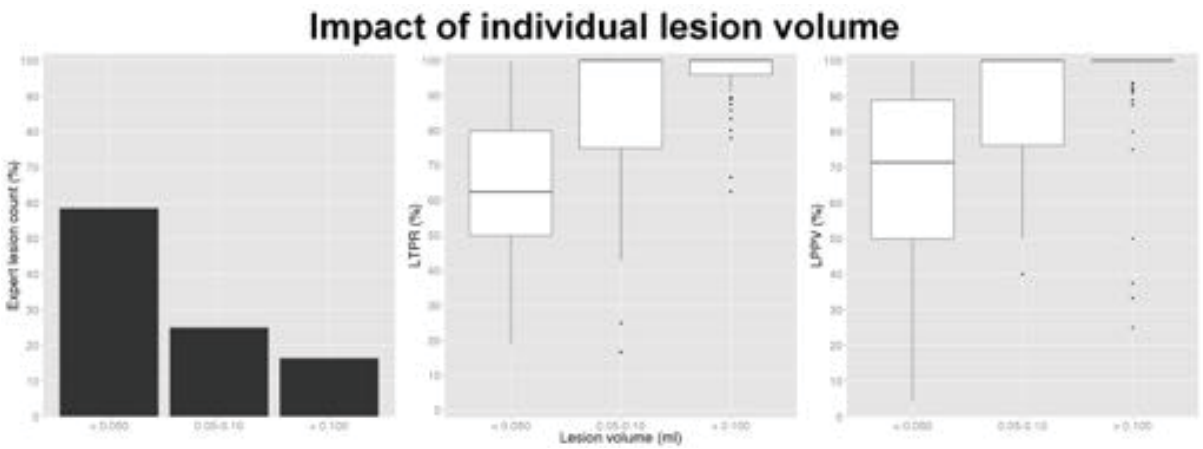


Figure 3.8: Expert segmentation lesion count, LTPR and LPPV per lesion size groups. The plot on the left shows the manually outline lesion count per each lesion volume group (<0.05, 0.05-0.10, > 0.10 ml), averaged across all subjects. The LTPR and LPPV measurement experiments were conducted with a patch radius of 1 voxel, a search area radius of 5 voxels and a pre-selection of 50 training subjects on the 108 RRMS subjects.

Impact of lesion spatial location and examples

In this section we present the RMNMS segmentation results with images to qualitatively describe its spatial behavior.

In Figure 3.9, the expert and the automatic RMNMS probability maps of the lesion segmentation show similar frequency and spatial distribution. While the TP and the FP follow the spatial prevalence of the peri-ventricular region, the spatial distribution of the FN is more uniform suggesting non-systematic segmentation errors.

Figure 3.10 shows images 3 RRMS patients with the highest, median and smallest lesion load with their respective RMNMS segmentation TP, FP and FN. One can appreciate the ability of the method to capture the presence of most of the lesion regardless of the amount and size of the subject's lesions.

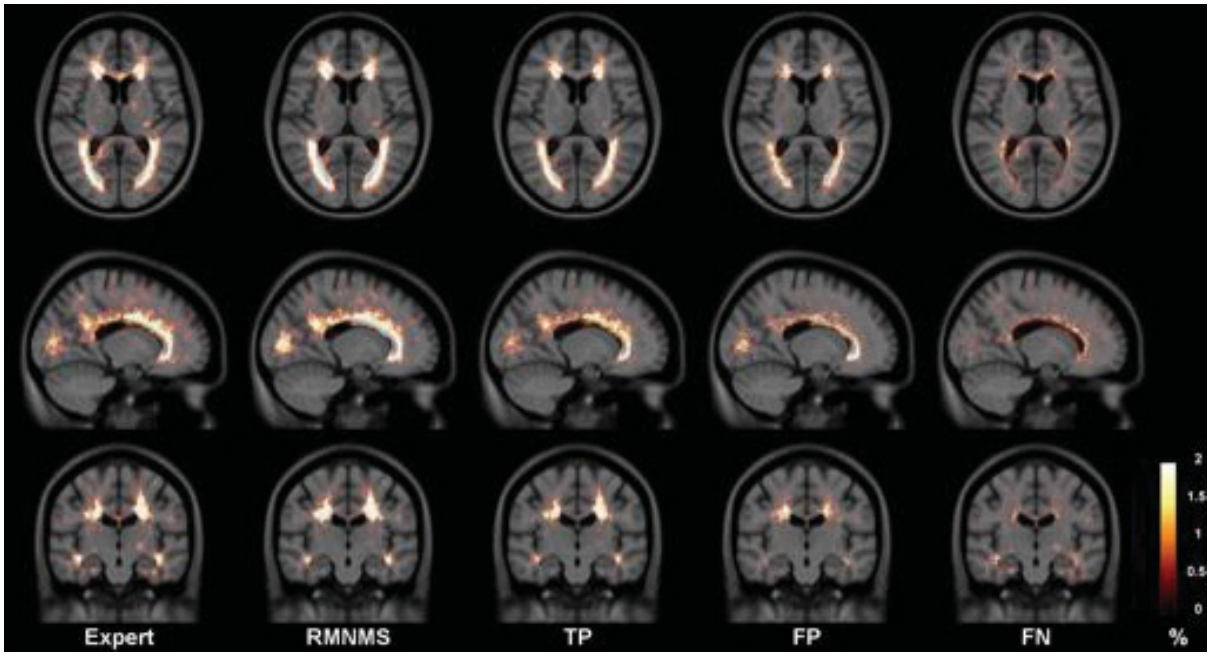


Figure 3.9: Expert and RMNMS, TP, FP and FN lesion segmentation probability maps for the 108 RRMS patients. All the maps are displayed within the same range and overlaid on the RRMS template T1W.

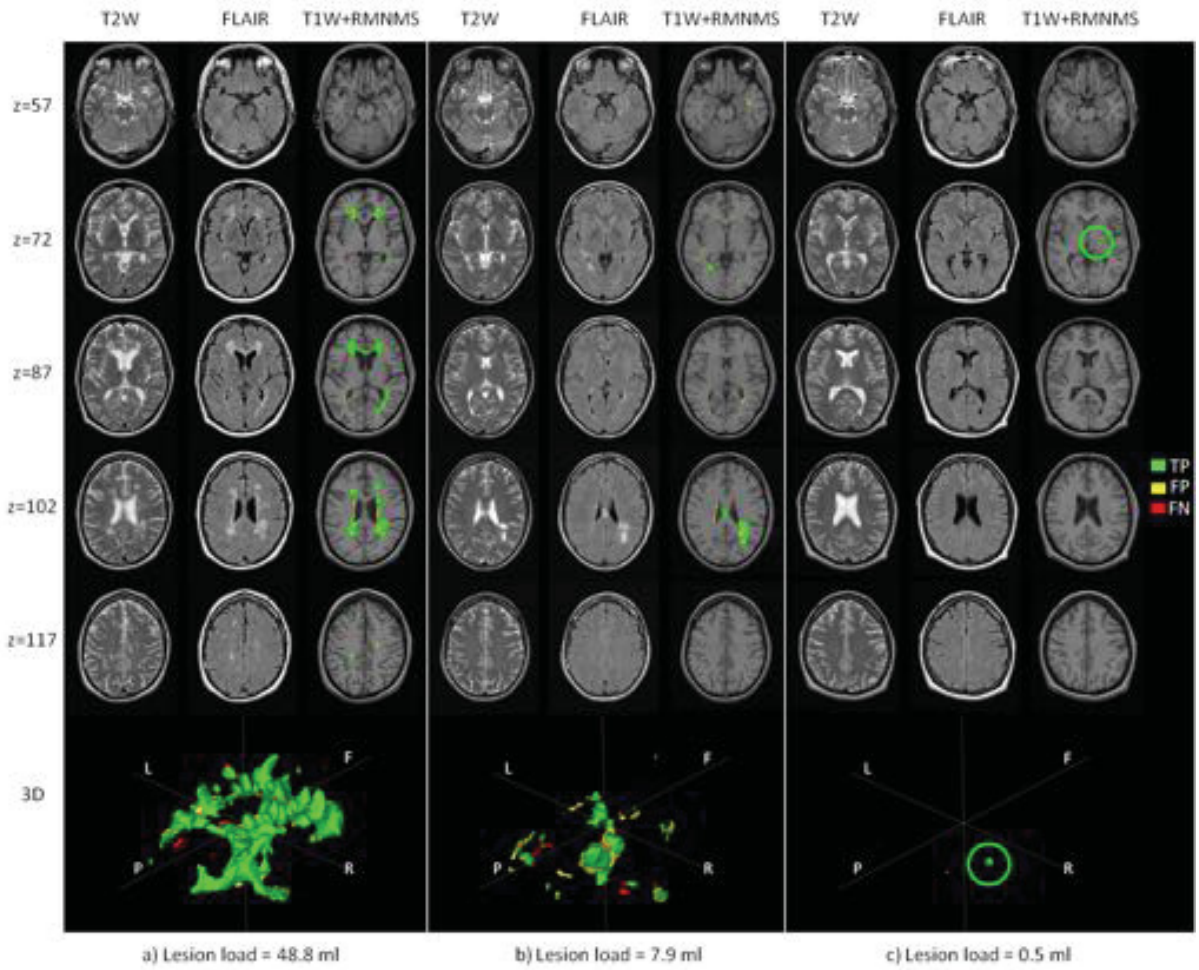


Figure 3.10: Segmentation results for 3 RRMS cases. a) the largest (48.8 ml), b) median (7.9 ml and c) the smallest (0.5 ml) lesion load of the cohort. The figure shows axial slices ("z" is the z-coordinate in mm in the MNI space) for T2W, FLAIR and T1W combined with the automatic RMNMS segmentation ("T1W+RMNMS") and 3D rendering of the segmentations (orientation is defined such as F=frontal, P=posterior, R= surgical right and L= surgical left). The overlapping voxels (TP) with the manual segmentation are represented in green, while the false positives (FP) are yellow and the false negatives (FN) are red. The green circle highlights the TP of the unique lesion for subject "c". The experiment was conducted with a patch radius of 1 voxel, a search area radius of 5 voxels and a pre-selection of 50 training subjects on the 108 RRMS subjects.

3.4.2. MSGC results

Images from the MSGC were pre-processed like the RRMS dataset. For segmentation, the training library consisted only of the MSGC training dataset. First, we present our leave-one-out cross validation results on the MSGC training set and then we compare our results on the testing set with other methods using an objective web-based system (Styner, Lee et al. 2008).

MSGC training dataset

The 20 MSGC training subjects RMNMS segmentations were evaluated in a leave-one-out cross-validation using ((20-1)x2=38 templates (including the mirrored images). We chose to use a bigger search area radius to compensate for the smaller number of training subject than was available in the RRMS validation. In order to capture the presence of lesions in a greater search area in the library the following parameters were used: patch size radius=3 and search area radius=7. It is interesting to note, that the DSC (43.8±16.03) results of RMNMS on this dataset are significantly smaller than for the RRMS dataset (62.3±14.6%). Similar comments can be made for the TPR of 43.9±19.1% and the PPV of 48.7±17.1%.

Given the decreased accuracy of RMNMS with the MSGC dataset we decided to compare the two manual gold standard labels using the same metrics. Comparing the two gold standard manual labels yields a median DSC of 23.7±13.5%, a TPR of 37.1±16.4% and a PPV of 20.2±19.5% confirming the low agreement between the raters.

MSGC testing dataset

The segmentation of the MSGC testing dataset was performed using the whole cohort of training subjects in the template library (20x2=40) with the same parameters as those used for the training experiment except for the pre-selection number that was set to 40. Our segmentation results were interpolated back to their original space and then uploaded to the MSGC website, where an objective independent automatic evaluation was performed. The MSGC provides a

results archive, allowing us to compare the performance of our method with other groups. The results are summarized in Table 3.3.

At the time of writing, RMNMS held the best result with an overall average summary score of 86.1 (note that 90 corresponds to a segmentation accuracy reaching human expert inter-rater variability). While RMNMS holds the best results for VolD and SurfD, this advantage is not statistically significant compared to Souplet et al. (2008), Geremia et al. (2011) and LesionTOADS; however RMNMS has a significantly lower FPR when compared to these methods.

Table 3.3: VolD, SurfD, TPR and FPR results on the MSGC testing dataset. Our method is compared to 3 methods. The best measures are in bold and the significant differences when comparing with RMNMS are in red for each rater (CHB and UNC).

Method	Rater	VolD (%)			SurfD (mm)			TPR (%)			FPR (%)			Score
		Mean	Std	*p*-value	Mean	Std	*p*-value	Mean	Std	*p*-value	Mean	Std	*p*-value	
Lesion TOADS	CHB	85.2	123.4	0.64	8.2	10.8	0.47	55.8	24.0	0.64	70.5	22.8	**<0.01**	79.96
	UNC	63.7	125.7	0.48	7.2	9.1	0.32	49.0	24.5	0.67	74.9	23.2	**<0.01**	
(Geremia, Clatz et al. 2011)	CHB	52.4	29.1	0.89	5.67	9.72	0.96	**59.0**	19.9	0.05	71.5	14.9	**<0.01**	82.07
	UNC	45.0	33.0	0.89	5.67	6.82	0.89	**51.2**	20.4	0.23	76.7	12.9	**<0.01**	
(Souplet, Lebrun et al. 2008)	CHB	86.4	107.3	0.10	8.40	11.1	0.13	58.2	23.5	0.41	70.6	18.1	**<0.01**	80.00
	UNC	57.9	30.8	0.14	7.54	8.43	0.22	49.1	16.1	0.66	76.3	17.4	**<0.01**	
(Tomas-Fernandez and Warfield 2011)	CHB	53.4	56.0	0.86	8.29	7.63	0.03	51.8	19.7	0.83	45.1	22.7	0.26	84.46
	UNC	**37.8**	28.3	0.34	7.03	5.75	0.20	42.0	16.0	0.19	44.1	23.0	0.87	
RMNMS	CHB	**51.3**	30.4	-	**5.49**	5.65	-	52.7	19.6	-	**41.96**	23.1	-	**86.11**
	UNC	46.3	25.7	-	**5.50**	4.22	-	47.0	19.6	-	**43.49**	20.6	-	

3.5. Discussion

In this article, we proposed a new method to detect MS lesions using a training library containing T2W and FLAIR images along with manual T2w lesion masks. Our adaptation of the increasingly popular NLM segmentation method to MS lesions identification with a new multi-contrast and RI distance measure has proven to be highly competitive in our internal validation and in an independent comparison. On a large clinical dataset of 108 RRMS patient, the best compromise between sensitivity, specificity and computation time using leave-one-out cross-validation was obtained with a patch radius of 1 voxel, a search area radius of 5 voxels and a pre-selection of 50 subjects (median results: DSC=60.1±16.4%, TPR=75.4±15.7%, PPV=55.0±20.1%, VolD=33.5±68.9%, LTPR=79.8±14.6% and LPPV=85.7±24.2%). Given the large RRMS cohort size and variability (e.g. lesion load, age, sex, MRI protocols and scanner brand, etc.), these results rank among the best in the MS segmentation literature (Lladó, Oliver et al. 2012). Furthermore, when compared to the state-of-the-art methods with the publicly available MSGC dataset used during the 2008 MICCAI challenge, the RMNMS yields highly competitive segmentation accuracy (best score, 86.11) and produced segmentations that are comparable to the inter-rater variability.

Our voxel-wise analysis showed promising result with respect to the ability to automatically define the volume and the boundary of the MS lesions. Moreover, our ability to segment MS lesions is relatively independent of the patient's lesion load and lesion location. We also investigated the ability of RMNMS to detect the presence of lesions as lesion-wise measures are often more clinically relevant. For example, lesion count is often used for diagnosis and the evaluation of treatment effect. In this aspect, RMNMS shows a great ability to detect the presence of lesions; it detects almost all lesions bigger than 0.05 ml and 62.5% of lesions smaller than 0.05ml. Furthermore, due to the ability of RMNMS to explore a large training set cohort with a large search radius, the probability of detecting MS lesions inside anatomical regions is still high within regions of infrequent MS lesions occurrence (i.e., non peri-ventricular lesions).

In both sets of experiments, results were obtained from a multi-center study, which highlights the robustness of our method in the face of inter-site variability. Whereas many methods require at least 3 MRI contrasts (T1W, T2W, PD or FLAIR) (Souplet, Lebrun et al. 2008, Geremia, Clatz et al. 2011), and others require even-more contrasts (FLAIR, diffusion tensor imaging fractional anisotropy and mean diffusivity...) (Morra, Tu et al. 2008), we use only two (T2W and FLAIR). This dual-contrast method presents multiple advantages. First, reducing the MRI acquisition time by reducing the number of contrasts can decrease the risk of corruption due to image artifacts, it reduces the financial cost and increases patient comfort. When compared to 3-contrast RMNMS (T2W+T1W+FLAIR), the dual-contrast RMNMS with T2W+FLAIR provides better results with shorter computational time.

NLM segmentation based on a single contrast image shows higher DSC results for the hippocampus (median DSC=88.4%) (Coupé, Manjón et al. 2011), brain (DSC=98.3%) (Eskildsen, Coupé et al. 2012), lateral ventricles (median DSC=96.1%) (Fonov, Coupé et al. 2012) and other structures of the brain (Rousseau, Habas et al. 2011). However, DSC is not an optimal similarity metric for small structure segmentation (Rohlfing, Brandt et al. 2004) and because of spatial scattering, anatomical variability and intensity variations, MS lesion segmentation is a much more complex problem. Indeed, our implementation of the standard NLM segmentation with multi-contrast algorithm (MNLM) only achieves a median LTPR=67.3%. Where, the multi-contrast RMNMS (T2W+FLAIR) significantly improves the detection of lesions (LTPR=79.8%) and significantly decreases the computational time. This demonstrates the importance of considering not only the voxel-by-voxel intensity similarity but also the importance of patch-based RI methods for the problem of lesion segmentation. Because of the important reduction in computational time, RMNMS enables the exploration of each training subject with a much wider search radius, which allows for capturing smaller lesions that can even be located in regions where there is low probability of lesion presence in the library. To further increase the presence of similar image in the training library and thus the presence of similar lesions, we used left-right mirrored images and showed the positive impact on the RMNMS segmentation results.

The NLM segmentation technique as applied to the anatomical structures mentioned above require a smaller set of pre-selected training subjects (20 subjects for hippocampus, lateral ventricles, brain) for optimal results while for MS lesions, RMNMS requires more than 40 training subjects to plateau. This difference can be easily explained by the characteristics of the structure to be segmented where spatial distribution, shape and size of MS lesions are not consistent and thus require a larger number of training subjects to capture this variability. Yet another advantage of the subject training pre-selection in the case of altered images is the selection of the "closest" subjects from the training library. Indeed, despite the presence of artifacts and abnormal intensity non-uniformity in the MSGC dataset (García-Lorenzo, Prima et al. 2008), RMNMS has proven to be highly accurate in part due to the pre-selection of the most representative training subjects.

The comparison of MS lesion segmentation algorithms is a difficult task as described by García et al. (2013) for multiple raisons: lack of publicly available datasets/methods, differing MRI contrasts, optimal parameters, and inter-rater segmentation variability. Indeed, variation of MS lesion manually defined on the same subject by different experts has been reported to vary greatly by Zijdenbos et al. (2002). The MSGC dataset (Styner, Lee et al. 2008) also shows significant inter-rater variability with VolD=68% and SurfD=4.85 mm. More importantly, the MSGC training set has an inter-rater reliability of 25% (DSC). One assumes that the MSGC testing set is similar. Despite these criticisms, the organizers of the MSGC are to be congratulated as the MSGC dataset is the first publicly available MS lesion dataset and independent platform for segmentation algorithm validation and comparison. That being said, the MSGC results need to be interpreted carefully with certain limitations in mind. First, the low agreement between the raters should be used as a reference. This can be done by mapping a 25% DSC to a 90% score to represent inter-rater variability when assessing methods. This poor inter-rater agreement may be due to the quality of the images and the presence of multiple artifacts as mentioned by García et al. (2008). The high inter-rater variability for the gold standard MSGC labels results in an upper bound on the quality metrics, as it is not possible to simultaneously agree with multiple manual raters that do not agree. For these reasons it is not surprising that

RMNMS obtained lower similarity measures on the MSGC than on the clinical RRMS dataset. Second, the MSGC provides pre-processed data (registration, interpolation…), which is not optimal for the different pre-processing steps specific to the different segmentation algorithms. Finally, the on-line validation metrics are only voxel-wise measures, but the MS segmentation problem cannot be only seen as a voxel-wise or volume difference problem. MS lesion segmentation is also a detection problem especially in the context of clinical studies where a method should capture the presence of all individual lesions. This is not reflected in the global DSC, VolD and SurfD measurements.

Despite these limitations, we compared our approach with state-of-art supervised and unsupervised methods (n>45) by submitting our segmentation results of the 23 MS test subjects to the MSGC website (Styner, Lee et al. 2008). While our RMNMS approach attained the first position at the time of writing with a score of 86.11, this result must be considered with the limitations described above. We feel that our evaluation with the multi-site clinical dataset is much more representative of quality and robustness of the RMNMS technique. We also compared our approach on our RRMS dataset to the popular and publicly available LesionTOADS approach (Shiee, Bazin et al. 2010). Compared to RMNMS, LesionTOADS is a topology preserving approach guided by probabilistic and topologic atlases. This approach was developed to segment T1W and FLAIR images and as any unsupervised approach it is less flexible to image variability that is not described by the underlying models. These differences could explain the better results obtained by RMNMS on both MS datasets.

Future work will focus on improving segmentation results for smaller lesions, further decrease the computational time with more advanced patch matching strategy (Ta, Giraud et al. 2014), investigate the performance and the pre-selection preferences with respect to scanner machine, site, gender and other clinical variables. Finally, we plan to make the RMNMS algorithm available online (http://www.bic.mni.mcgill.ca/RMNMS).

3.6. Conclusion

We have proposed a new method for segmenting MS lesions. Our method, RMNMS, is a multi-contrast and rotation-invariant distance adaptation of the non-local means operator. RMNMS presents highly competitive results compared to state-of-the-art supervised and unsupervised methods and provides segmentation quality near inter-rater variability for MS lesion segmentation. RMNMS, with multi-contrast and rotation-invariant patch distance, demonstrates that the non-local approach is able to detect structures that vary in size, shape and location such as MS lesions.

Acknowledgements: NeuroRx Research for providing the expert manual segmentations, Canadian Institutes of Health Research (MOP-111169 & 84360), les Fonds de Recherche Santé Québec and the MS Society of Canada (BioMed PhD studentship, 691 and an Operating grant to DLC). This study has been carried out with financial support from the French State, managed by the French National Research Agency (ANR) in the frame of the Investments for the future Programme IdEx Bordeaux (ANR-10-IDEX-03-02), Cluster of excellence CPU and TRAIL (HR-DTI ANR-10-LABX-57). This work has been supported also by the Spanish grant TIN2013-43457-R from the Ministerio de Economia y competitividad.

CHAPTER 4 FOCAL PATHOLOGY INPAINTING

Preface

In medical imaging, the presence of MS lesions can affect the accuracy of automated MRI processing tools such as registration, segmentation and cortical extraction, and thus confound the resulting analysis. In the past for registration, lesions were simply masked away and ignored. Here, we propose a method to replace lesion regions with normal appearing tissue, using a non-local mean technique. This new approach exploits the redundancy of the image to detect the most plausible healthy patches to synthesize appearing healthy tissue. Importantly, this inpainting technique does not require any pre-processing steps apart from the delineation of the region to be filled. Ultimately, the detection of pathology proposed in Chapter 3 will be used to estimate the location of the region to inpaint.

This chapter was derived from an earlier conference proceedings published at the 2013 endMS conference (Guizard, Nakamura et al. 2013) and a full version was published in (Guizard, Nakamura et al. 2015).

Non-local means inpainting of MS lesions in longitudinal image processing

Nicolas Guizard, Kunio Nakamura, Pierrick Coupé,

Douglas L. Arnold, D. Louis Collins

4.1. Abstract

In medical imaging, multiple sclerosis (MS) lesions can lead to confounding effects in automatic morphometric processing tools such as registration, segmentation and cortical extraction and subsequently alter individual longitudinal measurements. Multiple magnetic resonance imaging (MRI) inpainting techniques have been proposed to decrease the impact of MS lesions in medical image processing, however, most of these methods make the assumption that lesions only affect white matter. Here, we propose a method to fill lesion regions using the patch-based non-local mean (NLM) strategy. The method consists of a hierarchical concentric filling strategy after identification of the lesion region. The lesion is filled iteratively, based on the surrounding tissue intensity, using an onion peel strategy. This concentric technique presents the advantage of preserving the local information and therefore the continuity of the anatomy and does not require identification of any a priori normal brain tissues. The method is first evaluated on 20 healthy subjects with simulated artificial MS lesions where we assessed our technique by measuring the peak signal-to-noise ratio (PSNR) of the images with inpainted lesion and the original healthy images. Second, in order to assess the impact of lesion filling on longitudinal image analyses, we performed a power analysis with sample size estimation to evaluate brain atrophy and ventricular growth in patients with MS. The method was compared to two different publicly available methods (FSL lesion fill and Lesion LEAP) and a more classic method, which

fills the region with intensities similar to that of the surrounding healthy white matter tissue or mask the lesions. The proposed method was shown to exceed the other methods in reproducing the fidelity of healthy subject images where the lesions were inpainted. The method also improved the power to detect brain atrophy or ventricular growth by decreasing the sample size by 25% in the presence of MS lesions.

4.2. Introduction

Multiple sclerosis (MS) is a chronic autoimmune disease that affects the central nervous system (CNS) and presents different clinical variants but it usually starts with a relapsing remitting phase (RRMS). The underlying neuronal pathology of a relapse consists of attacks of the myelin and creates focal inflammation leading to lesions in both white matter (WM) and grey matter (GM) and can ultimately lead to demyelination, gliosis and axonal loss. Quantification of MS lesions, also known as plaques, is often used in clinical studies as a marker for disease burden because they are visible on conventional magnetic resonance imaging (MRI) (Fazekas, Barkhof et al. 1999). In addition, MRI enables the exploration of the morphological differences. In MS, structural segmentation (i.e. tissue classification (Zijdenbos, Forghani et al. 1998)) and voxel-wise analysis (i.e. voxel-based morphometry (VBM) (Prinster, Quarantelli et al. 2006, Lansley, Mataix-Cols et al. 2013) or deformation based-morphometry (DBM) (Tao, Datta et al. 2009)) have been used to measure these differences. These tools have been used to assess longitudinal changes of anatomical structures (Nakamura, Guizard et al. 2014) or normal appearing brain tissue (NABT) (Sanfilipo, Benedict et al. 2006). However, MS lesions can swell, shrink and disappear over weeks or months depending on the pathological activity and evolution of the disease (Rovira, Auger et al. 2013). These longitudinal changes affect their appearance on MRI and thus can potentially affect image processing tools such as registration (Brett 2001, Meier and Fisher 2005) and tissue classification (Nakamura and Fisher 2009, Chard, Jackson et al. 2010), and may lead to longitudinal inconsistencies.

In order to remove the variability due to MS lesions, various approaches have been proposed. Depending on the application and the final objective, after identification of the region of interest (ROI), it is possible either to remove ("Mask-out") or to replace these voxels with potential NABT intensity values. Masking-out MS lesion has shown some limitations in the context of longitudinal brain atrophy measurements (Battaglini, Jenkinson et al. 2012). Lesion filling or inpainting strategies consist in replacing or synthesizing voxel values within the region of the MS lesion by representative NABT values. A variety of approaches have been proposed in the literature. Sdika and Pelletier (2009) described three different inpainting strategies: *basic, local white matter (LWM) and global white matter inpainting*. *Basic inpainting* was inspired from Telea et al. (2004) and consists in propagating the local average of the outer region towards the inner region of the lesion mask equivalent to an *onion peel strategy*. *Local white matter inpainting* uses a prior tissue classification of the NABT to fill the lesion with the local normal appearing WM (NAWM) intensity average. *Global white matter inpainting* fills the lesion region with the global intensity average of the NAWM obtained from the tissue classification. Chard et al. (2010) proposed LEAP (LEsion Automated Preprocessing) which also uses NABT classification but extracts the NAWM histogram properties to obtain its intensity peak and noise properties to fill the lesion region. Later, Battagliani et al. (2012) proposed an approach implemented in FSL[4] which fills the lesion with random intensity values from the surrounding NABT distribution of WM and partial WM volumes. These methods focused on reducing the impact of white matter lesions and have been shown to improve results for cortical GM atrophy measurement (Ceccarelli, Jackson et al. 2012, Magon, Gaetano et al. 2014, Popescu, Ran et al. 2014) as well as for white matter atrophy estimation (Chard, Jackson et al. 2010). However, methods such as *basic inpainting* use the surrounding voxels to fill and propagate intensities and thus can potentially fill the lesion regions with undesired intensities. The main limitation of these methods is their assumption that only WM should contain lesions. Furthermore, these methods

[4] http://fsl.fmrib.ox.ac.uk/fsl/fslwiki

rely on tissue classification which can be challenging in presence of MS (Derakhshan, Caramanos et al. 2010) due to the underlying neuropathology affecting the NAWM intensity (Vrenken, Geurts et al. 2006).

In the computer vision community, the field of image inpainting has the goal of producing a plausible image after the removal of a region defined by an operator. Inpainting is often used to restore image deterioration (e.g. scratches, dust speckles…), remove or add elements (e.g. text elements, publicities, persons…) from the remaining information of the image. The main inpainting methods in the literature may be categorized as being sparsity-based, variational, and patch-based. Bertalmio et al. (2014) provides an interesting review of the inpainting literature. Here we describe a patch-based approach inspired from methods that were initially proposed for texture synthesis. During the last few decades, several paradigms have been used in computer vision. First, the method described in Efros et al. (1999) has proven to be effective, using an "onion-peel" strategy to fill the region from its outer surface to its inner core. Their method compares the available patches (small regions of the image) and fills the considered "empty" central voxel of a patch (a small nxn area, where typically n=5..15) with the central voxel intensity value of the most similar patch before moving to the next voxel to be filled. Later, Criminisi et al. (2004) proposed an *exemplar-based* approach which fills the whole patch instead of the central voxel for faster processing, while prioritizing the filling of edges first. Despite impressive visual results, several limitations remain for these inpainting algorithms. The main limitation is that by using only the best match sample chosen could be corrupted or not a perfect match. More recently, the Non-Local Mean (NLM) method, used to compare patch similarities initially proposed for image denoising (Buades, Coll et al. 2005), takes advantage of the image redundancy by using a large number of patches instead of the closest one and has been applied to 2D image inpainting (Wong and Orchard 2008).

Although using patch-based inpainting strategies has shown promising results in computer vision in natural and artificial scenes, it has yet not been fully exploited in medical imaging. This approach presents the enormous advantage of not requiring any tissue segmentation *a priori*, and allowing rough larger lesion delineations. Another advantage of not requiring tissue

classification is that the method does not depend on specific image contrasts. Indeed, our inpainting approach can be applied to any types of MRI acquisition protocols. Inspired by the computer vision inpainting techniques, we used an exemplar-based NLM inpainting strategy in the context of MS lesion filling in MRI (Guizard, Nakamura et al. 2013). The proposed method consists of a concentric filling strategy. After identification of the lesion region, the lesion is filled using an onion peel strategy where concentric layer's voxels of the lesion are successively replaced by the weighted average of the surrounding *normal* patches (detailed below). Inspired by our initial NLM lesion inpainting technique (Guizard, Nakamura et al. 2013), Prados et al. (2014) applied a similar approach with a different initialization strategy (they used the original voxel values) and smoothed the NLM result. Here, we propose to improve the initialization and the convergence using a hierarchical framework, which synthesizes the image intensity variability in the lesion mask.

In this article, we provide a thorough validation using simulated lesions on healthy subjects where we assessed the similarity of the inpainted lesion images and the original images using peak signal noise ratio (PSNR). We also performed power analysis on longitudinal MS patient data to detect changes over time. We compare our proposed method to three different publicly available MS inpainting methods: LWMI (Sdika and Pelletier 2009), LEAP (Chard, Jackson et al. 2010) and FSL lesion filling (Battaglini, Giorgio et al. 2009).

4.3. Methods

In the following section we first describe the NLM inpainting, the filling strategy and the proposed hierarchical approach. Here, given an image I and the lesion ROI (L), we define the inpainted image $\hat{I}$ at the voxel location i as to obtain the final image I^*, such as:

$$I^*(i) = \begin{cases} \hat{I}(i) \forall\, i \in L(i) \\ I(i)\ \forall i \notin L(i). \end{cases}$$ Equation 4.1

4.3.1. NLM inpainting

The propose NLM inpainting approach takes advantage of image redundancy to locally average similar realizations of the image. Indeed, the idea of the NLM was initially proposed for image denoising (Buades, Coll et al. 2005) to reduce the noise of the image by averaging the voxels of patches that would have the same intensity in the noise-free image. Similarly to denoising, our inpainting strategy exploits the redundancy of the image to fill the lesion.

The patch distance estimator (*dist*) used for denoising is here adapted for inpainting by comparing the patch *P(I(i))* centered on *i* (in red in Figure 4.1) with the patch *P*(*I(j)*) centered on *j* (in green in Figure 4.1) within a certain search area (*Ω*):

$$dist(P(I(i)), P(I(j))) = \sum_{x \in P(I(i)) \wedge y \in P(I(j)) | i \in L^* \wedge j \in \Omega} (I(x) - I(y))$$ Equation 4.2

where the voxel *i* belongs to the considered voxel of the lesion mask layer *L** (in yellow Figure 4.1).

This distance is then used to computed the weighted function, *w(i,j),* designed to attribute a smaller weight to greater distance measures of the corresponding patches *P(I(i))* and *P(I(j))*, such as:

$$w(i,j) = e^{-\frac{dist\left(P(\hat{I}(i)),P(\hat{I}(j))\right)}{h^2}}$$ Equation 4.3

where h^2 is a chosen smoothing parameter, discussed in the following section.

Once the corresponding distance with *i* of every voxel *j* belonging to Ω is estimated, the ROI, $L^*(x)$, is filled with the weighted average:

$$\hat{I}(i)_{i\in L^*} = \frac{\sum_{j\epsilon\Omega} w(i,j)I(j)}{\sum_{j\epsilon\Omega} w(i,j)}$$ Equation 4.4

4.3.2. Filling strategy

The filling strategy is important in image inpainting in order to preserve continuity of edges and visual consistency present in the true image. The proposed NLM inpainting strategy consists in a hierarchical inpainting of concentric layers, iterated for different smoothing parameters.

Concentric filling: The concentric technique presents the advantage of propagating the local information and therefore the continuity of the anatomy (Efros and Leung 1999).

Cubic patches of voxels from the outermost layer to be filled are compared to patches from the surrounding voxels not in the lesion mask *L*. After filling a layer of *L*, this process is repeated on the next interior layer of the new lesion mask, *L**, until reaching its core. Only voxels outside of *L*,* thus including the already inpainted voxels, are used during the patch distance estimation.

Hierarchical inpainting: In order to optimize the performance of the NLM inpainting and to reduce the ambiguity in the case of large lesion (Liu and Caselles 2013), we embed the filling strategy within a hierarchical multi-resolution framework.

Starting from the downsampled resolution and the outside layer of the lesion mask, the process fills the next interior layer until reaching the center of the lesion mask before moving to the next hierarchical level where this process is repeated. The original image and its lesion mask are interpolated at different resolution scales (k) using respectively tri-linear and nearest neighbor interpolations. Starting from the lowest resolution level, the inpainting results of the innermost concentric layer are then used to initialize the following level. The inpainted regions of lower k levels are interpolated using tri-linear interpolation to the *k-1 level* to replace the voxels filled at the previous iteration.

Smoothing parameter (h^2): Within the NLM approaches, h^2 is critical to attribute weight to the most similar patches. For our inpainting problem, decreasing h^2 attributes less weight to less similar patches while a bigger h^2 value tends to provide smoother inpainting results. Therefore, for each inpainted voxel at each hierarchical level and each concentric layer, we iterate the NLM inpainting with the following successive h^2 values [0.9, 0.7, 0.5, 0.3, 0.1]. Starting the inpainting of the considered voxel with a big h^2, we initiate the voxel filling with a smooth value with respect to the neighborhood (Ω). Then, successively decreasing h^2 to 0.1, is equivalent to searching for the most similar patch (i.e., the minimum intensity distance) in Ω, thus synthesizing the finer image textural details.

The concentric and hierarchical inpainting processes are graphically illustrated respectively by the "Layer" and the "Level" axes in Figure 4.1. In the following experiments, we used three (k=3) isotropic resolution levels (4, 2 and 1mm) with similar patch sizes (9x9x9 voxels).

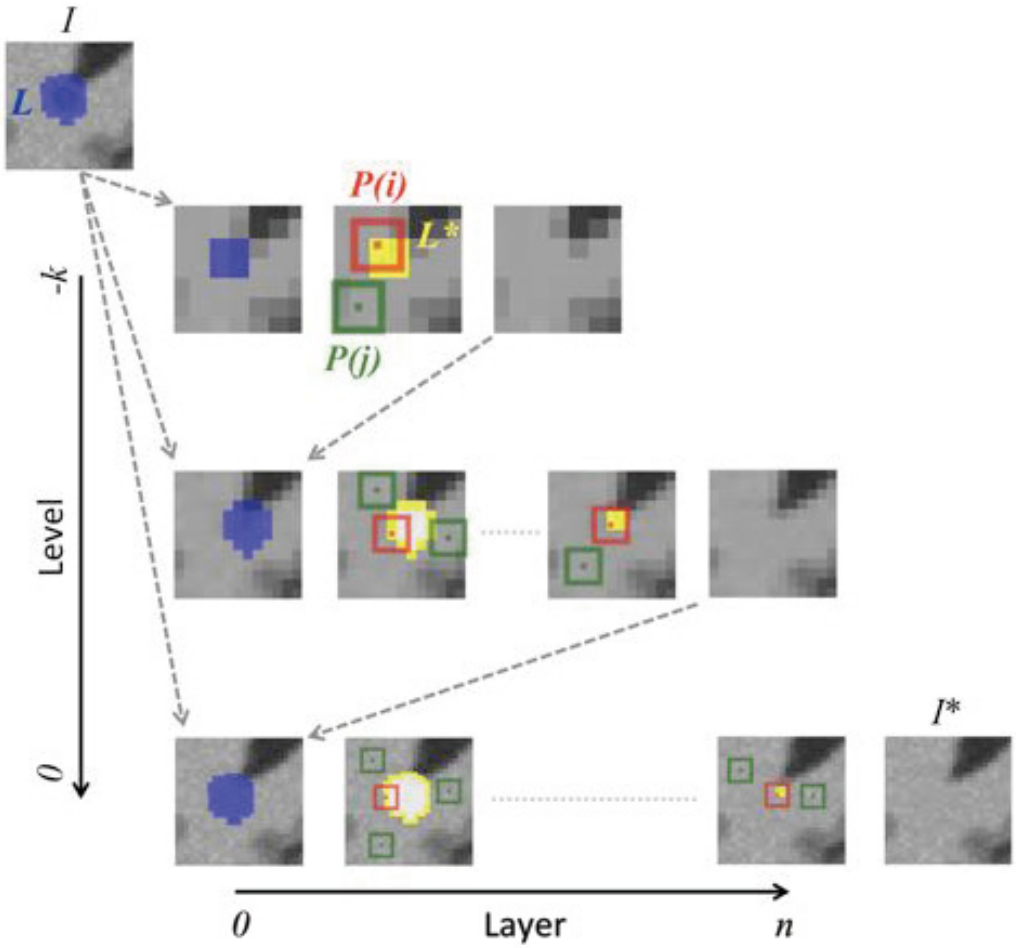


Figure 4.1: NLM lesion inpainting strategy. The inpainting process starts with the lesion mask (*L*) of the original image (*I*) in the downsampled space *k* to obtain the inpainted image of this level. Then, the inpainted region is upsampled into the image of the next hierarchical level. The inpainting itself consists in finding the most similar patches (*P(j)*, in green) in the "non-lesion" region with the considered patch *P(i)*. All voxels in white are not considered during the patch distance estimation. The concentric filling is described by the boundary of the current mask (*L** in yellow) shrinking by one voxel at the next "Layer". The original lesion mask *L* is reinitiated at the beginning of each hierarchical "Level".

4.4. Experiments

In the following section we describe 1) the data used in our experiments, 2) the simulated MS lesion data such that the original MRI intensity information can be used as a ground truth, 3) the longitudinal power analysis to detect brain atrophy and 4) the different methods evaluated.

4.4.1. Data

The Montreal Neurological Institute research ethics committee gave approval for this study and all subjects gave informed consent. To evaluate the proposed algorithm, two neuroimaging datasets were used anonymously:

- From a multi-site clinical study with 67 relapsing-remitting MS patients (RRMS, mean age 37.5 y, SD 10.0 y). Each patient underwent an MRI at two time points, baseline (*m00*) and 12 month (*m12*), that included sagittal T1W [TE=9-11 ms, TR=30-40 ms, flip angle=30°, in-plane resolution=0.977x0.977 mm^2, slice thickness=1.5 mm], T2W [TE=65-104 ms, TR=3666-8585 ms, flip angle=90°, in-plane resolution=0.977x0.977 mm^2, slice thickness=3 mm] and PD [TE=10-18 ms, TR=2200-3800 ms, flip angle=90°, in-plane resolution=0.977x0.977 mm^2, slice thickness=3 mm] images. The MRI data were acquired on 1.5T scanners from different manufacturers: GE (n=20), Philips (n=18) and Siemens (n=29).

- From this RRMS database, we randomly selected T1W images of twenty MS patients to simulate realistic MS lesions on BrainWeb simulation MRIs (http://brainweb.bic.mni.mcgill.ca/brainweb/) (Collins, Zijdenbos et al. 1998) from 20 healthy subjects (Aubert-Broche, Griffin et al. 2006).

Although the application of our inpainting method is not limited to a specific imaging modality, T1W images were chosen since they are acquired as part of many standard imaging protocols and are widely used to assess longitudinal volume changes in MRI. In addition, this modality was used by the other inpainting methods we wish to compare to in this analysis.

4.4.2. Artificial MS lesions validation

The different inpainting methods are evaluated using artificial numerical MS lesions that are simulated on healthy subject MRIs. Simulations are done using the strategy of Brett et al. (2001)., whereby MRI data from MS patients are used to simulate WM lesions on healthy subject MRIs. Here, the goal was to create T1W MS lesions for which we know the underlying ground truth (from the healthy subject data) such that we can compare inpainting results across different methods.

The simulation, illustrated in Figure 4.2, was performed on the healthy brain image (*H*) using real lesions from the MS patient image (*M*), and can be summarized as follows:

1. Pre-processing: i) intensity non-uniformity correction (Sled, Zijdenbos et al. 1998), ii) intensity normalization using linear histogram matching and iii) linear registration (Collins, Neelin et al. 1994) to the stereotaxic ICBM152 template.
2. Tissue classification of *H* and *M*: after an automatic segmentation of the WM, GM, cerebrospinal fluid (CSF) and T2W MS lesions (only on the patients) by a multi-spectral Bayesian classifier (Francis 2004) using the T1W, T2W and PD images. From prior probability model of the segmentation estimated from a training dataset, the M is segmented using Bayes' theorem, where the distribution of each tissue classes is used to estimate the parameters of their Gaussian distribution. The automatic T2W lesion outlines of *M* were superimposed on T1W, T2W and PD for manual reviews. Experts who underwent extensive training on similar MS patient MRI data carefully reviewed the MS lesion mask, *L*.
3. For each H:M pair: Compute the voxel-wise intensity ratio (*R*) of the healthy WM (obtained from stage 2) intensity average (WMa) and the T1W voxel intensity of lesion tissue ($T1W_M(i)$)from the corresponding manually-corrected mask (*L*) of *M* for a voxel *i*:

$$R_M(i) = \begin{cases} \frac{\mathrm{T1W_M}(i)}{\mathrm{WMa}} & \forall \mathrm{i} \in L \\ 1 & \forall \mathrm{i} \notin L \end{cases}$$ Equation 4.5

4. Estimate the non-linear transformation (NL_{reg}) between *M* and *H* (Avants, Epstein et al. 2008).
5. Using the transformation (NL_{reg}), interpolate spatially *R* and *L* into the *H* space and obtain *R'* and *L'*.
6. From the interpolated *R'* and *L'*, create a new image (*H'*) where the final image intensity voxels equal $R' \times H$ where the lesion (*L'*) is defined and *H* everywhere else.

The six steps are repeated for the 20 H:M pairs. The resulting simulated dataset allowed us to assess the impact of the patch search radius for the proposed NLM inpainting before comparing it to state of the art inpainting approaches.

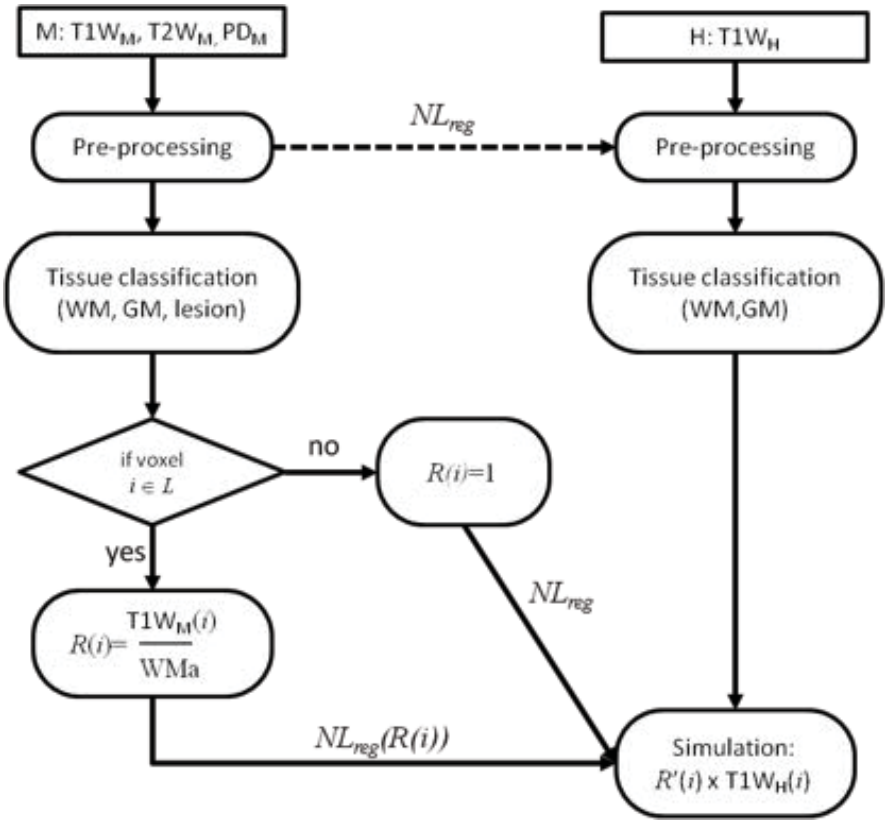


Figure 4.2: Flowchart of the lesion simulation

As a means to evaluate the inpainting algorithms, we can assess the fidelity of the restored image *I** by comparing it to the original image *I.* In the computer vision literature (Group 2003, Wong and Orchard 2008, Fadili, Starck et al. 2009), this is often done using the peak signal to noise ratio (PSNR) by measuring the ratio of the maximum possible power of the signal and the mean squared error (MSE) between the restored and the original image:

$$PSNR = 20.log_{10}\left(\frac{MAX_1}{\sqrt{MSE}}\right) \qquad \text{Equation 4.6}$$

where MAX_1 is the maximum possible pixel value (255 for 8 bits voxel storage) while MSE is estimated estimated within the lesion mask, *L*, between the original image *I* (before adding the lesion) and the inpainted version, *I**:

$$MSE = \frac{1}{n} \sum_{x=1 \in L}^{n} \|I^*(x) - I(x)\|^2 \qquad \text{Equation 4.7}$$

where *n* is the number of voxels. Thus, in this MS lesion simulation framework, we expect a smaller MSE, and thus a higher PSNR when the reconstruction is more similar to the original image.

We first evaluated the performance of our NLM inpainting approach for different search area radii, which is an important factor to find similar patches. The size of the search radius also influences the computational burden.

We then assessed the PSNR results of the different inpainting methods while simulating potential manual lesion segmentation variability and lesion mask misalignment. This is done by varying the original lesion mask boundary (β_0) with morphological operations on the ground truth lesion mask, through dilation by 1 or 2 voxel layers (β_1 and β_2) around the whole lesion mask volume. This will enable characterization of the methods with respect to smaller or larger lesions.

4.4.3. Longitudinal MS data validation

In the second set of experiments, our proposed inpainting technique is evaluated and compared to three different publicly available inpainting algorithms and a masking technique using real longitudinal MS patient data in order to determine the impact of the method on the power to detect longitudinal volume changes.

The longitudinal MRIs were pre-processed using steps 1) and 2) of section 0 in order to obtain the lesion segmentations for each subject's time-point. All inpainting techniques used to the same set of lesion labels for the comparison.

In order to compare the performance of the inpainting methods, the popular longitudinal atrophy measurement tool SIENA (Smith, Zhang et al. 2002) was used to measure the percent brain volume change (PBVC) as well as the percent ventricular volume change (PVVC) in the MS dataset. SIENA starts with brain and skull segmentation (Smith 2002) to perform skull-based registration and analysis in the half-way space of the subject. Then, the brain and non-brain boundary is estimated from tissue classification (Zhang, Brady et al. 2001) before computing the perpendicular displacement between the brain boundaries of the two time-points. Finally, the surface displacement is averaged to obtain a global estimate of PBVC, and the PVVC if ventricle masks are used instead of brain masks.

Statistical comparison of the inpainting approaches was conducted using power analysis where we estimated the sample size (*per arm*), n, required to detect pre-specified treatment effect without accounting for normal aging atrophy (Anderson, Bartlett et al. 2007), such as:

$$n = \frac{2[(a+b)]^2\sigma^2}{(\mu_1 - \mu_2)^2} \qquad \text{Equation 4.8}$$

where μ_1 and μ_2 are the mean rate of volume change in the placebo and treated groups respectively, and σ^2 the corresponding variance of the rate of change. Here, we only had a control MS group, we thus estimated sample sizes for 10, 30 and 50% treatment effects, so that $u_2 = u_1*(1-0.10)$, $u_2 = u_1*(1-0.30)$ and $u_2 = u_1*(1-0.50)$, respectively. The analysis was conducted with 80% power (a=0.842) and a significance level of 0.05 (b=1.96). The 95% confidence intervals were estimated by bootstrapping 10,000 times. The treatment effect are derived from previous clinical trial studies, where treatment effects on RRMS brain atrophy was around 50% (Rudick, Fisher et al. 1999).

4.4.4. Methods compared

We compared our method to 4 other methods that deal with MS lesions: 3 inpainting methods and 1 masking method:

- **LWM** (Sdika and Pelletier 2009) estimates the tissue classes of the NABT to fill the lesion with the intensity average of the surrounding NAWM. Because this method is not publically available, we implemented our own version.
- **LEAP** (Chard, Jackson et al. 2010) also uses tissue classification of the NABT but applies the intensity properties of the NAWM histogram to the region being filled. LEAP is available at: http://www.nmrgroup.ion.ucl.ac.uk/analysis/lesionfill.html
- **FSL** lesion filling (Battaglini, Giorgio et al. 2009) fills the lesion from random intensity values estimated in the surrounding NABT after estimating the tissue WM and partial WM volumes. The **FSL** lesion filling method is available at: http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/lesion_filling.
- **Masked-out**: We also evaluated the impact of removing the MS-lesion for the longitudinal analysis by masking the lesion out (or so called "**Masked-out**" approach (Battaglini, Jenkinson et al. 2012)).

4.5. Result

4.5.1. Artificial MS lesions

NLM inpainting search area

The NLM inpainting algorithm does not require *a-priori* knowledge of the NABT, GM or WM and searches for the most similar patches throughout the whole brain. However, as shown in Figure 4.3, the PSNR plateaus around a radius of 10 voxels (note the discontinuous x axes), precluding the need of doing a brain-wide search. As such, a search area radius of 10 voxels was used in the remaining experiments since it provides a good compromise between reconstruction fidelity and computational burden.

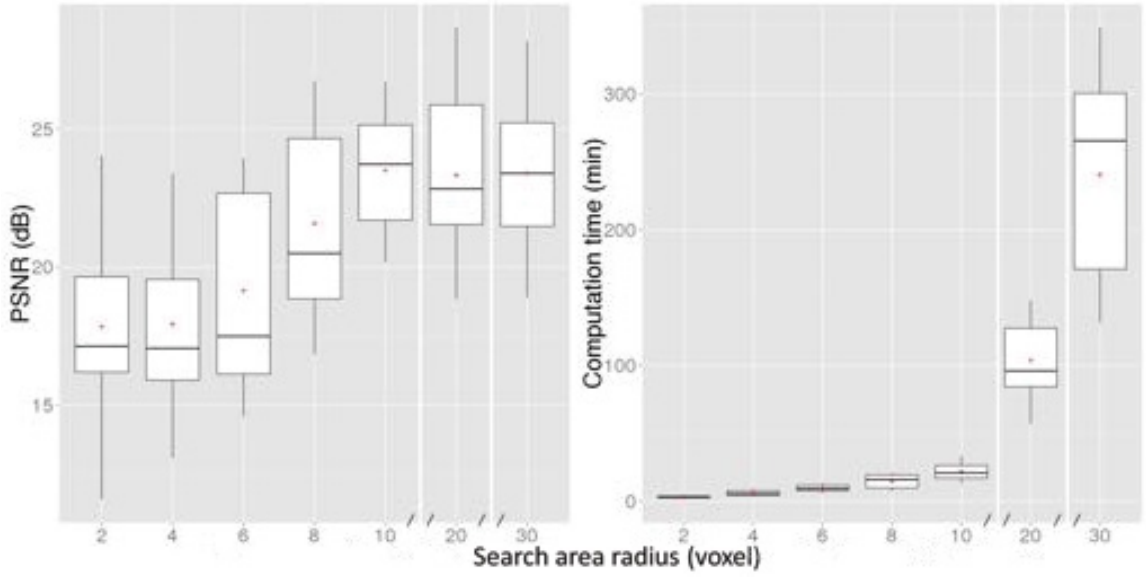


Figure 4.3: PSNR measure and computation time of the proposed NLM method for different search area radii. The boxes represent the lower and upper quartile with the median as the central black line and the mean with a red cross. The whiskers extend to the most extreme data point. Note that the x axis is discontinuous after 10 voxels.

Inpainting of simulated MS lesions

Here, we compare the NLM inpainting approach with LWM, FSL and LEAP, while incorporating segmentation variability by simulating different lesion boundaries from the original lesion segmentation (β_0). We do not compare to the Masking-out technique, as it does not attempt to model the original data.

Figure 4.4 presents the PSNR results of the inpainting strategies for 3 different levels of lesion mask boundaries. We can notice that NLM outperforms the other methods regardless of the lesion mask size (β_0, β_1 and β_2).

A one-way between subjects ANOVA was conducted to compare the effect of inpainting on the PSNR reconstruction measure [$F(4,135)=6.40$, $p<0.01$]. The Bonferroni-adjusted t-test analysis revealed that NLM is significantly better than LWM and LEAP ($p<0.01$) with β_0 and these results are summarized in Table 4.1.

NLM's PSNR is stable when β increases since this approach is not specific to WM intensity distribution which can be altered when the mask used to compute the PSNR becomes bigger than the actual simulated lesion.

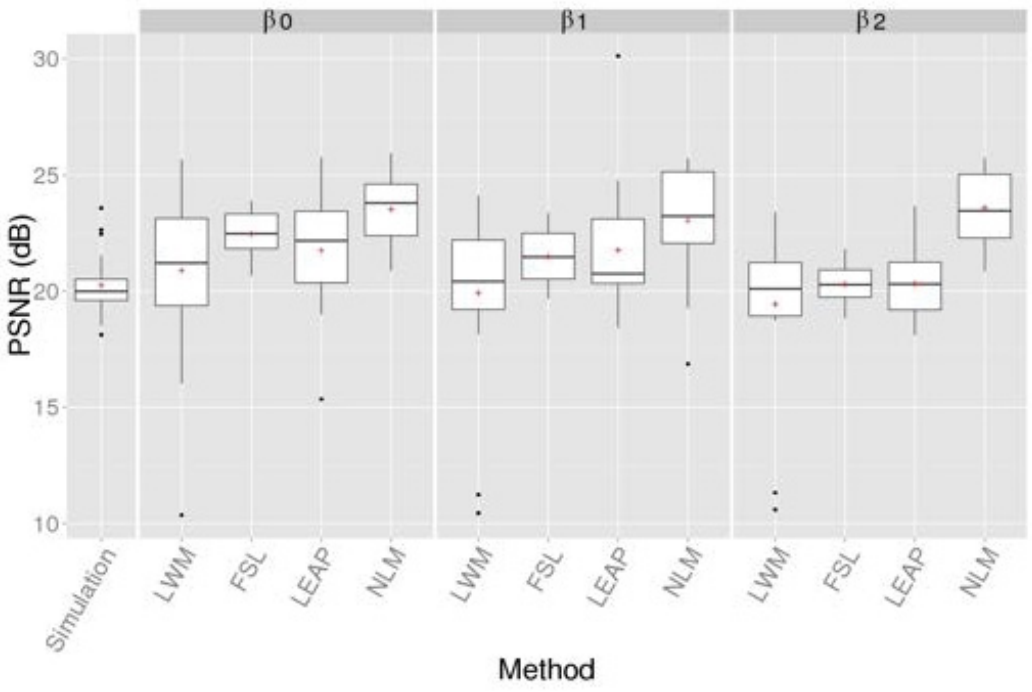


Figure 4.4: PSNR of the simulated (no-inpainting) and inpainted images with the 4 techniques when compared to the original images for different lesion mask boundaries (β). Statistical analysis at β0 is reported in Table 4.1.

Table 4.1: Mean average and standard deviation (SD) of the PSNR for the simulated images without inpainting (no-inpainting) and the different inpainting methods with the original lesion mask (β_0) and Bonferroni-adjusted multi-comparison t-test of the PSNR results.

	Simulation (no-inpainting)	**LWM**	**FSL**	**LEAP**	**NLM**
Mean PSNR	20.25	20.89	22.46	21.75	**23.52**
SD	1.38	3.47	1.01	2.33	**1.42**
t-test with NLM [(t-value, df), p-value]	[(8.68, 27), <0.01]	[(5.32, 27), <0.01]	[(2.56, 27), 0.03]	[(5.43, 27), <0.01]	

Figure 4.5 illustrates examples of the inpainting results for the different techniques based on original images and the simulation of 3 different lesion types. The 3 cases were chosen to visualize typical large (A), medium (B), and small (C) peri-ventricular MS lesions. Visual inspection of the lesion filling with NLM shows qualitatively more plausible contrast, intensity gradients, texture and anatomy compared to other methods. For example, in case A, the NLM inpainting recovers the curved contour of the lateral ventricles despite the fact that the lesion mask reaches the CSF boundary. This is not the case for the LWM and LEAP methods; both show some "bleeding" into the ventricles (red arrows in second row of Figure 4.5). In addition, on cases B and C, the WM/GM boundary gradient is more gradual with NLM and more faithfully reproduces the original contrast. Furthermore, the overall texture of the NLM reproduces the surrounding noise level, while LEAP tends to over-smooth and FSL seems to introduce noise (black dots highlighted with yellow arrows in cases A and B)

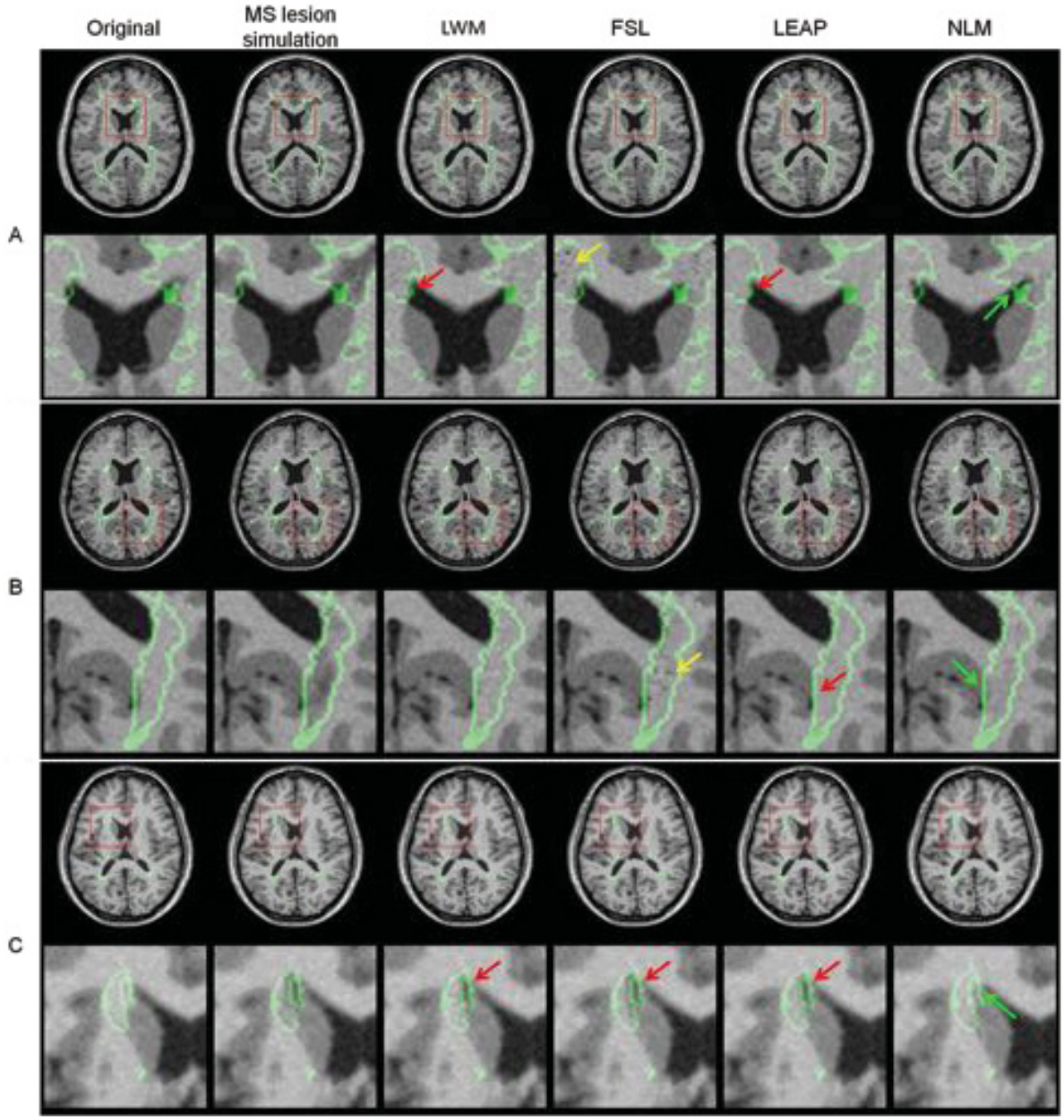


Figure 4.5: Lesion simulation examples with the original image, MS lesion simulation and the different lesion filling results LWM, FSL, LEAP and NLM for 3 different lesion sizes (A=large, B=medium and C=small). The red arrows point to anatomically improbable lesion filling results, the yellow arrows point to intensity problems while the green arrows point to plausible and realistic anatomy. The original lesion boundary and the red square on the axial image depicts the zoom-in image region shown below it. Note that these images were generated at β0.

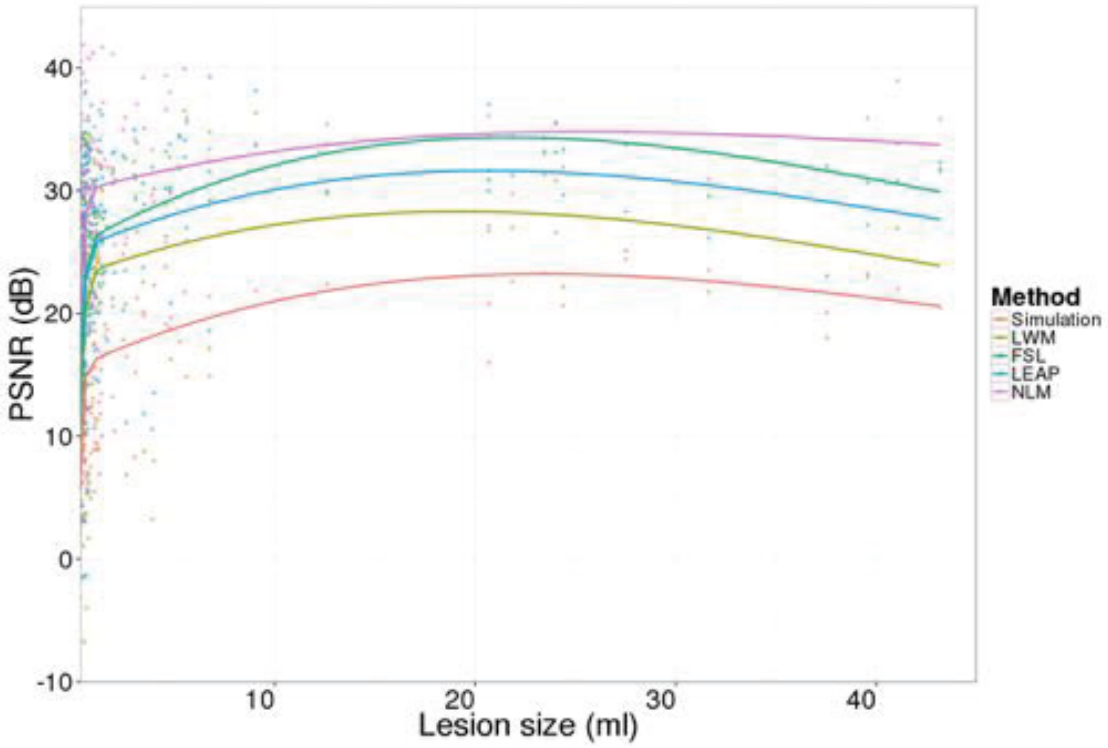


Figure 4.6: Lesion-wise PSNR results for the 20 simulated images (no-inpainting) and the inpainting approaches (LWM, FSL, LEAP and NLM).

4.5.2. Longitudinal MS data

Lesion inpainting on an individual MS longitudinal dataset

In Figure 4.7, we show examples of lesion filling on longitudinal data from an MS patient using LWM, FSL, LEAP and NLM. As can be seen on the original T1W images, the lesion boundary has changed between the two time-points. This is likely affecting the inpainting results, since as discussed previously and described in Figure 4.5, the performances of LWM, FSL and LEAP are more affected by lesion boundaries than the NLM method. This limitation can be appreciated by comparing the right and left panels of Figure 4.7, where the extent of the inpainting 'bleeding' into ventricles is different for the different time-points. Clearly, this would lead to erroneous longitudinal measures of ventricular enlargement, for example. In contrast,

NLM lesion filling seems to provide more plausible contrast and tissue boundaries gradient that are consistent between both time-points (panel *m00* and *m12*).

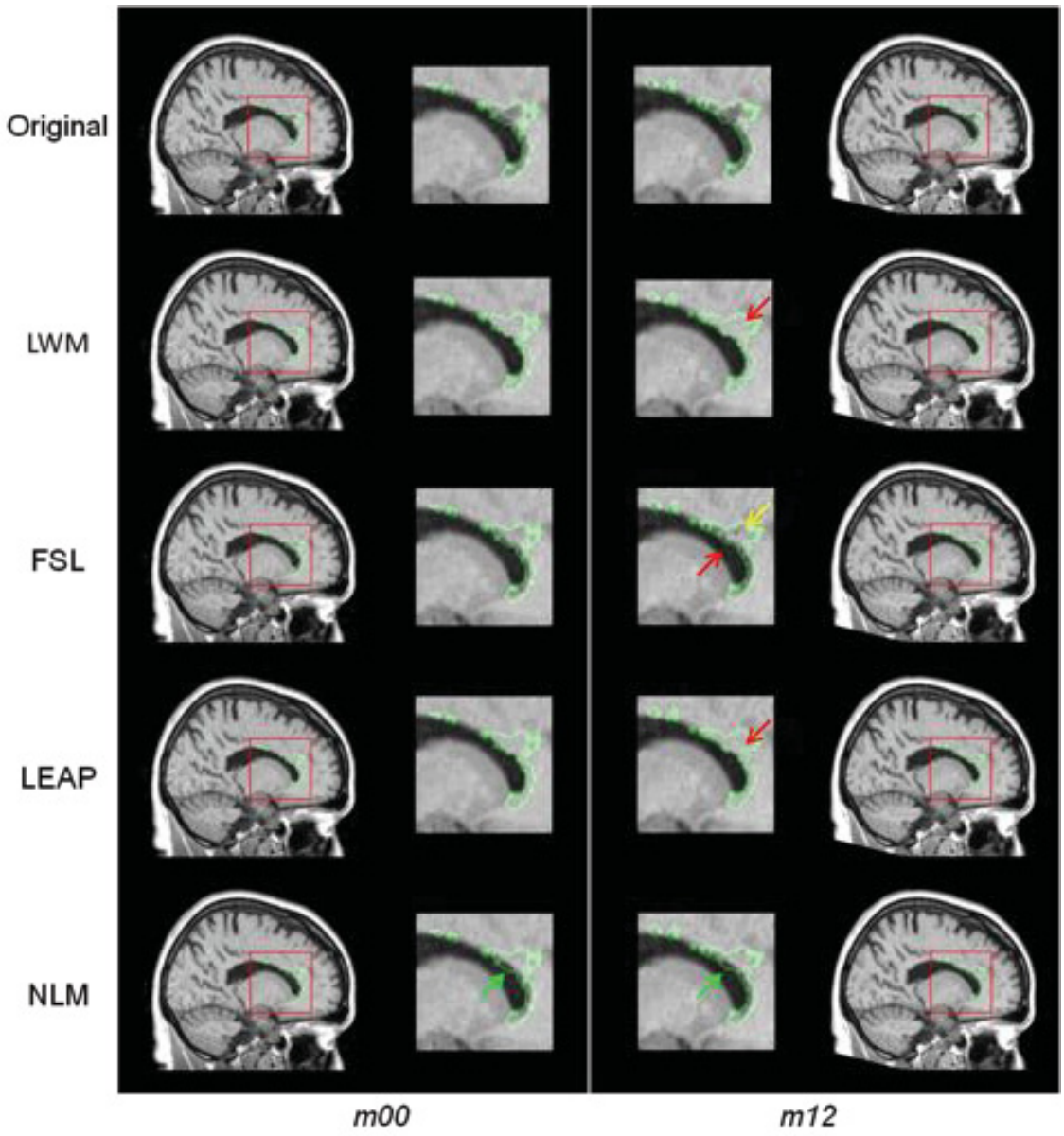


Figure 4.7: Example of lesion inpainting on real longitudinal MS data from one patient at baseline (*m00*) and one year later (*m12*) (first row) for the different methods (LWM, FSL, LEAP and NLM).

Power analysis of brain atrophy measures

The longitudinal analysis of brain atrophy (PBVC) and ventricular (PVVC) enlargement measurements for all 67 MS subjects using SIENA are summarized in Table 4.1. The inpainting (LWM, FSL, LEAP, NLM) and the masking strategies resulted in similar mean PBVC and PVVC changes of about -1.1% and 3.8% respectively. However, NLM has the smallest variability (PBVC SD=0.83% and PVVC SD=4.28%) thus leading to the smallest required sample sizes to detect changes across all assumed treatment effects (10, 30 and 50%). In fact, NLM inpainting leads to a reduction in the number of subjects by a factor of 14% to detect brain volume changes and 21% to detect ventricular enlargement, compared to the volume change estimation on the original data.

An example of the SIENA brain boundary change results for one subject can be seen in Figure 4.8. The figure shows unexpected focal boundary fluctuations (red arrows) in locations where lesions were present on the "original" image without lesion inpainting and with lesion masking but also with WML, FSL and LEAP lesion inpainting. These changes are particularly visible in regions of larger lesions (e.g., peri-ventricular region). The inpainting approaches reduce these fluctuations but the NLM inpainting results show the most homogenous changes across the boundaries. This likely contributes to the lower variability that this method provides across the whole dataset.

Table 4.2: PBVC and PVVC SIENA results and 10000 bootstrapping sample size estimation, with a power of 80% and a confidence interval of 95% for different treatment effects (10, 30 and 50%) between *m00* and *m12*. The smallest sample sizes are in bold font. Note, that PVVC cannot be estimated with the Masked-out approach.

Method	PBVC				PVVC			
	Mean % (SD)	Sample Size			Mean % (SD)	Sample size		
		10%	30%	50%		10%	30%	50%
Original	-1.12 (0.94)	1171 829-1740	130 93-196	47 33-71	3.85 (5.24)	583 329-1145	65 37-125	24 13-46
Masked-out	-1.15 (0.98)	1106 788-1638	123 98-183	44 31-66	-	-	-	-
LWM	-1.08 (0.95)	1117 806-1661	124 90-185	45 32-66	3.76 (4.81)	555 333-1070	62 37-114	22 13-42
FSL	-1.13 (0.94)	1153 829-1730	129 92-193	46 33-69	3.82 (4.79)	539 316-1026	60 35-114	22 13-42
LEAP	-1.12 (0.94)	1179 847-1769	130 92-193	47 34-71	3.97 (4.88)	506 308-956	56 34-108	20 12-38
NLM	-1.14 (0.83)	**999 763-1389**	**110 84-152**	**40 31-56**	3.93 (4.28)	**446 297-727**	**49 33-2**	**18 12-29**

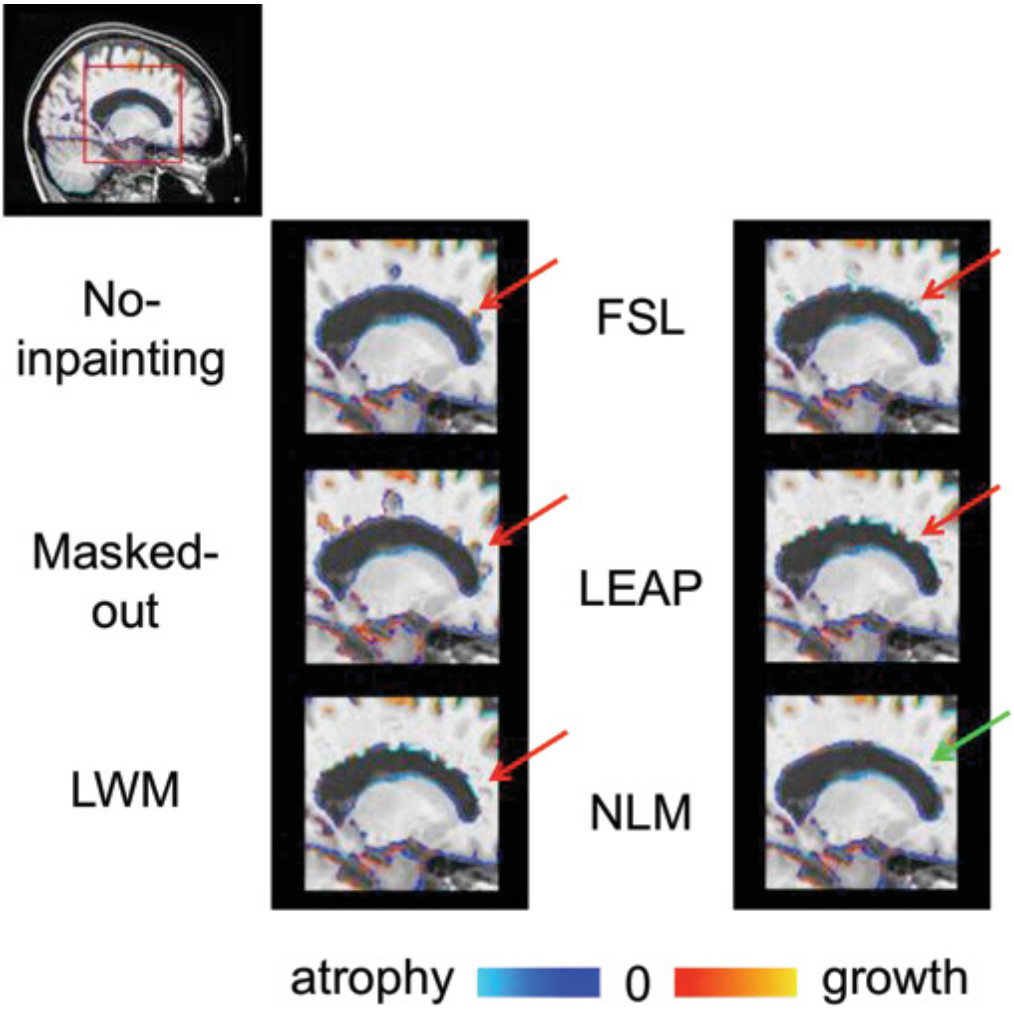


Figure 4.8: SIENA brain boundary changes (atrophy=blue and growth red) of the "original" images and with the different strategies to account for lesions (Masked-out, LWM, FSL, LEAP and NLM).

4.6. Discussion

In this work, we propose a new inpainting NLM method to replace MS lesion ROIs with intensities from surrounding normal-appearing brain tissue. We demonstrated the efficiency of our approach in the context of longitudinal image analysis. The proposed approach presents the advantages of not requiring any pre-processing (after lesion identification) and could be applied to any MR image contrast. With MS lesion simulations and with RRMS 1-year longitudinal

brain change measures, the results of this study show that the proposed method was superior to the most commonly used inpainting approaches. Furthermore, the qualitative visual results of the proposed approach are realistic and anatomically plausible.

On simulated MS lesions our inpainting approach allows us to reproduce with the best fidelity the original "lesion free" MRI images. Using the NLM operators allows replacing a lesion voxel with voxels having the most similar patches without any explicit *a priori* for healthy tissue classification of the healthy tissues. The different boundaries of the lesion masks confirmed this, where bigger masks of the actual lesion do not affect the fidelity of the reconstruction. Therefore, the definition of the lesion mask does not require an accurate definition covering only the affected white matter tissues but a rather "bigger" mask definition. Indeed, because of the propagation strategy using concentric layers, we suggest applying a morphological operator to dilate the lesion mask ROI in order to effectively avoid the propagation of affected tissue intensities.

The evaluation of the different methods in the context of longitudinal brain atrophy and ventricular enlargement measures qualitatively and quantitatively favor the proposed NLM inpainting algorithm. These results suggest that MS lesion inpainting in the context of clinical longitudinal MRI studies have substantial advantages to detect brain atrophy and have already proven to improve some longitudinal structural measurements (Nakamura and Fisher 2009, Magon, Gaetano et al. 2014). MS lesions are more frequently located in the peri-ventricular region of the brain (Narayanan, Fu et al. 1997). This spatial preference could explain the stronger improvement in power to detect ventricular enlargement for the NLM inpainting.

In this study, we only consider WM lesions and T1W images as do most of the inpainting approaches available in MS imaging (LWM, FSL and LEAP). These approaches require modification of their algorithm to fill GM lesions. But, in MRI studies, GM or cortical lesions have been found in the majority of the RRMS populations (64%) (Calabrese, Agosta et al. 2009). Our method, which does not depend on specific image contrasts, is more flexible in that it can

deal with any region where the intensities need to be replaced by intensities from nearby *normal* regions.

We initially proposed to use the NLM as an inpainting operator (Guizard, Nakamura et al. 2013). In this paper, we provide a more thorough validation as well as some improvements on our original approach. Our initial approach inspired Prados et al. (2014) to develop a similar inpainting method based on the NLM, however, their method is different as they only estimate the minimum intensity distance patch before applying a smoothing kernel. Here, we improve on our initial method by proposing a pyramidal hierarchical filling strategy, which enables to capture more structural information at a lower level, propagating this inpainting to the next level. While Prados et al. (2014) search for the most similar patch throughout the whole brain, we show here that inpainting results plateau after at a certain search area distance radius. We believe that the NABT intensities might not be similar across the whole brain, thus limiting the search to a certain distance from the lesion seems adequate. Furthermore, this limited search area reduces the computational burden in comparison to searching over the whole brain area.

On the clinical aspect experiments, the estimated sample size required with SIENA in the current study was smaller (130 with a 30% treatment effect on the original version of SIENA) than previously reported by Anderson et al. (2007) (191 with a 30% treatment effect). These differences could be explained by different factors such as the RRMS population treatment, difference of power (90% for the later) and the study design.

Future work will focus on combining automatic lesion segmentation (Guizard, Coupé et al. 2015) with the proposed inpainting approach to provide a fully automatic approach. We plan to assess the impact of lesion inpainting in the context of longitudinal non-linear registration and diffusion weighted imaging in order to assess the focal atrophy in the surrounding of the lesion without the confounds due to the presence of lesions.

4.7. Conclusion

We developed a technique to replace tissues of interest, such as MS lesion, with healthy appearing tissues in order to perform cross-sectional and longitudinal image analyses. The method is robust and can improve the statistical power of detecting brain atrophy in MS. Furthermore, the proposed approach does not require any other image pre-processing than the lesion masking.

Acknowledgements: NeuroRx Research for providing the expert manual segmentations, Canadian Institutes of Health Research (MOP-111169 & 84360), les Fonds de Recherche Santé Québec and the MS Society of Canada (BioMed PhD studentship, 691 and an Operating grant to DLC). This study has been carried out with financial support from the French State, managed by the French National Research Agency (ANR) in the frame of the Investments for the future Programme IdEx Bordeaux (ANR-10-IDEX-03-02), Cluster of excellence CPU and TRAIL (HR-DTI ANR-10-LABX-57).

CHAPTER 5 NON-LINEAR REGISTRATION SYMMETRY

Preface

Over the past decade, longitudinal image data has become more and more available, where multiple scans are collected over a period of time on the same subject. Longitudinal design presents the advantages of reducing the confounding effects of inter-individual morphological variability by using each subject as his or her own control. Subtle changes such as longitudinal atrophy require specific non-linear registration approach mainly due to the fact that traditional non-linear registration algorithms have focussed on cross-sectional registration. Indeed, it is important to obtain robust and unbiased measures of the atrophy that can identify subtle brain changes over time. The symmetry or inverse consistency of the registration is very important for longitudinal and cross-sectional subject registration, in order not to bias a specific deformation direction.

Chapter 5 appeared as a conference proceeding of *MICCAI Spatio-Temporal Image Analysis (STIA) workshop* in 2010.

Impact of Non-Linear Registration Symmetry in Longitudinal MRI Studies

Nicolas Guizard, Vladimir Fonov, Bérengère Aubert-Broche, D. Louis Collins and the Alzheimer's Disease Neuroimaging Initiative[5]

5.1. Abstract

The evaluation of brain atrophies in neuroimaging studies is important, especially for neurodegenerative diseases such as Multiple Sclerosis, Alzheimer's and Parkinson's where brain volume changes over time must be assessed. Deformation and tensor-based morphometry are popular methods to evaluate brain difference cross-sectionally and/or longitudinally in a voxel-wise manner. Both of these methods depend highly on the quality of the non-linear registration between pairs of images. The deformation field obtained may present error and/or inaccuracy which affect the detection of these changes. With this in mind, we propose a new approach to constrain the symmetry of the non-linear registration. The performances of the proposed approach and popular techniques were evaluated on cross-sectional (LPBA40) and longitudinal databases (ADNI). The results show the importance of symmetry in longitudinal MRI study and promising results in the prognosis of dementia.

---

[5] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. ADNI investigators include (complete listing available at http://www.loni.ucla.edu/ADNI/Data/ADNI Authorship List.pdf).

5.2. Introduction

Multiple sclerosis (MS), Alzheimer's Disease (AD) and Parkinson's Disease (PD) are neurodegenerative diseases that affect millions of people around the world. Magnetic resonance imaging (MRI) is a powerful tool, which facilitates diagnosis and prognosis. Longitudinal MRI studies have the potential to assess and help in understanding disease progression. Recent image processing tools, such as Deformation-based morphometry (DBM), look at the local deformation obtained by registering each subject to a reference or to the same subject over time. DBM analyses can be performed to assess population differences (Shen 2003) or longitudinal brain changes within an individual subject or compare to a population reference (Leporé, Brun et al. 2007). In large clinical population trials, longitudinal DBM can be used to evaluate the disease onset or progression.

Because of its internal design, DBM creates deformation maps that depend on the quality of the non-linear registration that might limit the sensitivity and accuracy (Christensen and Johnson 2001, Yushkevich, Avants et al. 2010). Today, many non-linear or non-rigid registration methods are available, mostly having a similar structure: a similarity measure to find how well the two images correspond, a transformation model to define how the image can be modified and an optimization algorithm which maximizes the similarity of the two images by changing the parameters of the transformation model. The symmetry or inverse consistency of the registration is very important for longitudinal and cross-sectional subject registration, in order not to bias a specific deformation direction. Christensen et al. (Christensen and Johnson 2001) suggested consistent image registrations and compared their approach qualitatively with other techniques (Thirion 1998). Symmetric non-linear registration has been widely investigated in medical imaging (Leow, Huang et al. 2005, Rogelj and Kovacic 2006, Yanovsky, Leow et al. 2009).

In this article, we propose a strategy to force symmetry within the publicly available ANIMAL method and this approach was compared with a symmetric optimization scheme technique (Avants, Epstein et al. 2008). For the evaluation, two different publicly available

databases were used: 1) the LPBA40 to evaluate non-linear registration quality and 2) the ADNI database to evaluate longitudinal non-linear bias. The metrics used to assess the methods were: region overlap metric (OM), intensity variance metric (IVM), inverse consistency (IC) and the transitivity (TM). Finally with the same ADNI cohort, a DBM analyses was performed to evaluate the power of the algorithms in the prediction of brain changes over time.

5.3. Materials and Methods

5.3.1. Unbiased symmetric non-linear registration

Throughout this article, the following notation will be used: in the domain Ω of 3, for two images I_i and $I_j \in \Omega$, the transformation T_{ij} is sought for each spatial coordinate x which matched both images such that $I_i \circ T_{ij}(x) = I_{ij}$ the interpolated image. The inverse of the deformation field is T_{ij}^{-1}. Symmetric optimization schemes such as (Leow, Huang et al. 2005) and (Avants, Epstein et al. 2008) minimize and regularize the forward and the backward transformation directly. Motivated by the work of (Guimond, Meunier et al. 2000, Joshi, Davis et al. 2004, Yanovsky, Leow et al. 2009), the registration was constrained to be symmetric by combining the forward and the backward registration in order to obtain $T_{ij} \approx T_{ji}^{-1}$. In the following equations, $S_{ij}(x)$, $T_{ij}(x)$ and $T_{ji}(x)$ are respectively the symmetric, the forward and the backward transformations of the image I_i to I_j. The symmetry of the registration was forced by combining both the forward and the backward registration separately, such that:

$$S_{ij}(x) = \frac{T_{ij}(x) + T_{ji}(x)^{-1}}{2}$$ Equation 5.1

If the registration is indeed symmetric, then the concatenation of the forward and backward transformations, $T_{ij}(T_{ji}(x))$, should be the "identity" transformation. The expression $T_{ij}(T_{ji}(x))$ will be simplified to $T_{ij}.T_{ji}$ in the remaining text. However, due to the optimization scheme and the regularization of the deformation fields, the "non-symmetry" of the registration may induce bias in the computation of the deformation. Our approach, inspired by (Thirion 1998) and (Christensen and Johnson 2001), adds the residual registration error $r(x)$, defined as $r_i(x) = T_{ij}(x).T_{ji}(x)$ to the forward and $r_j = T_{ji}.T_{ij}$ to the backward registration:

$$S_{ij}(x) = \frac{T_{ij}(x).r_i + \left(T_{ji}(x).r_j\right)^{-1}}{2}$$

$$S_{ij}(x) = \frac{T_{ij}(x).\left(T_{ji}(x).T_{ij}(x)\right) + (T_{ji}(x).\left(T_{ij}(x).T_{ji}(x)\right)^{-1}}{2}$$ Equation 5.2

To preserve the diffeomorphism which ensures the invertibility of the deformation field, the computation of the symmetric deformation S was done in the log-Euclidean space (Arsigny, Commowick et al. 2006).

The symmetrization described above is applicable to any non-linear registration algorithm, however it was implemented in our in-house algorithm, ANIMAL[6] (Collins, Neelin et al. 1994). ANIMAL uses a multi-scale vector deformation estimation with a normalized cross-correlation (CC) similarity measure. Local registration is achieved in a hierarchical manner with a Gaussian blurring. ANIMAL uses hierarchical iterations with different step sizes and blurring kernels. In

[6] Available at: http://www.bic.mni.mcgill.ca/software/

this article, ANIMALsym denoted the symmetrization of ANIMAL with the residual error correction. The symmetrization constraint was applied at each iteration of the hierarchical scheme, in addition to performing 4 supplementary iterations at each level to take into consideration the indirect regularization of the symmetrization.

Recently, Klein et al. (Klein, Andersson et al. 2009) compared 14 fully automatic non-linear algorithms. In this paper, the proposed algorithm was compared to the best symmetric optimization scheme from Klein's study: SyN[7] (Avants, Epstein et al. 2008). The SyN algorithm from Klein's study, with the identical configuration uses a bi-directional optimization scheme with a gradient descent, diffeomorphic forward and backward transformations and CC as a similarity measure. To allow comparisons, the following parameters for the registration were used (based on the supplementary section of Klein et al. (2009)):

- ANIMAL: "*minctracc -non-linear corrcoef -weight 1 -stifness 1 -similarity 0.3 -sub lattice 6 -iterations [20x20x20x20x20x10] -step [32x16x12x8x6x4x2] -lattice diam step∗3 step∗3 step∗3 -identity moving.mnc fixed.mnc transformation.xfm*" for each iteration, a respective blurring of 16x8x6x4x3x2x1mm was applied and "x" represents the successive iterations. Note the supplementary finest iteration performed at 2mm this is to address the problem from the Klein's paper where ANIMAL was used with a 4mm fit.

- SyN: "*ANTS 3 -m PR[fixed.mnc, moving.mnc, 1, 2] -o transformation.xfm -r Gauss[2,0] -t SyN[0.5] -i 30x99x11 -use-Histogram-Matching*"

The goal of this article was to compare ANIMAL and ANIMALsym with SyN to demonstrate that the ANIMAL-based techniques are at least as good as SyN when run with the appropriate parameters.

[7] Available at: http://picsl.upenn.edu/ANTS/index.php

5.3.2. Data used for the evaluation

LPBA40: The LPBA40 database[8] (Shattuck, Mirza et al. 2008) consists of 40 T1-weighted images of normal subjects with 56 manually delineated structures following the LPBA40 protocol[9] (Figure 5.1, A). The volumes were linearly registered to MNI Talairach-like MNI305 space (Evans, Collins et al. 1993). In order to compare with Klein et al. results, the preprocessed data were downloaded and the linear pair-wise registered data from the supplementary material available with the paper.

ADNI: For the longitudinal registration assessment, 60 subjects were selected from the ADNI database (Figure 5.1, B). The evaluation was limited to 3 time points (0, 12 and 24 months) as it is sufficient to evaluate the transitivity and inverse consistency of the algorithms. The T1-weighted images were acquired with a magnetization-prepared rapid acquisition with a gradient echo (MP-RAGE) sequence from different ADNI sites with a 1.5T MRI scanner with a resolution of 0.94x0.94x1.2 mm^3. The cohort was constituted of 20 normal controls (NC), 20 mild cognitive impaired (MCI) and 20 Alzheimer disease (AD) patients with their respective clinical information (visit, MCI conversion time to AD...). Each subject of the ADNI group went through the preprocessing pipeline which consisted of the following successive stages: 1) Intensity non-uniformity correction (Sled, Zijdenbos et al. 1998) 2) Image denoising (Coupé, Yger et al. 2008) 4) Brain masking (Smith 2002) 5) Intensity normalization (Nyúl and Udupa 1999) 6) Linear registration (Collins, Neelin et al. 1994) with 9 degrees of freedom (3 rotations, 3 translations and 3 shearing) of the month 12 and 24 to month 0 (baseline) image in native space 7) Computation of the linear and non-linear registration transformations (Collins, Neelin et al. 1994, Collins, Holmes et al. 1995) of each baseline to the stereotaxic space of ICBM152

[8] Available at: http://www.loni.ucla.edu/Atlases/LPBA40
[9] Available at: http://www.loni.ucla.edu/Protocols/LPBA40

(Mazziotta, Toga et al. 2001, Mazziotta, Toga et al. 2001) 8) Tissue classification (white matter, grey matter and cerebral spinal fluid) (Zijdenbos, Forghani et al. 1998)

Scan-Rescan: Because no anatomical brain changes should occur during the same day, 15 subjects have been scanned twice during the same day (Figure 5.1, C). The T1-weighted MRI images were acquired on a 1.5T SIEMENS MRI scanner with a 3D FLASH sequence and a resolution of 1x1x1mm. Following usual MRI study protocols, after the first set of image acquisition, the subject was taken out from the scanner before getting back in the rescan session.

5.3.3. Experiments

Cross-sectional registration evaluation: In order to compare with recent non-linear registration evaluation articles (Hellier, Barillot et al. 2003, Klein, Andersson et al. 2009), the LPBA40 database was used with similarity metrics similar to those used previously.

- Overlap Metric (*OM*): The different labeled structure overlap after the registration indicates the anatomical agreement of the registration. The overlap of the 56 segmented regions of the subject (or the moving image) (denoted *I*) and the target (or the fixed image) (denoted *F*). *OM* depends on the size of the structure and varies between 0-1 with 1 indicating a perfect match:

$$OM = \frac{label(I \cap F)}{F} \qquad \text{Equation 5.3}$$

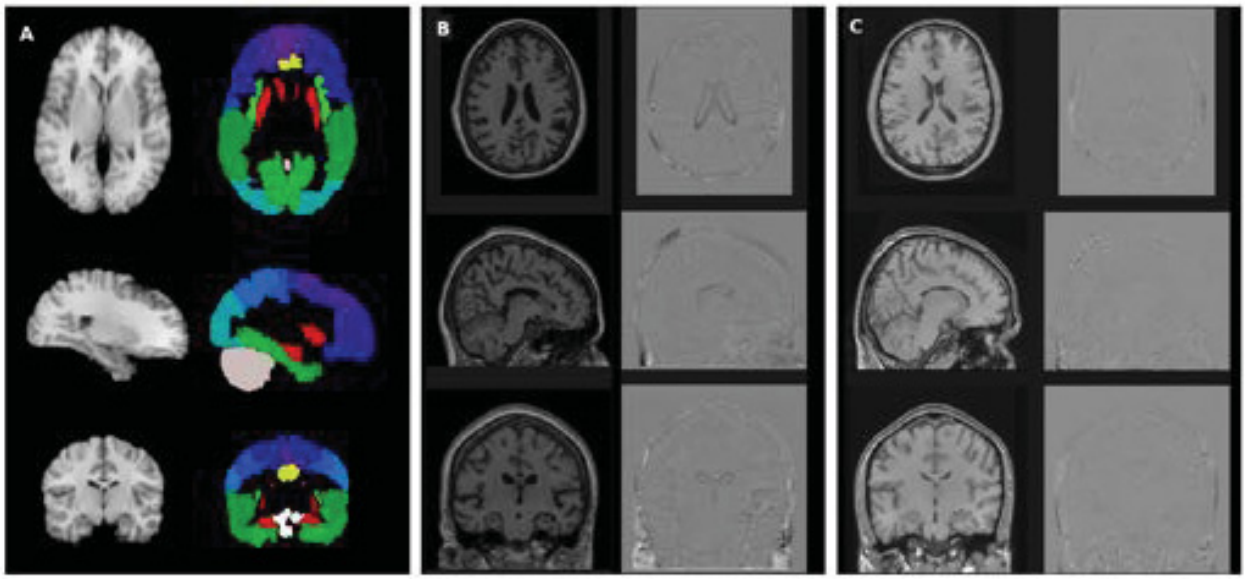


Figure 5.1: Example MRI subjects: LPBA40 subject (A), with the raw brain MR image (left) and the segmented brain (right), ADNI (B) subjects, with the raw brain MR image (left) and the subtracted 1 year scan to it (right) and scan-rescan subject (C), with the raw image on the left and the image difference of the scan and the following scan.

- Intensity Variance Metric (*IVM*): After registration the intensity difference between the source and the target should tend to 0. This metric evaluates the intensity variance of N registered subjects i to the target j. *IVM* depends on the intensity value of the images and tends to 0 if the match is perfect:

$$IVM = \frac{1}{(N-1)^2} \sum_{j=1, j\neq i}^{N} \sum_{i=1, i\neq j}^{N} \left(I_{ij}(x) - A_j(x) \right)^2 \qquad \text{Equation 5.4}$$

$$where \quad A_j = \frac{1}{N} \sum_{i=1}^{N} I_{ij}(x)$$

The same subject is used as a source and a target and thus makes the sample non-independent. To overcome this issue, a permutation test was performed (Menke and Martinez 2004, Fink, Klein et al. 2010). The significance of the results was evaluated with a Tukey-corrected p-value.

Longitudinal registration evaluation: In the context of longitudinal studies with the ADNI database, the inverse consistency (*IC*), the transitivity (*TM*) and the direct bias (*DB*) of the non-linear registration algorithms were evaluated. In the following section, l, *m* and *n* represent 3 successive time-points.

- Inverse Consistency Metric (*IC*): Regardless of the direction, non-linear algorithms should obtain similar results. To evaluate the symmetry of the registration (Christensen, Geng et al. 2006), the *IC* measures the residual distance of the concatenation of the forward T_{lm} and the backward T_{ml} transformations on a voxelwise level. In the case of perfect symmetry of the registration *IC* should equal 0:

$$IC = \sqrt{\left\| T_{lm}(T_{ml}(x)) - x \right\|^2} \qquad \text{Equation 5.5}$$

- Transitivity Metric (*TM*): In longitudinal studies with multiple time-points, an important aspect is the transitivity or the ability of the algorithm to combine the successive registrations. The *TM* is the difference of the concatenation of the forward registrations T_{lm} followed by T_{mn} and the direct transformation from the first time point to the last time points T_{ln} transformations of the same image on a voxel-wise level with regard to the displacement. *TM* tends to 0 if the combination of the transformation is not giving registration error:

$$TM = \sqrt{\left\| (T_{lm}.T_{mn}) - T_{ln} \right\|^2} \qquad \text{Equation 5.6}$$

- Direct Bias (*DB*): To assess the tendency of the algorithms to privilege atrophy or growth, the local log-transformed Jacobian (*J*) (Chung 2003) was computed for each pair (*N*) of the scan-rescan database:

$$DB = \frac{1}{2N}\sum_{1}^{N} J_{lm}(x) + J_{ml}(x) \qquad \text{Equation 5.7}$$
$$where \quad J_{lm}(x) = \log(\det(\partial dT_{lm}(x) / \partial dx)$$

- DBM analyses: Medial temporal lobe atrophy in Alzheimer disease differs from other dementia diseases as it has been shown in histopathology and MRI studies (Braak and Braak 1991, Barkhof 1999). The evaluation of the registration overall quality as well as the inverse consistency of the non-linear registration algorithm set the basics of unbiased DBM analysis. With the same ADNI cohort, a longitudinal DBM analysis was performed on the 3 time points (0, 12 and 24 months). The local log-transformed Jacobian (Eq.) was computed for the deformation field *T* for each time point to the baseline (0 month) yielding m12 and m24 transformations. The means of each subject Jacobian were compared by means of paired *t*-tests. The Jacobian map was also computed in the white matter (WM) of the temporal lobes obtained with the tissue classification to estimate the WM atrophy in Alzeihmer's disease.

This regional DBM analyses were then used to assess the power of prognosis of each method for the MCI population in order to predict their conversion to AD after their first visit (m0). From our ADNI cohort of 20 MCI subjects, 8 subjects converted to AD after a period of time between 12 to 24 months.

5.4. Results

Cross-Sectional evaluation results

The cross-sectional evaluation results of the 1560 registrations with the LPBA40 database are represented in Figure 5.2. Regardless of the metric, all techniques had similar overall results. The paired t-test ANOVA did not detect any significant difference in *OM* between the three

techniques. An interesting finding is that the overall *OM* is consistent with the *OM* per region (Figure 5.3).

ANOVA showed a statistical difference between methods (p-value < 0.0001), with ANIMALsym giving the best results for IVM.

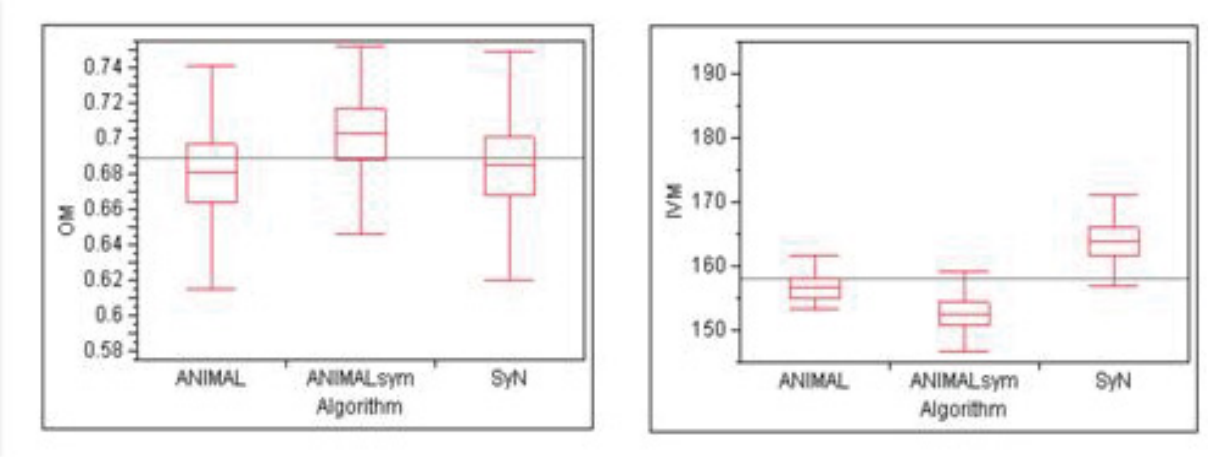


Figure 5.2: Overlap metric (OM) and intensity variance (IVM) for the different non-linear registration methods obtained with the LPBA40 database.

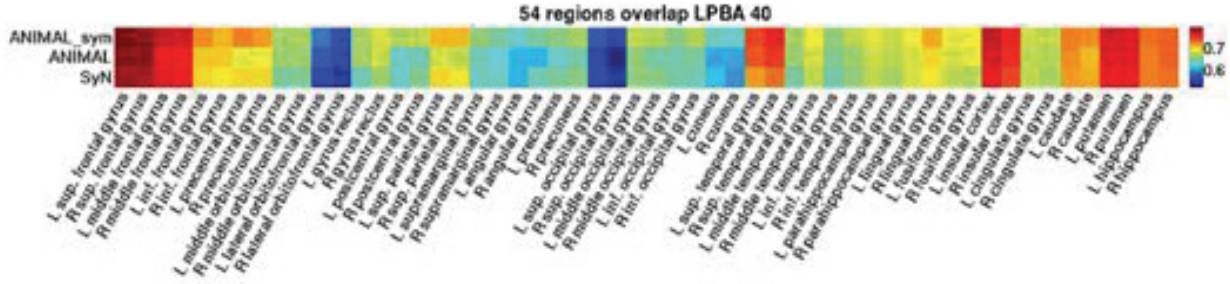


Figure 5.3: Overlap per brain region: this matrix represents the overlap of the 56 regions for SyN, ANIMAL and ANIMALsym. The color scale is from 0 (blue) to 1 (red).

Longitudinal evaluation results

The registration of the different time-point images from the ADNI database enabled the measurement of the symmetry of the algorithm (Figure 5.4). As expected the symmetrization improved the inverse consistency of the standard ANIMAL algorithm (p-value < 0.0001). With a mean IC of 0.154mm (±0.03), ANIMALsym achieved significantly better results than SyN (p-value < 0.0001).

Regardless of the symmetry, accurate registration should provide a better transitivity. ANIMAL and ANIMALsym yield significantly better transitivity of the registration compared to SyN (Figure 5.4), respectively: ANIMAL = 0.15(±0.05) (p-value < 0.0001) and ANIMALsym = 0.17(±0.05) (p-value < 0.00001).

The registration of the scan-rescan data demonstrated equal sensitivity of the 3 methods (Figure 5.5). All methods presented a mean *DB* close to 0 with a normal distribution of their log-transformed Jacobian.

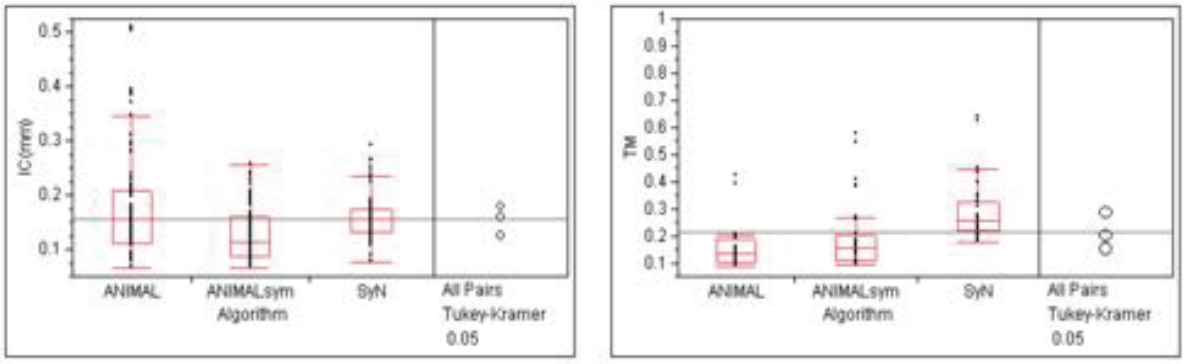


Figure 5.4: Inverse consistency (IC) and Transitivity metric (TM) results obtained for each non-linear algorithms with the ADNI database.

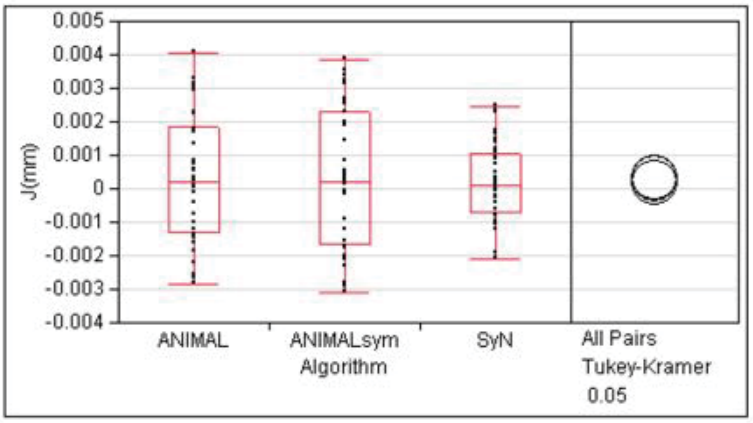


Figure 5.5: Direct bias of the scan-rescan dataset for each algorithm.

DBM analysis

The total brain log-transformed Jacobian of DBM analysis across each population and each method for the ADNI database is shown in (Figure 5.6). This figure highlights qualitatively the inherent difference of the deformation obtained. Compared to the other algorithms, ANIMALsym estimated smaller regional atrophy or growth (i.e. cortical). With regards to the atrophy in the medial temporal lobes (Figure 5.7), all the algorithms showed stable atrophy for the control subjects. ANIMAL, ANIMALsym and SyN detect an increase of atrophy for the MCI and the AD at m12 and m24. Despite the large distribution of the log-transformed Jacobian or ANIMALsym, the match-pairs t-test for the AD and MCI subject with ANIMAL ($t=4.2$, $p<0.001$) and ANIMALsym ($t=2.0, p<0.01$) are significant. The preliminary results of the log-transformed Jacobian of the MCI and the MCI converted to AD measured in the temporal lobes (Figure 5.8) showed the highest mean difference with ANIMALsym ($p < 0.01$).

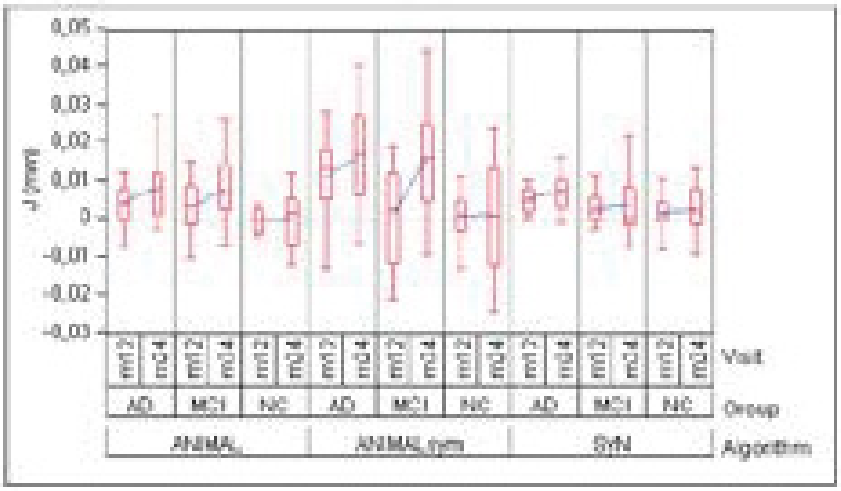


Figure 5.6: DBM analyses of the ADNI database: Representation of the log-transformed Jacobian (> 0 atrophy) for each registration of the subject follow-up visit to the baseline visit (m12 to m0 and m24 to m0).

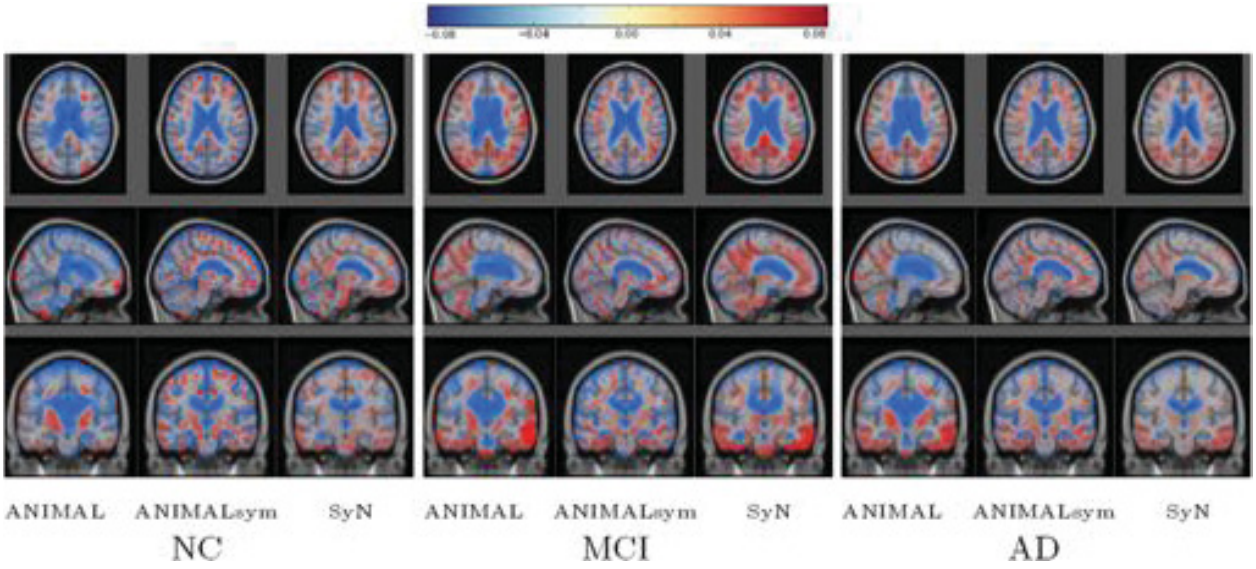


Figure 5.7: DBM analyses log-transformed Jacobian determinant average per population (NC, MCI, AD) at m12 visit registered to the baseline visit (m0): The DBM averaged results with each algorithms are represented in red represents atrophy and blue represents expansion overlaid on the ICBM152 template.

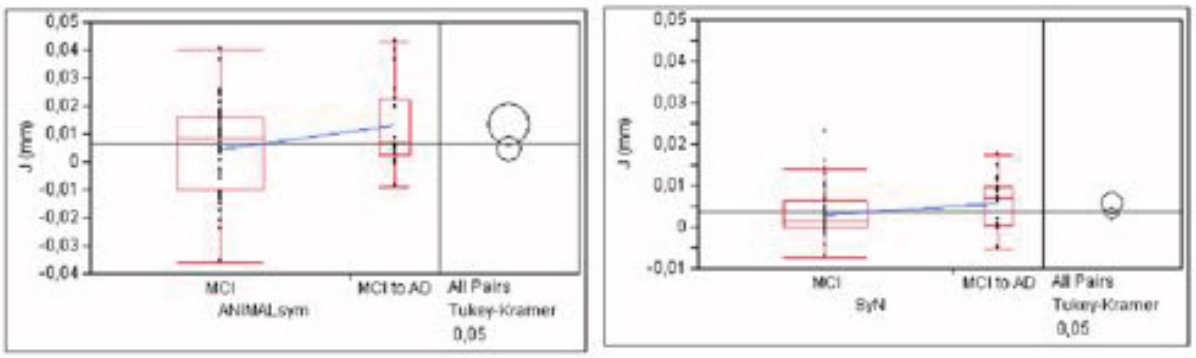


Figure 5.8: Converting and non-converting MCI DBM comparisons: Representation of the mean log-transformed Jacobian between the converted and non-converted MCI to AD.

5.5. Discussion

The main contributions of this article are the evaluation of different non-linear registration techniques and their accuracy/sensitivity to detect brain atrophy in the context of a longitudinal study.

In this article, three non-linear registration methods were compared, SyN, ANIMAL and its proposed symmetric version, ANIMALsym. The evaluation of the methods was performed on three different databases. To evaluate the quality of the overlap and the image matching, the LPBA40 database was used with its 56 manual labeled regions. Then, for the evaluation of the behavior of the algorithms in the context of longitudinal MRI study, 60 subjects were chosen from the ADNI database with 3 different time-points and 15 subjects with their scan-rescan images. Finally, a DBM analyses was performed on the ADNI subjects (20 NC, 20 MCI and 20 AD subjects) and assessed the ability of the registration method to predict the conversion of MCI subject to AD.

The region overlap is not high, probably because of manual labelling but they are consistent with the findings of Klein et al. (Klein, Andersson et al. 2009) for SyN. ANIMAL performs better with a finer iteration step at 2mm. The intensity variance reinforces the overlap results.

The cross-sectional results are interesting in evaluating the quality of the registration but do not insure the quality of the registration algorithm in the context of longitudinal registration. Therefore, the inverse consistency and the transitivity of the algorithm were evaluated and showed that the symmetrization constrain of ANIMAL does insure symmetry as well as the symmetric optimization scheme offered by SyN. The scan-rescan dataset showed that neither of the symmetric methods have a significant bias, which would produce false atrophy/growth discovery.

The quality and the inverse consistency of ANIMALsym and SyN, set the basis to perform a DBM analyses on the ADNI groups. Visual inspection of the DBM maps showed anatomical coherence of the atrophies and growths. In the temporal lobes, both ANIMAL and ANIMALsym detected atrophy progression for the MCI and the AD subjects between the m12 and m24 which are coherent with previous findings (Burton, Barber et al. 2009). The small number of subjects for the prediction of converting subject from MCI to AD limits our results, however they shown interesting preliminary results with ANIMALsym.

5.6. Conclusion

In longitudinal neuroimaging studies, symmetric registrations are important to evaluate the brain deformations over time. Non-symmetric registration results in error, bias and wrong estimation of brain atrophy and enlarging. This article compared the symmetry of well-established non-linear registration symmetric and non-symmetric method with constrained symmetrization strategies.

The constrained symmetrization, using bi-directional registration to force the symmetry of the transformations results in overall better results than the symmetric optimization algorithm. Subsequently, a longitudinal DBM analysis was performed to evaluate the sensitivity of our proposed method and obtained significant atrophy detection in Alzheimer's, mild cognitive

impaired and normal control subjects in the medial temporal lobes. With more investigation, the symmetric registration framework proposed in this article could be used as a prognosis tool to predict dementia or other pathology.

CHAPTER 6 LONGITUDINAL NON-LINEAR REGISTRATION

Preface

For longitudinal studies, numerous groups have investigated paired images and different approaches have been developed to evaluate the brain changes. By taking into consideration the temporal dimension of the data, the analyses of the anatomical morphology will be less sensitive to spatio-temporal image acquisition variability (patient repositioning, MRI sequence, scanner hardware, image resolution…), as well as punctual anatomical differences due to lesions or inflammation.

Chapter 6 was published in (Guizard, Fonov et al. 2015). Earlier forms of the work have been published in conference proceedings at the *MICCAI Spatio-Temporal Image Analysis (STIA) workshop* in 2012 where it was accepted for an oral presentation. A variant of this work, where we applied the proposed method to a longitudinal RRMS cohort was accepted for poster presentation at the *ISMRM White Matter Study Group* in 2013.

Spatio-temporal regularization for longitudinal registration to an unbiased 3D individual template

Nicolas Guizard, Vladimir S. Fonov, Daniel García-Lorenzo, Kunio Nakamura, Bérengère Aubert-Broche and D. Louis Collins

6.1. Abstract

Neurodegenerative diseases such as Alzheimer's disease present subtle anatomical brain changes before the appearance of clinical symptoms. Manual structure segmentation is long and tedious and although automatic methods exist, they are often performed in a cross-sectional manner where each time-point is analyzed independently. With such analysis methods, bias, error and longitudinal noise may be introduced. Noise due to MR scanners and other physiological effects may also introduce variability in the measurement. We propose to use 4D non-linear registration with spatio-temporal regularization to correct for potential longitudinal inconsistencies in the context of structure segmentation. The major contribution of this article is the use of *individual template creation with spatio-temporal regularization of the deformation fields* for each subject. We validate our method with different sets of real MRI data, compare it to available longitudinal methods such as FreeSurfer, SPM12, QUARC, TBM, and KNBSI, and demonstrate that *spatially local temporal regularization* yields more consistent rates of change of *global structures* resulting in better statistical power to detect significant changes over time and between populations.

6.2. Introduction

Longitudinal measures of brain volumetry are powerful tools to assess the anatomical changes underlying on-going neurodegenerative processes. In different neurological disorders, such as multiple sclerosis (MS), Alzheimer's disease (AD) and Parkinson's disease (PD), brain atrophy has been shown to be good surrogate marker of disease progression (Chard, Brex et al. 2003, Burton, McKeith et al. 2004, Ridha, Barnes et al. 2006). Magnetic resonance imaging (MRI) can provide reproducible 3D structural images of the brain, which can be used to assess its integrity. Furthermore, the emergence of freely available longitudinal MRI databases, (e.g., Alzheimer's Disease Neuroimaging Initiative (ADNI) (Mueller, Weiner et al. 2005), Open Access Series of Imaging Studies (OASIS) (Marcus, Fotenos et al. 2009) and others) provide the necessary data to develop and test new methods and investigate the longitudinal structural changes of healthy and pathological brains.

Image processing in MRI-based neuro-anatomical studies is often performed in a cross-sectional manner where each time-point is evaluated independently. Typically, brain morphometry comparisons can be done by matching paired images (template-to-subject or subject-to-subject), where the deformation field is used to map atlas regions or to compute voxel-wise comparisons of anatomical changes as in deformation-based morphometry (DBM). However, in the context of longitudinal datasets, the robust estimation of anatomical changes is still challenging (Thompson and Holland 2011). Indeed, in the case of neurodegeneration occurring in a short period of time (2-3 years), if we assume that longitudinal changes are smoothly varying, spatially local, and temporally monotonic processes, considering individual time-points independently can generate unnecessarily noisy longitudinal measurements due to the intrinsic noise associated with each visit. Different studies have shown the impact of the MRI acquisition protocol on structural measurements (Caramanos, Fonov et al. 2010) and cortical thickness (Fonov, Coupé et al. 2012). Therefore, methods that integrate constraints from the temporal dimension (i.e., 4D methods) should produce more accurate, robust and stable measures of the longitudinal anatomical changes resulting in a more realistic estimation of temporal

evolution. Different approaches have been proposed to overcome the complexity of anatomical 4D longitudinal data image analysis. We classify these methods in 2 major groups: "4D" and "longitudinal 3D". The 4D approaches treat the individual and/or group-wise longitudinal data as an ensemble and provide longitudinal models or measurements. They are mathematically sophisticated approaches that have been proposed in the context of modeling larger anatomical changes over time (i.e., growth over the span of childhood). For example, a 4D population model creation using Gaussian kernel regression has been suggested by Davis et al. (2007) where each image is registered independently to a moving average, avoiding the creation of an explicit parameterized model of the longitudinal changes (Figure 6.1a). Kernel regression has also been used in the framework of the Large Deformation Diffeomorphic Metric Mapping (LDDMM) for brain shapes (Durrleman, Pennec et al. 2009) (Figure 6.1b) and images (Durrleman, Pennec et al. 2009, Hart, Shi et al. 2010, Sadeghi, Prastawa et al. 2010). Regarding intra-subject 4D registration, Lorenzi et al. (2010) have proposed 4D non-linear registration via a global 4D deformation optimization scheme in the Demons registration framework. Finally, Wu et al. (2011) introduced an implicit mean-shape of the population which could be used for individuals. Their approach maximizes the spatio-temporal correspondence and continuity from a set of temporal fibre bundles (Figure 6.1c).

The longitudinal 3D approaches include the adaptation of popular 3D/cross-sectional methods with some longitudinal constraints or longitudinal pre-processing. For instance, in the context of clinical evaluation over a few years where anatomical changes are small and continuous, the use of 3D individual template targets have been proposed to perform non-linear registration (Kraemer and Thiemann 1987, Ashburner and Ridgway 2012, Reuter, Schmansky et al. 2012) or tensor-based analyses (TBM) (Hua, Hibar et al. 2013). Indeed, to compare anatomical differences, 3D population templates have proven their importance for different applications such as mapping function, structure, or vasculature (Thompson and Toga 2002) and group comparisons (Ashburner, Hutton et al. 1998). While different techniques exist to create unbiased population templates for multi-subject cross-sectional studies (Guimond, Meunier et al. 1998, Fonov, Evans et al. 2011), few of these techniques have been developed for the creation of

an individual 3D subject template. More recently, Reuter et al. (2012) created a subject-specific 3D template for longitudinal analysis by computing the median image of the linearly registered images of the same subject from different time-points and this method is implemented within the longitudinal version of FreeSurfer[10] (Reuter, Schmansky et al. 2012). In the continuity of their work on voxel-based morphometry (VBM) (Ashburner and Friston 2000, Baron, Chételat et al. 2001, Chetelat, Landeau et al. 2005), Ashburner et al. (2012) presented an unbiased "group-wise intra-subject" template with an iterative longitudinal non-uniformity correction, linear and non-linear diffeormorphic registration that is implemented in the Statistical Parametric Mapping 12 (SPM12)[11]. Aubert-Broche et al. (2013) also proposed to use robust non-linear individual templates to perform tissue classification and segmentation of pediatric images.

Inspired by previous work and the need for longitudinal analysis, we propose to include spatio-temporal constraints to analyze longitudinal MRI volumes, combining the advantages of both 4D longitudinal and 3D longitudinal approaches. An iterative algorithm is presented to create subject-specific templates for structural segmentation (Figure 6.1d). The decomposition of the longitudinal deformation fields, similar to a Taylor series, enables local spatial constraints as well as temporal regularization. While the spatial constraints aim to preserve the anatomical consistency in the image, the voxel-wise temporal regularization tackles the potential longitudinal alteration of the images. The temporal regularization is achieved with a local voxel-wise linear regression of the deformation components over time, resulting in a more consistent global longitudinal deformation. In this article, we first evaluate the stability and robustness of our method with a scan-rescan dataset, then, we assess its power to analyze a longitudinal cohort from the ADNI database. We show that a weak local spatial constraint over time can have significant positive global effects to significantly reduce inter-visit variability in the measurement of structure volumes such as the lateral ventricles, hippocampi and brain parenchyma.

10 http://surfer.nmr.mgh.harvard.edu/
11 http://www.fil.ion.ucl.ac.uk/spm/

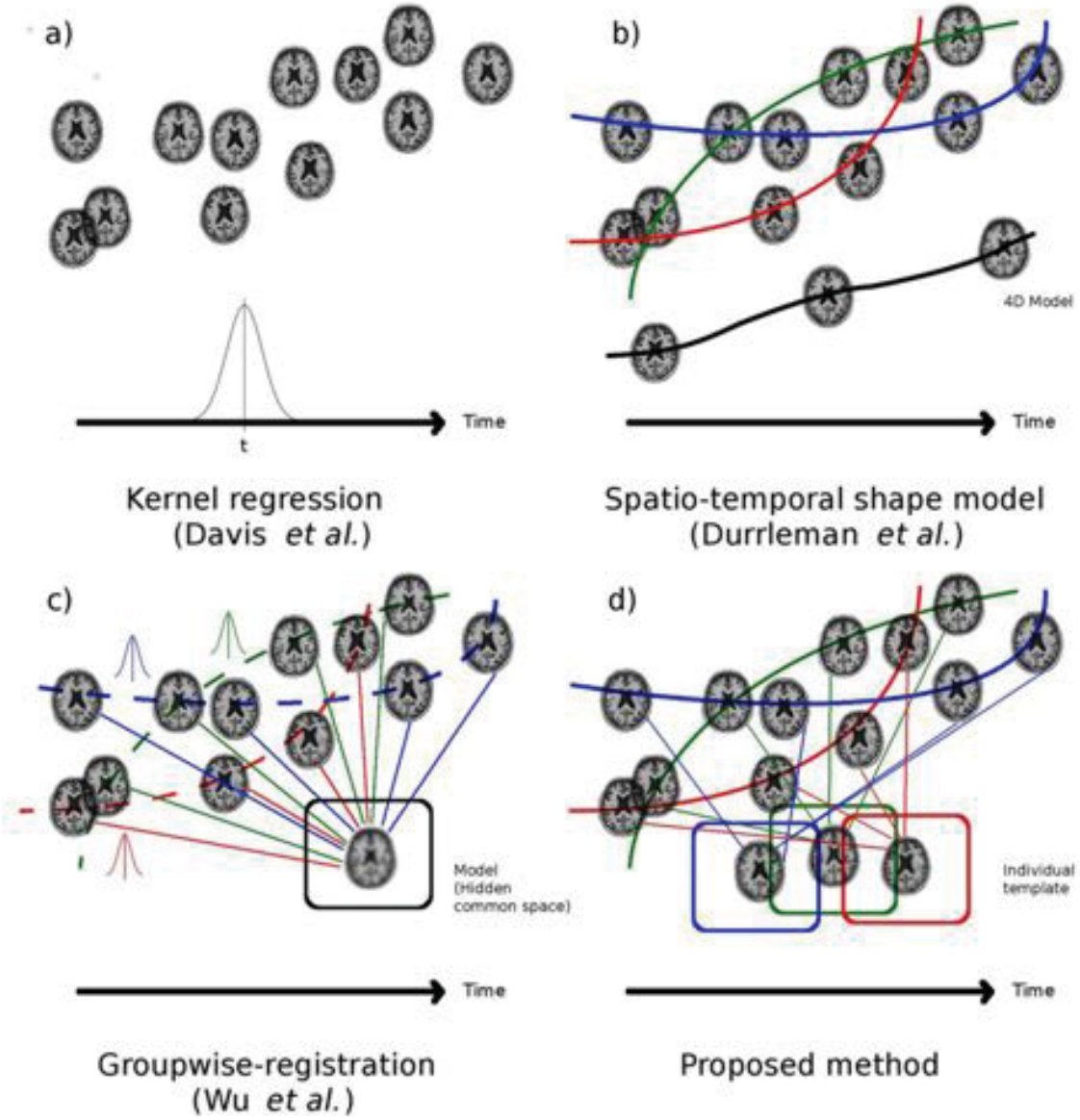


Figure 6.1: Longitudinal registration and template creation methods. Each vignette (a, b, c and d) represents different strategies proposed to overcome longitudinal MRI data analysis.

6.3. Methods

The objective of the template creation algorithm is to find the non-linear transformations that minimize the anatomical shape differences between images to create the most representative average of the subject's anatomy. Processing is achieved in two steps. First, all data is processed cross-sectionally to bring each volume into stereotaxic space. Second, this data is used to build a subject-specific individual template. The method and notation is inspired from Fonov et al. (2011) and Aubert-Broche et al. (2013), and described in the following sections. The nomenclature is summarized in Table 6.1.

Table 6.1: Notation

Symbol	Definition
v	Voxel position v varying from 0 to N
k	Iteration k
$I_t(v)$	Set of images for subject I from different time-points t
$\Phi_I^k(v)$	Individual template of subject I at voxel v
$\psi_{t,\Phi}^k(v)$	Deformation field of time-point t to template Φ at voxel v
$\Phi_{t,\Phi}^k(v)$	Bias free deformation field of time-point t to template Φ at voxel v
Ω_v	Neighborhood or patch surrounding voxel v
$T(v,t)$	Trajectory of voxel v at time t
$\Im T(v,t)$	Jacobian matrix of voxel v at time t
$\beta_t(v)$	Non-uniformity field at voxel v

6.3.1. Cross-sectional pre-processing

All MRI data are pre-processed to reduce the effects of artifacts and noise. The standard deviation of the MRI Rician noise is estimated automatically and image redundancy is used to reduce the noise using a non-local patch-based technique (Coupé, Manjón et al. 2010). A non-parametric estimation of the slow varying non-uniformity field corrects the intensity inhomogeneity produced by scanner radio-frequency coil variations (Sled, Zijdenbos et al. 1998). In addition, linear histogram matching is performed between each subject and a reference image to normalize the image intensities between subjects/scans to a range between 0.0 and 100.0. The reference image was created to represent the ageing population brain anatomy from the AD cohort using the unbiased template creation approaches proposed by Fonov et al. (2011). Finally, to correct for variation in head position, orientation and size, an initial 9 parameter linear registration (translation, rotation and scale) is computed to bring each subject into the ICBM152 template stereotaxic space (Collins, Neelin et al. 1994).

6.3.2. Longitudinal processing

The subject-specific template is based on the work of Guimond et al. (1998, 2001), Joshi et al. (2004) and Fonov et al. (2011) where a template is created in two steps, first using linear registration and second, using non-linear registration with a spatio-temporal regularization (section 2.2.3).

Linear individual template

In order to refine the alignment of individual images and estimate global whole brain scale factors between the consecutive visits, we perform a hierarchical iterative linear registration. Starting with the individual stereotaxic image average as the initial target, the linear individual

template ($\Phi_L(v)$) is then defined as the intensity average of the B-spline (order 4) interpolated individual visit scans after affine registration. For each subject, a twelve parameter affine registration (Collins, Neelin et al. 1994), based on an intensity cross-correlation similarity measure, is performed between the time-points' (t=[0..n]) and the subject-specific template volumes at 32, 16, 8 and 4mm hierarchical step sizes.

Non-linear minimum deformation individual template

A non-linear subject-specific template $\Phi_{NL}(v)$ is estimated with an iterative approach, similar to the linear template, but using non-linear registration in order to estimate the local deformation between the visits and the individual template. To create $\Phi_{NL}(v)$, a minimum deformation template (MDT) approach is used as described by Fonov et al. (2011). However, here the MDT estimation is modified to account for spatio-temporal regularity constraints (described in 2.2.3) and the implementation of the 4D constraints is done in the framework of a 3D non-parametric vector field estimator using the Automatic Non-linear Image Matching and Anatomical Labeling (ANIMAL) procedure (Collins and Evans 1997).

For the MDT, ANIMAL estimates the non-linear deformation field required to align two image volumes in a hierarchical manner, where the algorithm maximizes the local cross-correlation of the blurred image intensity of the source image with the equivalently blurred image intensity of a target image. Starting from down-sampled images, the displacement vectors that best match the two images are stored at the nodes of a 3D grid, producing a dense deformation field. Then, the deformation field is upsampled and used to initiate the deformation at the next hierarchical iteration where the blurring kernel is reduced, and the deformation field is refined. Details of the ANIMAL algorithm are described in (Collins, Holmes et al. 1995, Collins and Evans 1997).

To satisfy the intensity constraint condition (eq.1) and the deformation constraint condition of (eq.2), we use an iterative approach. At each iteration, ANIMAL is used to map the voxels v

from the MRI of a subject at time-point t=[0..n], $I_t(v)$, to the current evolving estimate of the template $\Phi_{NL}^k(v)$ at iteration *k* through a deformation transformation $\psi_{t,\Phi}^k(v)$. This is followed by the removal of the bias (or mean deformation $\sum_{t=0}^{n}\psi_{t,\Phi}^k(v)$) to obtain $\phi_{t,\Phi}^k(v)$(eq. 3) (thus enforcing the condition in eq.2) and calculating a new estimate of the template $\Phi_{NL}^{k+1}(v)$ (eq. 4).

$$\Phi_{NL}^k(v) = \underset{\Phi}{\operatorname{argmin}} \sum_{t=0}^{n} \int_{volume} \left(\Phi_I^k(v) - I_t\left(\psi_{t,\Phi}^k(v)\right)\right)^2 dv \qquad \text{Equation 6.1}$$

$$\Phi_{NL}^k(v) = \underset{\Phi}{\operatorname{argmin}} \sum_{t=0}^{n} \int_{volume} \left|\overline{\psi_{t,\Phi}^k(v)}\right|^2 dv \qquad \text{Equation 6.2}$$

$$\varphi_{t,\Phi}^k(v) = \psi_{t,\Phi}^k(v) \circ \overline{\sum_{t=0}^{n}\psi_{t,\Phi}^k(v)} \qquad \text{Equation 6.3}$$

$$\Phi_{NL}^{k+1}(v) = \frac{1}{n}\sum_{t=0}^{n} I_t\left(\varphi_{t,\Phi}^k(v)\right) \qquad \text{Equation 6.4}$$

In these equations, the operation $\circ$ denotes concatenation of transformations, and $\overline{X}$ denotes inversion of a transformation *X*.

The algorithm is initialized with the individual linear template ($\Phi_L(v)$). At each iteration *k*, $\psi_{t,\Phi}$ is the non-linear transformation required to map I_t to Φ_{NL}^k which was obtained using ANIMAL. It is spatially constrained with a linear elastic body model while it minimizes the intensity difference of the paired images (i.e., between template and time-point images). The linear elastic body constraints are justified in such intra-subject registrations where very large

deformations are not expected. The parameters of the hierarchical non-linear registration are chosen to ensure that the transformation defined by the vector field is smooth, bijective and invertible (Fonov, Evans et al. 2009). The details of the iterative hierarchical schedule and the non-linear registration parameters for the 3D grid step size, image blurring kernel and similarity measure neighborhood size are summarized in Table 6.2. The registration schedule parameters are similar to Fonov et al. (Fonov, Evans et al. 2009) and ANIMAL is robust to changes in parameters by a factor of 2 (Collins and Evans 1997, Chakravarty, Bertrand et al. 2006).

This subject-specific template creation process yields the non-linear deformations to map each of the subject time-points toward the template. By concatenating a forward transformation to the template and the inverse transformation toward a specific time-point, we can obtain the total non-linear transformation between two time-points transitively.

Table 6.2: ANIMAL non-linear registration schedule. For each iteration, we define a step size as the distance between control nodes for the free-form deformation recovered. The blurring kernel is the size of the full-width-half-maximum of the Gaussian kernel used to blur the source and target data. The local correlation which defines the local similarity is estimated in the neighborhood of diameter equals to the neighborhood size parameter.

Iteration	**Step size (mm)**	**Blurring kernel (mm)**	**Neighborhood size (mm)**
1	16	8	48
2-3	8	4	24
4-5	4	2	12
6-7	2	1	6
8-9	1	1	6

The MDT algorithm described above is modified to include an additional constraint for the non-linear transformations between time-points. It is implemented as an additional regularization

step which is performed at each iteration of the template creation in the spatio-temporal domain in order to obtain a smooth non-linear deformation over time, since we expect the anatomical changes to happen in a slow and continuous fashion. We replace the individual time-point non-linear registrations $\psi^{k}_{t,\Phi}$ with a continuous and smooth transformation field $T(v,t)=[\psi(v)^{k}_{t_0,\phi},\ldots,\psi(v)^{k}_{t_n,\phi}]$ where *T(v,t)* can be seen as the trajectory of voxel position *v* over time *t*. The proposed spatio-temporal regularization of the longitudinal deformation field is achieved through the following steps:

First, we decompose the longitudinal deformation component of the transformation into a simplified Taylor series expansion of order 1 in space, where the higher order terms are neglected, which allows for spatial regularization, such as:

$$T(v+\Delta v,t)\approx T(v,t)+\Im T(v,t)\cdot\Delta v \qquad \text{Equation 6.5}$$

This Taylor expansion presents the advantage of accounting for the longitudinal deformation (or temporal trajectory, $T(v,t)$) as well as the longitudinal local variation (Jacobian matrix, $\Im T(v,t)$).

Second, we want to regularize the trajectory ($T(v,t)$) to obtain smooth longitudinal deformations while preserving the longitudinal local variation of the Jacobian matrix ($\Im T(v,t)$), such as:

$$\Im T(v,t) = \begin{bmatrix} \frac{\partial T(v,t)_1}{\partial v_1} & \cdots & \frac{\partial T(v,t)_1}{\partial v_3} \\ \vdots & & \vdots \\ \frac{\partial T(v,t)_3}{\partial v_1} & \cdots & \frac{\partial T(v,t)_3}{\partial v_3} \end{bmatrix} \approx$$

$$\begin{bmatrix} \frac{T(v+\Delta v,t)_1 - T(v-\Delta v,t)_1}{2\cdot\Delta v_1} & \cdots & \frac{T(v+\Delta v,t)_1 - T(v-\Delta v,t)_1}{2\cdot\Delta v_3} \\ \vdots & & \vdots \\ \frac{T(v+\Delta v,t)_3 - T(v-\Delta v,t)_3}{2\cdot\Delta v_1} & \cdots & \frac{T(v+\Delta v,t)_3 - T(v-\Delta v,t)_3}{2\cdot\Delta v_3} \end{bmatrix} \quad \text{Equation 6.6}$$

To preserve the spatial consistency, we approximate the Jacobian matrix $\Im T(v,t)$ from equation 6, by averaging across finite differences, such as:

$$T(v+\Delta v,t) \approx T(v,t) + \frac{1}{|\Omega_v|}\sum_{u\in\Omega_v} \Im T(u,t)\cdot\Delta v \quad \text{Equation 6.7}$$

where Ω_v is the local neighbourhood centered on *v*. Thus, this approximation provides a spatially regularized longitudinal deformation and in our experiments, we found that a 3x3x3 local neighbourhood was a good comprise between spatial smoothing and computational time.

Simultaneously, we perform linear regression of the zeroth order term in eq. 5 in the temporal domain such as:

$$T(v,t) \approx T_0(v) + T_1(v)\cdot t \quad \text{Equation 6.8}$$

where $T_0(v)$ is the intercept and $T_1(v)$ is the slope of the linear regression.

Thus, we effectively perform spatio-temporal regularization of the set of deformations fields with a spatial regularization (eq. 6) and a temporal regression (eq. 8), such as:

$$T^*(v,t) = T_0(v) + T_1(v)\cdot t + \frac{1}{|\Omega_v|}\sum_{u\in\Omega_v} \Im T(u,t)\cdot \Delta v \qquad \text{Equation 6.9}$$

We use the resulting regularization procedure instead of eq. 6.3 in the MDT template creation.

This approach presents the advantage of taking into consideration the longitudinal deformation at each voxel and at the local neighbourhood level by the means of the local Jacobian matrix and the explicit local voxel-wise regularization of the deformation field components.

Individual template-based bias field correction

Intensity non-uniformity may vary between longitudinal scans due to differences in field inhomogeneity (B_1) and receiver coil sensitivity (Liang and Lauterbur 2000) as well as differences in the positioning of the subject inside the coil. As described by Holland et al. (2011) as well as Ashburner and Ridgway (2012), if uncorrected, these temporal intensity non-uniformities could be detected as atrophy or growth with intensity-based non-linear registration tools. Therefore, inspired by the differential intensity inhomogeneity correction proposed by Lewis et al. (2004), we propose to use the intensity difference of the subject-specific template and the warped time-point image to estimate the smooth longitudinal inhomogeneity correction field with N_3 (Sled, Zijdenbos et al. 1998). N_3 iteratively sharpens the histogram of the image intensity difference by de-convolving Gaussian fields from the true signal, while using splines to represent the estimated bias field. During the iterative process of the individual template creation and after the spatio-temporal regularization, the image intensity difference of the subject visit (I_t) and the current template ($\Phi_{NL}(v)$) is computed at each iteration after resampling I_t with

the transformation $\psi^{k}_{t,\Phi}$. The bias field for each visit (β_t^k) is estimated from the differential image (eq. 10 and 11).

$$\alpha_t^k = N_3\left(I_t^k\left(\psi^k_{t,\Phi}(v_i)\right) - \Phi_I^k(v_i)\right) \qquad \text{Equation 6.10}$$

$$\beta_t^k = \alpha_t^k \Big/ \exp\left(\frac{1}{n}\sum_{t=0}^{n}\log\left(\alpha_t^k\right)\right) \qquad \text{Equation 6.11}$$

Then the bias field is transformed back into the native time-point space to correct the residual longitudinal inhomogeneity of the source images for the following iteration (eq. 11).

$$I_t^{k+1}(v) = I_t^k(v)\cdot\beta_t^k\left(\overline{\psi^k{}_{t,\Phi}(v_i)}\right) \qquad \text{Equation 6.12}$$

Optimization and convergence

The non-linear template creation optimization is done at 5 hierarchical levels, starting with deformations estimated every 16, 8, 4, 2 and finally 1mm and the corresponding non-linear registration parameters are summarized in Table 2. At each level, the regularizations are performed consecutively in the order of equations 2, 9 and 3. An initial spatial regularization is applied to the subject visit-template deformation with a Gaussian kernel while for the spatio-temporal regularization, the whole time series deformation set is constrained (eq. 9). In our implementation, different parameters of the spatio-temporal regularization can be adjusted. The neighborhood size of the Jacobian matrix computation can be increased to obtain smoother deformations.

In previous cross-sectional template creation studies, we found that 9 iterations are enough for the convergence of the iterative process at each hierarchical level (Fonov, Evans et al. 2011). In the case of individual template creation, the additional longitudinal regularization could slow down convergence but it is compensated by the anatomical similarity of the images being registered. We found in our experiments that 9 iterations are thus also sufficient to converge. The template, longitudinal non-uniformity correction and deformation fields estimated at one hierarchical level are all used to initialize the procedure at the next hierarchical level.

6.3.3. Experiments

Data

Two neuroimaging datasets were used anonymously to evaluate the proposed algorithm. All subjects gave written informed consent at the time of enrollment for imaging and completed questionnaires approved by each participating site's Institutional Review Board (IRB).

Scan-rescan dataset

First, to evaluate stability and potential bias, a scan-rescan database of 20 healthy subjects scanned 4 times within the same week (twice during a first session and twice again over 2 different days) was used. Each subject was taken out from the scanner before getting back in for each rescan session. No volume change is expected for the subjects in this database. The T1-weighted MRI images were acquired on a 1.5T SIEMENS MRI scanner with a 3D spoiled gradient echo (GRE) sequence (TR=22ms, TE=9.2ms, flip angle=30^{o}, 1mm isotropic voxels).

ADNI dataset

Second, to evaluate the performance of the algorithm when changes over time are expected, we used data obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in

2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. More information about the ADNI investigators is given in the Acknowledgement section.

To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

From the website (www.adni.loni.ucla.edu/ADNI), ADNI-1 AD and normal controls (NC) subjects with 4 time-points (0, 6, 12 and 24 months) acquired on a 1.5T scanner that are part of the standardized set of subjects as described by Wyman et al. (2013) were selected. This selection yielded 155 NC (age average at baseline=76.0±4.9 years) and 98 AD patients (age average at baseline=75.3±7.3) that passed quality control (Jack, Bernstein et al. 2008). The 3D T1-MPRAGE images (TR=2300–3000, TE=/3–4 ms, flip angle=8–9°, section thickness=1.2 mm, 256 reconstructed axial sections) with the following image pre-processing: gradient non-linearity distortion correction (grad-wrap (Jovicich, Czanner et al. 2006)) and intensity non-uniformity (N3 (Sled, Zijdenbos et al. 1998)) were used for subsequent analysis.

Metrics

In order to evaluate the stability, regularity, continuity and bias of the proposed approach, we chose metrics based on ventricular, hippocampi and cerebral segmentations for each subject at each time-point. These structures were chosen since they have previously been used to represent the progression of neurodegenerative processes such as in MS or AD (Fisher, Rudick et al. 2002, Lewis and Fox 2004, Nestor, Rupsingh et al. 2008). For the methods described below, these structures were either (i) segmented locally using the patch-based technique proposed initially by Coupé et al. (2011) for hippocampus, for ventricles by Fonov et al. (2012) and for brain Eskildsen et al. (2012) combined with a Bayesian tissue classifier (Zijdenbos, Forghani et al. 1998) to remove cerebrospinal fluid (CSF) from the initial brain mask to conserve only brain tissue; or (ii) data was downloaded from the "MRI image analyses" section of the ADNI website[12] as indicated below.

Methods compared

The proposed method is compared to seven other methods. Like the proposed method, the first two are based on the ANIMAL non-linear registration framework, while the five others are based on publicly available methods that include FreeSurfer, SPM12, QUARC, TBM and KNBSI[13]. The eight techniques are identified as follows:

- LIT: The longitudinal individual template is the new method proposed in this paper, with spatio-temporal regularization with an individual template.
- IT: The individual template method is like LIT with longitudinal pre-processing using all time-points, but without applying the spatio-temporal regularization.

[12] www.loni.ucla.edu/ADNI
[13] http://sourceforge.net/projects/bsintegral/

- CS: The cross-sectional method is based on ANIMAL, and uses direct linear registration (Collins, Neelin et al. 1994) of each time-point independently to the common stereotaxic space (MNI template) after intensity non-uniformity correction.

CS, IT and LIT represent different levels of the pipeline stages as seen in Figure 6.2, thus enabling an evaluation of the contribution of the longitudinal processing and the spatio-temporal regularization steps.

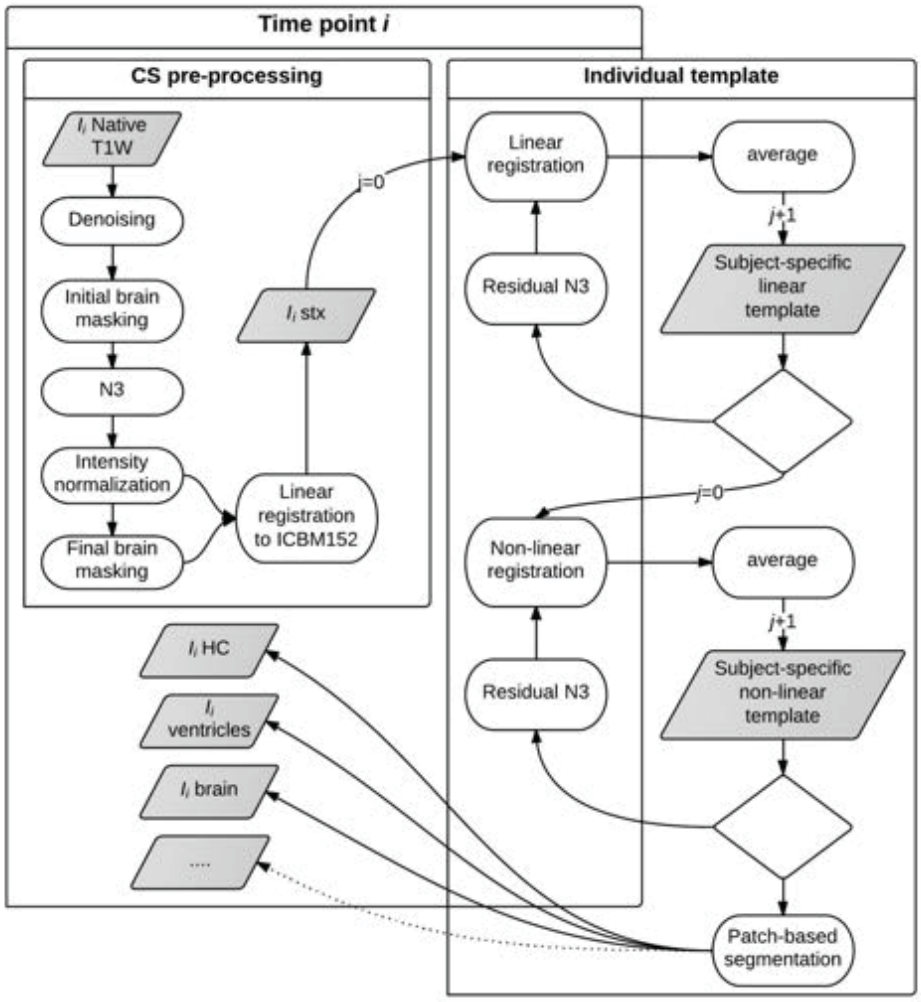


Figure 6.2: Longitudinal pipeline diagram. The different steps performed on each subject time-points are represented in the left part of the diagram, where the processes in the left small square represents the cross-sectional (CS) part of the pipeline. The individual template (IT) creation (linear and non-linear) is represented in the right side of the figure.

For CS, the structure segmentation was performed independently on the scan from each subject's time-point and the volume change was estimated by computing the volume difference between the visits. In the case of the IT and LIT longitudinal approaches, only the subject-specific template was segmented and the estimated non-linear transformations were used to transform the segmentation to each time-point and estimate the Jacobian determinant. The volume change was estimated by integrating the Jacobian determinant within the regional structure masks for whole brain, ventricles and hippocampi.

- FS: The longitudinal stream of FreeSurfer software (version 5.1) was chosen as it has shown better results for longitudinal analyses than the cross-sectional version, except for longitudinal whole brain measurement (Holland, McEvoy et al. 2012). FS provides structural segmentations of each subject time-point that are initialized by independent cross-sectional segmentations estimated from a linear individual template. For the scan-rescan analysis, the longitudinal version of FS was used to segment the hippocampi, lateral ventricles and brain. Briefly, FS processing included motion correction and averaging (Cheng, Edwards et al. 2010) of multiple volumetric T1 weighted images (when more than one was available), removal of non-brain tissue using a hybrid watershed/surface deformation procedure (Ségonne, Dale et al. 2004), automated Talairach transformation (Collins, 1994), intensity normalization (Sled, Zijdenbos et al. 1998) and segmentation of the subcortical white matter and deep gray matter volumetric structures (including hippocampi, amygdala, caudate, putamen, and ventricles) (Fischl, Salat et al. 2002, Fischl, Salat et al. 2004). For analysis of the ADNI data, we downloaded the appropriate values from the ADNI website (UCSF-Longitudinal FreeSurfer (5.1), 2014/05/01) as we felt that these would have been optimally run by the authors.

- SPM12: A unified model which combines intensity non-uniformity correction, linear registration and non-linear registration was proposed by Ashburner et al. (Ashburner and Ridgway 2012) and implemented in SPM12. Their method produces

a subject-specific template and uses the Jacobian determinants of the deformation map of the visit toward the template. As SPM12 does not create structure segmentations, our in-house segmentation tools were applied on the SPM12 subject-specific template and the volume change was estimated by integrating the Jacobian determinant within the regional structure masks for whole brain, ventricles and hippocampi. SPM12 was run locally for the scan-rescan and ADNI data.

- KNBSI: KNBSI (Leung, Clarkson et al. 2010) is based on the classic boundary shift integral (BSI) procedure (Freeborough and Fox 1997) and measures the tissue boundary displacement of a pair of images for the whole brain. KNBSI uses tissue specific normalization, k-means classifiers and specific parameters to account for large multi-site image intensity variability (in terms of SNR and tissue contrast differences). To account for the multiple tissue boundaries of the hippocampus, we used the double intensity windowing approach technique which estimates the boundary shift between CSF and grey matter as well as between grey and white matter (Leung, Barnes et al. 2010). For the scan-rescan data, KNBSI was run locally for all structures after applying our in-house differential bias correction as recommended by the author. For the ADNI data, KNBSI data was downloaded from the ADNI site for whole brain and ventricles (Fox Lab, 2014-01-31), again to have optimally run values. We ran double window KNBSI locally for the hippocampi, as these values were not available on the ADNI website.

- QUARC: Quantitative anatomical regional change (Holland, Dale et al. 2011) estimates the volume changes over a region defined in the baseline image where the deformation is estimated by combining pair-wise forward and backward non-linear transformations with intensity normalization. As QUARC is not publically available, we did not use it in the scan-rescan evaluation. However, for analysis of the ADNI data we downloaded QUARC results (UCSD, downloaded on 2014-06-01) from the ADNI web site.

- TBM: The tensor-based morphometry method proposed by Hua et al. (2013), first estimates the statistical properties of the Jacobian determinant of non-linear deformations used to map training subjects to a population template. Second, a group of voxels with a significant rate of atrophy as measured by the Jacobian ($p<0.001$) in the temporal lobes are defined as a region of interest ("stat-ROI"). Finally, a single measurement for each subject, of an independent testing set, is obtained by integrating the Jacobian determinant of the non-linear deformations to the identical population template within the stat-ROI. TBM is not publically available and was not evaluated with the scan-rescan data. TBM results for ADNI data were downloaded from the ADNI website (USC, 2013-11-17).

Each image processing pipeline has a different level of robustness, and MRIs that do not pass quality control could adversely affect the estimation of statistical power. Instead of a head-to-head comparison, we decided to keep only datasets that passed visual quality control. For the data downloaded from the ADNI website, quality control information was only available for FS, KNBSI, and QUARC data. Subjects who passed quality control with the following arguments were kept for the power analyses: FS: QVERALLQC = "Pass" or "Partial"; for the ventricular KNBSI: BSI VENTACCEPT = 1, REGRATING $\leq$ 3, for KNBSI: KMNREGRATING $\leq$ 3; and QUARC QCPASS = 1. The final cohort number for each method is summarized in Table 6.3.

Table 6.3: Number of ADNI-1 subjects available from the downloaded results or that passed quality control and use for the power analyses for the different methods.

Method	NC	AD
CS	153	95
IT	155	98
LIT	155	98
FS	152	96
SPM12	98	60
KNBSI	105	66
QUARC	131	73
TBM	115	73

6.3.4. Statistics

Variability and bias

For the scan-rescan dataset, the percent volume change (VC) and the absolute percent volume change (aVC) were used respectively to evaluate bias and variability of structure volume (V). For each structure of each subject at time-point t of the n visits and the structure average volume ($\frac{1}{n}\sum_{i=1}^{n} V_i$), VC and aVC were estimated with the following formulas:

$$VC_t = 100 * \left(1 - \frac{V_t}{\frac{1}{n}\sum_{i=0}^{n} V_i} \right) \text{ and } aVC_t = \left| VC_t \right| \qquad \text{Equation 6.13}$$

The significant differences between the match-paired segmentations were compared with a paired t-test for the VC comparison and a Wilcoxon sign-rank test for the aVC comparison. The Wilcoxon signed rank test was chosen over a paired t-test because the scan-rescan aVC values do not follow a normal distribution due to the use of the absolute value.

For the longitudinal dataset, the percent volume change measures atrophy or growth using the baseline volume (V_0) as a reference such as:

$$longVC_t = 100 * \left(1 - \frac{V_t}{V_0} \right) \qquad \text{Equation 6.14}$$

Power analyses

For the longitudinal results, we use power analyses to estimate the required sample size to assess the interaction of treatment and time in a longitudinal study where smaller longitudinal variability will enable better detection of a potential treatment effect. Here, the volume change was estimated using a linear mixed-effect model (LME). Indeed, linear mixed-effect modeling has shown to be a powerful statistical technique to analyze longitudinal data (Verbeke and Molenberghs 2000). In this study, we used a simple LME of the volume changes ($longVC$) consisting of a temporal, time-point t interval ($Interval_{It}$) and group ($Group$) fixed-effects while subject (I) was chosen as random effects, such as:

$$longVC_{It} = (\beta_1 + \beta_2 \times Group + b_I) \times Interval_{It} + \varepsilon_{It} \qquad \text{Equation 6.15}$$

Power analyses, as described by Diggle et al. (2013) and applied in Reuter et al. (2012), for longitudinal analysis were performed to estimate the sample size. From the LME model estimation, the common variance (unexplained variability in $longVC$), the correlation of the repeated observations, the number of time-points, the smallest meaningful difference in the rate of change between AD and NC to be detected (effect size), the power of the test (here we chose 80%) and the within-subject variance of the time-points were used to compute sample size. Using the Diggle et al. (2013) formula, power analysis was performed using the R software package[14] with the *lme4*[15] *and longpower*[16] library. The 95% confidence intervals of the estimated sample sizes were obtained from 1000 parametric bootstrappings of the LME model.

The stability of LME model is influenced by the variability of the data as well as the number of time-points. Similarly, the power of the method is more influenced by the baseline and final time-point. Thus, only subjects with 4 time-points successfully passing the quality control were included for the power analysis (Table 6.3).

[14] http://www.r-project.org
[15] http://cran.r-project.org/web/packages/lme4
[16] http://cran.r-project.org/web/packages/longpower

6.4. Results

Qualitatively, a general overview of the pipeline segmentation and individual template of one subject can be appreciated in Figure 6.3. Also, an example of individual template-based longitudinal non-uniformity intensity correction is depicted in Figure 6.4.

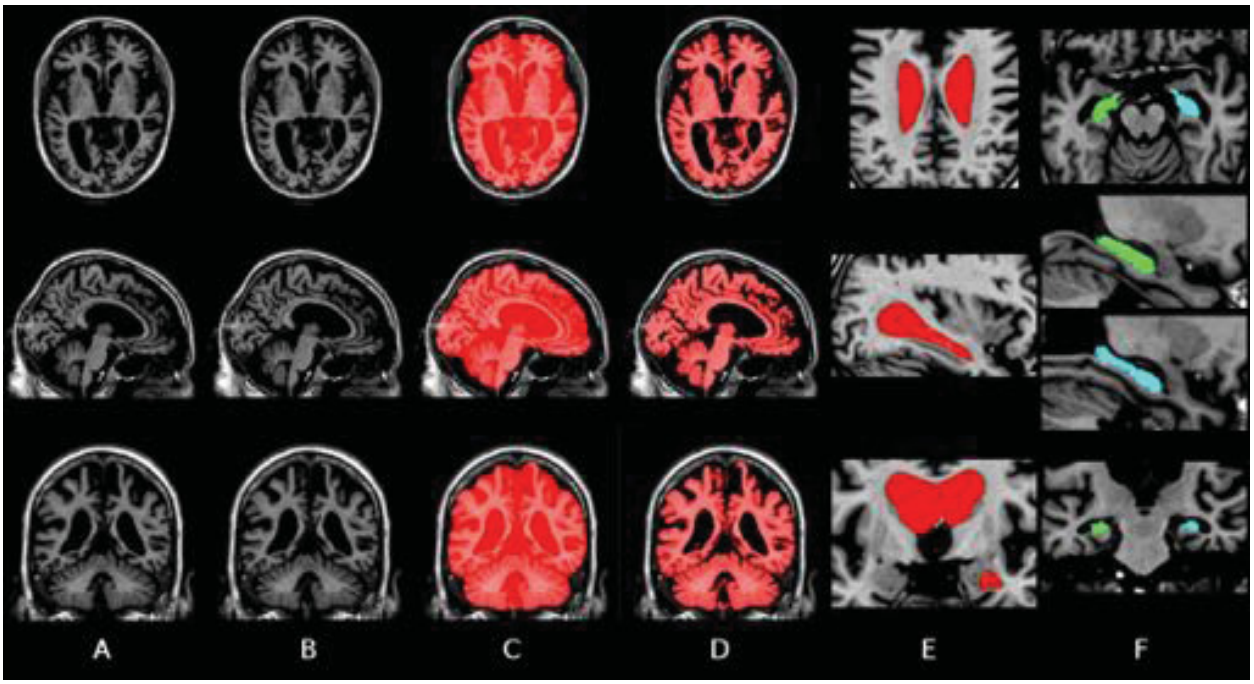


Figure 6.3: Individual LIT template segmentations of an AD subject from ADNI. Axial, sagittal and coronal slices are presented with from left to right: A) Linear individual template, B) non-linear individual template, C) BEaST skull-stripped mask, D) brain mask, E) lateral ventricle mask and F) right (blue) and left (green) hippocampus mask.

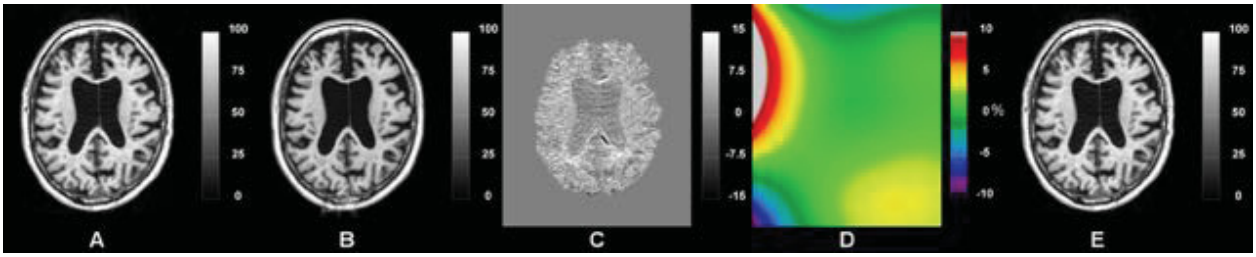


Figure 6.4: Individual longitudinal template-based bias field correction of an AD subject from the ADNI database. From left to right: A) baseline time-point, B) individual linear template, C) baseline time-point and individual template intensity difference image (or A-B), D) bias field of the difference image (C) and E) the baseline image after correction of the longitudinal bias field (D). (Note the different ranges on the color bars.)

6.4.1. Scan-rescan dataset

The scan-rescan dataset should show no anatomical variability since the 4 MRI scans were acquired during a week. Figure 6.5 shows the brain, ventricle and hippocampi volume changes (VC and aVC) for the cross-sectional (CS) and the longitudinal techniques (IT, LIT, FS, SPM12 and KNBSI) for the repeated sessions. For VC and aVC, the smallest structures present the highest volume variability. The method variability ranking is similar across structures excepted for FS and KNBSI, which show more variability for the lateral ventricles and the hippocampi measures.

No significant bias was found when comparing the mean VC of the different methods, as the mean VC values from all methods was centered on zero.

When looking at the aVC, the longitudinal methods (IT and LIT) significantly reduce ($p<0.02$) the variability for all the segmented structures compared to the cross-sectional (CS) method. The longitudinal pre-processing and registration methods such as SPM12 and LIT result in smaller variability between successive sessions compared to CS, IT, FS and KNBSI. Furthermore, SPM12 and LIT methods significantly reduce the aVC for the ventricle segmentation ($p<0.02$). The mean aVC respectively for the brain, ventricles and left/right

hippocampi with the LIT approach are (in percent change): 0.093 (±0.073), 0.355 (±0.387), 0.279 (±0.277) and 0.416 (±0.432).

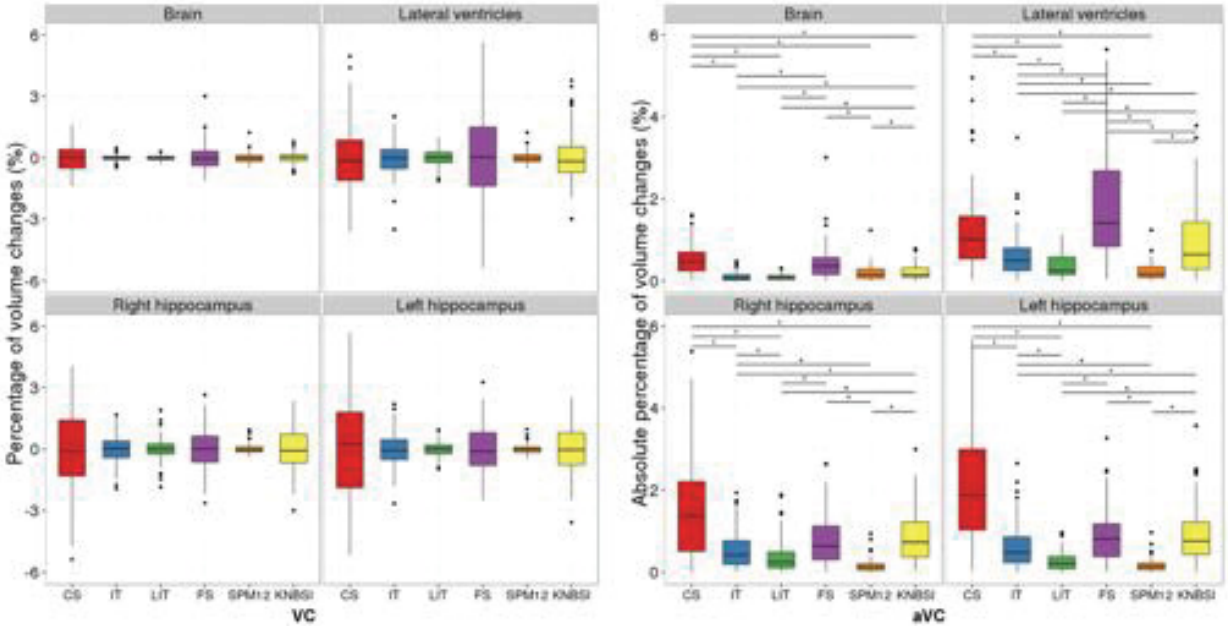


Figure 6.5: Brain, ventricle and left/right hippocampus percentage of volume change for the scan-rescan dataset for the different methods (CS, IT, LIT, FS, SPM12 and KNBSI). The significant difference ($p<0.01$) computed with a match paired Wilcoxon sign-rank are represented by a * where the pairs are represented by the thicker horizontal lines.

6.4.2. ADNI longitudinal dataset

The identical cohort of subjects from ADNI-1 was used to evaluate the different methods, however, subjects scans failing during pipeline processing or absent at the time of method result publication were removed from the analyses and the final number of subjects analysed is summarized in Table 6.3.

Smaller longitudinal variability should improve the statistical power to detect changes in an individual, and facilitate detection of group differences (treatment effects) and thus reduce the

number of subjects required for analysis in a clinical trial. In Table 6.4, we provide estimates of the different sample sizes required to detect a treatment effect that would reduce the annual AD atrophy rate by 25% for the different structures and methods. Table 6.4 also shows the annual atrophy rate for the different structures and methods. In general, amongst all methods compared, the LIT method requires a smaller number of subjects per arm for all treatment effect sizes for the brain and hippocampi measurements while KNBSI hold the smallest sample size for the lateral ventricles.

Figure 6.6 shows the individual longitudinal whole brain, ventricular and hippocampi changes (or cumulative atrophy) for each group (NC in blue and AD in red) and are described in more detail in the following sections.

Table 6.4: Sample size per arm needed to detect a 25% reduction in the annualized rate of brain, ventricular and hippocampus volume change at 80% power for the different methods, while taking into account the normal rate of atrophy. The smallest detectable difference in the rate of change between AD and NC (effect size) and the estimated annual atrophy rate for the different structures for normal controls (NC) and Alzheimer subjects (AD) are also provided with the range representing the 95% confidence interval obtained from parametric bootstrapping of 1000 times. (Table is on the next page)

Structure	Method	Sample size per arm [95% CI]		Effect size in %/year [95% CI]		Annual atrophy rate in % [95% CI]			
						NC		AD	
Brain	CS	>1000	[- -]	-0.11	[-0.21 -0.07]	-0.78	[-1.14 -0.43]	-1.24	[-1.71 -0.98]
	IT	146	[127 199]	-0.30	[-0.37 -0.23]	-0.68	[-0.90 -0.58]	-1.87	[-2.05 -1.62]
	LIT	98	[56 135]	-0.29	[-0.34 -0.25]	-0.66	[-0.79 -0.58]	-1.84	[-2.00 -1.71]
	FS	367	[248 551]	-0.17	[-0.19 -0.13]	-0.62	[-0.73 -0.49]	-1.29	[-1.38 -1.11]
	SPM12	312	[90 524]	-0.09	[-0.11 -0.06]	-0.18	[-0.23 -0.14]	-0.53	[-0.61 -0.47]
	KNBSI	117	[95 149]	-0.21	[-0.23 -0.18]	-0.70	[-0.76 -0.57]	-0.17	[-1.59 -1.42]
	QUARC	278	[98 529]	-0.17	[-0.22 -0.12]	-0.61	[-0.71 -0.47]	-1.29	[-1.44 -1.11]
	TBM	216	[98 320]	-0.14	[-0.16 -0.11]	-0.25	[-0.33 -0.17]	-0.79	[-0.87 -0.73]
Lateral ventricles	CS	173	[127 271]	1.39	[1.05 1.56]	4.46	[3.78 5.39]	10.03	[8.81 10.80]
	IT	214	[141 305]	1.18	[0.93 1.37]	3.90	[3.16 4.62]	8.64	[7.40 9.59]
	LIT	148	[80 190]	1.30	[1.00 1.60]	3.86	[3.19 4.33]	9.04	[8.21 9.72]
	FS	199	[108 287]	1.51	[1.12 1.84]	4.53	[3.68 5.64]	10.57	[9.66 11.51]
	SPM12	145	[125 186]	0.89	[0.79 0.93]	2.36	[2.12 2.68]	5.93	[5.41 6.26]
	KNBSI	199	[153 281]	1.50	[1.09 1.81]	4.46	[3.63 5.48]	10.47	[9.34 11.34]
	QUARC	167	[23 225]	1.84	[1.54 2.39]	4.67	[3.15 6.05]	12.02	[11.47 13.44]
Right hippocampus	CS	240	[123 353]	-0.58	[-0.68 -0.49]	-1.48	[-1.74 -1.18]	-3.81	[-4.19 -3.39]
	IT	131	[14 205]	-0.57	[-0.74 -0.46]	-1.09	[-1.28 -0.83]	-3.38	[-3.87 -3.01]
	LIT	70	[52 90]	-0.65	[-0.73 -0.57]	-0.82	[-1.05 -0.55]	-3.43	[-3.62 -3.18]
	FS	191	[70 294]	-0.71	[-0.91 -0.52]	-1.33	[-1.73 -1.08]	-4.17	[-4.93 -3.60]
	SPM12	>1000	[- -]	-0.04	[-0.10 0.02]	-0.17	[-0.29 0.05]	-0.35	[-0.56 -0.01]
	KNBSI	173	[67 280]	-1.26	[-1.62 -0.87]	-0.68	[-6.43 -4.90]	-5.73	[-0.25 0.54]
	QUARC	130	[93 166]	-0.59	[-0.65 -0.51]	-0.98	[-1.16 -0.76]	-3.32	[-3.55 -3.01]
Left hippocampus	CS	219	[190 317]	-0.62	[-0.68 -0.47]	-1.44	[-1.93 -1.16]	-3.94	[-4.27 -3.41]
	IT	91	[66 130]	-0.60	[-0.69 -0.45]	-1.10	[-1.35 -1.01]	-3.48	[-3.83 -3.10]
	LIT	67	[43 88]	-0.61	[-0.70 -0.52]	-0.93	[-1.18 -0.77]	-3.38	[-3.65 -3.20]
	FS	140	[71 167]	-0.84	[-0.99 -0.78]	-1.10	[-1.30 -0.78]	-4.46	[-4.82 -4.33]
	SPM12	>1000	[- -]	0.02	[-0.05 0.06]	-0.16	[-0.37 0.10]	-0.08	[-0.30 0.05]
	KNBSI	194	[111 266]	-1.07	[-1.28 -0.01]	-0.94	[-5.67 -4.84]	-5.23	[-0.07 0.35]
	QUARC	133	[83 203]	-0.51	[-0.60 -0.37]	-1.08	[-1.42 -0.75]	-3.12	[-3.32 -2.73]

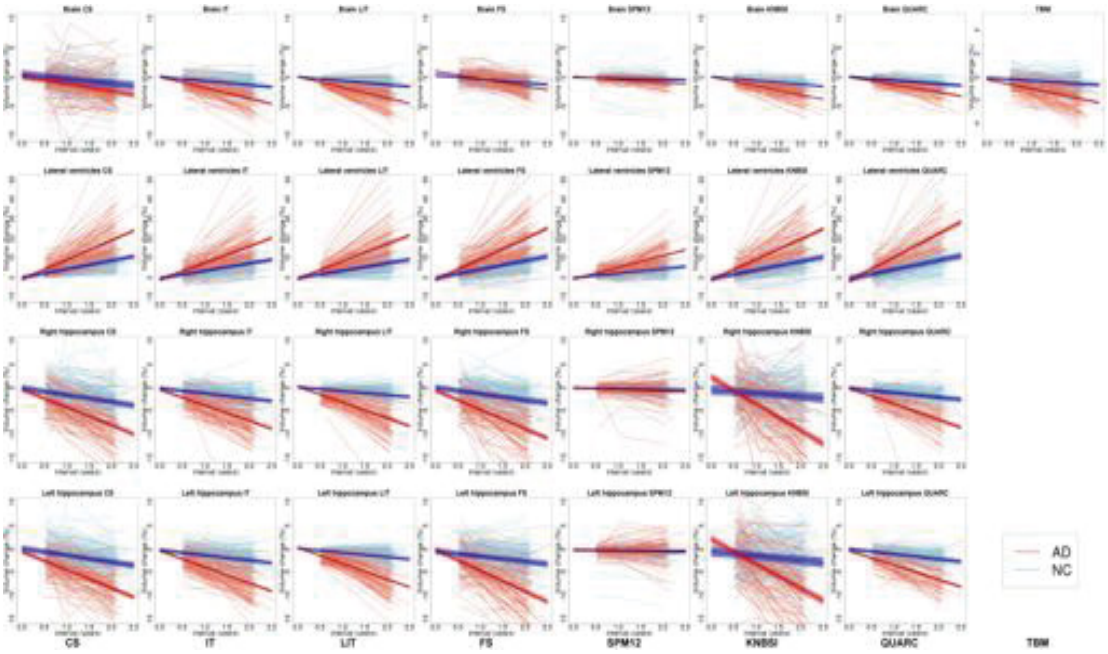


Figure 6.6: Longitudinal individual and linear mixed model with confidence intervals for the NC (blue) and AD (red) groups. Brain, ventricular and left/right hippocampi volume changes for CS, IT, LIT, FS, SPM12, KNBSI and TBM. Each thin full line represents an individual subject volume. Thicker lines represent the LME model for the respective groups while the shaded bands represent the 95% confidence interval on the mean model.

Whole brain measurements

With regard to the whole brain, the LIT method results in a sample size of 98 to detect a 25% change in brain atrophy, versus 146 for IT and more than 1000 subjects required for the cross-sectional approach (CS). Furthermore, the LIT sample size is smaller than KNBSI (117) and TBM (216 subjects per arm). The LIT sample size range (56-135) overlaps with the following approaches: IT (127-199), KNBSI (95-149), QUARC (98-529) and TBM (98-320), however, LIT and IT provide a stronger effect size (-0.29 and -0.34) than these other methods.

Regarding the individual trajectories seen in the spaghetti plots in Figure 6.6, LIT provides a more progressive and regular individual trend while preserving group differences. It is interesting to note that the local constraints on the Jacobian over time result in a structure-wide regularization. KNBSI and QUARC measurements show a reduced individual longitudinal variability as well, compared to CS, FS, IT and TBM.

Lateral ventricle measurements

Among the different techniques tested, SPM12 and LIT yield the best power to detect a 25% reduction in lateral ventricular enlargement with only 145 and 148 subjects required per arm, but SPM12 show the tightest range (125-186 and 80-190, respectively). The LIT effect size is stronger than SPM12 with a value of 1.30 versus 0.89 to discriminate the ventricular growth rate change between AD and NC. The CS approach of our pipeline yields better performance than the IT method (173 and 214 subjects, respectively), but the LIT reduces this number to 148 subjects.

When looking at the segmented lateral ventricle volumes in Figure 6.6, the trend of the observed ventricular enlargement is similar between the methods, but there is a net decrease of intra-subject variability for the longitudinal methods (IT, LIT, FS, SPM12, KNBSI and QUARC), as evidenced by spaghetti plots with more realistic, less chaotic changes over time. We can also appreciate with Figure 6.6 that the lateral ventricle volume changes are the strongest but also the more stable progression compared to other structures regardless of the method.

Hippocampus measurements

Among the different hippocampal methods tested, the LIT technique yields the best power to detect a 25% reduction in atrophy, with 67 subjects (left side) and 70 subjects (right side) required. When the temporal constraint is not applied to the deformations, the IT method requires 91 and 131 subjects (left and right side, respectively) to detect the same change. The other methods require more than 100 subjects to detect the same potential treatment effect. The estimation of the LME for SPM12 did not converge well enough to perform power analyses. FS shows the stronger effect size (-0.71±0.20 and -0.84±0.15 for the right and left hippocampi) but the effect size variability is much larger than for LIT (-0.65±0.08 and 0.61±0.09).

Figure 6.6 shows that the individual hippocampal trajectory variability is clearly decreased with the longitudinal methods, and in particular with IT, LIT and QUARC.

Jacobian determinant maps

The concatenation of the transformation allows us to assess the total deformation between two specific time-points. Following this idea, Figure 6.7 shows the Jacobian determinant of the deformations estimated for the longitudinal methods (IT and LIT) for an AD patient. The IT Jacobian maps have multiple punctuate shrinking and enlarging regions within the ventricles that are not consistent with the notion of gradual ventricular growth that is relatively homogenous throughout the ventricle. By using a subject-specific template and the 4D regularization with the LIT methods (rightmost images), there are focal and consistent deformations that overlap well with the anatomy that is assumed to change with AD. Indeed, one can appreciate stronger temporal lobe atrophy detected with the LIT approach.

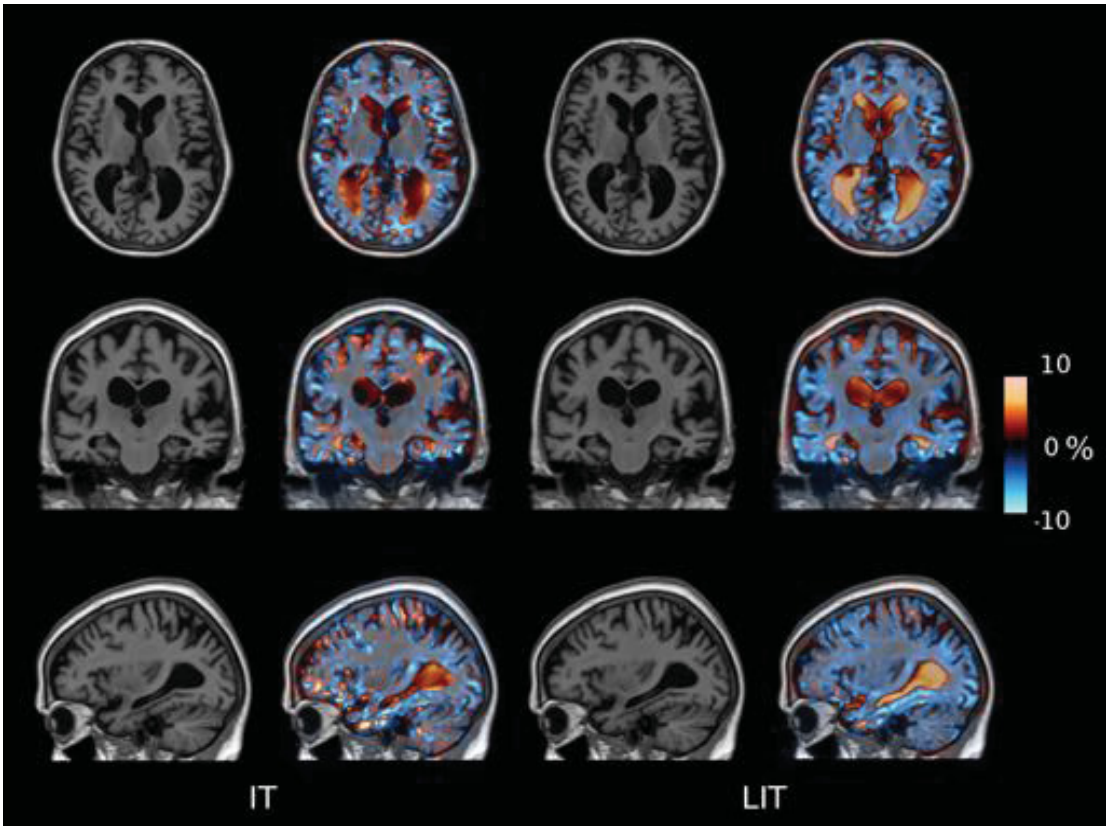


Figure 6.7: Longitudinal deformation fields. Deformation fields from baseline to the 12 month time-point for the longitudinal approaches (IT and LIT) where red represent growth and blue atrophy for a randomly chosen AD subject.

6.5. Discussion

In this article, we have presented a new approach for the estimation of individual longitudinal changes using individual subject-specific templates and spatio-temporal regularization. We also provide an unbiased framework for analysing longitudinal data where every time-point is processed with the same steps. A robust estimation of the deformations is obtained using an unbiased individual template approach, minimizing deformations between subject time-points. Meanwhile, a local spatio-temporal regularization is achieved with bi-linear regression of the deformation field and its Jacobian matrices. The regression of the decomposition enables a temporal regularization at a local voxel level. Furthermore, we compared our technique with a traditional cross-sectional approach, as well as recent powerful methods, FS, SPM12, KNBSI, QUARC and TBM. Longitudinal image analysis bias was assessed on a scan-rescan dataset, and power analysis to detect a potential treatment effect on an Alzheimer cohort was chosen.

Longitudinal image analyses can be subject to bias in particular due to non-linear registration when an arbitrary reference image is chosen (Thompson and Holland 2011) or due to interpolation asymmetry (Yushkevich, Avants et al. 2010). 4D Hammer (Shen and Davatzikos 2004) or non-linear registration such as Diffeomorphic Demons (Vercauteren, Pennec et al. 2009, Lorenzi, Ayache et al. 2010) and ANTs (Avants, Grossman et al. 2006) require a reference image to be defined, therefore introducing possible bias. Symmetric interpolation and registration might not be sufficient to correct for bias when there are more than 2 time-points. The use of an individual template, as suggested by Reuter et al. (2012) showed no bias and our approach exploits this strength and adds non-linear registration to obtain a more accurate anatomical correspondence between time-points. Then, the individual template can then be use to segment brain structures directly and not only to initialize the segmentation as it is done in FS. By definition, our approach, using a longitudinal pre-processing to remove interpolation bias and an individual template for non-linear registration, is symmetric and transitive as it is similar to a "half-way" space registration approach (Smith, De Stefano et al. 2001). Indeed, the longitudinal pre-processing applies the same number of interpolations and removes the intensity inter-visit

non-uniformity. Furthermore, the non-linear registration is performed towards a common target, producing unbiased deformation fields that can be combined to obtain robust and transitive non-linear deformations between time-points. However, the individual template estimation depends of the current set of subject's time-points, thus the addition of new time-points will require to re-estimate the individual template. In our experiments, we have demonstrated that the LIT method provides a more robust longitudinal measure on a scan-rescan dataset where no changes are expected. We have also found that by using individual subject-specific templates (IT, LIT, FS and SPM12), structure volume variability is decreased compared to the cross-sectional approaches like CS that uses a single common template (e.g., the ICBM152 model) for all subjects. Among the three longitudinal methods tested, the LIT and SPM12 demonstrated the least bias and the smallest variability in structure volumes which is expected are both methods apply a longitudinal regularization and therefore minimize the temporal variability.

Experiments on ADNI data reveal increased stability in estimating individual changes over time compared to standard cross-sectional approaches. Indeed, the cross-sectional approach was chosen as reference method and allowed us to show an important improvement in the measurement of longitudinal change thanks to the longitudinal pre-processing (IT) and temporal regularization constraint (LIT). However other strictly cross-sectional approaches with independent time-point measures have showed to perform better in a similar study such as FS in Holland et al. (2012).

The longitudinal regularization of the deformation *at a local level* reduces the longitudinal noise in volume estimation *at the global/structural level*, while the hierarchical iterative process produces a robust individual template that allows for better anatomical matching across time in an individual. An important aspect of longitudinal clinical and research studies is the cost of recruiting subjects and scanning them at multiple time-points. The proposed longitudinal analysis techniques will allow for better power to detect differences between groups, and thus will lead to the reduction of the number of subjects required for research and for clinical trials. Compared to the literature, where similar ADNI cohorts of AD and NC were used, our power analysis show similar sample sizes required to detect treatment effects for the FS, KNBSI, QUARC and TBM

approaches (Holland, McEvoy et al. 2012). The proposed temporal constraint (LIT), after longitudinal processing (IT), reduces the sample sized by a factor by approximately 50% for brain, 70% for the lateral ventricles and 50% for hippocampus. The temporal constraint from SPM12, developed to optimize longitudinal VBM, produces unbiased results on the scan-rescan dataset but might be over regularizing the longitudinal deformation to detect structural change on the ADNI cohort in this experiment. Indeed, it is important to mention that smoothing the temporal fluctuation could remove temporal artifacts, while it could also smooth real signal fluctuations. Within LIT, the bi-linear longitudinal constraint is only applied at a voxel level, i.e., the displacement of a point in the brain is constrained to move in a linear fashion over time. But the volume of the structure is not explicitly constraint the structure to continuously increase or decrease. This constraint results in global more continuous volume change which is not the result of an explicit constraint on the volume.

Another interesting finding is that longitudinal pre-processing and individual template creation does not affect the longitudinal measurements of anatomical structures in the same manner. Indeed, structures such as the lateral ventricles with high contrast and less sensitive to bias field and distortion, resulting in a similar sample size for both longitudinal and cross-sectional approaches. However, the spatio-temporal regularization is able to decrease the longitudinal variability of such structures and therefore reduce the sample size.

We limited our comparison to publicly available methods and/or results on the ADNI-1 cohort, but other methods have been developed and applied on real longitudinal data. The complexity and/or the computational cost of these methods (Durrleman, Pennec et al. 2009, Wu, Wang et al. 2011, Liao, Jia et al. 2012) may limit the application to large database such as ADNI. Wu et al. (2011) aligned all longitudinal images of a population toward a hidden common space equivalent to a template and it can be applied to a single subject. The individual longitudinal deformations or "temporal fibers" are estimated without any priors but regularized with a Gaussian kernel to preserve the continuity of the longitudinal deformation field. Similarly, Lorenzi et al. (2010) proposed to fit a linear model to constrain the longitudinal velocity fields of the subjects time-points in the Demons' framework with the baseline image used as a reference.

Despite the fact of their approaches being more general in modeling the deformations, the usual small number of time-points might limit the longitudinal continuity.

Finally, the main focus of this article was to compare longitudinal regularization versus longitudinal pre-processing and cross-sectional approaches. Although we focused on whole brain, lateral ventricles and hippocampi, any other structures can be analyzed longitudinally as far as the individual template can be segmented. The longitudinal Jacobian determinant maps show interesting results to measure voxel-wise deformation individually with the spatio-temporal regularization. The deformation maps present plausible anatomical atrophies such as in gray matter and in temporal lobes as well as a uniform ventricular enlargement. The results are encouraging and hold the potential of voxel-wise longitudinal DBM of neurodegenerative diseases.

6.6. Conclusion

This study evaluated a longitudinal framework with spatio-temporal regularization of deformation fields and the creation of an individual 3D template through non-linear registration in the context of longitudinal neuroimaging studies. The experiments were carried out on scan-rescan and ADNI datasets. In comparison with freely available and popular methods, the spatio-temporal regularization (LIT) shows competitive results in regard to robustness, power and stability while reducing the number of subjects required to show statistical differences between groups. In addition, the LIT approach showed promising results for longitudinal DBM analysis and can be easily adapted to investigate specific anatomical biomarkers.

Acknowledgements: Canadian Institutes of Health Research (MOP-111169 & 84360), les Fonds de Recherche Santé Québec and the MS Society of Canada (BioMed PhD studentship, 691) and CBRAIN (http://cbrain.mcgill.ca/, Principal Investigator: Alan C. Evans)(Frisoni, Redolfi et al. 2011), a Canada-wide platform for distributed processing, for computation. Data collection and sharing for this project were funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). The ADNI is funded by the National

Institute on Aging and the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics NV, Johnson & Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc., F. Hoffmann-La Roche, Schering-Plough, Synarc Inc., as well as nonprofit partners, the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to the ADNI are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study was coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles.

CHAPTER 7 DISCUSSION AND FUTURE WORK

The main goal of this Ph.D. thesis was to develop the necessary longitudinal MRI workflow to quantify in-vivo morphological changes in neurodegenerative disease such as MS and AD. We designed methods to account for longitudinal intra-subject pathological variability and took steps to minimize methodological variability in atrophy measures. To achieve this objective, we proposed the following:

- A method to detect and remove potential atrophic confounding such as focal pathology such as MS lesions which are sporadic and dispersed.
- A method to assess and correct the methodological bias in non-linear registration algorithms.
- A method to account for longitudinal image acquisition variability.

In this chapter, before concluding, we summarize our main findings and, discuss the strengths, weaknesses, novelty and scientific impact of each four manuscripts, and prospective future work based on these thesis contributions.

7.1. Automated focal pathology detection

The first study of this thesis (Chapter 3) focuses on detecting focal pathology automatically such as MS lesions and was assessed on a clinical RRMS patient database as well as on an independent validation dataset, and compared to state-of-the-art methods. The results of the

proposed RMNMS approach show highly competitive results to automatically detect small to large lesions, and also to estimate the total burden of lesion load.

One of the strengths of this work was to propose a simple and intuitive method to segment MS lesions from exemplar images. Manual MS lesion segmentation has been, and is still, considered as the gold standard approach to quantify progression of the focal pathology in clinical trials. Our approach proposed to simply exploit large dictionaries of manual segmented lesions to find similar local information within these library images when segmenting a new dataset. Furthermore, the simplicity of the algorithm allows a wide search for the most similar images in the anatomical library. This optimal search of most similar images and similar patches enables segmentation of images with different intensity features (i.e., different scanners, sites, or image acquisition parameters) which usually requires extensive training for unsupervised segmentation approach. Despite being acquired from different scanner manufacturers and different sites, the segmentation results showed homogenous results.

Another interesting aspect for large multi-site clinical trials is the number of image contrasts required. While most MS lesions segmentation approaches require T1W, T2W, PDW and FLAIR images, we demonstrated that if enough examples are present in the library of images, only two modalities (T2W and FLAIR) are sufficient. This smaller number of image contrasts can reduce the cost of MRI acquisition per patient. More importantly, this also reduces the scanning time for the patient, leading to the potential of better quality data as patients may have less fatigue and less movement artifacts. However, the strength of not requiring a model could be a weakness for the RMNMS method where large cohorts of pre-segmented subject data are not available. In our validation, we showed that 50 subjects is sufficient to capture the presence of most lesions. Therefore, exploiting the MSGC library of images, freely available from MICCAI 2008, should provide similar results.

While the article focused on MS lesions segmentation to demonstrate is capability, the algorithm could be directly applied to similar focal pathology segmentation such as white matter hyperintensities (WMH) lesion segmentation. Similarly to T2W-lesions in MS, WMH lesions are

bright on FLAIR images and are associated with degenerative changes of small vessels and implicated in the pathogenesis of cognitive decline and dementia (de Leeuw, de Groot et al. 2001). Such an application would require a training set of WMH lesions.

Our lesion segmentation focused on WM lesions in MS while GM pathology can be equally or more extensive (Vercellino, Plano et al. 2005). However, GM lesions have little contrast on conventional MRI and generally appear normal. Advanced imaging methods have been developed (i.e. double inversion recovery (DIR)(Geurts, Pouwels et al. 2005), and phase-sensitive inversion recovery (PSIR)(Nelson, Poonawalla et al. 2007)) and have improved the detection of cortical lesions by 150-500% with respect to conventional MRI. Therefore, including the detection the focal GM pathology could lead to a better assessment of focal pathology. Again, this application would require an appropriate training set of exemplar data as the only requirement to apply the proposed patch-based lesion segmentation technique.

Further research should be devoted to the detection of lesions longitudinally, in serial data, which will enable assessment of disease activity over time. It is important to detect the evolution of MS lesions as it can be used as a marker of disease activity and as a potential surrogate for relapses (Guttmann, Kikinis et al. 1999). More importantly the relationship of acute, chronic and remyelinating lesions with respect to focal and diffuse atrophy is still not fully understood. Multiple approaches have been developed to detect MS lesions in longitudinal MRI data (Thirion and Calmon 1999, Prima, Ayache et al. 2002, Bosc, Heitz et al. 2003), or to detect the appearance of new lesions (Rey, Subsol et al. 2002, Elliott, Arnold et al. 2013). The bottleneck of these methods is still in the accurate initial detection of chronic lesions. Exploiting the high sensitivity of our cross-sectional segmentation approach could provide a good baseline reference. In the same multi-modal and rotation invariant NLM segmentation framework, the baseline image segmentation could be used to detect the new or disappearing lesions in follow-up images within an *a contrario* framework. In image segmentation, the *a contrario* segmentation consist in identifying a structure when is presence is too unlikely to happen by chance (Desolneux, Moisan et al. 2003). The purpose of this model is to detect outliers (e.g. new MS lesions). The method is

composed of two main parts; namely a naive model (the baseline segmented image) and one or several measurements related to the structure to be detected.

7.2. Focal pathology inpainting

In the continuity of Chapter 3, we developed a new technique in Chapter 4 to replace affected tissues, such as MS lesion, with healthy appearing tissues in order which minimize their impact when estimating tissue deformation with the goal of improving the power to detect longitudinal MRI changes in MS. During the endMS 2013 conference (Guizard, Nakamura et al. 2013), we were the first group to propose using an exemplar-based MRI inpainting approach, which exploits the redundancy of normal appearing tissues to replace pathological tissues without any *a priori*. Previously, inpainting approaches in MS required tissue segmentation priors which can be difficult to obtain when diffuse tissue pathology is present, as is the case in MS.

Although the method presented in Chapter 4 is applied in the context of MS lesion inpainting, the method could be applied to other white matter lesions such as the FLAIR WMH in AD. Moreover, the hierarchical strategy to inpaint large regions with the non *a priori* strategy could allow filing of large regions with different tissue types, as could be the case for tumor, or for tissue with implanted electrodes (e.g., in epilepsy or Parkinson's disease) which can also affect standard image processing algorithms. Preliminary highly promising qualitative results can be appreciated in Figure 7.1, where the region delineated in green, has been removed and inpainted by our algorithm.

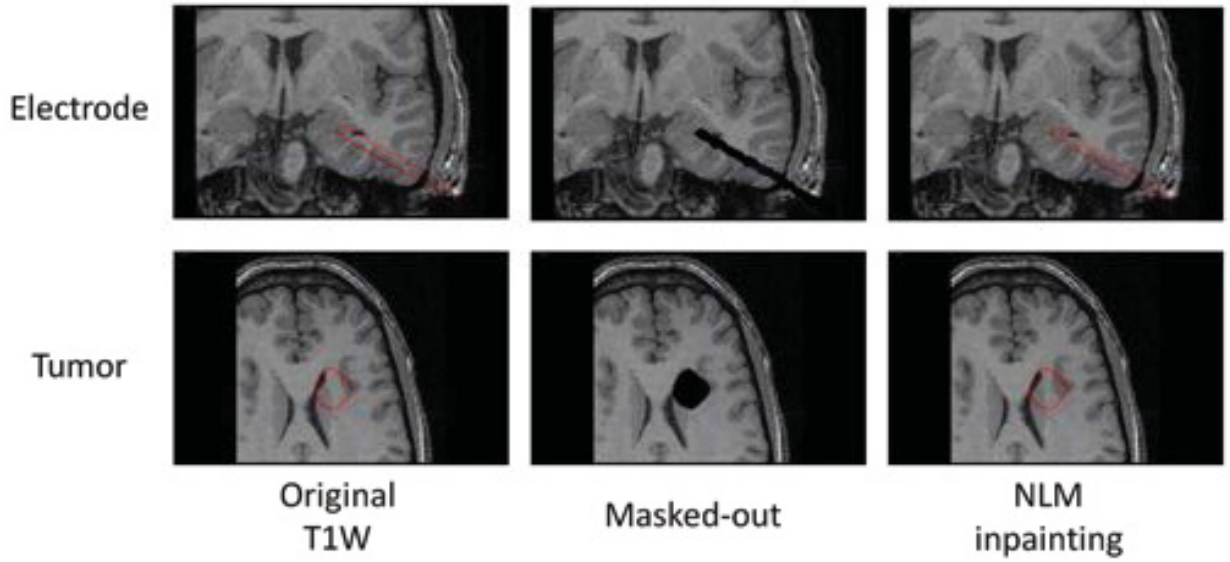


Figure 7.1: Example of simulated electrode and tumor inpainting results on T1W images with NLM inpainting algorithm. From left to right, the figure shows the original images, the region masked-out from the images and the NLM inpainting results. Note that the mask of the region being inpainted in overlaid on top of the image in red.

For longitudinal processing, the method can improve the power of the popular SIENA and VIENA image analyses tools, which measure changes affecting the boundary of a ROI. However, further research should be conducted to assess the impact of lesion inpainting in the evaluation of focal morphological changes of the NABT. Indeed, the weakness of our validation is the lack of voxel-wise analysis of the inpainting algorithm in a longitudinal study (e.g., in a DBM or VBM study). With respect to inpainting impact on non-linear registration, Sdika et al. (Sdika and Pelletier 2009) evaluated the impact of their LWM inpainting approach on cross-sectional non-linear registration from simulated inpainted lesions on healthy subject T1W images and compared the resulting deformation fields of the corresponding healthy subject T1W images. Their simulation is similar to the one used in our PSNR analysis, which could potentially be used to simulate MS lesions in longitudinal healthy subject MRI data. However, we believed that this kind of validation framework, inspired by Brett et al. (Brett 2001), is limited by the non-linear registration being used for the simulation where brain affected by lesions are registered to healthy brains. This limitation will be further amplified in longitudinal non-linear registration where other variability might affect the results as mentioned in Chapter 5 and Chapter 6.

Focusing on our ultimate goal to measure focal atrophy in a reliable manner, validating the impact of inpainting in longitudinal MRI data will require measuring the power to detect morphological changes at on a voxel-based level (Chételat, Desgranges et al. 2008). Such a validation was out of the scope of this manuscript, but preliminary results show a direct positive impact of lesion inpainting on the longitudinal non-linear deformation Jacobian determinant map of an individual (Figure 7.2). These promising focal atrophy measurement results should provide a better estimation of the longitudinal peri-lesional morphological changes.

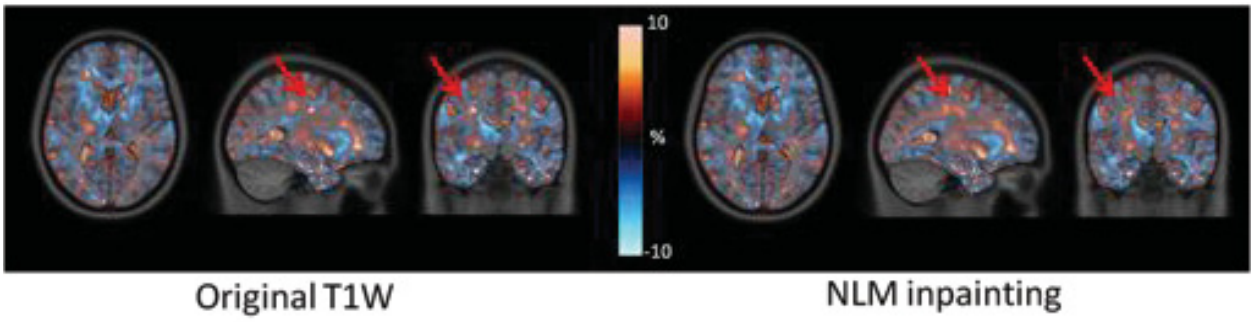


Figure 7.2: Longitudinal DBM analysis of an MS patient over a year interval with and without NLM inpainting of the MS lesions. Expansion is shown in orange and atrophy in blue. The red arrows point toward MS lesions location in the original image.

Apart from requiring tissue segmentation, the main limitation of the different inpainting approaches, including ours, is the necessity of providing the ROI to inpaint. Because our approach is not tissue specific and is robust to "over" segmentation, the combination of our RMNMS lesion segmentation and the NLM inpainting would provide a fully automatic lesion detection and segmentation.

7.3. Non-linear registration symmetry

In contrast with the manuscripts of Chapter 3 and Chapter 4, the focus of the study presented in Chapter 5 was to explore the intrinsic methodological bias and accuracy to detect brain atrophy in pair-wise non-linear registration. In this manuscript, we assessed the current most popular freely available method SyN and some variants of our in-house ANIMAL algorithm. The results demonstrated that symmetry is important to reduce deformation bias of the non-linear registration direction. In this study, we also proposed a symmetrization constraint applicable to any non-linear registration algorithm which provided unbiased non-linear registration.

Pair-wise registration algorithms have been evaluated intensively, but for the first time, we evaluated their performance on scan-rescan and longitudinal MRI data. For the longitudinal evaluation, we chose to use AD instead of MS population because longitudinal AD datasets are now freely available (ADNI, OASIS, MIRIAD…) and they provide the necessary support to share the results for comparison between other groups as mentioned in Chapter 6. Furthermore, the lower amount of visible lesions on T1W images in AD reduces confounds for non-linear registration due to focal pathology. Indeed, the detection of small image changes such as lesion changes is a difficult task, while the appearance or disappearance of lesions can breach the expected one-to-one correspondence of the images being registered. The transformation model or the regularization of the non-linear registration, needs to account for this potential lack of correspondence or the image itself can be adapted as proposed in Chapter 4.

ANIMAL regularization uses a Gaussian-based approximation of the linear-elastic deformation. While this regularization has shown its efficacy in numerous applications (segmentation, cross-sectional non-linear registration…), it can limit the convergence of the registration in the case of large deformations. In typical clinical study, the interval between baseline scan and the final follow-up for an individual is usually limited to a few years, thus large deformations are not expected in this time frame. However, future works should focus on integrating a more appropriate deformation model such as visco-elastic regularization to capture

potential large deformations. However, increasing the complexity of the deformation and improving the accuracy of the registration should maintain the symmetry and the regularity of the deformation fields. To obtain reliable DBM results, the deformation fields need to be plausible. Therefore, another approach could be to spatially weight the regularization kernel based on the current deformation fields (Commowick, Stefanescu et al. 2005), or based on the anatomy (or voxel's intensity) of the image being registered using adaptive anisotropic filtering (Forsberg, Andersson et al. 2010). Inspired by the hallucination approach used for resolution enhancement (Baker and Kanade 2000), we propose to spatially adapt the regularization from the "source" image information using the NLM algorithm to estimate weights from the source image and apply them on the deformation field. Using the patch-based NLM estimator will allow us to attribute similar weights to the deformation vectors of similar anatomical regions on the "source" image. This hallucination principle was applied to Jacobian determinant maps and their corresponding anatomical T1W images in the context cross-sectional MS subject registration (see Appendix 7.A).

The comparison different approaches on 60 subjects with baseline and 2 follow-up scans required a huge amount of computational resources to estimate the forward and backward non-linear registration between all possible individual pairs. Despite being set to assess non-linear registration approach, the validation framework did not study the impact of image pre-processing in longitudinal pair-wise non-linear registration. Consequently, I initiated the development of a longitudinal pipeline which was initially presented during the MICCAI 2012 novel imaging biomarkers for Alzheimer disease (NIBAD) workshop (Guizard, Fonov et al. 2012). This challenge assessed the ability to detect brain atrophy in AD patients and normal controls and resulted in a contribution to an article currently *in final revision* (Cash, Frost et al. Submitted) in NeuroImage. The longitudinal pipeline uses a subject-specific individual template as described in Chapter 6 and as discussed in the following section. The approach is therefore symmetric and transitive by construction and we can exploit the non-linear registrations toward this individual template to obtain the full trajectory between all time-points. The use of a template can be seen as the equivalent of the principal behind the pair-wise half-way space registration approach.

7.4. Longitudinal non-linear registration

The manuscript presented in Chapter 6 described a new approach to account for longitudinal variability due to intrinsic image variability. Embedded in the creation of an unbiased individual 3D template from multiple time-points, we described our longitudinal pipeline that can account for image acquisition temporal variability. Compared to pair-wise approaches, this longitudinal workflow and the spatio-temporal regularization provided more stable longitudinal morphological measures that in term allows to better identify brain atrophy. The reasoning being that the simultaneous analysis of all the subject's time-points as an ensemble should decrease the intra-subject longitudinal measure variability. Indeed, we found that our longitudinal approach has shown to exceed cross-sectional and pair-wire approaches to detect morphological changes.

The first finding of this manuscript is the importance of longitudinal pre-processing. In this manuscript, we corrected for different technical confounding factors in the image acquisition that could alter atrophy measurement. The proposed pipeline uses an iterative approach to create an individual template which aligns, removes distortion, normalizes images and removes image intensity bias of the subject time-points. Our pipeline has shown to produce accurate longitudinal measures and is currently used by many groups at the MNI (Dr. Andrea Bernasconi, Dr. Alain Dagher, Dr. Doug L. Arnold) but also internationally (Dr. Jan Krasensky, Charles University, Prague, Czech Republic).

Moreover, the proposed approach follows individual brain changes over time at a voxel level (i.e., voxel trajectories) and sets the basis for longitudinal DBM analysis. As mentioned above (Section 7.3), the deformation field should not only provide good image matching but provides plausible deformation maps. This affirmation is even truer in the case of longitudinal deformation where, continuity of the deformation should adequately describe the longitudinal morphological changes. Furthermore, in neurodegenerative disease, such as MS, an important confounding factor of atrophy is the “pseudoatrophy” effect. This effect has been observed in anti-inflammatory clinical trials (i.e., interferon) where the whole brain atrophy rate is greater in

the treated group than in the placebo group in the first year, before reversing in the second year (Rudick, Fisher et al. 2000). The volume decrease is presumably due to resolution of inflammatory edema. The inflammatory activity has shown to not only affect global brain atrophy measure but also more local WM atrophy when subject to inflammation activity (Tiberio, Chard et al. 2005). Therefore focal inflammation could affect focal atrophy measurements. We believe that the proposed spatio-temporal regularization could compensate this focal pathological variation. Moreover, to account for focal pathology, we have combined the methods to segment and inpaint lesions (described in Chapter 3 and Chapter 4) to further minimize their potential adverse impact on the estimation of local deformations. Preliminary results on RR MS longitudinal atrophy measures using our spatio-temporal regularization framework have shown promising results and were presented at the ISMRM WM study group of 2013 (see Appendix 7.B). To conclude on the longitudinal deformation field, Figure 7.3 provides an example of the potential of longitudinal DBM analysis performed with our in-house framework on AD, age-matched normal controls, and young MS subjects, where the annual atrophy of the three populations can be visualized.

Further research should exploit our proposed longitudinal pipeline to characterize the relationship of global and focal atrophy in neurodegenerative diseases over time. Research should explore if focal atrophy occurs initially around a new lesion in MS or in specific regions in AD for instance, and if a second and more diffuse atrophy occurs globally or in the projection areas of the affected regions of the brain. However, DBM is limited to capture the micro-structural changes thus limiting the sensitivity to accurately detect neurodegeneration dynamics. DBM of anatomical images should be associated with other MRI modalities such as MTR and DWI. For instance, using DWI could reveal brain tissues abnormalities as it is sensitive to micro-structural alterations and therefore it can be used to estimate the diffusion directionality. Diffusion directionality is believed to describe various aspects of axonal injury such as demyelination, gliosis, and gross axonal degeneration. Additionally, correlating these measures with clinical variables (patient age, disability, disease type and duration) will provide further insight on the neurodegenerative diseases.

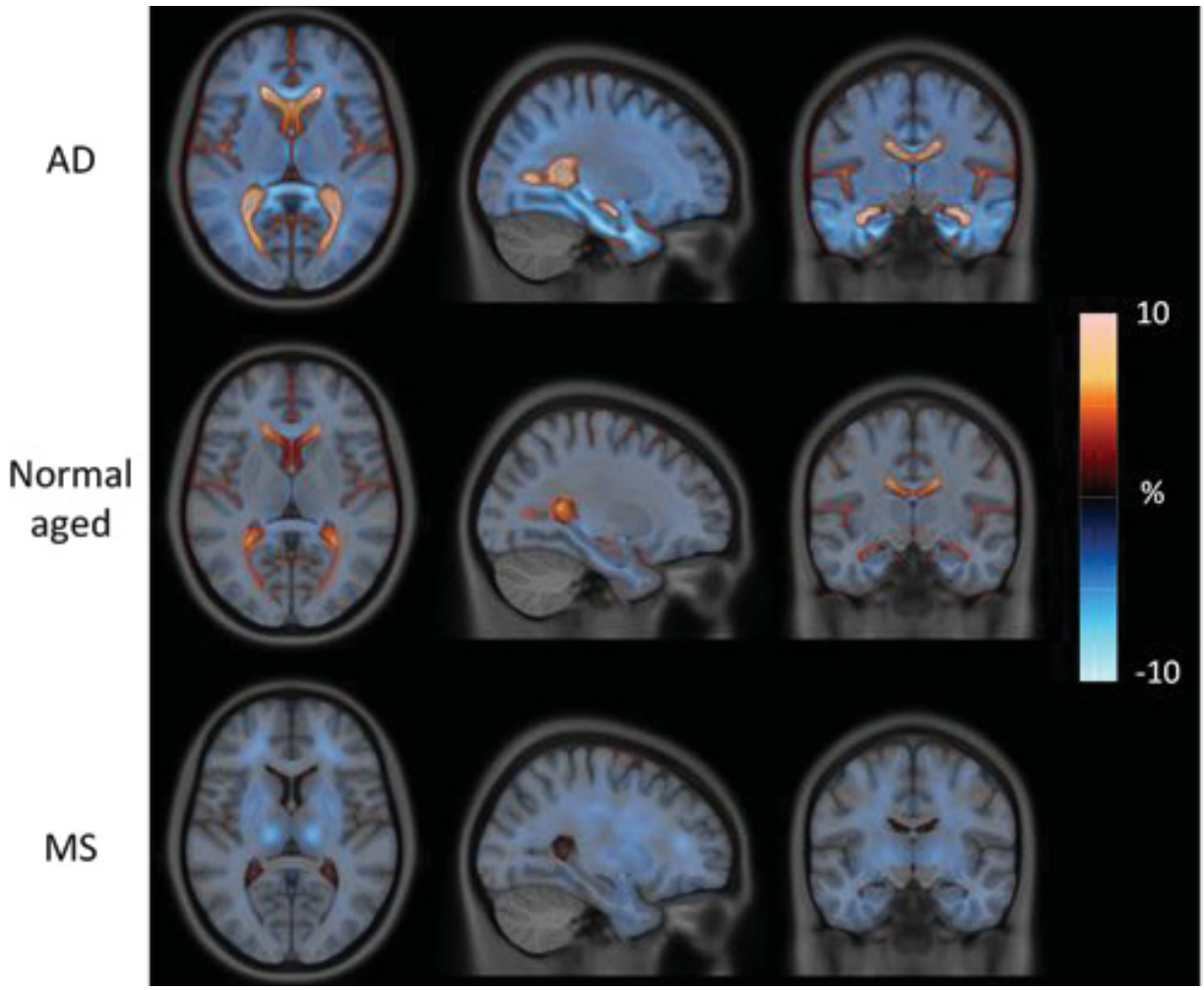


Figure 7.3: Longitudinal DBM annual atrophy rate of 103 AD, 161 age-matched normal control and 55 young MS subjects (red=expansion and blue = contraction or atrophy in percent). The annual atrophy rate was estimated over a period of 36 months and each individual atrophy rate was interpolated to the ICBM152 space before being averaged.

7.5. Conclusion

In conclusion, in the context of neurodegenerative disease, I have developed the necessary methods to detect and remove pathological variability, remove intrinsic image variability and improve the methods to measure brain atrophy. In this thesis, the algorithms developed have been validated and shown to be reliable and robust. Thus, combining the developed workflow to measure focal brain atrophy appears to be a highly promising imaging marker for clinical trials. This work provides the necessary tools to better understand the origin and the diffusion of neurodegeneration, which will in turn provide an invaluable insight of the devastating neurodegenerative diseases.

Appendix 7.A: Poster presented at the OHBM Conference (2011)

Robust fusion of Jacobian maps for Deformation-Based Morphometry

Nicolas Guizard, Pierrick Coupe, Vladimir S. Fonov, Douglas L. Arnold, D. Louis Collins

McConnell Brain Imaging Center, MNI, McGill University, Montreal, CANADA

n.guizard@gmail.com

Introduction

In population studies with Magnetic Resonance Imaging (MRI), Deformation-Based Morphometry (DBM) [1] has received more and more attention as a tool to investigate and identify anatomical differences between groups in cross-sectional studies or anatomical changes over time in longitudinal studies. In DBM methods, statistical analyses of the parameters required to normalize all subjects into a common space enables the localization of even subtle brain shape changes. However, DBM is highly dependant on the non-linear registration accuracy used for normalization. Registration errors due to anatomical singularities such as non-homologous gyri, the presence of lesions or image artifacts may result in outliers in the analysis. While smoothing may reduce the resulting adverse effects of these errors, it may also lead to over-smoothing of the results and reduce localization power.

Therefore, we propose a method to improve the robustness of the DBM method by using a robust patch-based aggregation of Jacobian maps in order to preserve fine local brain changes. The method is evaluated using MRI data from 20 patients with multiple sclerosis (MS).

Methods

- In this study, 20 T1W images (1mm³ resolution) from MS patients, went through the same processing stages: intensity non-uniformity correction [2], intensity normalization [3], linear-registration [4]. The log-transformed Jacobian was computed from the non-linear registration to the ICBM-152 template [5].
- In the context of medical imaging, acquisition artifacts, the presence of pathology as well as inaccuracies in image processing may result in non-normal distribution of errors in estimated parameters. Robust estimations such as median could be used instead of the mean in order to deal with these outliers. In this study, we adapt a robust estimator based on patches in a fashion similar to that recently proposed for MR image denoising and template creation [6,7]. Robust patch-based estimation utilizes similarity of a multi-voxel patch instead of single voxel intensity to detect outliers. Therefore, during the estimation of the population Jacobian, the patch-based robust estimation takes into account local image structure and preserves the fine image registration details.
- We initialize estimation of the population Jacobian $J^0(x)$ by computing the median at each voxel. Then iteratively, until convergence is reached, we estimate the 3D patch distance (L2-norm) between the Jacobian $J_i(x)$ of each subject i and the previous estimation of the population Jacobian $J^{t-1}(x)$; this distance is used to calculate a weighted average which is used as new estimate of the population Jacobian. A normalization factor $N(x)$ is calculated at each iteration as median patch distance. The method is mathematically described in Fig. 1.

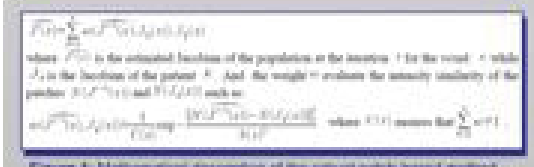


Figure 1: Mathematical description of the robust patch-based method.

Results

A robust DBM analysis is expected to produce a better defined representation of the morphological differences with the template-atrophy due to MS in the case the 20 patients analyzed here, with reduced variance due to elimination of outliers. We compare performance of the classical mean and individual Gaussian smoothing (2mm) average with the proposed robust technique (patch sizes: 1, 2, 3 and 4mm). We measure the mean and the standard deviation of the Jacobian aggregation for each method. Results are shown in Figs 2 and 3. Compared to the two standard approaches (classical mean and Gaussian smoothing), our proposed approach shows a more refined identification of local changes (i.e. ventricles, sulci, ...) with the smaller patches.

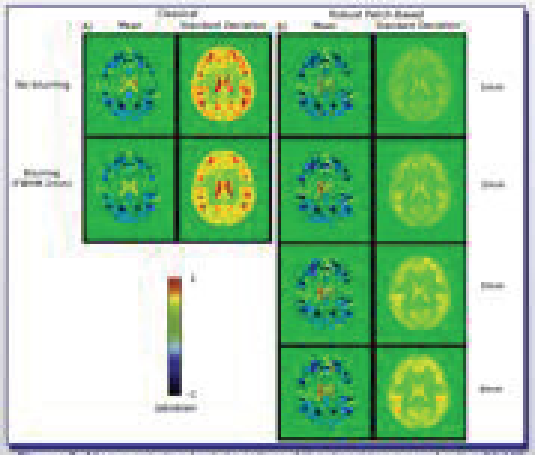


Figure 2: Mean and standard deviation of the Jacobian maps for the 20 MS patients: A) "Classical" mean approach without blurring and blurring (2mm Gaussian kernel) and B) Robust mean with 4 different patch sizes (1, 2, 3 and 4mm).

		Measured
Classical	No blurring	0.42 ± 0.14
	Blurring (FWHM 2mm)	0.38 ± 0.13
Robust Patch Based (radius in mm)	1	0.17 ± 0.06
	2	0.21 ± 0.09
	3	0.24 ± 0.09
	4	0.27 ± 0.10

Figure 3: Whole brain mean and standard deviation (std) results for the different Jacobian aggregation approaches.

Discussion and Conclusion

- We proposed to apply the local patch-based approach in the context of a standard DBM study.
- The estimation of the population Jacobian with our robust patch-based method provides more accurate results compared to the classical mean Jacobian computation and appears to identify finer anatomical differences with the reference image.

[1] Ashburner et al., 1998, "Identifying global anatomical differences: Deformation-based morphometry", Human Brain Mapping
[2] Sled et al., 1998, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data", IEEE TMI
Lui et al., 1999, "On standardizing the MR image intensity scale", Magnetic Resonance in Medicine
[4] Collins et al., 1997, "ANIMAL: Validation and Applications of Nonlinear Registration-Based Segmentation", IJPRAI
[5] Avants et al., 2004, "Geodesic estimation of large deformation anatomical shape averaging and interpolation", Neuroimage
[6] Coupe et al., 2008, "An Optimized Blockwise Non Local Means Denoising Filter for 3-D magnetic Resonance Images", IEEE TMI
[7] Coupe et al., 2010, "Template Construction using a Patch-based Robust Estimator", OHBM 2010 Annual Meeting

Appendix 7.B: Poster presented at the WM study group workshop, ISMRM (2013)

Individual template creation with spatio-temporal regularization for longitudinal atrophy measurement

Nicolas Guizard, Vladimir S. Fonov, Douglas L. Arnold, D. Louis Collins

McConnell Brain Imaging Center, MNI, McGill University, Montreal, CANADA

n.guizard@gmail.com

Introduction

Brain atrophy is a hallmark of neurodegeneration [1]. Manual segmentation whole brain volumes or of the lateral ventricles is inefficient and can suffer from inter- and intra-observer variability. Automatic segmentation methods exist, but are usually performed in a cross-sectional manner where each visit is analyzed independently. Therefore, bias, error and longitudinal noise are introduced. In multiple sclerosis (MS), focal lesions and inflammation can introduce additional noise in the volumetric measurement. To deal with these challenges, we propose 4D non-linear registration with spatio-temporal regularization to correct for potential longitudinal inconsistencies. We compare to previous methods and evaluate the effect of lesions on brain and ventricular volume quantification.

Methods

A total of 268 multi-centre longitudinal MRIs [FLASH: TE=4-10ms, TR=15-30ms, voxel=1.5×1.0×1.0 mm3] with 4 time-points over 3 years from 67 relapsing-remitting MS patients (RRMS) (mean age 37.5, SD 10.0) were used to compare different cross-sectional and longitudinal processing approaches to measure ventricular and brain volumes, segmented using a non-local patch-based approach [2].

Cross-sectional processing (CC): Prior to brain and ventricle segmentation, all MRI data underwent cross-sectional (CC) pre-processing for each time point, including denoising [3], nonuniformity correction with N3 [4] and stereotaxic registration to the ICBM 152 template space [5].

Longitudinal processing (LC, LM and LL): Starting with the CC images, an individual template image was created (Fig. 1) - first using iterative affine registration of each time-point to the previous average [6], and then using non-linear registration [7] (LC). Two additional variants of this longitudinal processing were evaluated. First, a lesion mask was used during the non-linear registration steps to avoid estimating deformations in lesions (LM) to reduce variability due to intensity changes associated with lesions. Second, we exploited temporal information to constrain non-linear deformation by applying a regressor locally at each time-point. This has the effect of improving the longitudinal consistency of the deformation (LL). For each longitudinal technique, the brain and ventricles are segmented on the subject-specific template, and then propagated to each time-point using the estimated non-linear transformations from the template-building step.

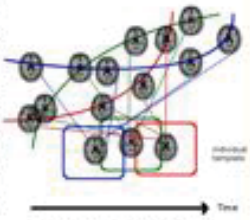


Figure 1: Individual template

Results

Longitudinal pre-processing reduced temporal variability when measuring brain and ventricular volumes compared to the cross sectional pipeline (Fig. 2). While lesion masking improves results in terms of effect size, the LL pipeline improves it even more with an Akaike information criteria (AIC) of 1293 showing that LL is the best pipeline of those compared. Despite a smaller slope with the LL approach, power analysis demonstrated it requires fewer subjects to estimate a 20% or 30% treatment effect with a 80% power (Table 1). Furthermore, deformation-based morphometry (DBM) shows more uniform and plausible anatomical atrophy patterns (Fig. 3).

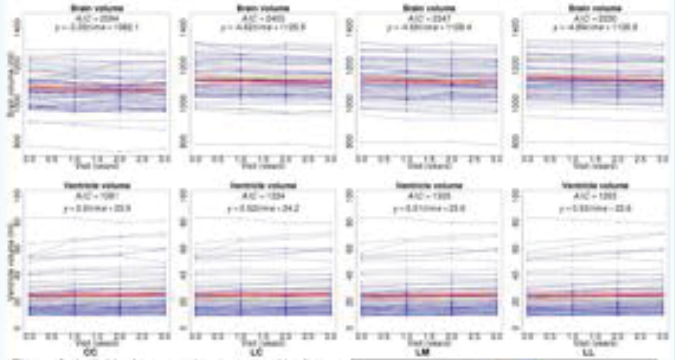


Figure 2: Individual temporal trajectory with linear mixed effects model and confidence interval (95%).

Treatment Effect (%)	20	30	20	30
Method	Brain		Ventricles	
CC	1800	800	426	189
LC	309	137	487	217
LM	262	116	449	199
LL	135	60	390	173

Table 1: Sample size estimation (80% power)

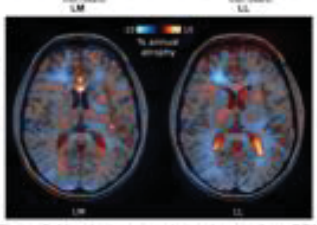


Figure 3: Voxel-wise annual atrophy rate from DBM

Conclusion

The proposed 4D constrained registration approach (LL) is less sensitive to temporal noise in estimating longitudinal brain and ventricular changes for MS patients than conventional cross-sectional segmentation approaches with and without lesion masking, reducing sample sizes required to detect a treatment effect on brain atrophy by an order of magnitude.

References: 1. Chard, 2003, JNNP. 2. Coupé, 2011, NeuroImage. 3. Coupé, 2008, TMI. 4. Sled, 1998, TMI. 5. Collins, 1994, JCAT. 6. Fonov, 2011, NeuroImage. 7. Collins, 1997, IJPRAI

Acknowledgements: We thank Teva Pharmaceutical for the use of MRI data.

#10

OTHER PUBLICATIONS

The following publications were published during the course of this thesis:

Journal articles:

David M. Cash, Chris Frost, Leonardo O. Iheme, Devrim Ünay, Melek Kandemir, Jurgen Fripp, Olivier Salvado, Pierrick Bourgeat, Martin Reuter, Bruce Fischl, Marco Lorenzi, Giovanni B. Frisoni, Xavier Pennec, Ronald Pierson, Jeffrey L. Gunter, Matthew L. Senjem, Clifford R. Jack Jr, **Nicolas Guizard**, Vladimir S. Fonov, D. Louis Collins, Marc Modat, M. Jorge Cardoso, Kelvin K. Leung, Hongzhi Wang, Sandhitsu R. Das, Paul A. Yushkevich, Nick C. Fox, Jonathan M. Schott, Sebastien Ourselin, "Assessing atrophy measurement techniques in dementia: Results from the MIRIAD Atrophy Challenge", NeuroImage, 2015, doi: 10.1016/j.neuroimage.2015.07.087

Rina Zelmann, Silvain Beriault, Morilo M Marinho, Kelvin Mok, Jeffery A. Hall, **Nicolas Guizard**, Claire Haegelen, Andre Olivier, G. Bruce Pike, D. Louis Collins. "Improving Recorded Volume in Mesial Temporal Lobe by Optimizing Stereotactic Intracranial Electrode Implantation Planning", IJCARS, 2015, doi:10.1007/s11548-015-1165-6

Yimming Xiao, Pierre Jannin, Tiziano D'Albis, **Nicolas Guizard**, Claire Haegelen, Florent Lalys, Marc Verin, D. Louis Collins. "Investigation of morphometric variability of subthalamic nucleus, red nucleus and substantia nigra in advanced Parkinson's disease patients using automatic segmentation and PCA-based analysis", HBM, 2014, doi: 10.1002/hbm.22478

Kunio Nakamura, **Nicolas Guizard**, Vladimir S. Fonov, Sridar Narayanan, D. Louis Collins, Douglas L. Arnold. "Jacobian integration method increases the statistical

power to measure gray matter atrophy in multiple sclerosis", NeuroImage, 2014, doi: 10.1016/j.nicl.2013.10.015

Bérengère M Aubert-Broche, Vladimir Fonov, Daniel Garcia-Lorenzo, Abderazzak Mouiha, **Nicolas Guizard**, Pierrick Coupé, Simon Eskildsen, D Louis Collins. "A new method for structural volume analysis of longitudinal brain MRI data and its application in studying the growth trajectories of anatomical brain structures in childhood", NeuroImage, 2013, doi: 10.1016/j.neuroimage.2013.05.065

Catherine Limperopoulos, Gevorg Chilingaryan, Nancy Sullivan, **Nicolas Guizard**, Richard L. Robertson, Adré J. du Plessis "Injury to the Premature Cerebellum: Outcome is Related to Remote Cortical Development", Cerebral Cortex, 2012, doi: 10.1093/cercor/bhs354

Claire Haegelen, Pierrick Coupé, Vladimir Fonov, **Nicolas Guizard**, Pierre Jannin, Xavier Morandi, D. Louis Collins. "Automated segmentation of basal ganglia and deep brain structures in MRI of Parkinson's disease", IJCARS, 2013, doi: 10.1007/s11548-012-0675-8

Simon F. Eskildsen, Pierrick Coupé, Kevin K. Leung, Vladimir Fonov, **Nicolas Guizard**, Shafik N. Wassef, Lasse R. Østergaard, D. Louis Collins, "BEaST: Brain Extraction based on non-local Segmentation Technique", NeuroImage, 2012, doi: 10.1016/j.neuroimage.2011.09.01

Peer-reviewed conference papers:

Bérengère Aubert-Broche, Vladimir Fonov, Daniel García-Lorenzo, Abderazzak Mouiha, **Nicolas Guizard**, Pierrick Coupé, Simon Eskildsen, Louis D. Collins. "A New Framework for Analyzing Structural Volume Changes of Longitudinal Brain MRI Data", Spatio-temporal Image Analysis for Longitudinal and Time-Series Image Data, 2012

Simon F. Eskildsen, Pierrick Coupé, Kevin K. Leung, Vladimir Fonov, **Nicolas Guizard**, Shafik N. Wassef, Lasse R. Østergaard, D. Louis Collins. "BEaST: Brain extraction using multiresolution non-local segmentation", Multi-Atlas Labeling and Statistical Fusion Workshop, Medical Image Computing and Computer Assisted Intervention Conference (MICCAI), Toronto, Canada, 2011

Nicolas Guizard, Pierrick Coupé, Nicolas Stifani, Stefano Stifani, D. Louis Collins, "Robust 3D Reconstruction and Mean-Shift Clustering of Motoneurons from Serial Histological Images", Medical Imaging and Augmented Reallity (MIAR), MICCAI, 2010, Proceedings, Beijing **(Best paper award)**

Conference abstracts:

Nicolas Guizard, Kunio Nakamura, Pierrick Coupé, Douglas L. Arnold, D. Louis Collins. "Non-local MS MRI lesion inpainting method for image processing", endMS conference, Saint-Sauveur, Canada, 2013

Kunio Nakamura, **Nicolas Guizard**, Vladimir S. Fonov, Sridar Narayanan, D. Louis Collins, Douglas L. Arnold . "Evaluation of Brain atrophy measurements methods: an MRI simulation study", endMS conference, Saint-Sauveur, Canada, 2013 (Best paper award)

Nicolas Guizard, Vladimir Fonov, Doug L. Arnold, D. Louis Collins. "Individual template creation with spatio-temporal regularization for longitudinal RRMS atrophy measure", ISMRM Multiple Sclerosis workshop, London, UK, 2013

Kunio Nakamura, **Nicolas Guizard**, Vladimir Fonov, Sridar Narayanan, D. Louis Collins, Douglas L. Arnold. "Pairwise Jacobian Integration Method to Measure Grey Matter Atrophy in Multiple Sclerosis.", ISMRM Multiple Sclerosis workshop, London, UK, 2013

Bérengère Aubert-Broche, Vladimir S. Fonov, **Nicolas Guizard**, Brenda Banwell, Sridar Narayanan, Douglas L. Arnold, D. Louis Collins. "Brain growth rate is reduced in paediatric-onset multiple sclerosis", European committees for treatment and research in multiple sclerosis, Lyon, France, 2012

Nicolas Guizard, Marc Verin, Vladimir Fonov, Douglas L. Arnold, D. Louis Collins, "4D non-linear registration for longitudinal analyses of ventricle enlargement in

RRMS patients", European and Americas committees for treatment and research in multiple sclerosis, Amsterdam, Netherland, 2011

Nicolas Guizard, Pierrick Coupé, Vladimir Fonov, Douglas L. Arnold, D. Louis Collins "Robust fusion of Jacobian maps for Deformation-Based Morphometry", Organization for of Human Brain mapping (OHBM), June 2011, Québec, Canada

Claire Haegelen, **Nicolas Guizard**, Pierrick Coupé, Florian Lalys, Pierre Jannin, Xaxier Morandi, D. Louis Collins. "Validation of basal ganglia segmentation on a 3T MRI Template", OHBM, June 2011, Québec, Canada

M. Bouyssi-Kobar, C. Clouchoux, N. Guizard, C. Limperopoulos. "4-D In Vivo Fetal MRI Brain Atlas", Young Europeen Scientist Meeting, Lisbon, 2010.

Nicolas Stifani, **Nicolas Guizard**, D. Louis Collins, Stefano Stifani, "Expression of the transcription factor RunX1 in spinal mortor neurons relevant to ALS", ALS Canada Sixth Annual Research Forum, 2010

Nicolas Guizard, Vladimir Fonov, D. Louis Collins, "Symmetric Optimization Scheme versus Constrained Symmetrization for Non-Linear Registrations", Medical Image Analysis on Multiple Sclerosis (MIAMS) Workshop, MICCAI, September 2009, London, UK

REFERENCES

(1998). "Burden of illness of multiple sclerosis: Part I: Cost of illness. The Canadian Burden of Illness Study Group." The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques **25**(1): 23-30.

(2012). A new way of looking at the impact of dementia in Canada, Alzheimer Society of Canada.

(2014). "World Alzheimer Report 2014."

Aljabar, P., R. A. Heckemann, A. Hammers, J. V. Hajnal and D. Rueckert (2009). "Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy." NeuroImage **46**(3): 726-738.

Amato, M., E. Portaccio, B. Goretti, V. Zipoli, A. Iudice, D. Della Pina, G. Malentacchi, S. Sabatini, P. Annunziata, M. Falcini, M. Mazzoni, M. Mortilla, C. Fonda and N. De Stefano (2010). "Relevance of cognitive deterioration in early relapsing-remitting MS: a 3-year follow-up study." Multiple sclerosis **16**(12): 1474-1482.

Anderson, V. M., J. W. Bartlett, N. C. Fox, L. Fisniku and D. H. Miller (2007). "Detecting treatment effects on brain atrophy in relapsing remitting multiple sclerosis: Sample size estimates." Journal of Neurology **254**(11): 1588-1594.

Arnold, D. L., P. M. Matthews, G. Francis and J. Antel (1990). "Proton magnetic resonance spectroscopy of human brain in vivo in the evaluation of multiple sclerosis: assessment of the load of disease." Magnetic Resonance in Medicine **14**(1): 154-159.

Arsigny, V., O. Commowick, X. Pennec and N. Ayache (2006). "A Log-Euclidean Framework for Statistics on Diffeomorphisms." Medical Image Computing and Computer-Assisted Intervention: MICCAI **4190**: 924-931.

Ashburner, J. and K. Friston (2000). "Voxel-Based Morphometry—The Methods." NeuroImage **11**(6): 805-821.

Ashburner, J. and K. Friston (2011). "Diffeomorphic registration using geodesic shooting and Gauss-Newton optimisation." NeuroImage **55**(3): 954-967.

Ashburner, J., C. Hutton, R. Frackowiak, I. Johnsrude, C. Price and K. Friston (1998). "Identifying global anatomical differences: Deformation-based morphometry." Human Brain Mapping **6**(5-6): 348-357.

Ashburner, J. and G. Ridgway (2012). “Symmetric diffeomorphic modeling of longitudinal structural MRI.” Frontiers in Neuroscience **6:197**.

Aubert-Broche, B., V. S. Fonov, D. García-Lorenzo, A. Mouiha, N. Guizard, P. Coupé, S. F. Eskildsen and D. L. Collins (2013). “A new method for structural volume analysis of longitudinal brain MRI data and its application in studying the growth trajectories of anatomical brain structures in childhood.” NeuroImage **82C**: 393-402.

Aubert-Broche, B., M. Griffin, G. B. Pike, A. C. Evans and D. L. Collins (2006). “Twenty new digital brain phantoms for creation of validation image data bases.” IEEE Transactions on Medical Imaging **25**(11): 1410-1416.

Audoin, B., G. Davies, L. Finisku, D. Chard, A. Thompson and D. Miller (2006). “Localization of grey matter atrophy in early RRMS.” Journal of Neurology **253**(11): 1495-1501.

Avants, B., M. Grossman and J. Gee (2006). Symmetric Diffeomorphic Image Registration: Evaluating Automated Labeling of Elderly and Neurodegenerative Cortex and Frontal Lobe. Biomedical Image Registration, Springer**:** 50-57.

Avants, B. B., C. L. Epstein, M. Grossman and J. C. Gee (2008). “Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain.” Medical Image Analysis **12**(1): 26-41.

Bağcı, U., J. K. Udupa and L. Bai (2010). “The role of intensity standardization in medical image registration.” Pattern Recognition Letters **31**(4): 315-323.

Bai, W., W. Shi, D. P. O'Regan, T. Tong, H. Wang, S. Jamil-Copley, N. S. Peters and D. Rueckert (2013). “A Probabilistic Patch-Based Label Fusion Model for Multi-Atlas Segmentation With Registration Refinement: Application to Cardiac MR Images.” IEEE Transactions on Medical Imaging **32**(7): 1302-1315.

Baker, S. and T. Kanade (2000). “Hallucinating faces.” International Conference on Automatic Face and Gesture Recognition: 83-88.

Bakshi, R., V. S. Dandamudi, M. Neema, C. De and R. A. Bermel (2005). “Measurement of brain and spinal cord atrophy by magnetic resonance imaging as a tool to monitor multiple sclerosis.” Journal of Neuroimaging **15**(4 Suppl).

Barkhof, F. (1999). “MRI in multiple sclerosis: correlation with expanded disability status scale (EDSS).” Multiple sclerosis **5**(4): 283-286.

Barkhof, F. J., M. Elton, J. Lindeboom, M. W. Tas, W. F. Schmidt, O. R. Hommes, C. H. Polman, A. Kok and J. Valk (1998). “Functional correlates of callosal atrophy in relapsing-remitting multiple sclerosis patients. A preliminary MRI study.” Journal of neurology **245**(3): 153-158.

Barnes, J., J. W. Bartlett, L. A. van de Pol, C. T. Loy, R. I. Scahill, C. Frost, P. Thompson and N. C. Fox (2009). "A meta-analysis of hippocampal atrophy rates in Alzheimer's disease." Neurobiology of aging **30**(11): 1711-1723.

Barnes, J., J. Foster, R. G. Boyes, T. Pepple, E. K. Moore, J. M. Schott, C. Frost, R. I. Scahill and N. C. Fox (2008). "A comparison of methods for the automated calculation of volumes and atrophy rates in the hippocampus." NeuroImage **40**(4): 1655-1671.

Barnes, J., A. K. Godbolt, C. Frost, R. G. Boyes, B. F. Jones, R. I. Scahill, M. N. Rossor and N. C. Fox (2007). "Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD." Neurobiology of Aging **28**(1): 20-28.

Baron, J. C., G. Chételat, B. Desgranges, G. Perchey, B. Landeau, V. de la Sayette and F. Eustache (2001). "In vivo mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease." NeuroImage **14**(2): 298-309.

Battaglini, M., A. Giorgio, M. Stromillo, M. Bartolozzi, L. Guidi, A. Federico and N. De Stefano (2009). "Voxel-wise assessment of progression of regional brain atrophy in relapsing-remitting multiple sclerosis." Journal of the Neurological Sciences **282**(1-2): 55-60.

Battaglini, M., M. Jenkinson and N. De Stefano (2012). "Evaluating and reducing the impact of white matter lesions on brain volume measurements." Human Brain Mapping **33**(9): 2062-2071.

Bazin, P.-L. and D. L. Pham (2008). "Homeomorphic brain image segmentation with topological and statistical atlases." Medical Image Analysis **12**(5): 616-625.

Bellon, E. M., E. M. Haacke, P. E. Coleman, D. C. Sacco, D. A. Steiger and R. E. Gangarosa (1986). "MR artifacts: a review." American journal of neuroradiology: AJNR **147**(6): 1271-1281.

Bendfeldt, K., S. Klöppel, T. E. Nichols, R. Smieskova, P. Kuster, S. Traud, N. Mueller-Lenke, Y. Naegelin, L. Kappos, E.-W. Radue and S. J. Borgwardt (2012). "Multivariate pattern classification of gray matter pathology in multiple sclerosis." NeuroImage **60**(1): 400-408.

Benedict, R., R. Zivadinov, D. Carone, B. Weinstock-Guttman, J. Gaines, C. Maggiore, J. Sharma, M.-A. Tomassi and R. Bakshi (2005). "Regional lobar atrophy predicts memory impairment in multiple sclerosis." American journal of neuroradiology: AJNR **26**(7): 1824-1831.

Bermel, R. and R. Bakshi (2006). "The measurement and clinical relevance of brain atrophy in multiple sclerosis." Lancet Neurology **5**(2): 158-170.

Bermel, R., J. Sharma, C. Tjoa, S. Puli and R. Bakshi (2003). "A semiautomated measure of whole-brain atrophy in multiple sclerosis." Journal of the Neurological Sciences **208**(1-2): 57-65.

Bertalmío, M., V. Caselles, S. Masnou and G. Sapiro (2014). Inpainting. Computer Vision. K. Ikeuchi, Springer US**:** 401-416.

Bitsch, A., T. Kuhlmann, C. Stadelmann, H. Lassmann, C. Lucchinetti and W. Brück (2001). "A longitudinal MRI study of histopathologically defined hypointense multiple sclerosis lesions." Annals of Neurology **49**(6): 793-796.

Bjartmar, C. and B. Trapp (2003). "Axonal degeneration and progressive neurologic disability in multiple sclerosis." Neurotoxicity research **5**(1-2): 157-164.

Bookstein, F. L. (1989). "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations." IEEE Transactions on Pattern Analysis and Machine Intelligence **11**(6): 567-585.

Borgefors, G. (1988). "Hierarchical chamfer matching: a parametric edge matching algorithm." IEEE Transactions on Pattern Analysis and Machine Intelligence **10**(6): 849-865.

Bosc, M., F. Heitz, J.-P. Armspach, I. Namer, D. Gounot and L. Rumbach (2003). "Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution." NeuroImage **20**(2): 643-656.

Braak, H. and E. Braak (1991). "Neuropathological stageing of Alzheimer-related changes." Acta Neuropathologica **82**(4): 239-259.

Bredesen, D. E., R. V. Rao and P. Mehlen (2006). "Cell death in the nervous system." Nature **443**(7113): 796-802.

Brett, M. (2001). "Spatial Normalization of Brain Images with Focal Lesions Using Cost Function Masking." NeuroImage **14**(2): 486-500.

Bridcut, R., T. Redpath, C. Gray and R. Staff (2001). "The use of SPAMM to assess spatial distortion due to static field inhomogeneity in dental MRI." Physics in Medicine and Biology **46**(5): 1357-1367.

Brinkmann, B. H., A. Manduca and R. Robb (1998). "Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction." IEEE Transactions on Medical Imaging **17**(2): 161-171.

Bro-Nielsen, M. and C. Gramkow (1996). Fast Fluid Registration of medical images. Visualization in Biomedical Computing. K. Höhne and R. Kikinis, Springer Berlin Heidelberg. **1131:** 265-276.

Broit, C. (1981). Optimal registration of deformed images, University of Pennsylvania, Philadelphia, United States.

Brookmeyer, R., E. Johnson, K. Ziegler-Graham and H. M. Arrighi (2007). "Forecasting the global burden of Alzheimer's disease." Alzheimer's & Dementia **3**(3): 186-191.

Brück, W., A. Bitsch, H. Kolenda, Y. Brück, M. Stiefel and H. Lassmann (1997). "Inflammatory central nervous system demyelination: correlation of magnetic resonance imaging findings with lesion pathology." Annals of Neurology **42**(5): 783-793.

Buades, A., B. Coll and J. M. Morel (2005). "A non-local algorithm for image denoising." Computer Vision and Pattern Recognition: CVPR **2**: 60-65.

Burton, E., I. McKeith, D. Burn, D. Williams and J. O'Brien (2004). "Cerebral atrophy in Parkinson's disease with and without dementia: a comparison with Alzheimer's disease, dementia with Lewy bodies and controls." Brain **127**(4): 791-800.

Burton, E. J., R. Barber, E. B. Mukaetova-Ladinska, J. Robson, R. H. Perry, E. Jaros, R. N. Kalaria and J. T. O'Brien (2009). "Medial temporal lobe atrophy on MRI differentiates Alzheimer's disease from dementia with Lewy bodies and vascular cognitive impairment: a prospective study with pathological verification of diagnosis." Brain **132**(1): 195-203.

Cabezas, M., A. Oliver, X. Lladó, J. Freixenet and M. Bach Cuadra (2011). "A review of atlas-based segmentation for magnetic resonance brain images." Computer Methods and Programs in Biomedicine **104**(3): e158-e177.

Cabezas, M., A. Oliver, E. Roura, J. Freixenet, J. C. Vilanova, L. Ramió-Torrentà, À. Rovira and X. Lladó (2014). "Automatic multiple sclerosis lesion detection in brain MRI by FLAIR thresholding." Computer Methods and Programs in Biomedicine **115**(3): 147-161.

Calabrese, M., F. Agosta, F. Rinaldi and et al. (2009). "Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis." Archives of Neurology **66**(9): 1144-1150.

Caramanos, Z., V. Fonov, S. Francis, S. Narayanan, B. Pike, L. Collins and D. Arnold (2010). "Gradient distortions in MRI: Characterizing and correcting for their effects on SIENA-generated measures of brain volume change." NeuroImage **49**(2): 1601-1611.

Caramanos, Z., S. J. Francis, S. Narayanan, Y. Lapierre and D. L. Arnold (2012). "Large, nonplateauing relationship between clinical disability and cerebral white matter lesion load in patients with multiple sclerosis." Archives of Neurology **69**(1): 89-95.

Cardenas, V. A., L. L. Chao, C. Studholme, K. Yaffe, B. L. Miller, C. Madison, S. T. Buckley, D. Mungas, N. Schuff and M. W. Weiner (2011). "Brain atrophy associated with baseline and longitudinal measures of cognition." Neurobiology of Aging **32**(4): 572-580.

Casanova, B., M. C. Martínez-Bisbal, C. Valero, B. Celda, L. Martí-Bonmatí, A. Pascual, L. Landente and F. Coret (2003). "Evidence of Wallerian degeneration in normal appearing white

matter in the early stages of relapsing-remitting multiple sclerosis: a HMRS study." Journal of Neurology **250**(1): 22-28.

Cash, D., C. Frost, L. O. Iheme, D. Ünay, M. Kandemir, J. Fripp, O. Salvado, P. Bourgeat, M. Reuter, B. Fischl, M. Lorenzi, G. B. Frisoni, X. Pennec, R. Pierson, J. L. Gunter, M. L. Senjem, J. C. R. Jack, N. Guizard, V. S. Fonov, D. L. Collins, M. Modat, M. J. Cardoso, K. K. Leung, H. Wang, S. R. Das, P. A. Yushkevich, N. C. Fox, J. M. Schott and S. Ourselin (Submitted). "Assessing atrophy measurement techniques in dementia: Results from the MIRIAD Atrophy Challenge." NeuroImage.

Ceccarelli, A., J. S. Jackson, S. Tauhid, A. Arora, J. Gorky, E. Dell'Oglio, A. Bakshi, T. Chitnis, S. J. Khoury, H. L. Weiner, C. R. Guttmann, R. Bakshi and M. Neema (2012). "The impact of lesion in-painting and registration methods on voxel-based morphometry in detecting regional cerebral gray matter atrophy in multiple sclerosis." American journal of neuroradiology: AJNR **33**(8): 1579-1585.

Ceccarelli, A., M. Rocca, E. Pagani, B. Colombo, V. Martinelli, G. Comi and M. Filippi (2008). "A voxel-based morphometry study of grey matter loss in MS patients with different clinical phenotypes." NeuroImage **42**(1): 315-322.

Chakravarty, M., G. Bertrand, C. Hodge, A. Sadikot and L. Collins (2006). "The creation of a brain atlas for image guided neurosurgery using serial histological data." NeuroImage **30**(2): 359-376.

Charcot, J. M. (1868). Histologie de la sclérose en plaques.

Chard, D. T., P. A. Brex, O. Ciccarelli, C. M. Griffin, G. J. Parker, C. Dalton, D. R. Altmann, A. J. Thompson and D. H. Miller (2003). "The longitudinal relation between brain lesion load and atrophy in multiple sclerosis: a 14 year follow up study." Journal of Neurology, Neurosurgery, and Psychiatry **74**(11): 1551-1554.

Chard, D. T., C. M. Griffin, G. J. Parker, R. Kapoor, A. J. Thompson and D. H. Miller (2002). "Brain atrophy in clinically early relapsing-remitting multiple sclerosis." Brain **125**(Pt 2): 327-337.

Chard, D. T., J. S. Jackson, D. H. Miller and C. A. M. Wheeler-Kingshott (2010). "Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes." Journal of Magnetic Resonance Imaging **32**(1): 223-228.

Charil, A., A. Dagher, J. Lerch, A. Zijdenbos, K. Worsley and A. Evans (2007). "Focal cortical atrophy in multiple sclerosis: Relation to lesion load and disability." NeuroImage **34**(2): 509-517.

Chaudhuri, A. (2013). "Multiple sclerosis is primarily a neurodegenerative disease." Journal of Neural Transmission **120**(10): 1463-1466.

Cheng, J., L. Edwards, M. Maldonado-Molina, K. Komro and K. Muller (2010). "Real longitudinal data analysis for real people: building a good enough mixed model." Statistics in medicine **29**(4): 504-520.

Chételat, G., B. Desgranges, B. Landeau, F. Mézenge, J. B. Poline, S. de la, F. Viader, F. Eustache and J. C. Baron (2008). "Direct voxel-based comparison between grey matter hypometabolism and atrophy in Alzheimer's disease." Brain **131**(1): 60-71.

Chetelat, G., B. Landeau, F. Eustache, F. Mezenge, F. Viader, V. de la Sayette, B. Desgranges and J. C. Baron (2005). "Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: a longitudinal MRI study." Neuroimage **27**(4): 934-946.

Chiaravalloti, N. and J. DeLuca (2008). "Cognitive impairment in multiple sclerosis." The Lancet. Neurology **7**(12): 1139-1151.

Christensen, G., X. Geng, J. Kuhl, J. Bruss, T. Grabowski, J. Allen, I. Pirwani, M. Vannier and H. Damasio (2006). Introduction to the Non-Rigid Image Registration Evaluation Project (NIREP).

Christensen, G. E. and H. J. Johnson (2001). "Consistent image registration." IEEE Transactions on Medical Imaging **20**(7): 568-582.

Christensen, G. E., R. D. Rabbitt and M. I. Miller (1996). "Deformable templates using large deformation kinematics." IEEE Transactions on Image Processing **5**(10): 1435-1447.

Chung, M. (2003). "Deformation-based surface morphometry applied to gray matter deformation." NeuroImage **18**(2): 198-213.

Chupin, M., A. R. Mukuna-Bantumbakulu, D. Hasboun, E. Bardinet, S. Baillet, S. Kinkingnéhun, L. Lemieux, B. Dubois and L. Garnero (2007). "Anatomically constrained region deformation for the automated segmentation of the hippocampus and the amygdala: Method and validation on controls and patients with Alzheimer's disease." NeuroImage **34**(3): 996-1019.

Clark, C. M., G. James, D. Li, J. Oger, D. Paty and H. Klonoff (1992). "Ventricular size, cognitive function and depression in patients with multiple sclerosis." The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques **19**(3): 352-356.

Coffey, C. E., W. E. Wilkinson, L. Parashos, S. A. R. Soady, R. J. Sullivan, L. J. Patterson, G. S. Figiel, M. C. Webb, C. E. Spritzer and W. T. Djang (1992). "Quantitative cerebral anatomy of the aging human brain." Neurology **42**(3): 527-527.

Collignon, A., F. Maes, D. Delaere, D. Vandermeulen, P. Suetens and G. Marchal (1995). "Automated multi-modality image registration based on information theory." Internation

Conference Information Processing in Medical Imaging; Computational Imaging and Vision: 263–274.

Collins, C. M., S. Li, Q. X. Yang and M. B. Smith (1997). "A Method for Accurate Calculation of B 1 Fields in Three Dimensions. Effects of Shield Geometry on Field Strength and Homogeneity in the Birdcage Coil." Journal of Magnetic Resonance **125**(2): 233-241.

Collins, D., J. Montagnat, A. Zijdenbos, A. Evans and D. Arnold (2001). Automated Estimation of Brain Volume in Multiple Sclerosis with BICCR**:** 141-147.

Collins, D. L. and A. C. Evans (1997). "ANIMAL: Validation and Applications of Non-Linear Registration-Based Segmentation." International Journal of Pattern Recognition and Artificial Intelligence **11**: 1271-1294.

Collins, D. L., P. Neelin, T. M. Peters and A. C. Evans (1994). "Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space." Journal of Computer Assisted Tomography **18**(2): 192-205.

Collins, D. L., A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes and A. C. Evans (1998). "Design and construction of a realistic digital brain phantom." IEEE Transactions on Medical Imaging **17**(3): 463-468.

Collins, L. and A. C. Evans (1997). "ANIMAL: Validation and Applications of Non-Linear Registration-Based Segmentation." International Journal of Pattern Recognition and Artificial Intelligence **11**: 1271-1294.

Collins, L., C. J. Holmes, T. M. Peters and A. C. Evans (1995). "Automatic 3-D model-based neuroanatomical segmentation." Human Brain Mapping **3**(3): 190-208.

Collins, L., P. Neelin, T. Peters and A. Evans (1994). "Automatic 3D Intersubject Registration of MR Volumetric Data in Standardized Talairach Space." Journal of Computer Assisted Tomography **18**: 192-205.

Collins, L. and J. Pruessner (2010). "Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion." NeuroImage **52**(2): 1355-1366.

Commowick, O., R. Stefanescu, P. Fillard, V. Arsigny, N. Ayache, X. Pennec and G. Malandain (2005). "Incorporating statistical measures of anatomical variability in atlas-to-subject registration for conformal brain radiotherapy." Medical Image Computing and Computer-Assisted Intervention: MICCAI: 927-934.

Compston, A. and A. Coles (2008). "Multiple sclerosis." Lancet **372**(9648): 1502-1517.

Confavreux, C. and S. Vukusic (2006). "Accumulation of irreversible disability in multiple sclerosis: from epidemiology to treatment." Clinical neurology and neurosurgery **108**(3): 327-332.

Confavreux, C. and S. Vukusic (2008). "The clinical epidemiology of multiple sclerosis." Neuroimaging Clinics of North America **18**(4).

Cordier, N., B. Menze, H. Delingette and N. Ayache (2013). "Patch-based Segmentation of Brain Tissues." MICCAI Challenge on Multimodal Brain Tumor Segmentation: 6 - 17.

Coupé, P., J. Manjón, M. Chamberland, M. Descoteaux and B. Hiba (2013). "Collaborative patch-based super-resolution for diffusion-weighted images." NeuroImage **83**: 245-261.

Coupé, P., J. Manjón, V. Fonov, J. Pruessner, M. Robles and L. Collins (2011). "Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation." NeuroImage **54**(2): 940-954.

Coupé, P., J. Manjón, E. Gedamu, D. Arnold, M. Robles and L. Collins (2010). "Robust Rician noise estimation for MR images." Medical Image Analysis **14**(4): 483-493.

Coupé, P., P. Yger, S. Prima, P. Hellier, C. Kervrann and C. Barillot (2008). "An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images." IEEE Transactions on Medical Imaging **27**(4): 425-441.

Criminisi, A., P. Perez and K. Toyama (2004). "Region filling and object removal by exemplar-based image inpainting." IEEE Transactions on Image Processing **13**(9): 1200-1212.

Crimins, J., A. Pooler, M. Polydoro, J. Luebke and T. Spires-Jones (2013). "The intersection of amyloid β and tau in glutamatergic synaptic dysfunction and collapse in Alzheimer's disease." Ageing research reviews **12**(3): 757-763.

Crum, W. R., R. I. Scahill and N. C. Fox (2001). "Automated Hippocampal Segmentation by Regional Fluid Registration of Serial MRI: Validation and Application in Alzheimer's Disease." NeuroImage **13**(5): 847-855.

Dalton, C., D. Chard, G. Davies, K. Miszkiel, D. Altmann, K. Fernando, G. Plant, A. Thompson and D. Miller (2004). "Early development of multiple sclerosis is associated with progressive grey matter atrophy in patients presenting with clinically isolated syndromes." Brain **127**(5): 1101-1107.

Damoiseaux, J. (2012). "Resting-state fMRI as a biomarker for Alzheimer's disease?" Alzheimer's Reseach & Therapy **4**(2).

Dastidar, P., T. Heinonen, T. Lehtimäki, M. Ukkonen, J. Peltola, T. Erilä, E. Laasonen and I. Elovaara (1999). "Volumes of brain atrophy and plaques correlated with neurological disability in secondary progressive multiple sclerosis." Journal of the Neurological Sciences **165**(1): 36-42.

Davatzikos, C. (1997). "Spatial Transformation and Registration of Brain Images Using Elastically Deformable Models." Computer Vision and Image Understanding **66**(2): 207-222.

Davis, B. C., P. T. Fletcher, E. Bullitt and S. Joshi (2007). "Population Shape Regression from Random Design Data." International Conference on Computer Vision: ICCV: 1-7.

Dawant, B. M., A. P. Zijdenbos and R. A. Margolin (1993). "Correction of intensity variations in MR images for computer-aided tissue classification." IEEE Transactions on Medical Imaging **12**(4): 770-781.

de Leeuw, F. E., J. C. de Groot, E. Achten, M. Oudkerk, L. M. P. Ramos, R. Heijboer, A. Hofman, J. Jolles, J. van Gijn and M. M. B. Breteler (2001). "Prevalence of cerebral white matter lesions in elderly people: a population based magnetic resonance imaging study. The Rotterdam Scan Study." Journal of Neurology, Neurosurgery & Psychiatry **70**(1): 9-14.

De Leon, M. J. and H. Braak (1999). An atlas of Alzheimer's disease, Parthenon Pub. Group.

De Stefano, N., A. Giorgio, M. Battaglini, M. Rovaris, M. P. Sormani, F. Barkhof, T. Korteweg, C. Enzinger, F. Fazekas, M. Calabrese, D. Dinacci, G. Tedeschi, A. Gass, X. Montalban, A. Rovira, A. Thompson, G. Comi, D. H. Miller and M. Filippi (2010). "Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes." Neurology **74**(23): 1868-1876.

Derakhshan, M., Z. Caramanos, P. S. Giacomini, S. Narayanan, J. Maranzano, S. J. Francis, D. L. Arnold and D. L. Collins (2010). "Evaluation of automated techniques for the quantification of grey matter atrophy in patients with multiple sclerosis." NeuroImage **52**(4): 1261-1267.

Desolneux, A., L. Moisan and J.-M. More (2003). "A grouping principle and four applications." IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(4): 508-513.

Dickerson, B., A. Bakkour, D. Salat, E. Feczko, J. Pacheco, D. Greve, F. Grodstein, C. Wright, D. Blacker, D. Rosas, R. Sperling, A. Atri, J. Growdon, B. Hyman, J. Morris, B. Fischl and R. Buckner (2009). "The cortical signature of Alzheimer's disease: regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals." Cerebral Cortex **19**(3): 497-510.

Diggle, P., P. Heagerty, K.-Y. Liang and S. Zeger (2013). Analysis of longitudinal data, Oxford University Press.

Donoho, D. and J. Johnstone (1994). "Ideal spatial adaptation by wavelet shrinkage." Biometrika **81**(3): 425-455.

Du, A. T., N. Schuff, J. H. Kramer, S. Ganzer, X. P. Zhu, W. J. Jagust, B. L. Miller, B. R. Reed, D. Mungas, K. Yaffe, H. C. Chui and M. W. Weiner (2004). "Higher atrophy rate of entorhinal cortex than hippocampus in AD." Neurology **62**(3): 422-427.

Dubois, B., H. Feldman, C. Jacova, S. Dekosky, P. Barberger-Gateau, J. Cummings, A. Delacourte, D. Galasko, S. Gauthier, G. Jicha, K. Meguro, J. O'Brien, F. Pasquier, P. Robert, M. Rossor, S. Salloway, Y. Stern, P. Visser and P. Scheltens (2007). "Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria." The Lancet. Neurology **6**(8): 734-746.

Duchesne, S., J. C. Pruessner and D. L. Collins (2002). "Appearance-Based Segmentation of Medial Temporal Lobe Structures." NeuroImage **17**(2): 515-531.

Durrenberger, P., F. Fernando, S. Kashefi, T. Bonnert, D. Seilhean, B. Nait-Oumesmar, A. Schmitt, P. Gebicke-Haerter, P. Falkai, E. Grünblatt, M. Palkovits, T. Arzberger, H. Kretzschmar, D. Dexter and R. Reynolds (2014). "Common mechanisms in neurodegeneration and neuroinflammation: a BrainNet Europe gene expression microarray study." Journal of Neural Transmission: 1-14.

Durrleman, S., X. Pennec, A. Trouvé, G. Gerig and N. Ayache (2009). "Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets." Medical Image Computing and Computer-Assisted Intervention: MICCAI **12**(Pt 1): 297-304.

Efros, A. A. and T. K. Leung (1999). "Texture synthesis by non-parametric sampling." International Conference on Computer Vision: ICCV **2**: 1033-1038 vol.1032.

Elliott, C., D. Arnold, L. Collins and T. Arbel (2013). "Temporally consistent probabilistic detection of new multiple sclerosis lesions in brain MRI." IEEE Transactions on Medical Imaging **32**(8): 1490-1503.

Eskildsen, S., P. Coupé, V. Fonov, J. Manjón, K. Leung, N. Guizard, S. Wassef, L. R. Østergaard and L. Collins (2012). "BEaST: brain extraction based on nonlocal segmentation technique." NeuroImage **59**(3): 2362-2373.

Evangelou, N., M. M. Esiri, S. Smith, J. Palace and P. M. Matthews (2000). "Quantitative pathological evidence for axonal loss in normal appearing white matter in multiple sclerosis." Annals of Neurology **47**(3): 391-395.

Evans, A. C., D. L. Collins, S. Mills, E. Brown, R. Kelly and T. M. Peters (1993). "3D statistical neuroanatomical models from 305 MRI volumes." Nuclear Science Symposium and Medical Imaging Conference: 1813-1817.

Fadili, M.-J., J.-L. Starck and F. Murtagh (2009). "Inpainting and zooming using sparse representations." The Computer Journal **52**(1): 64-79.

Fazekas, F., F. Barkhof, M. Filippi, R. I. Grossman, D. K. Li, W. I. McDonald, H. F. McFarland, D. W. Paty, J. H. Simon, J. S. Wolinsky and D. H. Miller (1999). "The contribution of magnetic resonance imaging to the diagnosis of multiple sclerosis." Neurology **53**(3): 448-456.

Ferguson, B., M. K. Matyszak, M. M. Esiri and V. H. Perry (1997). "Axonal damage in acute multiple sclerosis lesions." Brain **120 (Pt 3)**: 393-399.

Fink, F., J. Klein, M. Lanz, T. Mitrovics, M. Lentschig, H. Hahn and H. Hildebrandt (2010). "Comparison of Diffusion Tensor-Based Tractography and Quantified Brain Atrophy for Analyzing Demyelination and Axonal Loss in MS." Journal of Neuroimaging **20**(4): 334-344.

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

Fischl, B., D. H. Salat, A. J. van der Kouwe, N. Makris, F. Segonne, B. T. Quinn and A. M. Dale (2004). "Sequence-independent segmentation of magnetic resonance images." Neuroimage **23 Suppl 1**: S69-84.

Fisher, E., R. A. Rudick, J. H. Simon, G. Cutter, M. Baier, J. C. Lee, D. Miller, B. Weinstock-Guttman, M. K. Mass, D. S. Dougherty and N. A. Simonian (2002). "Eight-year follow-up study of brain atrophy in patients with MS." Neurology **59**(9): 1412-1420.

Fonov, V., P. Coupé, M. Styner and L. Collins (2012). Automatic lateral ventricle segmentation in infant population with high risk of autism. Organization for Human Brain Mapping, Beijing, China.

Fonov, V., A. Evans, K. Botteron, R. Almli, R. McKinstry and L. Collins (2011). "Unbiased average age-appropriate atlases for pediatric studies." NeuroImage **54**(1): 313-327.

Fonov, V., A. C. Evans, R. C. McKinstry, C. R. Almli and D. L. Collins (2009). "Unbiased nonlinear average age-appropriate brain templates from birth to adulthood." NeuroImage **47**: S102-S102.

Forsberg, D., M. Andersson and H. Knutsson (2010). "Adaptive anisotropic regularization of deformation fields for non-rigid registration using the morphon framework." International Conference on Acoustics Speech and Signal Processing: ICASSP: 473-476.

Fox, N., R. Black, S. Gilman, M. Rossor, S. Griffith, L. Jenkins and M. Koller (2005). "Effects of Aβ immunization (AN1792) on MRI measures of cerebral volume in Alzheimer disease." Neurology **64**(9): 1563-1572.

Fox, N., G. Ridgway and J. Schott (2011). "Algorithms, atrophy and Alzheimer's disease: Cautionary tales for clinical trials." NeuroImage **57**(1): 15-18.

Fox, N. C., S. Cousens, R. Scahill, R. J. Harvey and M. N. Rossor (2000). "Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects." Archives of Neurology **57**(3): 339-344.

Fox, R., F. Bethoux, M. Goldman and J. Cohen (2006). "Multiple sclerosis: advances in understanding, diagnosing, and treating the underlying disease." Cleveland Clinic journal of medicine **73**(1): 91-102.

Fox, R. J., T. Cronin, J. Lin, X. Wang, K. Sakaie, D. Ontaneda, S. Y. Mahmoud, M. J. Lowe and M. D. Phillips (2010). "Measuring Myelin Repair and Axonal Loss with Diffusion Tensor Imaging." American journal of neuroradiology: AJNR: ajnr.A2238.

Francis, S. J. (2004). Automatic Lesion Identification in MRI of Multiple Sclerosis Patients. Master, McGill University, Montreal, Quebec.

Freeborough, P. A. and N. C. Fox (1997). "The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI." IEEE Transactions on Medical Imaging **16**(5): 623-629.

Freeborough, P. A. and N. C. Fox (1998). "Modeling Brain Deformations in Alzheimer Disease by Fluid Registration of Serial 3D MR Images." Journal of Computer Assisted Tomography **22**(5): 838-843.

Freeborough, P. A., N. C. Fox and R. I. Kitney (1997). "Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans." Computer Methods and Programs in Biomedicine **53**(1): 15-25.

Frisoni, G., N. Fox, C. Jack, P. Scheltens and P. Thompson (2010). "The clinical use of structural MRI in Alzheimer disease." Nature reviews. Neurology **6**(2): 67-77.

Frisoni, G. B., A. Redolfi, D. Manset, M. E. Rousseau, A. Toga and A. C. Evans (2011). "Virtual imaging laboratories for marker discovery in neurodegenerative diseases." Nature reviews. Neurology **7**(8): 429-438.

García-Lorenzo, D., S. Francis, S. Narayanan, D. Arnold and L. Collins (2013). "Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging." Medical Image Analysis **17**(1): 1-18.

García-Lorenzo, D., S. Prima, D. L. Collins, D. L. Arnold, S. P. Morrissey and C. Barillot (2008). "Combining robust expectation maximization and mean shift algorithms for multiple sclerosis brain segmentation." MICCAI workshop on Medical Image Analysis on Multiple Sclerosis (validation and methodological issues)(MIAMS'2008): 82-91.

García-Lorenzo, D., S. Prima, S. P. Morrissey and C. Barillot (2008). "A robust Expectation-Maximization algorithm for Multiple Sclerosis lesion segmentation." The MIDAS Journal-MS Lesion Segmentation (MICCAI 2008 Workshop): 277.

Geremia, E., O. Clatz, B. Menze, E. Konukoglu, A. Criminisi and N. Ayache (2011). "Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images." NeuroImage **57**(2): 378-390.

Gerig, G., O. Kubler, R. Kikinis and F. A. Jolesz (1992). "Nonlinear anisotropic filtering of MRI data." IEEE Transactions on Medical Imaging **11**(2): 221-232.

Geurts, J. J., P. J. Pouwels, B. M. Uitdehaag, C. H. Polman, F. Barkhof and J. A. Castelijns (2005). "Intracortical Lesions in Multiple Sclerosis: Improved Detection with 3D Double Inversion-Recovery MR Imaging 1." Radiology **236**(1): 254-260.

Ghanei, A., H. Soltanian-Zadeh and J. P. Windham (1998). "Segmentation of the hippocampus from brain MRI using deformable contours." Computerized Medical Imaging and Graphics **22**(3): 203-216.

Gold, R., L. Kappos, D. L. Arnold, A. Bar-Or, G. Giovannoni, K. Selmaj, C. Tornatore, M. T. Sweetser, M. Yang and S. I. Sheikh (2012). "Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis." New England Journal of Medicine **367**(12): 1098-1107.

Good, C. (2001). "A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains." NeuroImage **14**(1): 21-36.

Goodman, A. D., T. R. Brown, L. B. Krupp, R. T. Schapiro, S. R. Schwid, R. Cohen, L. N. Marinucci, A. R. Blight and M. S. F. I. Fampridine (2009). "Sustained-release oral fampridine in multiple sclerosis: a randomised, double-blind, controlled trial." Lancet **373**(9665): 732-738.

Gousias, I. S., D. Rueckert, R. A. Heckemann, L. E. Dyet, J. P. Boardman, A. D. Edwards and A. Hammers (2008). "Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest." NeuroImage **40**(2): 672-684.

Group, V. Q. E. (2003). Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, Phase II (FR_TV2).

Gudbjartsson, H. and S. Patz (1995). "The Rician distribution of noisy MRI data." Magnetic resonance in medicine **34**(6): 910-914.

Guimond, A., J. Meunier and J.-P. Thirion (1998). "Automatic Computation of Average Brain Models." Medical Image Computing and Computer-Assisted Intervention: MICCAI: 631-640.

Guimond, A., J. Meunier and J.-P. Thirion (2000). "Average brain models: a convergence study." Comput. Vis. Image Underst. **77**(9): 192-210.

Guimond, A., A. Roche, N. Ayache and J. Meunier (2001). "Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections." IEEE Transactions on Medical Imaging **20**(1): 58-69.

Guizard, N., P. Coupé, V. S. Fonov, J. V. Manjón, D. L. Arnold and D. L. Collins (2015). "Rotation-invariant multi-contrast non-local means for MS lesion segmentation." NeuroImage: Clinical **8**(0): 376-389.

Guizard, N., V. Fonov, B. Aubert-Broche, D. García-Lorenzo, P. Coupé, S. Eskildsen and D. L. Collins (2012). Robust individual template pipeline for longitudinal MR images. MICCAI 2012 Workshop on Novel Biomarkers for Alzheimer's Disease and Related Disorders, Nice, France.

Guizard, N., V. S. Fonov, D. García-Lorenzo, K. Nakamura, B. Aubert-Broche and D. L. Collins (2015). "Spatio-Temporal Regularization for Longitudinal Registration to Subject-Specific 3d Template." PloS one **10**(8): e0133352.

Guizard, N., K. Nakamura, P. Coupé, D. L. Arnold and D. L. Collins (2013). Non-local MS MRI lesion inpainting method for image processing. The endMS Conference, Saint-Sauveur, Canada.

Guizard, N., K. Nakamura, P. Coupé, V. S. Fonov, D. L. Arnold and D. L. Collins (2015). "Non-local means inpainting of MS lesions in longitudinal image processing." Frontiers in Neuroscience **9**.

Guttmann, C. R., R. Kikinis, M. C. Anderson, M. Jakab, S. K. Warfield, R. J. Killiany, H. L. Weiner and F. A. Jolesz (1999). "Quantitative follow-up of patients with multiple sclerosis using MRI: reproducibility." Journal of Magnetic Resonance Imaging **9**(4): 509-518.

Hajnal, J. V. and D. L. G. Hill (2014). Medical Image Registration, Taylor & Francis.

Hammers, A., R. Heckemann, M. J. Koepp, J. S. Duncan, J. V. Hajnal, D. Rueckert and P. Aljabar (2007). "Automatic detection and quantification of hippocampal atrophy on MRI in temporal lobe epilepsy: A proof-of-principle study." NeuroImage **36**(1): 38-47.

Harmouche, R., L. Collins, D. Arnold, S. Francis and T. Arbel (2006). "Bayesian MS Lesion Classification Modeling Regional and Local Spatial Information." International Conference on Pattern Recognition: ICPR **3**: 984-987.

Hart, G., Y. Shi, H. Zhu, M. Sanchez, M. Styner and M. Niethammer (2010). DTI Longitudinal Atlas Construction as an Average of Growth Models. Spatio-Temporal Image Analysis for Longitudinal and Time-Series Image Data Workshop (MICCAI 2010), Beijing, China.

Hashimoto, M., H. Kazui, K. Matsumoto, Y. Nakano, M. Yasuda and E. Mori (2005). "Does Donepezil Treatment Slow the Progression of Hippocampal Atrophy in Patients With Alzheimer's Disease?" American Journal of Psychiatry **162**(4): 676-682.

Heckemann, R. A., J. V. Hajnal, P. Aljabar, D. Rueckert and A. Hammers (2006). "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion." NeuroImage **33**(1): 115-126.

Held, K., E. Rota Kops, B. J. Krause, W. M. Wells, R. Kikinis and H. W. Müller-Gärtner (1997). "Markov random field segmentation of brain MR images." IEEE transactions on medical imaging **16**(6): 878-886.

Hellier, P. (2003). "Consistent intensity correction of MR images." International Conference on Image Processing: ICIP **1**: 9-12.

Hellier, P., C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D. L. Collins, A. Evans, G. Malandain, N. Ayache, G. E. Christensen and H. J. Johnson (2003). "Retrospective evaluation of intersubject brain registration." IEEE Transactions on Medical Imaging **22**(9): 1120-1130.

Holden, M. (2008). "A Review of Geometric Transformations for Nonrigid Body Registration." IEEE Transactions on Medical Imaging **27**(1): 111-128.

Holland, D., A. M. Dale and I. Alzheimer's Disease Neuroimaging (2011). "Nonlinear registration of longitudinal images and measurement of change in regions of interest." Medical Image Analysis **15**(4): 489-497.

Holland, D., L. McEvoy, A. Dale and I. the Alzheimer's Disease Neuroimaging (2012). "Unbiased comparison of sample size estimates from longitudinal structural measures in ADNI." Human Brain Mapping **33**(11): 2586-2602.

Horwitz, A. R. (2012). "The origins of the molecular era of adhesion research." Nature reviews. Molecular Cellular Biology **13**(12): 805-811.

Hu, S. and L. Collins (2007). "Joint level-set shape modeling and appearance modeling for brain structure segmentation." NeuroImage **36**(3): 672-683.

Hua, X., D. P. Hibar, C. R. Ching, C. P. Boyle, P. Rajagopalan, B. A. Gutman, A. D. Leow, A. W. Toga, C. R. Jack, Jr., D. Harvey, M. W. Weiner, P. M. Thompson and I. Alzheimer's Disease Neuroimaging (2013). "Unbiased tensor-based morphometry: improved robustness and sample size estimates for Alzheimer's disease clinical trials." Neuroimage **66**: 648-661.

Hua, X., A. Leow, N. Parikshak, S. Lee, M.-C. Chiang, A. Toga, C. Jack, M. Weiner and P. Thompson (2008). "Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: an MRI study of 676 AD, MCI, and normal subjects." NeuroImage **43**(3): 458-469.

Huber, S. J., R. A. Bornstein, K. W. Rammohan, J. A. Christy, D. W. Chakeres and R. B. McGhee (1992). "Magnetic resonance imaging correlates of neuropsychological impairment in multiple sclerosis." Journal of Neuropsychiatry Clinic Neuroscience **4**(2): 152-158.

Hurwitz, B. J. (2009). “The diagnosis of multiple sclerosis and the clinical subtypes.” Annals of Indian Academy of Neurology **12**(4): 226.

Huynh, J. L. and P. Casaccia (2013). “Epigenetic mechanisms in multiple sclerosis: implications for pathogenesis and treatment.” Lancet Neurology **12**(2): 195-206.

Jack, C., D. Knopman, W. Jagust, L. Shaw, P. Aisen, M. Weiner, R. Petersen and J. Trojanowski (2010). “Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade.” The Lancet. Neurology **9**(1): 119-128.

Jack, C., M. Shiung, J. Gunter, P. O’brien, S. Weigand, D. Knopman, B. Boeve, R. Ivnik, G. Smith and R. Cha (2004). “Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD.” Neurology **62**(4): 591-600.

Jack, C., M. Slomkowski, S. Gracon, T. Hoover, J. Felmlee, K. Stewart, Y. Xu, M. Shiung, P. O’Brien and R. Cha (2003). “MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD.” Neurology **60**(2): 253-260.

Jack, C. R., Jr., M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, L. W. J, C. Ward, A. M. Dale, J. P. Felmlee, J. L. Gunter, D. L. Hill, R. Killiany, N. Schuff, S. Fox-Bosetti, C. Lin, C. Studholme, C. S. DeCarli, G. Krueger, H. A. Ward, G. J. Metzger, K. T. Scott, R. Mallozzi, D. Blezek, J. Levy, J. P. Debbins, A. S. Fleisher, M. Albert, R. Green, G. Bartzokis, G. Glover, J. Mugler and M. W. Weiner (2008). “The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods.” Journal of Magnetic Resonance Imaging **27**(4): 685-691.

Jack, C. R., M. M. Shiung, J. L. Gunter, P. C. O'Brien, S. D. Weigand, D. S. Knopman, B. F. Boeve, R. J. Ivnik, G. E. Smith, R. H. Cha, E. G. Tangalos and R. C. Petersen (2004). “Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD.” Neurology **62**(4): 591-600.

Jäger, F., Y. Deuerling-Zheng, B. Frericks, F. Wacker and J. Hornegger (2006). “A new method for MRI intensity standardization with application to lesion detection in the brain.” Procs **1010**: 269-276.

Jagust, W. and J. Noseworthy (2000). “Brain atrophy as a surrogate marker in MS.” Neurology **54**(4): 782-783.

Joshi, S., B. Davis, M. Jomier and G. Gerig (2004). “Unbiased diffeomorphic atlas construction for computational anatomy.” NeuroImage **23**: S151-S160.

Jovicich, J., S. Czanner, D. Greve, E. Haley, A. van der Kouwe, R. Gollub, D. Kennedy, F. Schmitt, G. Brown, J. Macfall, B. Fischl and A. Dale (2006). “Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data.” Neuroimage **30**(2): 436-443.

Jovicich, J., S. Czanner, X. Han, D. Salat, A. van der Kouwe, B. Quinn, J. Pacheco, M. Albert, R. Killiany, D. Blacker, P. Maguire, D. Rosas, N. Makris, R. Gollub, A. Dale, B. C. Dickerson and B. Fischl (2009). "MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths." NeuroImage **46**(1): 177-192.

Kalpouzos, G., G. Chételat, J.-C. Baron, B. Landeau, K. Mevel, C. Godeau, L. Barré, J.-M. Constans, F. Viader, F. Eustache and B. Desgranges (2009). "Voxel-based mapping of brain gray matter volume and glucose metabolism profiles in normal aging." Neurobiology of Aging **30**(1): 112-124.

Kamber, M., R. Shinghal, D. L. Collins, G. S. Francis and A. C. Evans (1995). "Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images." IEEE transactions on medical imaging **14**(3): 442-453.

Kantarci, K., R. C. Petersen, B. F. Boeve, D. S. Knopman, S. D. Weigand, P. C. O'Brien, M. M. Shiung, G. E. Smith, R. J. Ivnik, E. G. Tangalos and C. R. Jack (2005). "DWI predicts future progression to Alzheimer disease in amnestic mild cognitive impairment." Neurology **64**(5): 902-904.

Karimaghaloo, Z., M. Shah, S. J. Francis, D. L. Arnold, D. L. Collins and T. Arbel (2012). "Automatic Detection of Gadolinium-Enhancing Multiple Sclerosis Lesions in Brain MRI Using Conditional Random Fields." IEEE Transactions on Medical Imaging **31**(6): 1181-1194.

Kaye, J., M. Moore, A. Dame, J. Quinn, R. Camicioli, D. Howieson, E. Corbridge, B. Care, G. Nesbit and G. Sexton (2005). "Asynchronous regional brain volume losses in presymptomatic to moderate AD." Journal of Alzheimer's Disease **8**(1): 51-56.

Kezele, I. B., D. L. Arnold and D. L. Collins (2008). "Atrophy in white matter fiber tracts in multiple sclerosis is not dependent on tract length or local white matter lesions." Multiple sclerosis **14**(6): 779-785.

Khayati, R., M. Vafadust, F. Towhidkhah and M. Nabavi (2008). "A novel method for automatic determination of different stages of multiple sclerosis lesions in brain MR FLAIR images." Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society **32**(2): 124-133.

Kidd, D., J. W. Thorpe, A. J. Thompson, B. E. Kendall, I. F. Moseley, D. G. MacManus, W. I. McDonald and D. H. Miller (1993). "Spinal cord MRI using multi-array coils and fast spin echo. II. Findings in multiple sclerosis." Neurology **43**(12): 2632-2637.

Kikinis, R., C. R. Guttmann, D. Metcalf, W. M. Wells, G. J. Ettinger, H. L. Weiner and F. A. Jolesz (1999). "Quantitative follow-up of patients with multiple sclerosis using MRI: technical aspects." Journal of magnetic resonance imaging : JMRI **9**(4): 519-530.

Kincses, Z. T., S. Ropele, M. Jenkinson, M. Khalil, K. Petrovic, M. Loitfelder, C. Langkammer, E. Aspeck, M. Wallner-Blazek, S. Fuchs, M. Jehna, R. Schmidt, L. Vécsei, F. Fazekas and C. Enzinger (2011). "Lesion probability mapping to explain clinical deficits and cognitive performance in multiple sclerosis." Multiple sclerosis **17**(6): 681-689.

Klein, A., J. Andersson, B. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. Christensen, L. Collins, J. Gee, P. Hellier, H. Joo, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. Woods, J. Mann and R. Parsey (2009). "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration." NeuroImage **46**(3): 786-802.

Kraemer, H. and S. Thiemann (1987). How Many Subjects?: Statistical Power Analysis in Research, Sage Publications, Inc.

Langlois, S., M. Desvignes, J. M. Constans and M. Revenu (1999). "MRI geometric distortion: a simple approach to correcting the effects of non-linear gradient fields." Journal of magnetic resonance imaging : JMRI **9**(6): 821-831.

Lansley, J., D. Mataix-Cols, M. Grau, J. Radua and J. Sastre-Garriga (2013). "Localized grey matter atrophy in multiple sclerosis: A meta-analysis of voxel-based morphometry studies and associations with functional disability." Neuroscience & Biobehavioral Reviews **37**(5): 819-830.

Lao, Z., D. Shen, D. Liu, A. Jawad, E. Melhem, L. Launer, N. Bryan and C. Davatzikos (2008). "Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine." Academic radiology **15**(3): 300-313.

Lassmann, H. (2008). "The Pathologic Substrate of Magnetic Resonance Alterations in Multiple Sclerosis." Neuroimaging Clinics of North America **18**(4): 563-576.

Lautenschlager, N. T., L. A. Cupples, V. S. Rao, S. A. Auerbach, R. Becker, J. Burke, H. Chui, R. Duara, E. J. Foley, S. L. Glatt, R. C. Green, R. Jones, H. Karlinsky, W. A. Kukull, A. Kurz, E. B. Larson, K. Martelli, A. D. Sadovnick, L. Volicer, S. C. Waring, J. H. Growdon and L. A. Farrer (1996). "Risk of dementia among relatives of Alzheimer's disease patients in the MIRAGE study: What is in store for the oldest old?" Neurology **46**(3): 641-650.

Lauterbur, P. C. (1973). "Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance." Nature **242**(5394): 190-191.

Lee, I. T. and C.-M. Yang (2012). "Role of NADPH oxidase/ROS in pro-inflammatory mediators-induced airway and pulmonary diseases." Biochemical Pharmacology **84**(5): 581-590.

Leese, J. A., C. S. Novak and B. B. Clark (1971). "An Automated Technique for Obtaining Cloud Motion from Geosynchronous Satellite Data Using Cross Correlation." Journal of Applied Meteorology **10**(1): 118-132.

Leow, A., S.-C. Huang, A. Geng, J. Becker, S. Davis, A. Toga and P. Thompson (2005). "Inverse Consistent Mapping in 3D Deformable Image Registration: Its Construction and Statistical Properties." Information Processing in Medical Imaging: IPMI: 493-503.

Leporé, N., C. Brun, X. Pennec, Y.-Y. Chou, O. Lopez, H. Aizenstein, J. Becker, A. Toga and P. Thompson (2007). "Mean Template for Tensor-Based Morphometry Using Deformation Tensors." Medical Image Computing and Computer-Assisted Intervention: MICCAI **4792**: 826-833.

Leung, K. K., J. Barnes, G. R. Ridgway, J. W. Bartlett, M. J. Clarkson, K. Macdonald, N. Schuff, N. C. Fox, S. Ourselin and I. Alzheimer's Disease Neuroimaging (2010). "Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease." Neuroimage **51**(4): 1345-1359.

Leung, K. K., M. J. Clarkson, J. W. Bartlett, S. Clegg, C. R. Jack, Jr., M. W. Weiner, N. C. Fox, S. Ourselin and I. Alzheimer's Disease Neuroimaging (2010). "Robust atrophy rate measurement in Alzheimer's disease using multi-site serial MRI: tissue-specific intensity normalization and parameter selection." Neuroimage **50**(2): 516-523.

Lewis, E. B. and N. C. Fox (2004). "Correction of differential intensity inhomogeneity in longitudinal MR images." Neuroimage **23**(1): 75-83.

Liang, Z. P. and P. C. Lauterbur (2000). Principles of Magnetic Resonance Imaging: a Signal Processing Perspective. Bellingham, Washington: New York: 2000.

Liao, S., H. Jia, G. Wu, D. Shen and I. Alzheimer's Disease Neuroimaging (2012). "A novel framework for longitudinal atlas construction with groupwise registration of subject image sequences." Neuroimage **59**(2): 1275-1289.

Liu, C., S. Edwards, Q. Gong, N. Roberts and L. D. Blumhardt (1999). "Three dimensional MRI estimates of brain and spinal cord atrophy in multiple sclerosis." Journal of Neurology, Neurosurgery, and Psychiatry **66**(3): 323-330.

Liu, Y. and V. Caselles (2013). "Exemplar-based image inpainting using multiscale graph cuts." IEEE Transactions on Image Processing **22**(5): 1699-1711.

Lladó, X., A. Oliver, M. Cabezas, J. Freixenet, J. Vilanova, A. Quiles, L. Valls, L. Ramió-Torrentà and À. Rovira (2012). "Segmentation of multiple sclerosis lesions in brain MRI: A review of automated approaches." Information Sciences **186**(1): 164-185.

Lockshin, R. A. and C. M. Williams (1964). "Programmed cell death—II. Endocrine potentiation of the breakdown of the intersegmental muscles of silkmoths." Journal of Insect Physiology **10**(4): 643-649.

Lorenzi, M., N. Ayache, G. Frisoni and X. Pennec (2010). 4D registration of serial brain's MR images: a robust measure of changes applied to Alzheimer's disease. Spatio-Temporal Image Analysis for Longitudinal and Time-Series Image Data Workshop (MICCAI 2010), Beijing, China.

Lublin, F. D. and S. C. Reingold (1996). "Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis." Neurology **46**(4): 907-911.

Luckhaus, C., M. Jänner, M. Cohnen, M. O. Flüss, S. J. Teipel, M. Grothe, H. Hampel, J. Kornhuber, E. Rüther, O. Peters, T. Supprian, W. Gaebel, U. Mödder and H. J. Wittsack (2010). "A novel MRI-biomarker candidate for Alzheimer's disease composed of regional brain volume and perfusion variables." European Journal of Neurology : the Official Journal of the European Federation of Neurological Societies **17**(12): 1437-1444.

Lukas, C., A. Minneboo, V. de Groot, B. Moraal, D. Knol, C. Polman, F. Barkhof and H. Vrenken (2010). "Early central atrophy rate predicts 5 year clinical outcome in multiple sclerosis." Journal of Neurology, Neurosurgery & Psychiatry **81**(12): 1351-1356.

Magon, S., L. Gaetano, M. M. Chakravarty, J. Lerch, Y. Naegelin, C. Stippich, L. Kappos, E.-W. Radue and T. Sprenger (2014). "White matter lesion filling improves the accuracy of cortical thickness measurements in multiple sclerosis patients: a longitudinal study." BMC Neuroscience **15**(1): 106.

Manjón, J., P. Coupé, A. Buades, D. Louis Collins and M. Robles (2012). "New methods for MRI denoising based on sparseness and self-similarity." Medical Image Analysis **16**(1): 18-27.

Manjón, J., S. Eskildsen, P. Coupé, J. Romero, L. Collins and M. Robles (2014). "Nonlocal Intracranial Cavity Extraction." International Journal of Biomedical Imaging **2014**: 1-11.

Manjón, J., N. Thacker, J. Lull, G. Garcia-Martí, L. Martí-Bonmatí and M. Robles (2009). "Multicomponent MR Image Denoising." International Journal of Biomedical Imaging **2009**: 1-10.

Manly, J., M.-X. Tang, N. Schupf, Y. Stern, J.-P. Vonsattel and R. Mayeux (2008). "Frequency and course of mild cognitive impairment in a multiethnic community." Annals of Neurology **63**(4): 494-506.

Mansfield, P. and P. K. Granell (1973). "NMR'diffraction'in solids." Journal of Physics **6**(22): L422.

Marcus, D., A. Fotenos, J. Csernansky, J. Morris and R. Buckner (2009). "Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults." Journal of Cognitive Neuroscience **22**(12): 2677-2684.

Mazziotta, J., A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, D. L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima and B. Mazoyer (2001). "A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM)." Philosophical transactions of the Royal Society of London **356**(1412): 1293-1322.

Mazziotta, J., A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, J. Feidler, K. Smith, D. Boomsma, H. Pol, T. Cannon, R. Kawashima and B. Mazoyer (2001). "A four-dimensional probabilistic atlas of the human brain." Journal of the American Medical Informatics Association **8**(5): 401-430.

McKhann, G., D. Drachman, M. Folstein, R. Katzman, D. Price and E. Stadlan (1984). "Clinical diagnosis of Alzheimer's disease." Neurology **34**(7): 939-939.

McKhann, G., D. Knopman, H. Chertkow, B. Hyman, C. Jack, C. Kawas, W. Klunk, W. Koroshetz, J. Manly, R. Mayeux, R. Mohs, J. Morris, M. Rossor, P. Scheltens, M. Carrillo, B. Thies, S. Weintraub and C. Phelps (2011). "The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." Alzheimer's & Dementia **7**(3): 263-269.

McVeigh, E. R., M. J. Bronskill and R. M. Henkelman (1986). "Phase and sensitivity of receiver coils in magnetic resonance imaging." Medical physics **13**(6): 806-814.

Meier, D. and E. Fisher (2005). "Atlas-based anatomic labeling in neurodegenerative disease via structure-driven atlas warping." Journal of neuroimaging **15**(1): 16-26.

Meier, D. and C. Guttmann (2003). "Time-series analysis of MRI intensity patterns in multiple sclerosis." NeuroImage **20**(2): 1193-1209.

Meier, D., H. Weiner and C. G. Guttmann (2007). "Time-series modeling of multiple sclerosis disease activity: A promising window on disease progression and repair potential?" Neurotherapeutics **4**(3): 485-498.

Meier, D., H. Weiner, S. Khoury and C. Guttmann (2004). "Magnetic resonance imaging surrogates of multiple sclerosis pathology and their relationship to central nervous system atrophy." Journal of Neuroimaging **14**(3 Suppl).

Menke, J. and T. R. Martinez (2004). "Using permutations instead of student's t distribution for p-values in paired-difference algorithm comparisons." International Joint Conference on Neural Networks: 1331-1335.

Miller, A. E. (2011). “Multiple Sclerosis: Where Will We Be in 2020?” Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine **78**(2): 268-279.

Miller, B. L., R. A. Moats, T. Shonk, T. Ernst, S. Woolley and B. D. Ross (1993). “Alzheimer disease: depiction of increased cerebral myo-inositol with proton MR spectroscopy.” Radiology **187**(2): 433-437.

Miller, D., F. Barkhof, J. Frank, G. Parker and A. Thompson (2002). “Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance.” Brain **125**(Pt 8): 1676-1695.

Miller, M., A. Trouve and L. Younes (2002). “On the metrics and euler-lagrange equations of computational anatomy.” Annual Review of Biomedical Engineering **4**(1): 375-405.

Mitchell, A. J. and M. Shiri-Feshki (2009). “Rate of progression of mild cognitive impairment to dementia – meta-analysis of 41 robust inception cohort studies.” Acta Psychiatrica Scandinavica **119**(4): 252-265.

Mohandas, E., V. Rajmohan and B. Raghunath (2009). “Neurobiology of Alzheimer's disease.” Indian Journal of Psychiatry **51**(1): 55-61.

Morra, J., Z. Tu, A. Toga and P. Thompson (2008). “Automatic Segmentation of MS Lesions Using a Contextual Model for the MICCAI Grand Challenge.” The MIDAS Journal-MS Lesion Segmentation (MICCAI 2008 Workshop).

Mueller, S., M. Weiner, L. Thal, R. Petersen, C. Jack, W. Jagust, J. Trojanowski, A. Toga and L. Beckett (2005). “The Alzheimer's disease neuroimaging initiative.” Neuroimaging Clinics of North America **15**(4): 869–877.

Musse, O., F. Heitz and J. P. Armspach (2001). “Topology preserving deformable image matching using constrained hierarchical parametric models.” IEEE Transactions on Image Processing **10**(7): 1081-1093.

Myers, R. H. (2004). “Huntington's Disease Genetics.” NeuroRX **1**(2): 255-262.

Nakamura, K. and E. Fisher (2009). “Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients.” NeuroImage **44**(3): 769-776.

Nakamura, K., N. Guizard, V. Fonov, S. Narayanan, L. Collins and D. Arnold (2014). “Jacobian integration method increases the statistical power to measure gray matter atrophy in multiple sclerosis.” NeuroImage: Clinical **4**: 10-17.

Naldi, P., L. Collimedaglia, D. Vecchio, M. G. Rosso, F. Perl, A. Stecco, F. Monaco and M. A. Leone (2011). “Predictors of attack severity and duration in multiple sclerosis: a prospective study.” The Open Neurology Journal **5**: 75-82.

Narayana, P. A., W. W. Brey, M. V. Kulkarni and C. L. Sievenpiper (1988). "Compensation for surface coil sensitivity variation in magnetic resonance imaging." Magnetic Resonance Imaging **6**(3): 271-274.

Narayanan, S., L. Fu, E. Pioro, N. De Stefano, D. L. Collins, G. S. Francis, J. P. Antel, P. M. Matthews and D. L. Arnold (1997). "Imaging of axonal damage in multiple sclerosis: spatial distribution of magnetic resonance imaging lesions." Annals of Neurology **41**(3): 385-391.

Nelson, F., A. Poonawalla, P. Hou, F. Huang, J. Wolinsky and P. Narayana (2007). "Improved identification of intracortical lesions in multiple sclerosis with phase-sensitive inversion recovery in combination with fast double inversion recovery MR imaging." American Journal of Neuroradiology **28**(9): 1645-1649.

Nestor, S., R. Rupsingh, M. Borrie, M. Smith, V. Accomazzi, J. Wells, J. Fogarty, R. Bartha and I. the Alzheimer's Disease Neuroimaging (2008). "Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database." Brain **131**(9): 2443-2454.

Niccoli, T. and L. Partridge (2012). "Ageing as a Risk Factor for Disease." Current Biology **22**(17): R741-R752.

Noblet, V., C. Heinrich, F. Heitz and J.-P. Armspach (2008). "Symmetric nonrigid image registration: application to average brain templates construction." Medical Image Computing and Computer-Assisted Intervention: MICCAI: 897-904.

Noblet, V., C. Heinrich, F. Heitz and J. P. Armspach (2006). "Retrospective evaluation of a topology preserving non-rigid registration method." Medical Image Analysis **10**(3): 366-384.

Noseworthy, J. H., C. Lucchinetti, M. Rodriguez and B. G. Weinshenker (2000). "Multiple sclerosis." The New England Journal of Medicine **343**(13): 938-952.

Nyúl, L. and J. Udupa (1999). "On standardizing the MR image intensity scale." Magnetic Resonance in Medicine **42**(6): 1072-1081.

Nyúl, L. G., J. K. Udupa and X. Zhang (2000). "New variants of a method of MRI scale standardization." IEEE transactions on medical imaging **19**(2): 143-150.

O'Connor, P., J. S. Wolinsky, C. Confavreux, G. Comi, L. Kappos, T. P. Olsson, H. Benzerdjeb, P. Truffinet, L. Wang, A. Miller and M. S. Freedman (2011). "Randomized Trial of Oral Teriflunomide for Relapsing Multiple Sclerosis." New England Journal of Medicine **365**(14): 1293-1303.

Owens, T. (2003). "The enigma of multiple sclerosis: inflammation and neurodegeneration cause heterogeneous dysfunction and damage." Current Opinion in Neurology **16**(3): 259-265.

Pelletier, J., L. Suchet, T. Witjas, M. Habib, C. R. Guttmann, G. Salamon, O. Lyon-Caen and A. A. Chérif (2001). "A longitudinal study of callosal atrophy and interhemispheric dysfunction in relapsing-remitting multiple sclerosis." Archives of Neurology **58**(1): 105-111.

Petersen, R., R. Roberts, D. Knopman, B. Boeve, Y. Geda, R. Ivnik, G. Smith and C. Jack (2009). "Mild cognitive impairment: ten years later." Archives of Neurology **66**(12): 1447-1455.

Polman, C., S. Reingold, B. Banwell, M. Clanet, J. Cohen, M. Filippi, K. Fujihara, E. Havrdova, M. Hutchinson, L. Kappos, F. Lublin, X. Montalban, P. O'Connor, M. Sandberg-Wollheim, A. Thompson, E. Waubant, B. Weinshenker and J. Wolinsky (2011). "Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria." Annals of Neurology **69**(2): 292-302.

Popescu, V., N. C. G. Ran, F. Barkhof, D. T. Chard, C. A. Wheeler-Kingshott and H. Vrenken (2014). "Accurate GM atrophy quantification in MS using lesion-filling with co-registered 2D lesion masks." NeuroImage: Clinical **4**: 366-373.

Poser, C. M., D. W. Paty, L. Scheinberg, W. I. McDonald, F. A. Davis, G. C. Ebers, K. P. Johnson, W. A. Sibley, D. H. Silberberg and W. W. Tourtellotte (1983). "New diagnostic criteria for multiple sclerosis: guidelines for research protocols." Annals of Neurology **13**(3): 227-231.

Prados, F., M. J. Cardoso, D. MacManus, C. Wheeler-Kingshott and S. Ourselin (2014). "A modality-agnostic patch-based technique for lesion filling in multiple sclerosis." Medical Image Computing and Computer-Assisted Intervention: MICCAI **17**(Pt 2): 781-788.

Prima, S., N. Ayache, A. Janke, S. Francis, D. Arnold and L. Collins (2002). "Statistical Analysis of Longitudinal MRI Data: Applications for Detection of Disease Activity in MS." Medical Image Computing and Computer-Assisted Intervention: MICCAI **2488**: 363-371.

Prineas, J. W. and F. Connell (1978). "The fine structure of chronically active multiple sclerosis plaques." Neurology **28**(9 Pt 2): 68-75.

Prinster, A., M. Quarantelli, G. Orefice, R. Lanzillo, A. Brunetti, C. Mollica, E. Salvatore, V. B. Morra, G. Coppola, G. Vacca, B. Alfano and M. Salvatore (2006). "Grey matter loss in relapsing–remitting multiple sclerosis: A voxel-based morphometry study." NeuroImage **29**(3): 859-867.

Proskuryakov, S. Y. a., A. G. Konoplyannikov and V. L. Gabai (2003). "Necrosis: a specific form of programmed cell death?" Experimental Cell Research **283**(1): 1-16.

Przedborski, S. and M. Vila (2001). "MPTP: a review of its mechanisms of neurotoxicity." Clinical Neuroscience Research **1**(6): 407-418.

Ramanan, V. and A. Saykin (2013). “Pathways to neurodegeneration: mechanistic insights from GWAS in Alzheimer's disease, Parkinson's disease, and related disorders.” American journal of neurodegenerative disease **2**(3): 145-175.

Ramirez, J., A. A. McNeely, C. J. Scott, D. T. Stuss and S. E. Black (2014). “Subcortical hyperintensity volumetrics in Alzheimer's disease and normal elderly in the Sunnybrook Dementia Study: correlations with atrophy, executive function, mental processing speed, and verbal memory.” Alzheimer's Reseach & Therapy **6**(4): 49.

Ranginwala, N., L. Hynan, M. Weiner and C. White (2008). “Clinical Criteria for the Diagnosis of Alzheimer Disease: Still Good After All These Years.” The American Journal of Geriatric Psychiatry **16**(5): 384-388.

Reuter, M., N. Schmansky, D. Rosas and B. Fischl (2012). “Within-subject template estimation for unbiased longitudinal image analysis.” NeuroImage **61**(4): 1402-1418.

Rey, D., G. Subsol, H. Delingette and N. Ayache (2002). “Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis.” Medical Image Analysis **6**(2): 163-179.

Richert, N. D., T. Howard, J. A. Frank, R. Stone, J. Ostuni, J. Ohayon, C. Bash and H. F. McFarland (2006). “Relationship between inflammatory lesions and cerebral atrophy in multiple sclerosis.” Neurology **66**(4): 551-556.

Ridha, B., J. Barnes, J. Bartlett, A. Godbolt, T. Pepple, M. Rossor and N. Fox (2006). “Tracking atrophy progression in familial Alzheimer's disease: a serial MRI study.” Lancet Neurology **5**(10): 828-834.

Ridha, B. H., M. R. Symms, D. J. Tozer, K. C. Stockton, C. Frost, M. M. Siddique, E. B. Lewis, D. G. MacManus, P. A. Boulby, G. J. Barker, M. N. Rossor, N. C. Fox and P. S. Tofts (2007). “Magnetization transfer ratio in Alzheimer disease: comparison with volumetric measurements.” American journal of neuroradiology: AJNR **28**(5): 965-970.

Rodriguez, M. and B. Scheithauer (1994). “Ultrastructure of multiple sclerosis.” Ultrastructural pathology **18**(1-2): 3-13.

Rogelj, P. and S. Kovacic (2004). “Spatial deformation models for non-rigid image registration.” Computer Vision Winter Workshop: CVWW: 79-88.

Rogelj, P. and S. Kovacic (2006). “Symmetric image registration.” Special Issue on The Second International Workshop on Biomedical Image Registration (WBIR'03) **10**(3): 484-493.

Rohlfing, T., R. Brandt, R. Menzel and C. Maurer (2004). “Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains.” NeuroImage **21**(4): 1428-1442.

Rousseau, F., P. Habas and C. Studholme (2011). "A supervised patch-based approach for human brain labeling." IEEE transactions on medical imaging **30**(10): 1852-1862.

Rovira, A., C. Auger and J. Alonso (2013). "Magnetic resonance monitoring of lesion evolution in multiple sclerosis." Therapeutic advances in neurological disorders **6**(5): 298-310.

Roy, S., Q. He, A. Carass, A. Jog, J. Cuzzocreo, D. Reich, J. Prince and D. Pham (2014). "Example based lesion segmentation." SPIE Medical Imaging **9034**: 90341Y-90341Y-90348.

Rudick, R., J.-C. Lee, J. Simon, R. Ransohoff and E. Fisher (2004). "Defining interferon β response status in multiple sclerosis patients." Annals of Neurology **56**(4): 548-555.

Rudick, R. A., E. Fisher, J.-C. Lee, J. T. Duda and J. Simon (2000). "Brain atrophy in relapsing multiple sclerosis: relationship to relapses, EDSS, and treatment with interferon β-1a." Multiple sclerosis **6**(6): 365-372.

Rudick, R. A., E. Fisher, J. C. Lee, J. Simon and L. Jacobs (1999). "Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing-remitting MS. Multiple Sclerosis Collaborative Research Group." Neurology **53**(8): 1698-1704.

Rudick, R. A., E. Fisher, J. C. Lee, J. Simon, L. Jacobs and G. the Multiple Sclerosis Collaborative Research (1999). "Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing-remitting MS." Neurology **53**(8): 1698-1698.

Rueckert, D., P. Aljabar, R. Heckemann, J. Hajnal and A. Hammers (2006). "Diffeomorphic Registration Using B-Splines." Medical Image Computing and Computer-Assisted Intervention: MICCAI **4191**: 702-709.

Sadeghi, N., M. Prastawa, J. Gilmore, W. Lin and G. Gerig (2010). "Towards analysis of growth trajectory through multimodal longitudinal MR imaging." SPIE Medical Imaging: 76232U-76232U-76238.

Sahraian, M. A., E. W. Radue, S. Haller and L. Kappos (2010). "Black holes in multiple sclerosis: definition, evolution, and clinical correlations." Acta Neurologica Scandinavica **122**(1): 1-8.

Sailer, M., N. A. Losseff, L. Wang, M. L. Gawne-Cain, A. J. Thompson and D. H. Miller (2001). "T1 lesion load and cerebral atrophy as a marker for clinical progression in patients with multiple sclerosis. A prospective 18 months follow-up study." European Journal of Neurology : the Official Journal of the European Federation of Neurological Societies **8**(1): 37-42.

Salomone, S., F. Caraci, G. M. Leggio, J. Fedotova and F. Drago (2012). "New pharmacological strategies for treatment of Alzheimer's disease: focus on disease modifying drugs." British Journal of Clinical Pharmacology **73**(4): 504-517.

Sanfilipo, M., R. Benedict, B. Weinstock-Guttman and R. Bakshi (2006). "Gray and white matter brain atrophy and neuropsychological impairment in multiple sclerosis." Neurology **66**(5): 685-692.

Sawhney, H. and R. Kumar (1997). "True Multi-Image Alignment and its Application to Mosaicing and Lens Distortion Correction." IEEE Transactions on Pattern Analysis and Machine Intelligence.

Scahill, R., J. Schott, J. Stevens, M. Rossor and N. Fox (2002). "Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI." National Academy of Sciences of the United States of America **99**(7): 4703-4707.

Scalfari, A., A. Neuhaus, A. Degenhardt, G. Rice, P. Muraro, M. Daumer and G. Ebers (2010). "The natural history of multiple sclerosis: a geographically based study 10: relapses and long-term disability." Brain **133**(Pt 7): 1914-1929.

Scheltens, P., D. Leys, F. Barkhof, D. Huglo, H. C. Weinstein, P. Vermersch, M. Kuiper, M. Steinling, E. C. Wolters and J. Valk (1992). "Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates." Journal of Neurology, Neurosurgery & Psychiatry **55**(10): 967-972.

Schmidt, P., C. Gaser, M. Arsic, D. Buck, A. Förschler, A. Berthele, M. Hoshi, R. Ilg, V. Schmid, C. Zimmer, B. Hemmer and M. Mühlau (2012). "An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis." NeuroImage **59**(4): 3774-3783.

Schott, J. M., S. L. Price, C. Frost, J. L. Whitwell, M. N. Rossor and N. C. Fox (2005). "Measuring atrophy in Alzheimer disease." Neurology **65**(1): 119-124.

Sdika, M. and D. Pelletier (2009). "Nonrigid registration of multiple sclerosis brain images using lesion inpainting for morphometry or lesion mapping." Human Brain Mapping **30**(4): 1060-1067.

Ségonne, F., A. M. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn and B. Fischl (2004). "A hybrid approach to the skull stripping problem in MRI." NeuroImage **22**(3): 1060-1075.

Shah, M., Y. Xiao, N. Subbanna, S. Francis, D. Arnold, L. Collins and T. Arbel (2011). "Evaluating intensity normalization on MRIs of human brain with multiple sclerosis." Medical image analysis **15**(2): 267-282.

Shannon, C. E. (1948). "A Mathematical Theory of Communication." Bell System Technical Journal **27**(3): 379-423.

Shattuck, D., M. Mirza, V. Adisetiyo, C. Hojatkashani, G. Salamon, K. Narr, R. Poldrack, R. Bilder and A. Toga (2008). "Construction of a 3D probabilistic atlas of human cortical structures." NeuroImage **39**(3): 1064-1080.

Shen, D. (2003). "Very High-Resolution Morphometry Using Mass-Preserving Deformations and HAMMER Elastic Registration." NeuroImage **18**(1): 28-41.

Shen, D. and C. Davatzikos (2004). "Measuring temporal morphological changes robustly in brain MR images via 4-dimensional template warping." NeuroImage **21**(4): 1508-1517.

Shen, D., S. Moffat, S. M. Resnick and C. Davatzikos (2002). "Measuring Size and Shape of the Hippocampus in MR Images Using a Deformable Shape Model." NeuroImage **15**(2): 422-434.

Shi, W., X. Zhuang, L. Pizarro, W. Bai, H. Wang, K.-P. Tung, P. Edwards and D. Rueckert (2012). "Registration Using Sparse Free-Form Deformations." Medical Image Computing and Computer-Assisted Intervention: MICCAI **7511**: 659-666.

Shiee, N., P.-L. Bazin, A. Ozturk, D. Reich, P. Calabresi and D. Pham (2010). "A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions." NeuroImage **49**(2): 1524-1535.

Simon, J. H. (2006). "Brain atrophy in multiple sclerosis: what we know and would like to know." Multiple sclerosis **12**(6): 679-687.

Simon, J. H., L. D. Jacobs, M. K. Campion, R. A. Rudick, D. L. Cookfair, R. M. Herndon, J. R. Richert, A. M. Salazar, J. S. Fischer, D. E. Goodkin, N. Simonian, M. Lajaunie, D. E. Miller, K. Wende, A. Martens-Davidson, R. P. Kinkel, Munschauer and C. M. Brownscheidle (1999). "A longitudinal study of brain atrophy in relapsing multiple sclerosis." Neurology **53**(1): 139.

Sled, J. G. and G. B. Pike (1998). "Standing-wave and RF penetration artifacts caused by elliptic geometry: an electrodynamic analysis of MRI." IEEE transactions on medical imaging **17**(4): 653-662.

Sled, J. G., A. P. Zijdenbos and A. C. Evans (1998). "A nonparametric method for automatic correction of intensity nonuniformity in MRI data." IEEE Transactions on Medical Imaging **17**(1): 87-97.

Smith, S. (2002). "Fast robust automated brain extraction." Human Brain Mapping **17**(3): 143-155.

Smith, S., N. Destefano, M. Jenkinson and P. Matthews (2000). "SIENA — Normalised accurate measurement of longitudinal brain change." NeuroImage **11**(5): S659-S659.

Smith, S., Y. Zhang, M. Jenkinson, J. Chen, P. M. Matthews, A. Federico and N. De Stefano (2002). "Accurate, robust, and automated longitudinal and cross-sectional brain change analysis." NeuroImage **17**(1): 479-489.

Smith, S. M., N. De Stefano, M. Jenkinson and P. M. Matthews (2001). "Normalized accurate measurement of longitudinal brain change." Journal of Computer Assisted Tomography **25**(3): 466-475.

So, R. W. K., T. W. H. Tang and A. C. S. Chung (2011). "Non-rigid image registration of brain magnetic resonance images using graph-cuts." Pattern Recognition **44**(10–11): 2450-2467.

Sotiras, A., C. Davatzikos and N. Paragios (2013). "Deformable Medical Image Registration: A Survey." IEEE Transactions on Medical Imaging **32**(7): 1153-1190.

Souplet, J., C. Lebrun, N. Ayache and G. Malandain (2008). "An Automatic Segmentation of T2-FLAIR Multiple Sclerosis Lesions." The MIDAS Journal-MS Lesion Segmentation (MICCAI 2008 Workshop).

Sperling, R., P. Aisen, L. Beckett, D. Bennett, S. Craft, A. Fagan, T. Iwatsubo, C. Jack, J. Kaye, T. Montine, D. Park, E. Reiman, C. Rowe, E. Siemers, Y. Stern, K. Yaffe, M. Carrillo, B. Thies, M. Morrison-Bogorad, M. Wagster and C. Phelps (2011). "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." Alzheimer's & Dementia **7**(3): 280-292.

Styner, M., J. Lee, B. Chin, M. Chin, O. Commowick, H. Tran, S. Markovic-Plese, V. Jewells and S. Warfield (2008). "3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation." The MIDAS Journal-MS Lesion Segmentation (MICCAI 2008 Workshop).

Sweeney, E., R. Shinohara, N. Shiee, F. Mateen, A. Chudgar, J. Cuzzocreo, P. Calabresi, D. Pham, D. Reich and C. Crainiceanu (2013). "OASIS is Automated Statistical Inference for Segmentation, with applications to multiple sclerosis lesion segmentation in MRI." NeuroImage: Clinical **2**: 402-413.

Ta, V.-T., R. Giraud, Collins and P. Coupé (2014). "Optimized PatchMatch for Near Real Time and Accurate Label Fusion." Medical Image Computing and Computer-Assisted Intervention: MICCAI **8675**: 105-112.

Talairach, J. and P. Tournoux (1988). Co-planar Stereotaxic Atlas of the Human Brain: 3-dimensional Proportional System : an Approach to Cerebral Imaging, G. Thieme.

Tao, G., S. Datta, R. He, F. Nelson, J. Wolinsky and P. Narayana (2009). "Deep gray matter atrophy in multiple sclerosis: a tensor based morphometry." Journal of the Neurological Sciences **282**(1-2): 39-46.

Telea, A. (2004). "An image inpainting technique based on the fast marching method." Journal of Graphics Tools **9**(1): 23-34.

Thevenaz, P. and M. Unser (1998). “An efficient mutual information optimizer for multiresolution image registration.” International Conference on Image Processing: ICIP **1**: 833-837.

Thirion, J. P. (1998). “Image matching as a diffusion process: an analogy with Maxwell's demons.” Medical Image Analysis **2**(3): 243-260.

Thirion, J. P. and G. Calmon (1999). “Deformation analysis to detect and quantify active lesions in three-dimensional medical image sequences.” IEEE Transactions on Medical Imaging **18**(5): 429-441.

Thompson, P. and A. Toga (2002). “A framework for computational anatomy.” Computing and Visualization in Science **5**(1): 13-34.

Thompson, P. M., K. M. Hayashi, G. I. de Zubicaray, A. L. Janke, S. E. Rose, J. Semple, M. S. Hong, D. H. Herman, D. Gravano, D. M. Doddrell and A. W. Toga (2004). “Mapping hippocampal and ventricular change in Alzheimer disease.” NeuroImage **22**(4): 1754-1766.

Thompson, W. and D. Holland (2011). “Bias in tensor based morphometry Stat-ROI measures may result in unrealistic power estimates.” NeuroImage **57**(1): 1-4.

Tiberio, M., D. Chard, D. Altmann, G. Davies, C. Griffin, W. Rashid, J. Sastre-Garriga, A. Thompson and D. Miller (2005). “Gray and white matter volume changes in early RRMS A 2-year longitudinal study.” Neurology **64**(6): 1001-1007.

Tincher, M., C. R. Meyer, R. Gupta and D. M. Williams (1993). “Polynomial modeling and reduction of RF body coil spatial inhomogeneity in MRI.” IEEE Transactions on Medical Imaging **12**(2): 361-365.

Tomas-Fernandez, X. and S. K. Warfield (2011). “A new classifier feature space for an improved Multiple Sclerosis lesion segmentation.” International Symposium on Biomedical Imaging: From Nano to Macro 1492-1495.

Tomas-Fernandez, X. and S. K. Warfield (2015) “A Model of Population and Subject (MOPS) Intensities with Application to Multiple Sclerosis Lesion Segmentation.” IEEE transactions on medical imaging DOI: citeulike-article-id:13547869.

Tomasi, C. and R. Manduchi (1998). “Bilateral filtering for gray and color images.” International Conference on Computer Vision: ICCV: 839-846.

Tong, T., R. Wolz, P. Coupé, J. V. Hajnal, D. Rueckert and A. s. D. N. Initiative (2013). “Segmentation of MR images via discriminative dictionary learning and sparse coding: Application to hippocampus labeling.” NeuroImage **76**: 11-23.

Trapp, B., J. Peterson, R. Ransohoff, R. Rudick, S. Mörk and L. Bö (1998). "Axonal Transection in the Lesions of Multiple Sclerosis." New England Journal of Medicine **338**(5): 278-285.

Turner, B., N. Ramli, L. D. Blumhardt and T. Jaspan (2001). "Ventricular enlargement in multiple sclerosis: a comparison of three-dimensional and linear MRI estimates." Neuroradiology **43**(8): 608-614.

Tustison, N., B. Avants, P. Cook, Y. Zheng, A. Egan, P. Yushkevich and J. Gee (2010). "N4ITK: improved N3 bias correction." IEEE transactions on medical imaging **29**(6): 1310-1320.

Urschler, M., J. Bauer, H. Ditt and H. Bischof (2006). SIFT and Shape Context for Feature-Based Nonlinear Registration of Thoracic CT Images. Computer Vision Approaches to Medical Image Analysis. R. Beichel and M. Sonka, Springer Berlin Heidelberg. **4241:** 73-84.

Van Leemput, K., F. Maes, D. Vandermeulen, A. Colchester and P. Suetens (2001). "Automated segmentation of multiple sclerosis lesions by model outlier detection." IEEE Transactions on Medical Imaging **20**(8): 677-688.

van Walderveen, M. A. A., W. Kamphorst, P. Scheltens, J. van Waesberghe, R. Ravid, J. Valk, C. H. Polman and F. Barkhof (1998). "Histopathologic correlate of hypointense lesions on T1-weighted spin-echo MRI in multiple sclerosis." Neurology **50**(5): 1282-1288.

Vanderbrug, G. J. and A. Rosenfeld (1977). "Two-Stage Template Matching." IEEE Transactions on Medical Imaging **26**(4): 384-393.

Verbeke, G. and G. Molenberghs (2000). Linear Mixed Models for Longitudinal Data, Springer New York.

Vercauteren, T., X. Pennec, A. Perchant and N. Ayache (2007). "Non-parametric Diffeomorphic Image Registration with the Demons Algorithm." Medical Image Computing and Computer-Assisted Intervention: MICCAI **4792**: 319-326.

Vercauteren, T., X. Pennec, A. Perchant and N. Ayache (2009). "Diffeomorphic demons: efficient non-parametric image registration." NeuroImage **45**(1 Suppl): S61-S72.

Vercellino, M., F. Plano, B. Votta, R. Mutani, M. T. Giordana and P. Cavalla (2005). "Grey matter pathology in multiple sclerosis." Journal of Neuropathology and Experimental Neurology **64**(12): 1101-1107.

Vinitski, S., C. F. Gonzalez, R. Knobler, D. Andrews, T. Iwanaga and M. Curtis (1999). "Fast tissue segmentation based on a 4D feature map in characterization of intracranial lesions." Journal of magnetic resonance imaging : JMRI **9**(6): 768-776.

Viola, P. and Wells (1997). "Alignment by Maximization of Mutual Information." International Journal of Computer Vision **24**(2): 137-154.

Vrenken, H., J. J. G. Geurts, D. L. Knol, C. H. Polman, J. A. Castelijns, P. J. W. Pouwels and F. Barkhof (2006). "Normal-Appearing White Matter Changes Vary with Distance to Lesions in Multiple Sclerosis." American Journal of Neuroradiology **27**(9): 2005-2011.

Vrenken, H., M. Jenkinson, M. A. Horsfield, M. Battaglini, R. A. van Schijndel, E. Rostrup, J. J. G. Geurts, E. Fisher, A. Zijdenbos, J. Ashburner, D. H. Miller, M. Filippi, F. Fazekas, M. Rovaris, A. Rovira, F. Barkhof and N. de Stefano (2013). "Recommendations to improve imaging and analysis of brain lesion load and atrophy in longitudinal studies of multiple sclerosis." Journal of Neurology **260**(10): 2458-2471.

Vrenken, H., S. Rombouts, P. J. W. Pouwels and F. Barkhof (2006). "Voxel-Based Analysis of Quantitative T1 Maps Demonstrates That Multiple Sclerosis Acts throughout the Normal-Appearing White Matter." American journal of neuroradiology: AJNR **27**(4): 868-874.

Wang, L., H. M. Lai, G. J. Barker, D. H. Miller and P. S. Tofts (1998). "Correction for variations in MRI scanner sensitivity in brain studies with histogram matching." Magnetic Resonance in Medicine **39**(2): 322-327.

Wang, L., J. S. Swank, I. E. Glick, M. H. Gado, M. I. Miller, J. C. Morris and J. G. Csernansky (2003). "Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging." NeuroImage **20**(2): 667-682.

Weinshenker, B. G. (1996). "Epidemiology of multiple sclerosis." Neurologic Clinics **14**(2): 291-308.

Weiss, N., D. Rueckert and A. Rao (2013). "Multiple sclerosis lesion segmentation using dictionary learning and sparse coding." Medical Image Computing and Computer-Assisted Intervention: MICCAI **16**(Pt 1): 735-742.

Wells, W. M., W. L. Grimson, R. Kikinis and F. A. Jolesz (1996). "Adaptive segmentation of MRI data." IEEE Transactions on Medical Imaging **15**(4): 429-442.

Weygandt, M., H.-M. Hummel, K. Schregel, K. Ritter, C. Allefeld, E. Dommes, P. Huppke, J. D. Haynes, J. Wuerfel and J. Gärtner (2015). "MRI-based diagnostic biomarkers for early onset pediatric multiple sclerosis." NeuroImage: Clinical **7**(0): 400-408.

Wong, A. and J. Orchard (2008). "A nonlocal-means approach to exemplar-based inpainting." International Conference on Image Processing: 2600-2603.

Wu, G., Q. Wang, D. Shen and A. s. D. N. Initiative (2011). "Registration of Longitudinal Brain Image Sequences with Implicit Template and Spatial-Temporal Heuristics." NeuroImage **59**(1): 404-421.

Wyman, B. T., D. J. Harvey, K. Crawford, M. A. Bernstein, O. Carmichael, P. E. Cole, P. K. Crane, C. DeCarli, N. C. Fox, J. L. Gunter, D. Hill, R. J. Killiany, C. Pachai, A. J. Schwarz, N.

Schuff, M. L. Senjem, J. Suhy, P. M. Thompson, M. Weiner, C. R. Jack, Jr. and I. Alzheimer's Disease Neuroimaging (2013). "Standardization of analysis sets for reporting results from ADNI MRI data." Alzheimer's & Dementia **9**(3): 332-337.

Xiao, Y., V. Fonov, S. Beriault, I. Gerard, A. Sadikot, Pike and Collins (2014). "Patch-based label fusion segmentation of brainstem structures with dual-contrast MRI for Parkinson's disease." International Journal of Computer Assisted Radiology and Surgery: 1-13.

Yanovsky, I., A. Leow, S. Lee, S. Osher and P. Thompson (2009). "Comparing registration methods for mapping brain change using tensor-based morphometry." Medical Image Analysis **13**(5): 679-700.

Younes, L. (2006). "Combining geodesic interpolating splines and affine transformations." IEEE Transactions on Image Processing **15**(5): 1111-1119.

Yuan, Y., Z. X. Gu and W. S. Wei (2009). "Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural MR imaging for prediction of rapid conversion to Alzheimer disease in patients with mild cognitive impairment: a meta-analysis." American journal of neuroradiology: AJNR **30**(2): 404-410.

Yushkevich, P. A., B. B. Avants, S. R. Das, J. Pluta, M. Altinay, C. Craige and ADNI (2010). "Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: an illustration in ADNI 3 T MRI data." NeuroImage **50**(2): 434-445.

Zhang, D., Q. Guo, G. Wu and D. Shen (2012). Sparse Patch-Based Label Fusion for Multi-Atlas Segmentation. Multimodal Brain Image Analysis. P.-T. Yap, T. Liu, D. Shen, C.-F. Westin and L. Shen, Springer Berlin Heidelberg. **7509:** 94-102.

Zhang, Y., M. Brady and S. Smith (2001). "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm." IEEE Transactions on Medical Imaging **20**(1): 45-57.

Zhou, J. and J. C. Rajapakse (2005). "Segmentation of subcortical brain structures using fuzzy templates." NeuroImage **28**(4): 915-924.

Zhuge, Y. and J. Udupa (2009). "Intensity Standardization Simplifies Brain MR Image Segmentation." Computer Vision and Image Understanding **113**(10): 1095-1103.

Ziemssen, T. (2009). "Multiple sclerosis beyond EDSS: depression and fatigue." Journal of the Neurological Sciences **Suppl 1**: 277.

Zijdenbos, A., R. Forghani and A. Evans (1998). "Automatic Quantification of MS Lesions in 3D MRI Brain Data Sets: Validation of INSECT." Medical Image Computing and Computer-Assisted Intervention: MICCAI: 439.

Zijdenbos, A. P., B. M. Dawant, R. A. Margolin and A. C. Palmer (1994). "Morphometric analysis of white matter lesions in MR images: method and validation." IEEE Transactions on Medical Imaging **13**(4): 716-724.

Zijdenbos, A. P., R. Forghani and A. C. Evans (2002). "Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis." IEEE Transactions on Medical Imaging **21**(10): 1280-1291.

Zivadinov, R. and T. Leist (2005). "Clinical-magnetic resonance imaging correlations in multiple sclerosis." Journal of Neuroimaging **15**(4 Suppl): 10S-21S.