



UNIVERSITA' DEGLI STUDI DI PAVIA
FACOLTA' DI INGEGNERIA
DIPARTIMENTO DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE
CORSO DI LAUREA MAGISTRALE IN BIOINGEGNERIA

TESI DI LAUREA

Investigating Network Spreading Models of disease in Multiple Sclerosis: the role of the cerebellum

Candidato: Arianna Mancini

Relatore: Prof. Claudia Gandini

Correlatore: Dr. Alexandra Young

A.A. 2024/2025

*Alla mia famiglia,
per aver reso possibile la realizzazione dei miei sogni.
Questo traguardo è anche vostro.*

*A chi sta lottando contro la malattia,
nella forza della speranza e nel continuo progresso della ricerca:
non siete soli.*

Contents

Abstract	V
Abstract (Italian version)	IX
1 Introduction	2
1.1 Introduction to Multiple Sclerosis (MS)	2
1.1.1 Magnetic Resonance Imaging in MS	4
1.1.2 Utility of model prediction in studying MS	5
1.1.3 The Cerebellum in MS	6
1.2 Purpose of the Thesis	8
2 Magnetic Resonance Imaging	11
2.1 Image formation	12
2.2 Relaxometry	14
2.3 Diffusion Imaging	16
2.3.1 The Diffusion Process	17
2.3.2 Diffusion MRI sequences	18
2.3.3 Diffusion Tensor Imaging	18
2.3.4 Neurite orientation dispersion and density imaging	20
2.3.5 Diffusion tractography	21
2.3.6 The Cerebellar-cortical loop	23
3 Network Spreading Models	31
3.1 Pathophysiological models	32
3.1.1 Prion-like mechanism	32
3.1.2 Network Diffusion Model	33
3.1.3 Fisher-Kolmogorov-Petrovsky-Piscounov model (KFPP)	36
3.1.4 Weighted-FKPP model	39
3.2 The Network Spreading Model Toolbox	40

4	Materials and Methods	43
4.1	Dataset analyses	44
4.1.1	Human Connectome Project	44
4.1.2	Connectomes	45
4.2	Network Spreading Model Toolbox application	48
4.2.1	Installation and environment setting	49
4.2.2	Network Diffusion Model application	49
4.3	Model outputs as brain-maps	54
4.3.1	Brain-maps pipeline	57
4.3.2	3d Rendering	59
4.4	Validation of NSM results in MS	60
4.4.1	Advanced MRI parameters in MS	61
4.4.2	MS dataset	63
4.4.3	MNI-space registration	65
4.4.4	Voxel-wise z-score maps calculation	66
4.4.5	Statistical tests	70
5	Results	76
5.1	Network Diffusion Model outputs	76
5.2	Brain maps	80
5.2.1	Disease spreading maps from cerebellar subregions	80
5.2.2	Disease spreading maps from the Thalamus	84
5.2.3	Disease spreading maps from Motor and Visual Cortices	87
5.3	Voxel-wise z-score maps	89
5.4	Principal Component Analyses results	94
6	Discussion and conclusions	102
6.1	Discussion	102
6.2	Considerations and future work	108
6.3	Conclusions	110
	Bibliography	112

Abstract

Multiple Sclerosis (MS) is a chronic autoimmune disease of the central nervous system (CNS) characterized by inflammation, demyelination, gliosis, and neuronal loss. Although disease progression has been extensively investigated thanks to advances in Magnetic Resonance Imaging (MRI) technologies, the mechanisms underlying disease onset remain unclear. The cerebellum, with its approximately 100 billion neurons, is commonly affected in MS, leading to both acute and chronic symptoms. Recent studies demonstrate clear cerebellar involvement at disease onset, with visible demyelinated inflammatory lesions and recognized clinical cerebellar symptoms, which are associated with a worse prognosis and increased disability.

In this context, complementary tools such as predictions from computational models can be used to gain deeper insights into disease mechanisms at onset and to overcome intrinsic MRI limitations. Network Spreading Models (NSM) are a set of connectome-based models of disease propagation used to probe mechanisms of pathology spread in neurodegenerative diseases such as Alzheimer or other forms of dementia.

The core of this thesis is the study of a particular category of Network Spreading Models, the Network Diffusion Model (NDM), which can capture disease progression through a diffusive mechanism mediated by the brain white matter network, i.e. the brain structural connectome. The models used for the simulations are implemented in Python within the Network Spreading Model Toolbox available on GitHub. Starting from a healthy structural connectome of the whole-brain, the model can reveal spatial patterns of pathological diffusion across gray matter regions when spread starts from an initially affected area - a seed region that can be defined within the model.

Given the importance of the cerebellum in the involvement of disease, this model is used to simulate the spread of the disease when it is assumed that the onset originates in the cerebellum, treating it as a pathological trigger. The model results

are then validated by comparison with a real dataset acquired at University College London, UK, consisting of patients and healthy controls at the first demyelination event (CIS stage). In addition to the cerebellum, other regions, such as the thalamus, which is known to be particularly vulnerable in the early stages of MS, were also selected as seeds for comparison.

The model simulations revealed distinct spatial patterns, analyzed through temporal graphs and brain maps computed in Python, depending on the assumed origin of the pathology. As expected, when spread begins from the cerebellum, regions more structurally connected to the seed are affected first, such as other cerebellar subregions, Brain-Stem and deep gray matter structures, including the thalamus and the globus pallidus, all of which are known to be involved in MS. Validation was performed through voxel-wise analyses of z-score maps obtained from quantitative MRI parameters reflective of microstructural tissue properties, comparing CIS patients to healthy controls, followed by statistical tests of regional average values. Although the analyses identified clusters of voxels deviating from healthy controls, these largely overlapped with white-matter lesions, which are not incorporated into the model. Statistical analyzes aimed at identifying gray-matter regions showing differences between the two groups and produced some similarities with the model predictions when seeding the model in the cerebellum, although further investigation, such as increasing the sample size or using an additional dataset, is required.

In conclusion, NDM demonstrated its ability to simulate disease spreading through a diffusive mechanism that propagates across gray matter regions from a seed, revealing distinct spatial patterns depending on the selected seed region. However, the current formulation does not incorporate the impact of white matter lesions nor takes into account alternative routes such as cerebrospinal fluid or blood spread. Future extensions should therefore integrate white-matter integrity measures to better align simulations with the biological features of MS. Incorporating these mechanisms will be crucial in developing a framework that more accurately reflects the complexity of MS and its early pathological triggers, potentially improving our understanding of the disease.

This thesis was presented as a poster at the *10th Course of the International School of Brain Cells and Circuits “Camillo Golgi”*, entitled *“A Decade of Discovery: The Erice Camillo Golgi School Modeling the Brain from Neurons to Large-Scale Signals and Digital Twins in Health and Disease”*, directed by Professors Egidio D’Angelo, Claudia Gandini

Wheeler-Kingshott and V. Jirsa, and held in Erice from December 2 to 7, 2025. The work was awarded second place in the posters award competition for the “Application” category.

Abstract (italian version)

La Sclerosi Multipla (MS) è una malattia autoimmune cronica del sistema nervoso centrale caratterizzata da infiammazione, demielinizzazione, gliosi e perdita neuronale. Sebbene la malattia sia stata ampiamente studiata grazie ai progressi nelle tecnologie di Risonanza Magnetica (MRI), i meccanismi alla base dell'esordio rimangono tuttora poco chiari. Il cervelletto, con circa 100 miliardi di neuroni, è comunemente coinvolto nell'MS, contribuendo alla comparsa di sintomi sia acuti che cronici. Studi recenti dimostrano un chiaro coinvolgimento cerebellare fin dalle fasi iniziali, con la presenza di lesioni infiammatorie demielinizzanti e sintomi clinici cerebellari specifici, associati a una prognosi peggiore e ad un aumento della disabilità.

In questo contesto, strumenti complementari come le predizioni di modelli computazionali possono fornire una comprensione più approfondita dei meccanismi patogenetici all'esordio e superare alcune limitazioni intrinseche dell'MRI. I Network Spreading Models sono una classe di modelli basati su dati da Diffusion MRI, e sono stati utilizzati per studiare malattie neurodegenerative come Alzheimer o altre forme di demenza.

Questa tesi si concentra sullo studio di una categoria particolare di Network Spreading Models, il Network Diffusion Model (NDM), che può simulare la progressione della malattia mediante un meccanismo di diffusione mediato dalla rete di materia bianca cerebrale, per esempio dal connettoma strutturale del cervello. Questi modelli sono implementati in Python all'interno del Network Spreading Model Toolbox, che è disponibile su GitHub. A partire da un connettoma strutturale sano dell'intero cervello, il modello può identificare pattern spaziali di diffusione della patologia nelle regioni di sostanza grigia quando la propagazione ha origine da un'area specifica - una seed region, definibile nel modello.

Dato il coinvolgimento patologico del cervelletto in MS, il modello è stato utilizzato per simulare la diffusione della malattia assumendo un'origine cerebellare,

trattandolo come un possibile innesco. I risultati del modello sono stati quindi validati confrontandoli con un dataset reale acquisito presso lo University College London (UK), composto da pazienti al primo evento di demielinizzazione (stadio CIS) e controlli sani. Oltre al cervelletto, anche altre regioni, come il talamo - noto per essere particolarmente vulnerabile nelle fasi iniziali di MS - sono state selezionate come seed.

Le simulazioni hanno rivelato pattern spaziali distinti, analizzati tramite grafici temporali e mappe cerebrali generate in Python, a seconda della regione di origine. Come previsto, quando la diffusione si propaga a partire dal cervelletto, le aree maggiormente connesse alla seed risultano le prime ad essere colpite, come per esempio le altre regioni cerebellari, il Brain-Stem e strutture di sostanza grigia profonda quali talamo e globus pallidus, tutte note per essere coinvolte nell'MS. La validazione è stata effettuata mediante analisi voxel-wise basate sul calcolo di mappe di z-score ottenute da parametri quantitativi di MRI, rappresentativi delle proprietà microstrutturali dei tessuti, confrontando i pazienti CIS con i controlli sani; seguita da test statistici sui valori medi regionali. Sebbene tali analisi abbiano identificato alcuni cluster di voxels devianti dal gruppo di controllo, essi risultavano in gran parte sovrapposti a lesioni della sostanza bianca, che però non vengono trattate nel modello. Le analisi statistiche, che avevano come obiettivo quello di individuare regioni di grigia che mostrassero differenze tra i due gruppi, hanno prodotto risultati coerenti con quelli del modello; tuttavia richiedono ulteriori approfondimenti - ad esempio ampliando la numerosità del campione o utilizzando un altro dataset.

In conclusione, NDM ha dimostrato di saper simulare la propagazione della patologia attraverso un meccanismo di diffusione, vincolato dal connettoma strutturale, che parte da un seed e si propaga nelle regioni di sostanza grigia, rivelando pattern spaziali distinti in base alla regione iniziale selezionata. Tuttavia, l'attuale formulazione non integra l'impatto delle lesioni nella sostanza bianca né considera vie alternative di diffusione, come la propagazione mediata dal liquido cerebrospinale o dal sistema vascolare. Futuri sviluppi dovrebbero quindi includere questi aspetti per allineare meglio le simulazioni alle caratteristiche biologiche dell'MS. L'integrazione di questi meccanismi potrà permettere lo sviluppo di un framework che ne rispecchi più accuratamente la sua complessità e i suoi trigger patologici precoci, contribuendo a migliorare la comprensione della malattia.

**Questa tesi è stata presentata alla poster session che si è tenuta al
10th Course of the International School of Brain Cells and Circuits**

“Camillo Golgi”, intitolato *“A Decade of Discovery: The Erice Camillo Golgi School Modeling the Brain from Neurons to Large-Scale Signals and Digital Twins in Health and Disease”*, diretto dai Professori Egidio D’Angelo, Claudia Gandini Wheeler-Kingshott e V. Jirsa e tenutosi a Erice dal 2 al 7 dicembre 2025. Il lavoro ha ottenuto il secondo posto nella poster award competition per la categoria “Application”.

Chapter 1

Introduction

This thesis aimed to tackle the issue of MS pathogenesis, exploring modeling approaches of disease spreading able to capture the early trigger mechanisms underlying the initiation of MS.

1.1 Introduction to Multiple Sclerosis (MS)

Multiple sclerosis (MS) is a chronic autoimmune disease of the central nervous system (CNS), characterized by inflammation, demyelination, gliosis, and neuronal loss (see details in Section 4.4.1).

Although the progression of the disease has been extensively investigated thanks to advances in technologies such as **magnetic resonance imaging (MRI)**, the mechanisms underlying its onset remain unclear. Recent studies suggest that a combination of genetic susceptibility and environmental exposures, acting from early development through young adulthood, contributes to the development of the disease. Several factors have been identified as having a strong influence on genetic predisposition to MS and disease risk, including vitamin D deficiency, season of birth, tobacco use, and exposure to the Epstein-Barr virus [9].

MS affects approximately 2.3 million people worldwide and shows a higher prevalence in females than in males. The disease is most commonly diagnosed between 20 and 50 years of age.

The primary targets of MS are the **myelinated axons** of the CNS, resulting in variable degrees of damage to both myelin and axons. A **clinical relapse** is defined as the appearance of new or worsening neurological symptoms lasting for at least 24 consecutive hours, in the absence of fever or infection, and followed by a period of clinical stability or improvement of at least 30 days.

The destruction of myelinated axons in the CNS is primarily mediated by im-

munogenic **T cells** that secrete pro-inflammatory cytokines, promoting a **T helper 1 (Th1)**-driven immune response. These T helper cells play a central role in initiating myelin damage in MS [9].

Some of the clinical manifestations of MS are Ataxia, Diplopia, Cognitive dysfunction, Sensory symptoms, Depression, Fatigue, Optic Neuritis, Spasticity, Anxiety and others.

There are various type of treatments which includes managing severe attacks, disease-modifying therapies that lessen MS biological activity and treatment for symptoms. Currently there are no therapies that supports remyelination or brain repair, but a number of potential strategies are being intensively researched.

The International Advisory Committee on Clinical Trials in MS categorized clinical course descriptors in 1996, with a revision in 2013 [14]. This Committee defines three main clinical courses:

- **Relapsing-remitting disease (RRMS)** - characterized by acute attacks followed by recovery;
- **Primary-Progressive (PPMS)** - characterized by a gradual worsening from onset;
- **Secondary-Progressive (SPMS)** - characterized by relapsing-remitting at onset but gradual worsening later in the disease course.

Furthermore, **Clinically Isolated Syndrome (CIS)** has been included among the clinical descriptors of MS and is recognized as the first clinical manifestation of the disease. CIS is characterized by a single episode of neurological symptoms consistent with inflammatory demyelination and suggestive of MS [14].

Despite the use of clinical descriptors to classify MS, accumulating evidence suggests that patients with MS share qualitatively similar (but quantitatively different) pathological features, regardless of clinical course. These features include inflammation and neurodegeneration, both of which are already present at disease onset. Variability in disease expression is influenced by patient-specific factors such as age, sex, genetic background, disease duration, and social and environmental exposures.

As a result, disability progression in MS can be viewed as the outcome of multiple interacting mechanisms that vary both between patients and within individual patients over time. For this reason, clinical characterization and treatment selection should increasingly take into account underlying pathophysiological mechanisms, in addition to traditional clinical descriptors [13].

Currently, the diagnostic process is based on the demonstration of at least two clinically relevant episodes and evidence of disease involvement in at least two anatomically distinct white matter (WM) regions of the CNS. Neurological symptoms must persist for more than 24 hours and occur in separate episodes separated by a minimum interval of one month [9]. The diagnosis of MS is currently established according to the **McDonald criteria**, which have evolved over the years to consider the brain but also the spinal cord and the optic nerve.

The primary tool for diagnosis and disease monitoring in both clinical practice and scientific research in MS is **MRI**, from which the understanding and management of MS have greatly benefited.

According to the McDonald criteria, a diagnosis of MS can be established at the first clinical demyelinating event (CIS) when MRI of the brain demonstrates demyelinating lesions with dissemination in both space and time [9].

1.1.1 Magnetic Resonance Imaging in MS

MRI (see technical details in Chapter 2) has, since its early introduction in the 1980s, become an essential tool for supporting the *diagnosis, longitudinal monitoring, assessment of therapeutic response, and scientific investigation* of MS [12].

The clinical use of MRI has increased in parallel with technological advances, and the widespread adoption of clinically routine MRI at 1.5 T has enabled both qualitative and quantitative assessment of CNS inflammatory demyelinating lesions and tissue atrophy. Clinical routine MRI is commonly referred to as **conventional MRI**.

Conventional MRI comprises a set of widely available, well-characterized, and highly standardized imaging protocols, including **T2-weighted**, **Fluid-Attenuated Inversion Recovery (FLAIR)**, **Short Tau Inversion Recovery (STIR)**, and **T1-weighted** sequences acquired at a magnetic field strength of 1.5 T in both the brain and spinal cord. These sequences are routinely used to guide clinical decision-making.

Despite its high diagnostic sensitivity, conventional MRI lacks specificity for the underlying MS pathology and is limited in the quantitative metrics required for robust clinical validation and prognostication. Indeed, conventional MRI is often unable to detect or distinguish ongoing pathological changes in normal-appearing WM (NAWM) [12].

For this reason, another category of MRI has been introduced in research studies: **advanced MRI**. Advanced MRI pulse sequences acquired at higher magnetic field strengths (3 T or 7 T), such as *magnetization transfer imaging, magnetic resonance*

spectroscopy, and **diffusion-weighted imaging (DWI)**, offer the opportunity to increase diagnostic precision with respect to the MS pathological processes and to improve clinical correlation as well as prediction of disability accumulation. **DWI-based tractography**, in particular, has emerged as a highly attractive tool among diffusion-based metrics, providing insights into the mechanisms underlying the development of both physical and cognitive impairment in cross-sectional and longitudinal studies (see technical details in Chapter 2.3).

Unfortunately, the clinical implementation of advanced MRI techniques remains limited, owing to variability in hardware availability, acquisition protocols, and other technical factors across institutions [12].

1.1.2 Utility of model prediction in studying MS

Beyond technical advancements in MRI for studying the evolution of pathology, other complementary tools, such as the **prediction from computational models**, are increasingly being used to gain a deeper insight into disease mechanisms and can be used to address questions related to disease onset, thereby overcoming some of the intrinsic limitations of MRI and of classical statistical approaches.

For example, an unsupervised machine learning algorithm has been applied to brain MRI scans to classify MS subtypes based on pathological features [6]. This algorithm, named **Subtype and Staging Inference (SuStain)**, identifies subgroups of subjects within the training dataset, where each subtype is characterized by a distinct pattern of changes across a set of features, such as MRI-derived abnormalities.

The resulting model is able to estimate the degree of similarity between an unseen patient’s MRI and each identified subtype and disease stage. In this study, the algorithm was used to redefine MS subtypes through a data-driven assessment of pathological changes observable on MRI scans, rather than relying on the evolution of clinical symptoms. Using this approach, three subtypes were identified: *cortex-led*, *normal-appearing white matter (WM)-led*, and *lesion-led*.

This model has demonstrated the ability to identify MRI-based subtypes that provide information on the underlying pathobiological mechanisms of MS and, based on MRI data, predict disease activity, disability progression, and treatment response more accurately than conventional clinical phenotypes. As a result, such an approach may be used to define more homogeneous patient groups for interventional clinical trials.

In another study, a **data-driven event-based model** was employed to infer the temporal sequence in which gray-matter (GM) regions become atrophic in MS and to investigate its association with disability accumulation [7]. Although GM atrophy

is known to be present from the earliest stages of MS, its temporal ordering had remained poorly understood. Moreover, GM atrophy in MS is spatially heterogeneous, with certain regions being more susceptible to degeneration than others.

By applying a data-driven model to a large MS cohort, the study demonstrated that GM atrophy progressively spreads over time to involve an increasing number of brain regions. Importantly, the inferred sequence of atrophy progression can be used to stage patients along the disease course automatically, providing a quantitative measure of neurodegenerative burden that is associated with disability accumulation.

Despite the availability of numerous tools, including imaging-based models to study disease pathology, the onset of MS remains unclear. While the evolution of pathological processes has been extensively investigated using MRI and data-driven models, critical information regarding the initial origin of the disease is still lacking.

Alternative modeling approaches may have the potential to capture the early *trigger* mechanisms underlying the initiation of MS. For instance, such models could aim to identify specific brain regions or biological conditions that act as initial triggers, from which inflammatory processes originate and subsequently propagate, leading to tissue damage and the emergence of clinical symptoms.

These models, known as Network Spreading Models (NSM), are the core of this thesis (see Chapter 3), with a particular focus on the Network Diffusion Model (NDM) (see details in Section 3.1). This model has been shown, especially in neurodegenerative diseases, to capture disease progression through a diffusion-like mechanism mediated by the brain’s WM network. These computational models, which are based on the structural connectome derived from diffusion MRI (dMRI), can simulate how MS may start and spread through the brain.

1.1.3 The Cerebellum in MS

While the cerebellum represents only about 10% of the total brain volume, it contains more than **100 billion neurons**, comprising approximately the same number of neural cells as the cerebrum (see details in Section 2.3.6). Generally, lesions in the cerebellum cause qualitative dysfunction rather than a complete loss of function, resulting in disorganized movements, loss of fine-tuning, and incomprehensible speech in patients [21]. MS commonly affects the cerebellum, leading to both acute and chronic symptoms. Indeed, cerebellar signs contribute significantly to clinical disability. Cerebellar pathology may cause limb, gait, and truncal ataxia, depending on the precise lesion site, as well as other features such as gaze-evoked nystagmus, dysarthria, and tremor. All these symptoms are particularly difficult to treat [1].

Consequently, patients with MS not only suffer from the more clinically obvious cerebellar motor dysfunction but also from specific cognitive deficits, likely related to cerebellar involvement, performing less well on tests of information processing speed, verbal fluency, and memory tasks, for example.

Clear involvement of the cerebellum at the onset of the disease, i.e., with visible demyelinated inflammatory lesions and traditionally recognized clinical cerebellar symptoms, has also been shown to be associated with a worse prognosis, and cerebellar relapses are associated with poor recovery, which in turn is associated with earlier onset of progressive disease. In general, cerebellar dysfunction is a feature of both RRMS and progressive MS, leading to a range of neurological manifestations, and its involvement is associated with increased disability and poorer prognosis [1].

The cerebellum is a predilection site for lesion development in MS; in fact, both GM and WM demyelination are observed, although the former has been shown to be more extensive and comparable to that in the spinal cord[8]. The extent of tissue demyelination in the cerebellum appears to correlate with the overall severity of MS-related pathology. However, histopathological reports demonstrating significant widespread neuronal and axonal degeneration in the cerebellum in MS suggest that diffuse, generalized cerebellar pathology is at least partially independent of focal lesions [21].

Furthermore, it has been demonstrated that “normal-appearing” cerebellar WM can be damaged at a very early stage of RRMS, even when no gross cerebellar lesions are visible using conventional MRI sequences. In fact, by employing a technique based on **Diffusion tensor imaging (DTI)**, cerebellar microstructural alterations can be detected even in patients who do not exhibit clinical signs of cerebellar dysfunction. This suggests that cerebellar microstructural changes are more prevalent than indicated by conventional MRI and may be associated, in progressive disease, with cerebellar atrophy [5].

Regardless of neuronal pathology in the MS cerebellum, some studies show a reduction in **Purkinje cell** density within lesions and alterations in neurofilament phosphorylation states. Additional changes in Purkinje cells have been documented in MS, including modifications in **ion channel** expression and receptor profiles. The **Na_v 1.8** sensory neuron-specific sodium channel, normally expressed at very low levels in Purkinje cells, is upregulated in MS, along with annexin light chain, which facilitates the functional expression of this ion channel. An interesting phenomenon of Purkinje cell heterokaryon formation has been observed in the MS cerebellum, which may represent a protective mechanism against injury. In MS, binucleated Purkinje cells are significantly increased compared to controls, suggesting a possible

adaptive process in response to the disease [1].

Despite recent studies, much remains to be understood about the early pathology of the cerebellum in MS and its potential role in the disease spreading. The cerebellum is a common site of MS-related pathology; however, its high connectivity with other regions of the CNS makes it challenging to precisely determine its role in clinical disability and its specific correlation with pathology, particularly in the early stages of the disease [21]. A better understanding of this relationship may help in designing future therapies.

1.2 Purpose of the Thesis

The purpose of this thesis is to investigate the possible role of the cerebellum in MS pathogenesis as a trigger of disease spreading along neuronal pathways, through the study of **Network Spreading Models** (see Chapter 3). In particular, the focus will be on the **Network Diffusion Model**, which is implemented in Python within the Network Spreading Model Toolbox available on GitHub [26]. Starting from diffusion-tractography images of healthy subjects, the model can simulate the propagation of pathology when the disease originates from specific regions of interest (ROIs).

Given the interest in the **cerebellum**, cerebellar subregions will be selected as the initial sites of damage in the model, allowing the spatial pattern of disease propagation across GM regions to be examined when the pathology is hypothesized to start from this structure. At the current stage, the model simulates damage in GM only; WM lesions are not yet included. The propagation mechanism follows the structural connectivity architecture of the brain, modeling disease as a diffusion-like process along pathways linking GM regions.

To assess whether the spatial patterns generated by the model resemble those observed in MS, the second stage of this thesis involves analyzing a real dataset acquired at the time of the first demyelinating event (**CIS stage**) (see Section 4.4.4). The objective is to determine whether the simulated patterns show any correspondence with the distribution of damage in patients. The broader aim is to explore computational models that may contribute to a deeper understanding of the earliest stages of MS, with a specific focus on the cerebellum as a potential initiator of the pathological process.

Although the NDM has been used to study the progression of several neurodegenerative conditions, to the best of my knowledge it has not previously been applied to investigate neuroinflammatory disorders such as MS and using a whole-brain connec-

tome which includes the cerebellum. Alongside the cerebellum, other regions, such as the thalamus, have also been used as seeds in the model simulations. The thalamus is known to be particularly vulnerable in MS, partly due to its proximity to the ventricles, which may facilitate the spread of inflammation through the cerebrospinal fluid (CSF), and partly because of its strong involvement in the cerebro-cerebellar loop and its extensive WM connections.

Damage to the thalamus observed in the model could therefore support the hypothesis that cerebellar involvement may act as an upstream trigger, initiating inflammation that ultimately contributes to the onset of MS. The comparison between the model simulations and the MS dataset is performed through visual inspection of brain maps and statistical analyses.

Chapter 2

Magnetic Resonance Imaging

MRI is a technique that takes advantage of the fact that hydrogen nuclei (protons) behave like tiny magnets when placed in a magnetic field. MRI is particularly sensitive to proton density, which reflects the amount of water in tissues. It should be noted that up to 60% of the adult human body is composed of water, and the brain alone contains about 73% water, which means our tissues are rich in protons.

MRI is also sensitive to how protons behave within a magnetic field. Inside the magnetic field of the scanner, protons reach an *equilibrium state* and are then excited. After this *excitation phase*, a *relaxation period* follows, during which signals are emitted and captured. These signals are used to construct images during the *acquisition phase*.

One major advantage of MRI is that it uses **electromagnetic radiation** in the *radio-frequency* range. Unlike other imaging techniques such as X-rays, Positron Emission Tomography (PET), or Computed Axial Tomography (CT), MRI does not use ionizing radiation. Radio-frequency waves are commonly used in everyday radio transmissions and are generally considered safe. Therefore, MRI has minimal side effects, as any tissue heating caused by the magnetic field can be effectively controlled.

Regardless of the broader electromagnetic spectrum, the frequency band used by clinical MRI scanners typically ranges from around 30 MHz to several hundred MHz, depending on the magnetic field strength. This falls within the radiofrequency (RF) portion of the electromagnetic spectrum.

Another important aspect is that MRI images can be made sensitive to specific tissue properties. This means that the same anatomical structure can appear very different depending on the imaging parameters (or "weighting"). These properties are often associated with tissue damage, so pathological changes can alter image contrast and help diagnose.

2.1 Image formation

The MRI system is based on three main components:

- Static magnetic field
- Radiofrequency (RF) coil
- Magnetic field gradients along the x, y, and z axes

It is important to remember that a proton has three key properties which are *electric charge*, *magnetic characteristics*, and the ability to *spin* around its own axis. The magnetic field is essential for exploiting these electromagnetic properties. Specifically, the proton can be modeled as a charged particle spinning around its axis, and its magnetic behavior is described by the concept of spin or **intrinsic angular momentum**, represented by the equation below:

$$P = \frac{h}{2\pi} \sqrt{I(I+1)} \quad (2.1)$$

where P = spin or intrinsic angular momentum, I = spin quantum number (in case of the proton $I = \frac{1}{2}$), h = Planck constant.

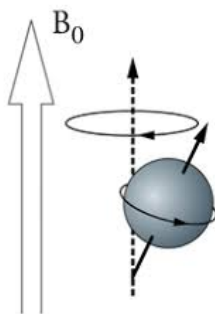


Figure 2.1: Proton behavior when placed in a magnetic field B_0 ¹

Figure 2.1 illustrates how a proton, which can be thought of as a tiny bar magnet, behaves when placed in a constant **magnetic field** B_0 with a defined direction and magnitude. The proton precesses at a specific frequency known as the **Larmor frequency**:

$$\omega_L = \gamma * B_0 \quad (2.2)$$

that depends on the magnetic field B_0 and the gyromagnetic ratio γ .

¹Available at: <https://gioaboffa.wordpress.com/2018/03/31/molecole-che-passione-8-settembre-2015-tumori>

Given the intrinsic angular momentum P , each nucleus has a **magnetic dipole moment** μ , which has both magnitude and spatial orientation depending on its spin state:

$$\mu = \gamma * P \quad (2.3)$$

From this we can define **magnetization** M , which is the vector sum of all magnetic dipole moments per unit volume:

$$M = \frac{1}{V} * \sum_i \vec{u}_i = 0 \quad (2.4)$$

When no external magnetic field is present, the dipole moments are randomly oriented, resulting in $M=0$. In the presence of B_0 , the nuclei begin to precess around the direction of B_0 at the Larmor frequency. This leads to the formation of a net magnetization M_0 , which is a positive equilibrium value. Spins can align either parallel (lower energy) or antiparallel (higher energy) to B_0 . In any given volume V , there is always a slight excess of spins in the lower energy state, resulting in a net magnetization M aligned with B_0 .

To measure M_0 , it must be rotated into the plane transverse to B_0 . This requires the generation of an electromagnetic field $B_1(t)$ with the following properties:

- It must oscillate at the Larmor frequency
- It must be applied perpendicular to B_0
- It must have a specific amplitude and duration to rotate the magnetization $M(t)$ into the transverse plane defined by B_0 and $B_1(t)$

$B_1(t)$ can be defined using the following formula:

$$B_1(t) = B_1 * \cos(\omega * t) \quad (2.5)$$

Radio-frequency coils are used to generate this electromagnetic field. The transmission coil emits the RF pulse $B_1(t)$ at the Larmor frequency, which interacts with M_0 , rotating the magnetization by 90° . The receiver coil then detects the signal from the transverse magnetization $M_{xy}(t)$, which precesses in the xy -plane around B_0 . This precession induces a varying magnetic flux in the receiver coil.

MRI also uses **magnetic field gradients**, which are variations in the static magnetic field B_0 along specific spatial directions. These gradients allow the precession frequency to vary with position, enabling spatial localization. Typically, the static magnetic field has a strength of $1.5T$, $3T$, or $7T$. Gradients can modify B_0

by 1 Gauss (10^{-4}) per centimeter along a given direction. This creates a frequency distribution that allows spatial differentiation along one axis (e.g., the x -axis), using a process called *frequency encoding*, which relies on the Fourier Transform.

To encode spatial information along the y -axis, a *phase encoding* gradient is applied. This introduces phase shifts that can also be decoded using the Fourier Transform, allowing signals to be assigned to their correct spatial positions.

Finally, gradients along the z -axis enable *slice selection*. By exciting only a specific frequency range with $B_1(t)$, a particular slice of tissue can be imaged.

Image formation is achieved by applying a **2D Inverse Fourier Transform** to the **K-space**, which contains all the acquired spatial information.

The combined effects of the three magnetic field gradients and encoding processes fill the K-space, enabling accurate image reconstruction.

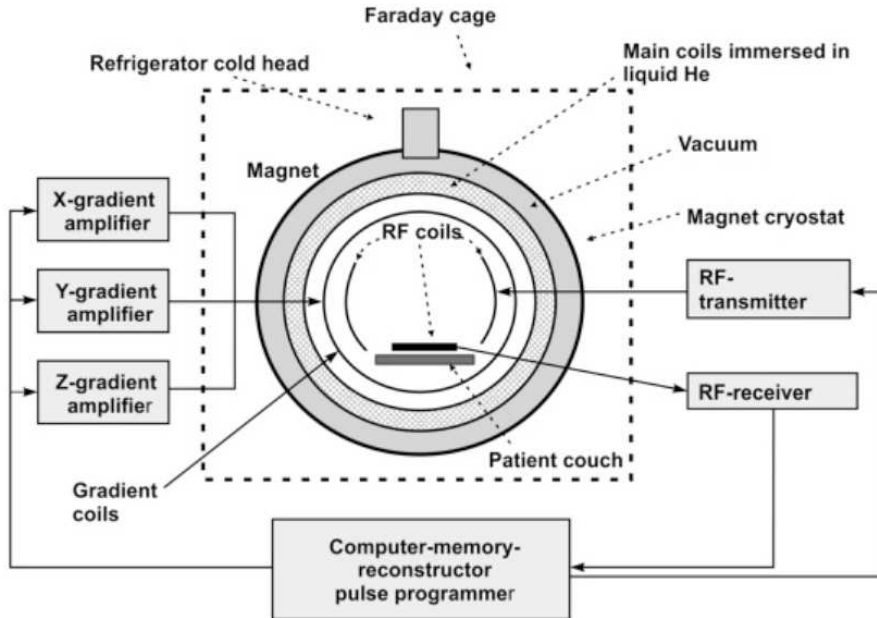


Figure 2.2: Basic components and architecture of an MR system

[17]

2.2 Relaxometry

As mentioned at the beginning of the chapter, following the excitation phase, a relaxation period occurs during which signals are emitted and captured. This *relaxation process* takes place as the system returns to equilibrium. Over time, M_{xy} returns to M_0 , aligned with B_0 . The recovery of the transverse magnetization M_{xy} induces a signal in the receiver coils, known as **Free Induction Decay (FID)**, which is governed by two main mechanisms:

- *Spin-lattice relaxation*
- *Spin-spin relaxation*

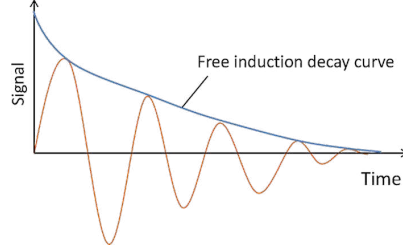


Figure 2.3: Free Induction Decay signal²

The longitudinal spin-lattice relaxation refers to the process by which the magnetization $M(t)$ realigns with the direction of the static magnetic field. This process depends on the microstructure and is described by an increasing exponential function characterized by the time constant T_1 , which represents the time required for the magnetization to return to equilibrium.

The transverse spin-spin relaxation, on the other hand, is caused by the loss of coherence among the spins. This dephasing leads to the disappearance of the transverse magnetization. Like spin-lattice relaxation, spin-spin relaxation is characterized by a time constant, T_2 , which corresponds to the time needed for M_{xy} to be reduced by 63% of its initial value. Unlike the former, this process follows a decreasing exponential function.

Every tissue-type, such as fluids, water-based tissues and fat-based tissues, is characterized by its own parameters:

- Proton density
- Spin-lattice relaxation time T_1
- Spin-spin relaxation time T_2

By configuring specific **sequences**, it is possible to control the influence of each property on the signal and the contrast of the final image. A sequence is a program provided to the scanner to obtain the desired signal, characterized by parameters such as **Echo Time (TE)** and **Repetition Time (TR)**, which can be adjusted by the operator. The TE is the time interval between the application of the RF pulse and the peak of the signal (echo) detected by the MRI scanner. A longer TE

²Sarah Abdulla, *T1 and T2 Signal*, Radiology Cafe, October 2021. Available at: <https://www.radiologycafe.com/frcr-physics-notes/mr-imaging/t1-and-t2-signal/>

emphasizes T_2 -weighted contrast, while a shorter TE reduces the influence of T_2 decay.

The TR refers the time interval between successive RF excitation pulses applied to the same slice in an MRI sequence. Shorter TR values emphasize T_1 -weighted contrast, while longer TR values allow more complete recovery of longitudinal magnetization, reducing T_1 effects and enabling other types of contrast (e.g., T_2 or proton density).

This allows the acquisition of weighted images, meaning that the signal intensity reflects primarily one property and becomes independent of the others. For example:

- In PD-weighted images, the water content is enhanced;
- T_2 -weighted images are typically used for pathology scans;
- T_1 -weighted images are commonly used for anatomical scans.

Sometimes, *contrast agents* are used to improve the specificity of MRI by providing additional weighting to help discriminate between different pathologies. In such cases, the most commonly used contrast agent is **Gadolinium** (Gd), which can penetrate the *blood-brain barrier*. For example, in inflammatory lesions caused by MS, these areas are enhanced. Furthermore, in highly vascularized tumors, Gd penetrates regions where the blood-brain barrier is disrupted, enhancing the areas invaded by the pathology.

Beyond the main tissue properties, MR can also be made sensitive to other characteristics such as water diffusion mechanisms, magnetization transfer, blood oxygenation, and more. This makes MRI a powerful tool for:

- probing tissue structures (*tractography*),
- exploring different water compartments (*myelin imaging*),
- studying brain function (*fMRI*),
- investigating *in vivo* chemistry (*spectroscopy*),
- assessing the integrity of the vascular system (*angiography*).

2.3 Diffusion Imaging

MRI can be sensitized to the motion of protons within biological tissues, thereby enabling the characterization of **tissue microstructure** through a specialized technique known as *Diffusion Imaging*.

2.3.1 The Diffusion Process

The diffusion process refers to the general transport of matter, whereby molecules intermingle due to thermal agitation in a stochastic manner. As each molecule moves independently of the others, collisions are inevitable, resulting in a random displacement of individual molecules, a phenomenon commonly referred to as the *random walk*.

Two principal forms of diffusion are typically considered:

- *Isotropic* diffusion;
- *Anisotropic* diffusion.

Diffusion is termed *isotropic* when water molecules are free to move equally in all directions, without encountering structural barriers that constrain their motion. Under such conditions, the spatial distribution of diffusing molecules assumes a spherical geometry.

This process can be quantitatively described using the "**root mean square displacement**" (**r.m.s.d.**), which increases over time. The rate of increase is governed by a proportionality constant known as the *diffusion coefficient*, denoted by **D**.

The Einstein relation for r.m.s.d. is expressed as:

$$\text{r.m.s.d.} = \sqrt{r^2} = \sqrt{6Dt} \quad (2.6)$$

Different substances exhibit distinct diffusion coefficients **D**. For instance, water possesses a relatively high diffusion coefficient due to the rapid mobility of its molecules.

Regardless of *anisotropic* diffusion, this process occurs when barriers and tissue microstructure restrict or hinder the diffusion of water, so water can move more freely along certain directions than others. For this reason, the propagation of water molecules can be described as an ellipsoid. This typology of diffusion is typical in biological tissues such as brain WM, skeletal muscle, kidney and cardiac muscle, which are characterized by the presence of heterogeneous but spatially ordered macromolecular, membranous, and cellular compartments [2]. In this case, Einstein's equation is no longer valid because the measured value **D** can change during the observation time. Since the real **D** is unknown, hence the actual diffusion curve, a new term can be defined: *apparent diffusion coefficient* (**ADC**).

2.3.2 Diffusion MRI sequences

Since MRI must be sensitive to the diffusion process, it becomes necessary to use **diffusion magnetic field gradients** to encode the position of the spins that move during the experiment. If spins move during the process, they change their *Larmor frequency*.

The *spin echo sequence* allows for the acquisition of a signal that measures the diffusion of protons. It consists of a $90^\circ B_1(t)$ RF pulse followed, after a $TE/2$ time interval, by a $180^\circ B_1(t)$ refocusing RF pulse that rewinds the FID. This process enables the formation of the **echo signal**. The first diffusion gradient creates a phase distribution of the spins, which is inverted by the second diffusion gradient. During phase evolution, if the spins are diffusing along that direction, their phases will only partially reverse, resulting in an attenuated signal. Specifically, the greater the signal attenuation, the greater the proton diffusion.

The *b-factor* is an important parameter that controls the amount and direction of diffusion weighting, and in particular determines the sensitivity to the diffusion process. It is defined by setting the amplitude, duration, separation, and direction of the diffusion gradients.

$$b = (\gamma G \delta)^2 \left(\Delta - \frac{\delta}{3} \right) \quad (2.7)$$

where γ is the *gyromagnetic ratio*, G is the magnitude of the diffusion gradient pulses, δ is the width of the diffusion gradient pulses and Δ is the time interval between the two.

The measure of the diffusion coefficient **ADC** is not *rotationally invariant* because it depends on the position of the patient relative to the magnet. Therefore, depending on the b-factor setting and thus the direction of the diffusion gradient, the image contrast is adjusted along that direction, reflecting the diffusion occurring in that direction. Therefore, anisotropy measures that use ADC and are based on DWIs are not quantitative because they lack objectivity [2].

2.3.3 Diffusion Tensor Imaging

DTI is a MR imaging modality which provides unique microstructural and physiological information [2]. The diffusion tensor **D**, defined below, can be calculated by measuring diffusion in at least six non-collinear directions and provides the most

accurate description of diffusion.

$$D = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix} \quad (2.8)$$

There are two ways to extract and present this information: one approach is to construct three-dimensional fiber maps and diffusion ellipsoid images, which emphasize the three-dimensional nature of diffusion in tissues; alternatively, the information contained in the six independent elements of the tensor can be summarized by deriving a set of scalar quantities [2].

The tensor can be diagonalized to give:

- **Eigenvectors** - $\epsilon_1, \epsilon_2, \epsilon_3$, representing the principle directions of diffusion;
- **Eigenvalues** - $\lambda_1, \lambda_2, \lambda_3$, representing the magnitude of diffusion along these directions.

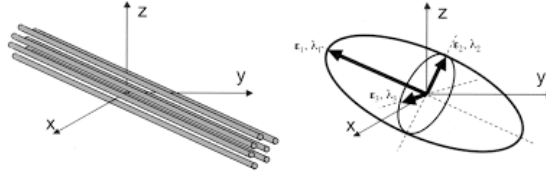


Figure 2.4: Graphic representation of the diffusion tensor: the three-dimensional tensor is characterized by three eigenvectors and three eigenvalues which represent the direction and magnitude of diffusion along the three principal directions (x,y,z).³

The diffusion tensor can be characterized by its shape - ranging from oval to round - which reflects high or low anisotropy respectively, and by its color, which indicates the dominant orientation of the tensor.

As introduced earlier, it is also possible to calculate rotationally invariant indices from the diffusion tensor [2], such as:

- **Trace** - $TR = \lambda_1 + \lambda_2 + \lambda_3 = 3ADC_{mean}$;
- **Mean Diffusivity** - $MD = \frac{trace}{3} = ADC_{mean}$;
- **Fractional Anisotropy** - FA, which quantifies how directional the diffusion of water is within biological tissues.

³Garic et al., *In vivo Restricted-Diffusion Imaging (RDI) is sensitive to differences in axonal density in typical children and adults*, 2020. doi: 10.1101/2020.08.17.254656.

The derivation and visualization of these quantitative scalar parameters, directly obtained from the diffusion tensor, is referred to as quantitative diffusion-tensor MRI. This approach enables the characterization of intrinsic features of tissue microstructure and microdynamics without being influenced by the choice of the laboratory coordinate system [2].

The diffusion tensor is a mathematical representation of a highly complex reality, which is why other techniques, such as **tractography**, are also used for analysis of diffusion.

2.3.4 Neurite orientation dispersion and density imaging

Neurite orientation dispersion and density imaging (NODDI) [31] is a practical dMRI technique for estimating the microstructural complexity of dendrites and axons *in vivo* on clinical MRI scanners. Neurite morphology offers a window into the structural basis of brain function; in fact, it is a key marker of brain development and aging, and its alterations are implicated in several neurological disorders, including MS.

NODDI indices provide more direct and specific markers of brain tissue microstructure compared to standard DTI metrics. Although DTI is sensitive to microstructural features, it lacks specificity for individual components of tissue architecture. Metrics such as FA or MD are widely used as surrogate measures of microstructural change, but they are inherently non-specific. As a consequence, variations in these measures cannot be unequivocally attributed to specific alterations in tissue microstructure.

NODDI is based on a tissue model with three types of microstructural environment which are the *intra-cellular*, the *extra-cellular* and cerebrospinal fluid (CSF) compartment. Each affects water diffusion within the environment in a unique way and gives rise to a separate normalized MR signal. The three compartment are described as follows:

- *Intracellular model*: it refers to the space bounded by the membrane of neurites, and it is modeled as a set of **sticks**, a cylinder of zero radius. The orientation distribution of sticks can be from highly parallel or highly dispersed;
- *Extra-cellular model*: it refers to the space around the neurites which is occupied by various types of glial cells and somas in GM. The diffusion of water molecules is hindered by the presence of neurites but not restricted, so it is modeled with gaussian anisotropic diffusion;
- *CSF compartment*: it models the space occupied by CSF as isotropic gaussian diffusion.

The NODDI model is characterized by a set of parameters that include the **intra-cellular volume fraction** (V_{fintra}), the **isotropic volume fraction** (V_{fiso}), and the **orientation dispersion index** (ODI), which are used as metrics for the voxel-wise analyses (see details in Chapter 4.4.4). FA has been shown to correlate with the NODDI-derived microstructural metrics V_{fintra} and ODI.

In addition, neurite density measures have been shown to correlate with the intensity of myelin staining and can therefore be used as markers for demyelinating disorders, while ODI quantify, in WM, the bending and fanning of axons and, in GM, the pattern of sprawling dendritic processes and provides a marker of microstructural complexity.

2.3.5 Diffusion tractography

One of the most commonly used diffusion reconstruction approaches is *Diffusion tractography*, which utilizes the diffusion tensor and represents it as a tract, corresponding to the dominant orientation of the ellipsoid. This process enables the estimation of the potential distribution of fibers in WM, and thus, the reconstruction of fiber bundles in the brain.

There are two main types of fiber reconstruction:

- **Streamline Method** (Deterministic);
- **Probabilistic Method**.

The streamline method, which is used to segment tracts and to extract FA and MD values within them, is simpler and faster, but it is also more sensitive to noise. On the other hand, the probabilistic method, which provides a connectivity value at each voxel, is more robust to noise and accounts for uncertainty in the diffusion direction.

Therefore, tractography is useful for obtaining structural connectivity measures between brain regions (*structural connectivity matrices*), which, when combined with *functional connectivity matrices*, can be used to analyze brain dynamics.

From the brain connectome to the structural connectivity matrix

To obtain a connectivity matrix, it is important to have two main pieces of information, which are synthesized into *nodes*, obtained from a **brain parcellation**, and *edges*, obtained from **brain connectivity**. The brain parcellation is based on *atlases*, which are created by combining information from many subjects in a common space, such as MNI152. Each label in an atlas specifies a brain region. This parcellation

can be based on anatomy, functionality, or both. The other component is brain connectivity, which is essential for identifying which regions are interconnected. To calculate a structural connectome, tractography is used to extract information about the edges. The pipeline is summarized in the following figure:

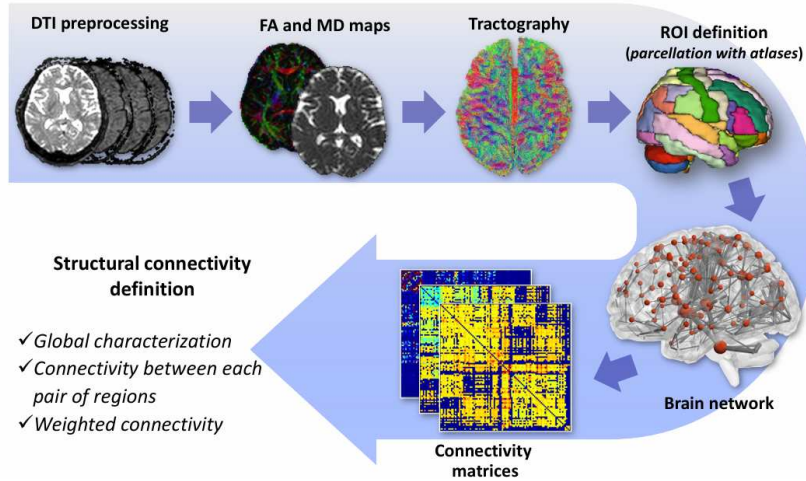


Figure 2.5: Structural connectivity: tractography maps calculated from DTI are used to extract streamlines connecting pairs of GM voxels and that, together with brain parcellation obtained from Atlases, are used to obtain connectivity matrices. The strength of the connection between two GM regions is given for example by the count of the number of streamlines connecting voxels belonging to both. ⁴

A brain connectivity matrix is a representation of a connected system in the form of a matrix. Each row and column represent a node (a specific brain region), while each intersection point represents an edge (the connection between two nodes). Edges are associated with **weights**, such as the normalized number of streamlines, which define the long-range connectivity of the brain.

Limitations of Tractography

Despite the advantages of tractography, it is important to highlight some technical limitations associated with this method:

- Identification of *crossing fibers*;
- Tracking errors;
- Lack of validation of the reconstructed tracts.

⁴Courtesy of Anita Monteverdi, PhD, Department of brain and behavioral sciences, University of Pavia

Some of these issues can be addressed using advanced diffusion methods, particularly the problem of crossing fibers, which is discussed in detail in Section 2.3.6. Nevertheless, tractography is the only available method for reconstructing axonal bundles in humans *in vivo*.

An important application of tractography involves its use in diseases with focal damage in order to observe the effects of lesions on the reconstructed data, such as in MS.

2.3.6 The Cerebellar-cortical loop

As discussed in Section 2.3.5, the diffusion tensor model has intrinsic limitations, most notably its inability to directly resolve crossing fiber structures. As a result, tractography methods based on diffusion tensor properties cannot achieve a complete brain reconstruction. In particular, this limitation allows only a partial reconstruction of cerebellar WM tracts and therefore, provides limited insight into the complex anatomical circuits of the cerebellum.

The introduction of a combination of two advanced diffusion techniques, which are a tract reconstruction based on **constrained spherical deconvolution (CSD)** and a super-resolution maps based on **track-density imaging (TDI)**[19] [18] enables the reconstruction of the cerebellar-cortical closed loop, effectively addressing the crossing fiber problem, as is discussed in the following sections. Before the description of the method used for pathway reconstruction, a brief introduction to the cerebellum is provided.

Cerebellum: anatomy and functions

The cerebellum is a structure of the CNS located at the back of the brain, immediately inferior to the occipital and temporal lobes, and within the *posterior cranial fossa*. It is separated from these lobes by a layer of dura matter called the *tentorium cerebelli* and lies at the same level as the *pons*, from which it is separated by the fourth ventricle.

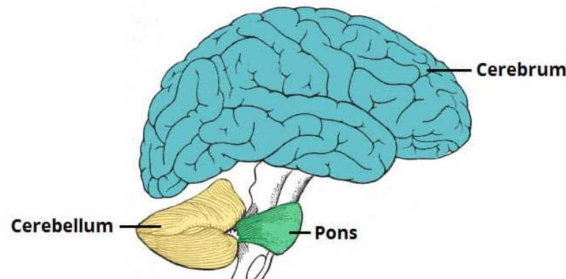


Figure 2.6: Cerebellum anatomical position⁵

Like other structures in the CNS, the cerebellum consists of GM and WM:

- **gray matter** - located on the surface of the cerebellum, forming the cerebellar cortex;
- **white matter** - located underneath the cerebellar cortex where the four deep cerebellar nuclei (*dentate, emboliform, globose, and fastigial nuclei*) are embedded.

The cerebellum can be distinguished in three anatomical lobes, which are the *anterior lobe*, the *posterior lobe* and the *flocculonodular lobe*, divided by two fissures - the *primary fissure* and *posterolateral fissure*. In the midline of the cerebellum lies the vermis. On either side of the vermis there is the intermediate zone, and lateral to that are the lateral hemispheres.

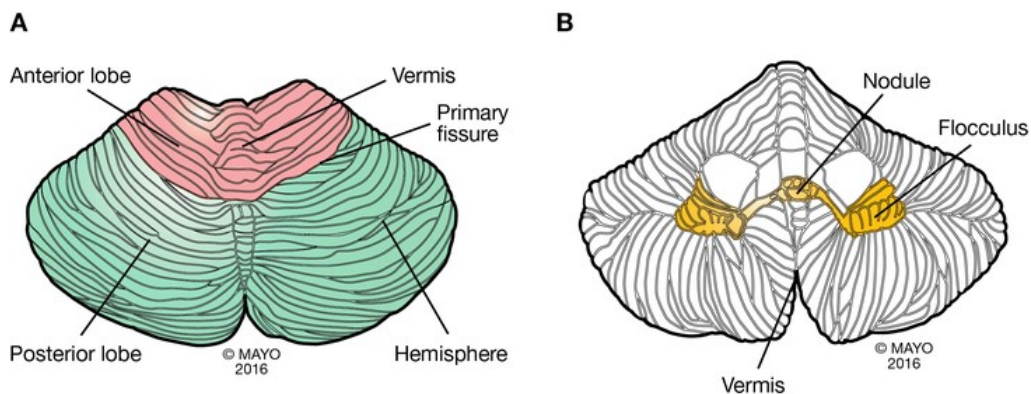


Figure 2.7: Cerebellum. A, Dorsal view. B, Ventral view[10]

Information is carried to and from the cerebellum through the cerebellum peduncles, which are the *middle cerebellar peduncle* that carries information to the cerebellum, while the *superior peduncle* from the cerebellum. The *inferior cerebellar peduncle* carries both afferent and efferent information.

The cerebellum can also be characterized in three functional areas:

- *Cerebrocerebellum*
- *Spinocerebellum*
- *Vestibulocerebellum*

The *Cerebrocerebellum*, based on the two lateral hemispheres, receives inputs from the *cerebral cortex* and *pontine nuclei*, while sends outputs to the *thalamus* and *red*

⁵Sara Venturini, *The cerebellum*, TeachMeAnatomy, October 2025. Available at: <https://teachmeanatomy.info/neuroanatomy/structures/cerebellum/>.

nuclei. This area is particularly involved in planning movements and motor learning, but it also has a role in cognition, language and limbic system.

On the the other side, the *Spinocerebellum*, that comprise the vermis and intermediate zone of the cerebellar hemispheres, is involved in regulating body movements and receives proprioceptive information.

Lastly, the *Vestibulocerebellum* is the functional equivalent to the flocculonodular lobe and receives inputs from the vestibular system and sends outputs back to the vestibular nuclei. In particular, it is involved in controlling balance and ocular reflexes, such as fixation on a target.

It is well known that the cerebellum plays an important role in motor control. In particular, it is involved in planning, coordination, precision, executing and timing of movements, as well as in motor learning. It compares actual movements with intended motor performance and optimizes the timing of motor function. It is also important in the adaptation of movement and posture [10].

Besides these functions, a crucial role of the cerebellum in cognitive processes and behavior has recently been discovered, thanks to its connections with associative areas of the cerebral cortex. This was first demonstrated in investigations on non-human primates and, thanks to developments in MRI technology, in humans as well. However, evidence in humans has been more limited due to the intrinsic limitations of tractography, as discussed before.

Nevertheless, advanced diffusion MRI[19, 18] has allowed the reconstruction of **cerebellar-cortical pathways**, which are essential for a better understanding of the involvement of cerebellum functions.

Reconstruction method

The method that allowed the reconstruction of both:

- Efferent Cerebello-Thalamo-Cortical (**CTC**) pathway;
- Afferent Cortico-Ponto-Cerebellar (**CPC**) pathway;

is based on two advanced diffusion techniques:

- **CSD** algorithm - it can model multiple fiber populations within a voxel and is able to resolve connection decussation;
- **Super-resolution maps** based on TDI - it allowed accurate seed and target region placement.

The main steps of the method are summarized below:

- *Generation of whole-brain tractography* to produce TDI maps using **MRtrix**, with an algorithm that combines the CSD technique and probabilistic streamline tractography,
- *Creation of TDI maps* from the streamlines, calculated as the total number of streamlines passing through each element of a user-defined super-resolution grid,
- *Reconstruction of pathways* by combining the CSD algorithm with probabilistic tractography, tracking bundles passing through two ROIs (the seed region and the target region),
- *Placement of seed/target ROIs*: masks were placed using the TDI images.

TDI maps have been shown to enable more anatomically accurate reconstruction of pathways compared to traditional approaches, such as **non-diffusion-weighted images**[19]. The reproducibility of these pathways over time was high and the characteristics, such as diffusion metrics and ROI locations, were consistent between subjects and independent of the timing of the definition of the ROI of the seed/target.

Contralateral cerebello-thalamo-cortical pathway

The main connections of the cerebellum-cortical circuit are shown in Figure 2.8.

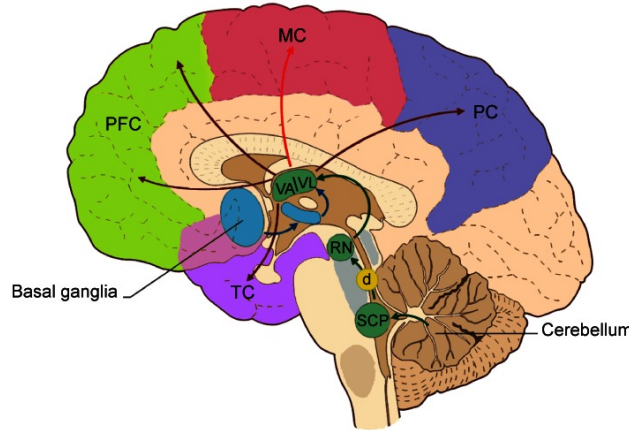


Figure 2.8: Cerebello-cortical connections [18]

Projections from the basal ganglia, via the subthalamic nucleus (STN), primarily reach the thalamic nuclei (VA/VL). The cerebellum sends its output through the superior cerebellar peduncle (SCP), the contralateral red nucleus (RN), and the VA/VL nuclei of the thalamus, projecting to various cerebral areas including the

motor cortex, prefrontal cortex, parietal cortex, and temporal cortex. The decussation (d) of the cerebello-thalamo-cortical pathway is indicated by the *yellow* circle [18].

To obtain the pathways, the seed ROI is defined as a sphere with a radius of 2 mm centered on the SCP in each cerebellar hemisphere and is identified in the coronal plane, while the target ROI is defined in the RN and recognized as a very hypotense region.

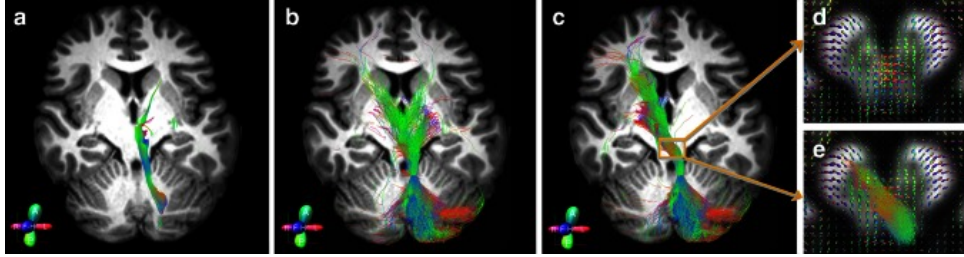


Figure 2.9: CTC pathway: comparison between the reconstruction using DTI and streamline method (a) and the reconstruction using the combination of CSD algorithm and probabilistic tractography (b and c). The DTI method operates only ipsilaterally while the other method reconstruct pathways both ipsilaterally and contralaterally (b). The contralateral connections are isolated adding a target region (c). [18]

The results shown in Figure 2.9 demonstrate that the combination of CSD and probabilistic tractography can successfully reconstruct the pathway [18]. In particular, DTI failed to reconstruct contralateral connections, as it is an approach that operates only ipsilaterally (Figure 2.9.a). In contrast, with the other method, it is possible to map the correct pathway connecting the cerebellar cortex through the SCP, RN, and thalamus, despite the challenge posed by decussation. Indeed, the CSD algorithm produced streamlines running both ipsilaterally and contralaterally (Figure 2.9.b). To isolate the contralateral connections, it is necessary to add the target region (Figure 2.9.c).

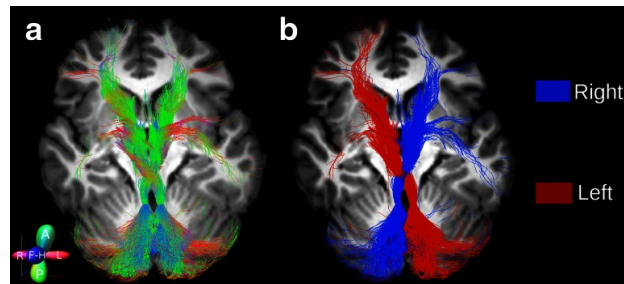


Figure 2.10: Left and right CTC pathways. The color-coded tracts follow the anatomy and directionality of both left and right tracts.[18]

Contralateral cortico-ponto-cerebellar pathway

The contralateral CPC pathway can be reconstructed using the same pipeline used for the CTC pathway [19] to gather critical information about the cerebro-cerebellar loops. The combination of CSD and probabilistic tractography made it possible to successfully reconstruct both left and right seeded contralateral CPC pathways. In this case, the pathway was reconstructed using a seed ROI placed in the *middle cerebellar peduncle* (MCP) and a target ROI in the *contralateral cerebral peduncle*.

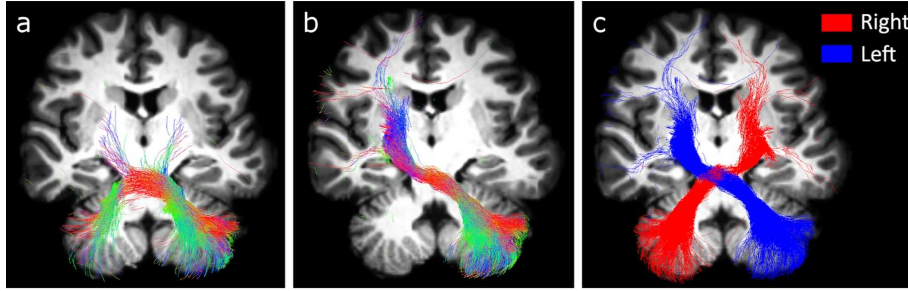


Figure 2.11: Reconstructed CPC pathway (a. Left and right pathways, b. Left pathway is selected) in a representative subject. The colour-coded tracts follow the anatomy and directionality of both left and right tracts (c)[19]

Cerebro-cerebellar closed-loops

This method successfully reconstructs both pathways, which are referred to as **cerebro-cerebellar closed-loop**, through which the cerebellum exerts its function in close communication with the cerebral cortex [19][18].

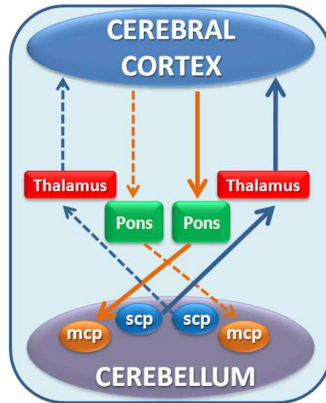


Figure 2.12: Schematic diagram of the cerebro-cerebellar loop: the CPC pathway connects the cerebrum to the cerebellum (in *orange*) via the pons, while the CTC pathway links the cerebellum to the cerebrum (in *blue*) through the Thalamus. [18]

In Figure 2.12, the cerebro-cerebellar connections are illustrated. Specifically, the CPC pathway connects the cerebrum to the cerebellum via the pons and the

contralateral MCP, while the CTC pathway links the cerebellum to the cerebrum through the SCP and the contralateral thalamus. Thus, the cerebro-cerebellar connections form closed loops, with the feedforward and feedback branches primarily passing through the CPC and CTC pathways. Through these loops, the cerebellum is thought to function as an information processing system that modulates cerebral activity.

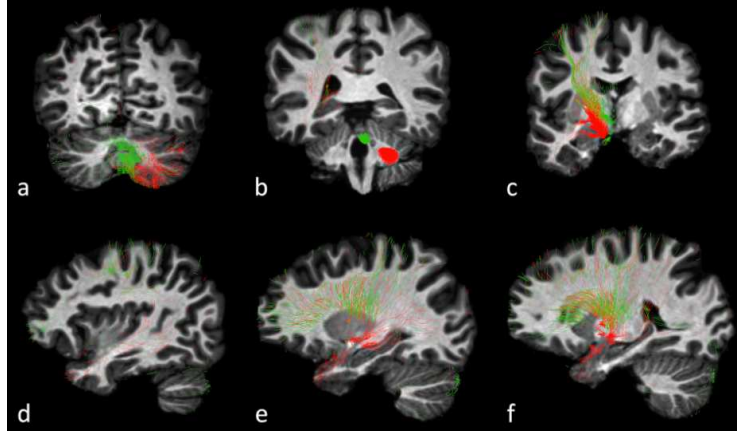


Figure 2.13: Left-seeded CPC tract (red) and CTC tract (green): visualization of the tracts through different slices[18]

Figure 2.13 provides visual evidence of the mutual connectivity between the two pathways. Both pathways predominantly occupy the lateral regions of *Crus I-II* and *lobules VIIb/VIII* (panel a). However, the output projections from the cerebellum differ (panel b). In the cerebral cortex, the CPC and CTC pathways show partial overlap within the frontal lobe (panels b, d, e). The temporal lobe is more densely innervated by CPC streamlines compared to CTC ones, whereas the prefrontal cortex receives a greater number of CTC streamlines than CPC (panels d, e, f). This result represented the first quantitative evidence based on MRI of cerebro-cerebellar connectivity in vivo [19].

Regarding the intrinsic limitations of MRI tractography, beyond the fact that the issue of crossing fibers has been effectively addressed through advanced diffusion methods, other limitations can also be mitigated. Specifically, the integration of anatomical information and a priori knowledge as constraints can help overcome the technique's inability to distinguish between afferent and efferent pathways, as well as improve the discrimination between monosynaptic and trans-synaptic connections.

In conclusion, this method enables the generation of **whole-brain tractography** which allows for a more in-depth study of the brain, including the cerebellum.

Chapter 3

Network Spreading Models

Data-driven disease progression models are a set of computational tools developed to learn long-term disease biomarker trajectories at a fine-grained temporal resolution for chronic diseases, using short-term data and without requiring prior knowledge of an individual’s disease stage [30]. These models rely on two key features:

- **Construction of a data-driven disease timeline:** modeling the temporal evolution of a disease indexed by a data-driven time axis;
- **Direct use of measured data:** the use of *in vivo* biomarkers enables the generation of quantitative disease timelines that offer biological insights.

These models are essential for mapping the progression of neurodegenerative diseases such as Alzheimer’s disease (AD), Parkinson’s disease, frontotemporal dementia, and Huntington’s disease. These conditions typically evolve over many years and are associated with stereotypical changes in disease biomarkers over time. Without these models, mapping disease timelines would be prohibitively expensive and impractical, as measurements often under-sample the critical pre-symptomatic phase. In addition, population heterogeneity and the complexity of disease mechanisms further complicate this task.

Moreover, data-driven progression models have broad applicability across a range of long-term chronic conditions. For instance, many other neurological and non-neurological diseases, including MS, also exhibit temporal evolution, including a preclinical stage and stereotypical biomarker changes.

Data-driven disease progression models can be broadly categorized into two main types:

- *Phenomenological models*
- *Pathophysiological models*

While *phenomenological models* aim to capture common trajectories of disease biomarkers without considering the underlying biological mechanisms, *pathophysiological models* seek to explain disease timelines in terms of biological and physical processes, such as the mechanisms driving pathology spread. Phenomenological models typically use weaker constraints based on expected trajectory shapes, which require complex fitting strategies. In contrast, pathophysiological models are more strongly constrained, allowing simpler fitting approaches.

This thesis studies a particular category of pathophysiological models, called **Network Spreading Models**.

3.1 Pathophysiological models

Pathophysiological models estimate the characteristics of biological processes that generate disease timelines which best predict biomarker measurements, with a focus on biological hypothesis testing and insight than on disease phenomenology. They typically take as input an approximation of a topological property of the brain, such as a connectivity map, and produce as output a set of expected biomarker trajectories along a data-driven disease timeline, along with the estimated parameters optimized by the model.

These models can be used to simulate the emergence or spread of pathology, including *prion-like diffusion* from an epicenter, constrained along WM fibers within and between brain regions.

3.1.1 Prion-like mechanism

Initial symptoms of cognitive decline in neurodegenerative diseases typically become noticeable only after one or two decades following the onset of pathological abnormalities in the brain. In fact, disease progression is characterized by a long and clinically silent incubation period during which prions grow and spread, followed by a brief and fatal clinical phase [29].

Prions, *proteinaceous infectious particles*, are infectious agents composed of misfolded proteins that act as corruptive seeds, initiating a chain reaction of misfolding, growth, and propagation. This means that misfolded proteins can induce the misfolding of adjacent same-species proteins, which then cascade along neuronal pathways [22].

Although the clinical and pathological manifestations of neurodegenerative disorders such as AD, Parkinson’s disease, and amyotrophic lateral sclerosis (ALS) may differ, they share common mechanisms involving the growth and spread of misfolded

proteins, which are named prions. Therefore, the *prion paradigm* may hold the key to developing models that explain the stereotypical features of neurodegenerative disorders.

In the following sections, three models that have been used to describe the spread of neurodegenerative diseases will be presented and analyzed. These models are:

- *Network Diffusion Model (NDM)*
- *Fisher-Kolmogorov-Petrovsky-Piscounov (FKPP) Model*
- *Weighted-KFPP Model (w-KFPP)*

3.1.2 Network Diffusion Model

The **NDM** [22] is capable of capturing the characteristics of *prion-like* disease progression through a diffusive mechanism - specifically, a model of random dispersion driven by concentration gradients. Notably, the behavior of this diffusive propagation can be mediated by the brain structural connectivity, derived from tractography using dMRI scans.

This model implies the existence of **eigenmodes** which represent distinct spatial patterns. These patterns bear a strong resemblance to known distributions associated with various forms of dementia [22]. Therefore, the macroscopic consequences of prion-like diffusive propagation across a healthy brain network may be predictive of large-scale patterns observed in different dementias.

The primary clinical application of this model is its ability to predict patterns of brain atrophy based on baseline scans. In other words, the model can forecast future atrophy for an individual, providing estimates of the neuroanatomical state of the patient at any given point in the future.

Development of the model

The model appears to capture the essential patterns of dementia-related neurodegeneration through a **first-order approximation**. Although both atrophy and the brain network dynamically evolve over time due to degeneration, the model assumes a static network to avoid non-linearity, which would render the eigenmodes uninterpretable.

Disease progression is modeled using a diffusion process [22] on a brain network $\mathcal{G} = \{v, \varepsilon\}$, where the nodes $v_i \in v$ represent the i -th cortical or subcortical GM structure, and the edges $e_{i,j} \in \varepsilon$ represent WM fiber pathways connecting structures i and j .

NDM eigenmodes are derived from structural healthy brain network, obtained from dMRI scans followed by a whole brain tractography.

The model [22] can be described using the *Network Heat Equation*, defined as follows:

$$\frac{d\mathbf{x}(t)}{dt} = -\beta H\mathbf{x}(t) \quad (3.1)$$

where H is the **graph Laplacian**, defined as:

$$H_{i,j} = \begin{cases} -c_{i,j} & \text{if } c_{i,j} \neq 0 \\ \sum_{i,j': e_{i,j'} \in \mathcal{E}} c_{i,j'} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

In this model, $\mathbf{x}(t) = \{x(v, t), v \in \mathcal{V}\}$ is a vector representing the disease factor at time t at each node in the network, and β is a constant controlling the propagation speed. The structures v_i are derived from brain MRI parcellation, while the connection strengths $c_{i,j}$ are measured via fiber tractography.

Since the model uses relatively large and distinct structures as brain network nodes, the effect of localized transmission is predominantly intranodal. Disregarding the limited effect of local intranodal leakage, the Laplacian H does not depend on self-connectivity within a node.

Cortical atrophy in a region k is modeled as the accumulation of the disease process in k , represented by the integral:

$$\varphi_k(t) = \int_0^t x_k(\tau) d\tau \quad (3.3)$$

For the entire model, this becomes:

$$\Theta(t) = \int_0^t \mathbf{x}(\tau) d\tau \quad (3.4)$$

Equation 3.1 is solved by:

$$\mathbf{x}(t) = e^{-\beta H t} \mathbf{x}_0 \quad (3.5)$$

where $\mathbf{x}_0$ is the initial disease pattern and $e^{-\beta H t}$ is the *diffusion kernel*, acting as a spatio-temporal blurring operator.

The computation of the above equation is performed via *eigenvalue decomposition* $H = U\Lambda U'$, where $U = [\mathbf{u}_1, \dots, \mathbf{u}_N]$, yielding:

$$\mathbf{x}(t) = U e^{-\Lambda \beta t} U' \mathbf{x}_0 = \sum_{i=1}^N \left(e^{-\beta \lambda_i t} (\mathbf{u}_i^\dagger \mathbf{x}_0) \right) \mathbf{u}_i \quad (3.6)$$

Most eigenmodes $\mathbf{u}_i$ correspond to large eigenvalues that decay rapidly due to exponentiation, leaving only the small eigenmodes, whose absolute values are referred to as **persistent modes**, to contribute meaningfully.

The time evolution of the hypothesized atrophy is a linear combination of eigenmodes, as shown in the equation below:

$$\Phi(t) = \int_0^t \sum_{i=1}^N \left(e^{-\beta \lambda_i t} \mathbf{u}_i^\dagger \mathbf{x}_0 \right) \mathbf{u}_i dt = \sum_{i=1}^N \frac{1}{\beta \lambda_i} \left(1 - e^{-\beta \lambda_i t} \right) \mathbf{u}_i^\dagger \mathbf{x}_0 \mathbf{u}_i \quad (3.7)$$

This consists of two components: a deterministic exponential part and a case-dependent component determined by the initial configuration $\mathbf{x}_0$ of the disease.

Eigenmodes

The application of this model demonstrates the resemblance between predicted and measured atrophy patterns [22], as shown in the following figure.

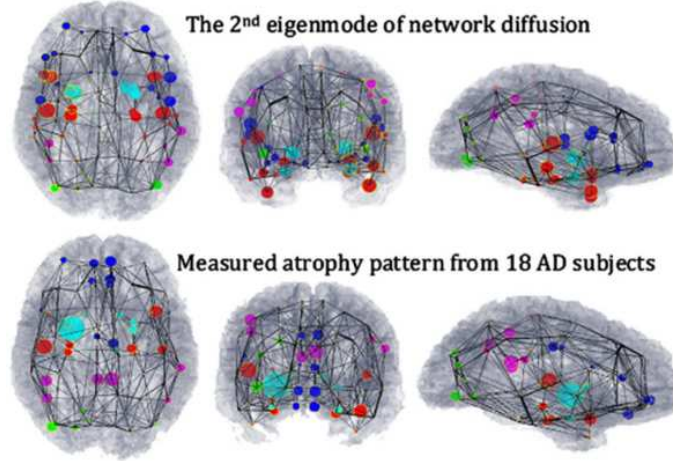


Figure 3.1: Visual correspondence between predicted and measured Alzheimer's atrophy patterns. (top) Predicted distribution: the value of the second eigenmode at each node is represented by the size of the corresponding ball. (bottom) Measured atrophy (t-statistic): the size of the ball represents the amount of atrophy measured in the corresponding region [22].

In Figure 3.1, a wire-and-ball mesh plot is used to visualize the brain atrophy pattern: each region in the brain is represented by a node (depicted as a ball), and the connectivity between two regions is depicted by a wire whose thickness denotes connection strength. A close homology is observed between the predicted and measured atrophy patterns.

The comparison between persistent modes, specifically the 2nd eigenmode, and the measured atrophy from the AD cohort supports the hypothesis that persistent

modes are homologous to known patterns of atrophy in neurodegenerative diseases.

A similar result has been demonstrated using the 3rd eigenmode of network diffusion, which shows a close homology to the measured atrophy pattern in subjects with behavioral variant frontotemporal dementia.

These findings have suggested the potential role of network diffusion eigenmodes in the automated and differential diagnosis of various dementias[22].

3.1.3 Fisher-Kolmogorov-Petrovsky-Piscounov model (KFPP)

The growth and spreading of misfolded proteins in neurodegenerative diseases can also be explained by a **physics-based reaction-diffusion model**, starting from the classical *Fisher-Kolmogorov equation* for population dynamics with an anisotropic diffusion model [29].

The *Fisher-Kolmogorov* equation is described as follows:

$$\frac{dc}{dt} = \text{Div}(\mathbf{D} \cdot \nabla c) + \alpha c[1 - c] \quad (3.8)$$

In the case of prion-like diseases, c represents the relative concentration of misfolded protein, α characterizes the growth of the concentration, while $\mathbf{D}$ its spreading. Once seeded, misfolded protein can spread across the brain by two distinct mechanisms:

- *Extracellular diffusion* - isotropic diffusion through the extracellular space
- *Axonal transport* - anisotropic diffusion along the local axonal direction

$$\mathbf{D} = d^{ext}\mathbf{I} + d^{axn}\mathbf{n} \otimes \mathbf{n} \quad (3.9)$$

It is generally assumed that axonal transport d^{axn} is faster than extracellular diffusion d^{ext} , $d^{axn} \geq d^{ext}$.

Since the Fisher-Kolmogorov equation is highly sensitive to its parameters, the success of the prion-like spreading simulation critically depends on the right balance between them:

- Extracellular diffusion d^{ext} ;
- Axonal transport d^{axn} ;
- Growth α ;
- Axonal orientation $\mathbf{n}$;
- Seeding region $\mathbf{B}$.

In particular, increasing the growth rate α triggers a faster spreading of pathogenic protein concentration c and reduces the activation time, while increasing extracellular diffusion and axonal transport causes a faster spreading out of the initial seeding region. This means that the simulation activation pattern is insensitive to the growth rate, which primarily affects the timing of activation, but highly sensitive to both extracellular diffusion and axonal transport, which affect both activation time and the emerging concentration pattern [29].

Prion-like spreading simulation

Figure 3.2 shows the results of the simulations of prion-like spreading of amyloid- β deposits in AD, tau inclusions in AD, α -synuclein inclusions in Parkinson's disease, and TDP-43 inclusions in ALS, over a total period of 40 years[29]. It represents three stages of neurodegeneration, where the first column refers to the initial conditions for simulations: the initial concentration of misfolded protein is zero across the brain except for the local seeding regions.

These results[29] display excellent agreement with the characteristic progression of specific proteinaceous lesions in neurodegenerative diseases over time inferred from post-mortem studies of the human brain.

So, once calibrated appropriately, the Fisher-Kolmogorov equation can predict the prion-like features of growth and spreading in various neurodegenerative diseases.

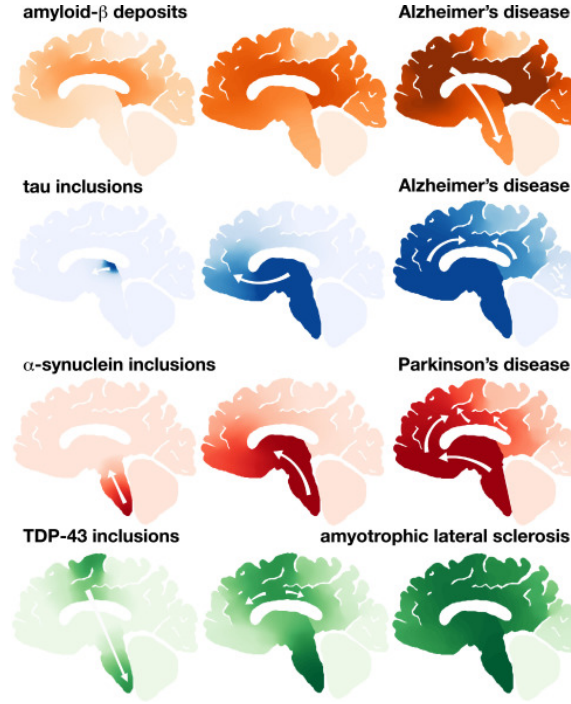


Figure 3.2: Prion-like spreading across a sagittal section. Amyloid- β deposits in AD, tau inclusions in AD, α -synuclein inclusions in Parkinson's disease, and TDP-43 inclusions in ALS. Three stages of progressive neurodegeneration, with white arrows indicating the spread of the lesions [29].

Biomarker abnormality

From the simulation of prion-like spreading it is also possible to obtain the biomarker abnormality curves [29], as shown in Figure 3.3. These curves result from integrating the total concentration of misfolded proteins across the brain.

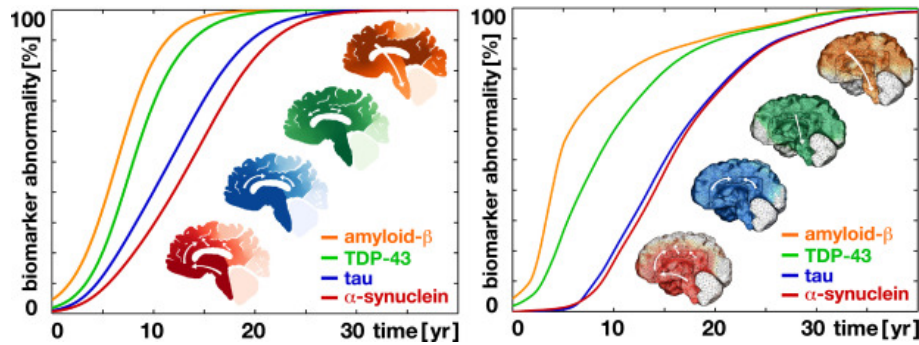


Figure 3.3: Biomarker abnormality from prion-like spreading of amyloid- β deposits, tau inclusions, α -synuclein inclusions, and TDP-43 inclusions across a sagittal section on the left and across the whole brain on the right. All biomarkers follow a similar smooth sigmoidal form which agrees with clinical models and observations: activation is slow in the first five to ten years after initial seeding, then increases for another ten years, and peaks approximately at 20 years.[29].

All biomarkers follow a similar *smooth sigmoidal* form, which agrees with clinical biomarker models for neurodegeneration and clinical observations. In all cases, activation is slow in the first five to ten years after initial seeding, then increases for another ten years, and peaks approximately at 20 years. This follows the common understanding of neurodegeneration, where a long and clinically silent incubation period, during which the disease amplifies, is followed by a brief and fatal clinical disease.

So, the Fisher-Kolmogorov equation captures the characteristic timeline of individual biomarkers in various types of neurodegeneration [29].

3.1.4 Weighted-FKPP model

The complexity and heterogeneity of neurodegenerative diseases suggests that multiple mechanisms may contribute synergistically through complex interactions, with a degree of contribution varying among individuals.

The *w-FKPP* equation [11] provides an alternative modeling framework that non-linearly couples the effects of spontaneous appearance and pathology spreading. In contrast to previous models, which make specific assumptions about the underlying mechanisms of the physical process, this approach estimates the mechanistic profiles of pathologies, including interactions between model components and quantification of uncertainty, through a Bayesian model selection framework.

Consequently, disease heterogeneity is captured by fitting the model at an individual level, allowing factors such as epicenters, progression rates, and the weights of network metrics to vary across subjects. This has been shown to yield combinations of mechanisms that better explain observations for each subject and to identify subgroups of patients sharing similar mechanistic profiles [11].

Model definition

The baseline model of disease spreading is the **Fisher-Kolmogorov equation**, which does not account for the synergistic effect of multiple mechanisms. In particular, it assumes a constant local production rate across regions and the diffusion of prions along the structural network from an epicenter.

What is new in the coupled model is that it weights the local production process using a combination of P network metrics $M = [m_1, \dots, m_p]$, employed as interaction components, and introduces a weighting vector $w = [w_1, \dots, w_p]$ to represent the contribution of each characteristic. This allows the incorporation of pathology appearance proportional to various brain network topologies into the disease spreading

process.

The coupled model is expressed as follows:

$$\frac{dc}{dt} = -k[\mathbf{H}c(t)] + \alpha \mathbf{M}\mathbf{w} \odot c(t) \odot [v - c(t)] \quad (3.10)$$

where $\mathbf{H}$ is the graph Laplacian matrix computed from the structural connectome, and k is the disease diffusion rate. The second term represents the production and aggregation process, which exhibits an increasing trend converging to a plateau level v at a rate α , and includes weighting by the combination of network metrics.

Metrics	formula	weights	mechanism
$\mathbf{m}_1$	$\mathbf{1} - \text{BetweennessCentrality}(\text{StructureC})$	w_1	Structural Segregation
$\mathbf{m}_2$	$\mathbf{1} - \text{ClusterCoefficient}(\text{StructureC})$	w_2	Structural Dispersion
$\mathbf{m}_3$	$\text{ClosenessCentrality}(\text{StructureC})$	w_3	Structural Centrality
$\mathbf{m}_4$	$\text{WeightedDegree}(\text{InvGeodist})$	w_4	Geodesic proximity
$\mathbf{m}_5$	$\text{BetweennessCentrality}(\text{FunctionC})$	w_5	Functional centrality
$\mathbf{m}_6$	$\mathbf{1}$	w_6	Even distribution

Table 3.1: Network metrics used as interaction components in the spreading models. **StructureC**, **InvGeodist**, **FunctionC** represent the structural connectome, the matrix of the inverse of geodesic distance along the cortical surface and the functional matrix. [11].

The integration of network metrics into the dynamic system of disease spreading improves the fit to observed pathology patterns and provides a potential approach for subtyping subjects according to their mechanistic profiles [11].

3.2 The Network Spreading Model Toolbox

The **NSM Toolbox** [26] is an open-source toolbox for NSMs that enables comparison of model fits across different models and parameter settings. In particular, it can be applied to connectome-based models of disease propagation, which are used to investigate mechanisms of pathology spread in neurodegenerative diseases. The models currently available and implemented in the toolbox are those described in the previous sections: the **NDM**, **FKPP**, and **weighted-FKPP**.

This toolbox was first applied to assess whether regional amyloid levels affect the production of pathological Tau in AD by comparing the results obtained from the three models [26]. The toolbox revealed that incorporating regional amyloid levels through the w-FKPP model improved the fit of pathology spreading models

to measured tau-PET data, particularly at later stages of AD. This finding supports the hypothesis that amyloid promotes tau aggregation.

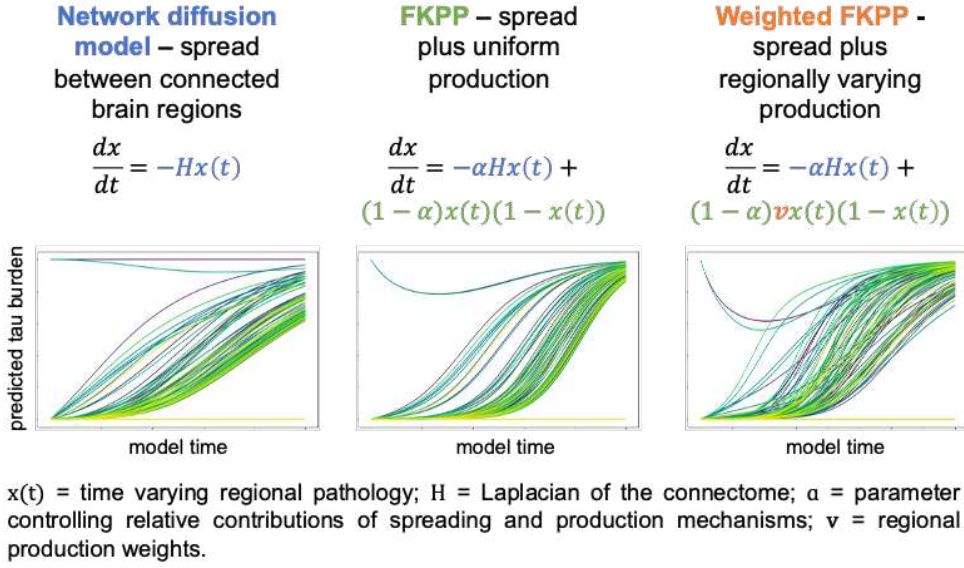


Figure 3.4: Illustration of the NSM toolbox and the models currently available. The equations are given for each model [26].

Chapter 4

Materials and Methods

This chapter describes how the **NSM Toolbox** (Section 3.2) is used to simulate the spreading of disease from a dataset consisting of ten **structural connectivity matrices** of ten healthy individuals, previously calculated from a set of dMRI data provided by the **Human Connectome Project (HCP)** (see Section 4.1.1).

The NDM has been applied to the mean structural connectivity matrix to simulate disease spreading and, subsequently, the results are examined using different methods, such as temporal graphs and brain maps, as explained in Sections 4.2.2 and 4.3.

The model outputs obtained are then compared with a MS dataset at baseline, in order to evaluate whether the resulting patterns of disease progression show any similarities with MS at the *first demyelination event*. To do this, a **voxel-wise** analysis is calculated based on the calculation of z-score maps (see Section 4.4.4) which includes, for each subject (patients and healthy controls), a series of metrics maps listed as follows:

- Fractional Anisotropy (FA);
- Mean Diffusivity (MD);
- Orientation dispersion index (ODI) - NODDI model;
- Intracellular Volume Fraction (V_{fintra}) - NODDI model;
- Isocellular Volume Fraction (V_{fiso}) - NODDI model.

In addition to voxel-wise analyses, statistical tests such as the Welch test and Principal Component Analysis (PCA) were also applied to the patient and healthy control groups to identify brain regions that already show some differences at baseline (see Section 4.4.5).

4.1 Dataset analyses

The dataset of the healthy subjects used to simulate disease spreading is based on structural connectivity matrices derived from dMRI images [20] (see Section 4.1.2). All analyses were conducted using **Python** programming language.

First, the origin of the dataset to which the reconstruction method was applied will be briefly discussed.

4.1.1 Human Connectome Project

The HCP [28] is a five-year collaborative initiative aimed at mapping human brain connections and their variability in healthy adults. A consortium of HCP investigators, led by Washington University, the University of Minnesota, and Oxford University, studied a cohort of 1,200 healthy individuals using multiple imaging modalities, complemented by extensive behavioral and genetic data.

MRI in the HCP includes four modalities which are structural MRI, resting-state fMRI, task-based fMRI and **dMRI**, all acquired at high resolution for a large-scale *in vivo* study. All subjects were scanned on a customized $3T$ scanner, while 200 of the same subjects were additionally scanned on a $7T$ scanner. These data are *freely* available to the scientific community and include not only raw imaging data but also processed datasets, reflecting multiple levels of pre-processing required for analysis and interpretation.

A major advancement in pulse sequence development was the implementation and optimization of slice-accelerated multiband acquisition for dMRI, which enabled a substantial increase in spatial resolution. Specifically, in diffusion imaging a resolution of 1.25 mm was achieved, allowing high-quality *in vivo* whole-brain imaging at $3T$, as clearly illustrated in Figure 4.1.

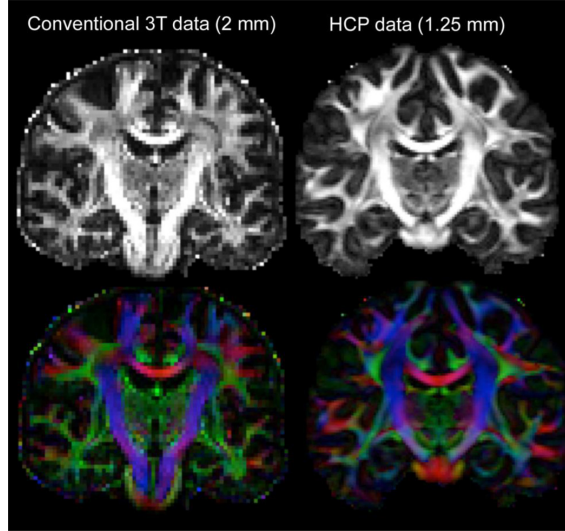


Figure 4.1: Representative FA and color-encoded principal diffusion direction images from the HCP dMRI data, compared with a conventional 2 mm images. The improvement in anatomical details is evident [28].

4.1.2 Connectomes

As mentioned before, the dataset consists of ten whole-brain structural connectivity matrices obtained from a **HCP** MRI dataset that includes DW-data [20]. In particular, minimally pre-processed data were downloaded from Connectome DB with the following characteristics: 1.25 mm isotropic resolution, $b = 1000, 2000, 3000$ s/mm², 90 isotropically distributed directions per b -value, 18 b0 images, and 3D T1-weighted images (resampled to the same resolution as the DW-data).

Healthy subjects included 3 males and 7 females aged 22-35 years, acquired using the *customized Siemens 3T Connectome Skyra scanner* with a 32-channel receive head coil.

Ad hoc Brain Atlas

The atlas used for the brain parcellation is based on an ad-hoc template comprising 126 regions and created in MNI152 space. The cerebral parcellation consists of 93 regions, defined by combining the **Automated Anatomical Labeling (AAL)** template for cortical regions with **First** for deep GM structures. The 33 cerebellar parcellations correspond to those defined by the **SUIT** template.

Structural connectivity matrices pipeline

The pipeline[20] to obtain the structural connectivity matrices is based on the following steps. First, 3D T1-weighted images were segmented using FSL into WM,

GM, subcortical GM, and CSF. Then, from DW-data, fiber orientation distributions were calculated for each tissue using the multi-shell, multi-tissue CSD algorithm in MRTrix3. Whole-brain tractography was performed with 30 million streamlines using probabilistic streamline tractography. The whole-brain reconstruction method is detailed in Section 2.12.

Finally, whole-brain tractography and the DW-atlas were combined to extract streamlines (edges) between regions (nodes). The weight of each edge was normalized to values in the range $[0, 1]$ with respect to the total number of streamlines belonging to the network.

The 10 **whole-brain structural connectivity matrices** dataset given is shown in Figure 4.2. The color bar at the bottom of the figure indicates that blue is associated with lower weights, while red higher connectivity values. The maps also include the cerebellum and its pathways. Specifically, the cerebellar submatrix is located in the bottom-right portion of the matrix and consists of a 33×33 block.

Brain Connectomes from MRI

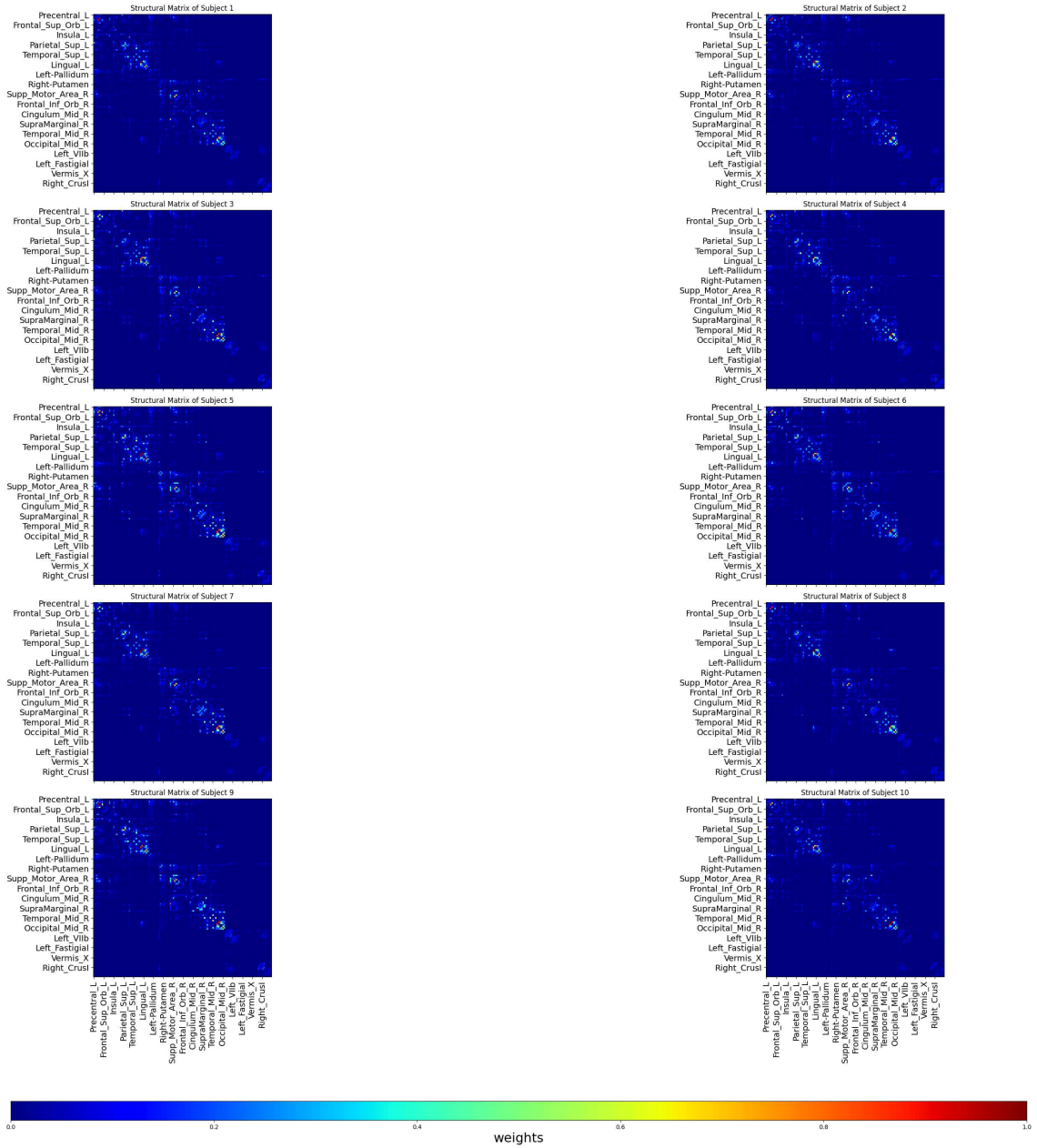


Figure 4.2: Whole-brain structural connectivity matrices of 10 healthy subjects obtained from a HCP dMRI dataset. The weights are normalized to values in the range $[0, 1]$ with respect to the total number of streamlines of the network

Only some of the labels are shown in Figure 4.2, but all regions of the combined atlas are included in the maps.

Mean structural connectome

Statistical matrices calculated from the individual connectomes in Figure 4.2 are shown below. The mean matrix assigns to each voxel the mean connectivity value, representing the average fiber density connecting two brain regions, while the other matrix shows the corresponding standard deviation, which helps to visualize how much the weights differ across individual matrices.

Statistical Matrices

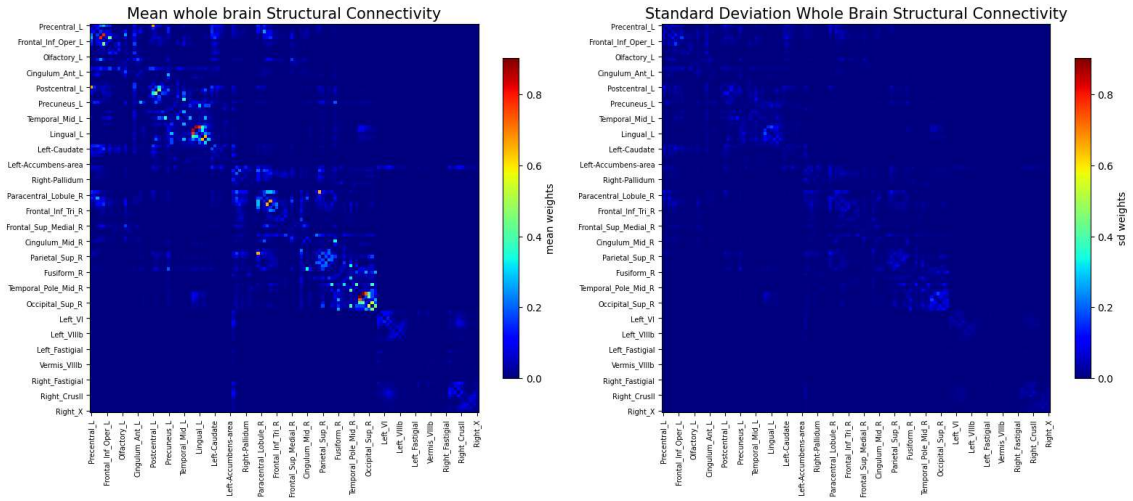


Figure 4.3: Structural connectivity matrices: (left) mean matrix and (right) standard deviation matrix calculated as mean and standard deviation of 10 healthy subjects

Since the **NSM Toolbox** is intended to be applied to the mean matrix, the standard deviation matrix is calculated to verify whether the mean connectome can be considered representative of the ten healthy subjects. The std values range from 0 to 0.175, which appears to be low. It is therefore reasonable to assume that the mean matrix is representative of the 10 healthy subjects, making it suitable for application of the toolbox.

4.2 Network Spreading Model Toolbox application

In this section, the focus is on the **NSM Toolbox**, which was introduced in Section 3.2. First, a brief overview of its functionality and installation process will be provided, followed by a description of its application to the connectome, detailing the input data and the resulting outputs.

This thesis specifically addresses the implementation of the simplest model, the **NDM**, which simulates the diffusion of pathology across the entire brain along

neuronal pathways. The simulation shows how damage spreads over time, when it starts from a specific brain region - the first one affected - and then propagates across the entire brain in order to examine the spatial pattern depending on the initial location of the damage. Damage is only considered in GM regions, as WM lesions are not included in the model. It is also important to note that, at this stage, the damage represents a generic impairment of a brain region and is not associated with any biologic disease factor.

4.2.1 Installation and environment setting

As previously mentioned, **NSM Toolbox** is an open-source tool available on **Git-Hub**. The first step in using it was its installation. The entire package, accessible at https://github.com/ucl-pond/network_spreading_models.git, was downloaded. Since the toolbox is developed in **Python**, the setup packages were installed within a Python environment, specifically through the terminal in **Visual Studio Code**.

The toolbox includes several folders containing the scripts for the implemented models, a **results** folder, where the model outputs are saved, and a **simulated_data** folder, where the model inputs can be stored.

Within the **simulated_data** folder, the mean connectome computed earlier and the label list containing all regions of the combined atlas were first loaded.

The toolbox also provides the possibility to compare model fits across the three implemented models and their respective parameter settings. In this regard, the scripts folder contains specific scripts that enable parameter optimization by generating optimized model input parameters. This process is achieved through the loading of a **target_data** file, which represents a structural connectome obtained from patient data and is used to fit the model outputs calculated from the healthy connectome provided as input.

Since the available material does not include a **target_data** file with the same resolution and atlas as the healthy connectome used for analysis, this thesis implements the NDM without employing the automatic optimization process. Instead, an alternative validation method will be discussed later (see Section 4.4).

4.2.2 Network Diffusion Model application

After installing and configuring the toolbox in Python, a script was developed to apply the NDM (see Section 3.1) to the mean structural connectome. Before running the simulation of the *disease spreading* across the entire brain starting from a specific brain region (**seed region**), it is essential to define the model inputs. These inputs

are listed as follows.

Model inputs

Apart from the previously discussed inputs such as the *connectome* and the *labels list* of the brain regions in the atlas, other model inputs have to be defined. In particular, these are the *seed region* and the number of *time points*. The *seed region*, as introduced before, consists of the initially affected area, from which the damage spreads to other regions.

This specific brain region can be set directly within the model, by putting it as input of the NDM *function* implemented in the toolbox. To do that, a vector x_0 has been initialized. This vector is defined as follows:

$$\mathbf{x}_0 = \begin{bmatrix} 0 & \dots & 1 & \dots & 0 \end{bmatrix} \in \mathbb{R}^n, \quad (4.1)$$

where $n = \text{len}(\text{label_list})$, hence the number of brain regions, and the entry equal to 1 corresponds to the seed region index.

Since the thesis aims to study the role of the *cerebellum* in disease, the cerebellar subregions were set as seeds in the simulation to observe how the spreading occurs across the brain. To achieve this, a *for-loop* was implemented to apply the model by considering each cerebellar region as a seed individually. Although the toolbox allows multiple regions to be set as seeds simultaneously, it was decided to use one region at a time to also investigate potential differences in the model outputs among the cerebellar subregions.

The cerebellum was not the only region studied as a seed. Other brain regions were also selected to visualize disease spreading. Specifically, since the validation method includes testing whether the model can predict MS pathways, some regions known to be particularly vulnerable in the disease were chosen. These include the *Thalamus*, *Motor Cortex*, and *Visual Cortex*. For the *Motor Cortex*, the region selected as seed based on the label list was the *Precentral* region, while for the *Visual Cortex* the *Calcarine* region.

Another model input that must be specified to run the simulation is the number of *time points*. The temporal evolution provided by the model is expressed in *time points* and does not correspond to real time units. In this work, it was set to *10,000 time points* to allow visualization of the model output curves from the initial stage up to the saturation point.

The diffusive constant β , which controls the propagation speed, was already set to 0.01 and kept fixed for all simulations.

Network Diffusion Model application code

In this paragraph, the Python code generated to run the NDM is presented and explained. Specifically, the code used to apply the model with cerebellar regions as seeds is illustrated. The same code can be modified depending on the chosen seed region.

```

1  t = np.arange(0, 10000, 1)
2  x0= np.zeros(126)
3
4  # Create a dictionary to store each model output
5  model_outputs = {}
6
7  for i in range(len(x0[93:])):
8      region_index = 93+i #cerebellar seed region index
9      region_name = ref_list[region_index]
10     x0[region_index]=1 #set seed region
11
12     # Model function
13     ndm = NDM(connectome_fname = connectome_fname, #connectome
14                beta = 0.01,
15                t = t,
16                ref_list=ref_list, #label list
17                x0=x0
18                )
19
20     model_output = ndm.run_NDM()
21     model_outputs[region_name] = model_output
22
23     plot_times_series(model_output,t,cmap='viridis')
24     plt.title(f'Seed_region:{region_name}')
25
26     x0[region_index] = 0 # reset seed

```

The `ref_list` is the list of regions, which must be in the same order as the one used in the connectome matrix. The `plot_time_series` function is available in the toolbox and it was modified to plot the model output as a temporal graph, as explained below.

Network Diffusion Model code

Before analyzing the model output, the code used to run the NDM is also shown[26].

```

1 def run_NDM(self):
2     H = self.get_Laplacian()
3     x_t = np.empty([len(H),self.Nt])
4     x_t[:,0] = 0
5
6     x_t[:,0] = self.get_initial_conditions()
7
8     for kt in range(1,self.Nt):
9         x_t[:,kt] = x_t[:,kt-1] + self.NDM_dx(H,x_t[:,kt-1])* self.
            ↪ dt
10
11     return x_t[self.cortical_idx]/np.max(x_t,axis=0)

```

The function `run_NDM` is implemented within the NDM class, which encapsulates all model parameters and the functions required to simulate the NDM. In this implementation, the following auxiliary functions are invoked:

- `get_Laplacian()` - calculates the Laplacian matrix from the connectome;
- `get_initial_conditions()` - gets the initial conditions if pre-defined;
- `NDM_dx()` - calculates the NDM formula[22].

The algorithm begins by computing the graph Laplacian from the brain connectome and then iteratively models the spatiotemporal diffusion of pathological burden along the structural network using a forward Euler integration scheme. Although no autonomous local growth is explicitly modeled, the diffusion process leads to an apparent accumulation of pathology in brain regions connected to initially affected areas. The code is commented in detail below.

1. Compute the Laplacian matrix H from the connectome;
2. Initialize a matrix x_t of size $n \times T$, where n is the number of regions and T is the number of time points;
3. Set the first column of x_t to the initial conditions x_0 ;
4. For each time step k from 1 to $T - 1$, the state of the system is updated according to:

$$x(t + \Delta t) = x(t) + \text{NDM_dx}(H, x(t)) \cdot \Delta t$$

reflecting the diffusion dynamics of the NDM;

5. Normalize the results at each time point by the maximum pathological load across all brain regions.

Model outputs

The function described above returns a normalized matrix representing the predicted regional damage levels at each time step. These results are stored in a dictionary using the *application code* presented in the previous paragraph and are later used to generate brain maps (see Section 4.3).

One way to visualize the model output is by plotting them as time series using the `plot_time_series` function mentioned earlier. This approach produces a plot showing the changes in damage levels over time for each brain region when it starts from a specific seed.

An example of such a model output is shown below:

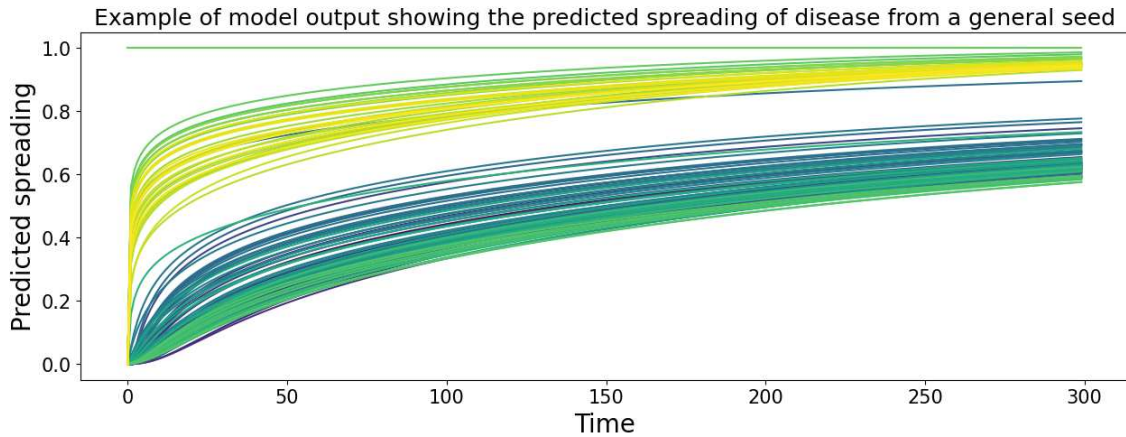


Figure 4.4: Model output (example with $t_points = 300$). Each curve, one for each brain region, illustrates how the disease factor (i.e., the predicted spreading of disease to that region, on the y -axis) evolves over time points (on the x -axis), ranging from 0 to 1, when the damage starts from a specific seed. A predicted spreading of 0 indicates no involvement of the region in the disease, while the increasing of the spreading in a region corresponds to an increasing of the disease factor there. The curve that remains constant at a value of 1 corresponds to the seed region.

Figure 4.4 illustrates an example of how the disease spreading evolves over time across different regions, starting from a specific seed. The plot displays a set of time-varying curves (one for each brain region) depicting their increasing damage with a specific seed region. All damage values are normalized to a range between 0 and 1. The curve corresponding to the seed region is easily identifiable, as it remains

at a predicted spreading value of 1 throughout the entire time interval. Each of the other curves refers to a different brain region, corresponding to the labels in the atlas.

What we reasonably expect from the model output is that diffusion will occur more rapidly in brain regions that are more strongly connected to the seed region. In other words, regions with a higher number of fibers linking them to the seed should exhibit faster spreading.

Since a simple temporal plot does not allow for a meaningful comparison of differences across regions, an alternative approach to analyze results is to generate a **brain map** at a selected time point t (see Section 4.3).

4.3 Model outputs as brain-maps

As mentioned previously, it is difficult to compare damage levels due to disease spreading from a time-evolution damage plot between different brain regions and between different seeds. For this reason, the next step in the thesis will focus on possible solutions to obtain brain maps of the model outputs at significant time points.

Since target data are not available, it is not possible to determine the optimal time point; however, t can be "intelligently" chosen at the point corresponding to the "belly of the curve", where regional values exhibit the greatest differences. At this selected time point, a damage value for each region is extracted and all of them are then mapped onto the brain.

The idea is therefore to take, for a single time point, all the damage values corresponding to all brain regions and represent them on a brain map. An example is shown below: each intersection point corresponds to a damage value of a specific brain region that must be stored for subsequent mapping (Figure 4.5).

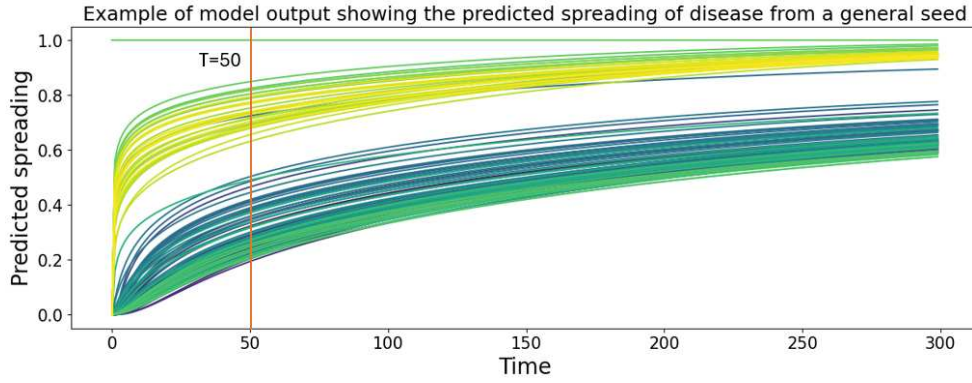


Figure 4.5: Example of a significant time point of the model output ($t_{\text{point}} = 50$), manually selected in order to select all the predicted spreading values at that time point and map them onto a brain image to investigate the spatial pattern of the spreading.

Two different methods are considered and analyzed to calculate brain-map rendering in order to select the most suitable one. These are listed as follows:

- **ggseg** package;
- **NIFTI libraries** in Python.

ggseg package

The *ggseg* is an open-source package available in both R and Python, with the Python version offering a simplified implementation. The Python module, which enables *ggseg*-like visualizations, can be installed directly through the Visual Studio Code terminal, whereas the original R version must be managed within the R environment. The *ggseg* package provides a framework for visualizing brain parcellations and neuroanatomical atlases in two-dimensional space using the grammar of graphics implemented in *ggplot2*. It supports the representation of cortical and subcortical brain regions, includes pre-defined atlases, and allows user-defined parcellations, thereby enabling reproducible and high-quality visualization of brain data.

Both versions were considered and applied to assess whether the resulting maps could effectively represent the outputs. The Python version was explored as a potential solution for brain mapping because it offered greater ease of use, as the model output dictionaries did not need to be exported into R-Studio. In contrast, the R version provides a more comprehensive set of functions and operations, enabling higher-quality brain visualizations. However, the implemented atlases in both versions include:

- Desikan-Killiany (DK) atlas for cortical regions;

- FreeSurfer aseg atlas for subcortical regions.

This means that the atlas did not coincide with the connectome regions (see 4.1.2), and therefore did not provide a way to map cerebellar regions. For this reason, to obtain brain maps using this package, a manual association between labels of different atlases was considered a possible solution, acknowledging that cerebellar subregions could not be mapped and merging different atlases was not feasible. The names of the subcortical regions were relatively easy to assign because they were anatomically similar; however, the main issue was that some regions could not be mapped. In particular, the Python version does not implement the **Brain-Stem** and the **Accumbens-area**, whereas the R version can map all subcortical regions except the **Accumbens-area**.

Another consideration is that a simplified version of the **AAL** atlas is supported in the R version. The problem in this case was that many cortical regions were not included in the atlas, such as several parietal, temporal, and occipital regions, while all frontal regions were supported.

Given these considerations, it was deemed more appropriate to return to the original idea, namely, to use the DK-Atlas and manually associate cortical regions in order to adapt the two atlases. The manual association of regions from the two atlases (**AAL** and **DK-Atlas**) was performed by overlaying them using the 3D visualization software **3d Slicer**. An anatomical correspondence approach was adopted, with the aim of achieving the best possible match between the regions of the **DK-Atlas** and those of the **AAL** atlas.

Despite the different strategies applied to produce brain maps, none of these methods offered a viable solution to optimize the maps: the atlases were incompatible for certain regions, merging them was not possible and the cerebellum could not be included. For this reason, another package was considered and used for comparison. This will be discussed in the next paragraph.

***Nifti* libraries on Python**

The Python libraries considered useful for mapping the results and adopted to address the mapping issue are listed as follows:

- **NiBabel**;
- **Nilearn**.

The **NiBabel** library provides access to common neuroimaging file formats such as **NIFTI**, and its primary purpose is to read, write, and manipulate neuroimaging

data stored in these formats. In contrast, **Nilearn** is designed for statistical analysis and machine learning on neuroimaging data and offers high-level tools to visualize statistical maps using brain atlases (**plotting** package was specifically adopted). Both libraries were employed to generate maps of the *entire brain*, including the cerebellum. Specifically, this required loading the combined atlas in NIfTI format.

The pipeline followed to obtain these maps will be discussed in detail in Section 4.3.1.

4.3.1 Brain-maps pipeline

Brain-maps of the model outputs are obtained following the next steps:

- Application of the NDM to the mean connectome and saving the model outputs into a dictionary (see code 4.2.2);

- Load the atlas in NifTI format;

```
1 atlas_img = nib.load(data_dir/ "AAL90_SUIT32_FIRST-1mm.
    ↪ nii.gz")
2 atlas_data = atlas_img.get_fdata()
```

- Load the label list containing ID and REGION_NAME;

```
1 labels_df = pd.read_csv(data_dir/ "Labels.csv")
```

- Select a time sample (**t_sample**);

- For each **seed region**, create a **Data Frame** with two columns: Region Name and Damage Value for the chosen time sample;

```
1 data = model_outputs[seed_region][:, t_sample]
2 df = pd.DataFrame({
3     'Region Name': ref_list,
4     'Damage Level': data
5 })
```

- Initialize a matrix for each seed region model output and fill it using the corresponding Data-Frame: use a *for loop* to assign the damage value of each region from the Data-Frame to the corresponding region in the empty map.

```
1 damage_map = np.zeros_like(atlas_data)
2 for _, row in labels_df.iterrows():
3     region_id = row['ID']
```

```

4     region_name = row['Regione']
5     match = df[df['Region Name'] == region_name]
6     if not match.empty:
7         damage_value = match['Damage Level'].values[0]
8         damage_map[atlas_data == region_id] =
            ↪ damage_value

```

- Create a NIfTI image from the resulting map and save it so that it can be opened in an imaging software such as 3d Slicer. Apply a values filter to make ROIs more distinguishable.

```

1 damage_map_filtered = np.where(damage_map < 0.09,
    ↪ damage_map, 0) #filtering
2 damage_img_filtered = nib.Nifti1Image(damage_map_filtered
    ↪ , affine=atlas_img.affine)
3 filename=f"damage_filtered_{t_sample}_{seed_region}.nii.
    ↪ gz"
4 filepath = saving_png/filename
5 nib.save(damage_img_filtered, filepath) # save as NIFTI
    ↪ image

```

- Visualize the *axial*, *sagittal* and *coronal* view of the brain-map in Python [x-slice y-slice z-slice] with a color legend indicating the level of damage:

```

1 plotting.plot_stat_map(damage_img_filtered, title=f"
    ↪ Damage map t={t_sample} Seed Region: {seed_region}"
    ↪ , cmap="viridis")

```

An example of the visualization of the three main views of the brain map is shown below:

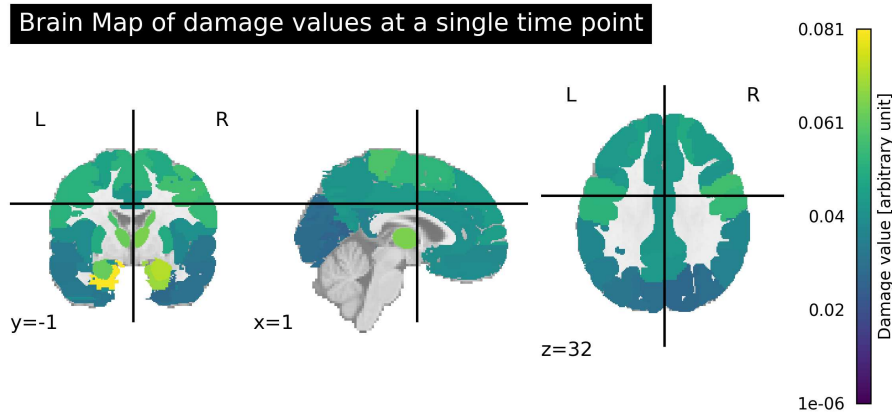


Figure 4.6: Example of a damage-level brain map (axial, sagittal, coronal view) at a single time point using NIfTI libraries in Python, displayed with the *viridis* colormap and filtered to include only damage levels below 0.1 in order to highlight cortical regions.

For each seed region, brain maps were obtained at manually selected early time points. The selection criteria for these time samples focused on time points near the beginning of the curves, since the objective is to analyze brain damage at the onset of the disease. For this reason, $\tau_{\text{sample}} = 500$ was used to generate brain maps at a single time point. With an adaptation of this pipeline, the temporal evolution of disease spreading across the brain could also be mapped. In this case, four time points were selected to extract a map for each: $\tau = [300, 500, 800, 1000]$. These time points were chosen near the start and kept close to each other to better visualize the progression of spread during the initial stages of the disease.

4.3.2 3d Rendering

As shown in Section 4.3.1, the Python libraries used above make it possible to generate and save brain maps of damage levels in NIfTI format, which can be opened and visualized in imaging software such as 3d Slicer. This software was used to produce a 3d rendering of the maps by loading the NIfTI image corresponding to a single time point for a selected seed region.

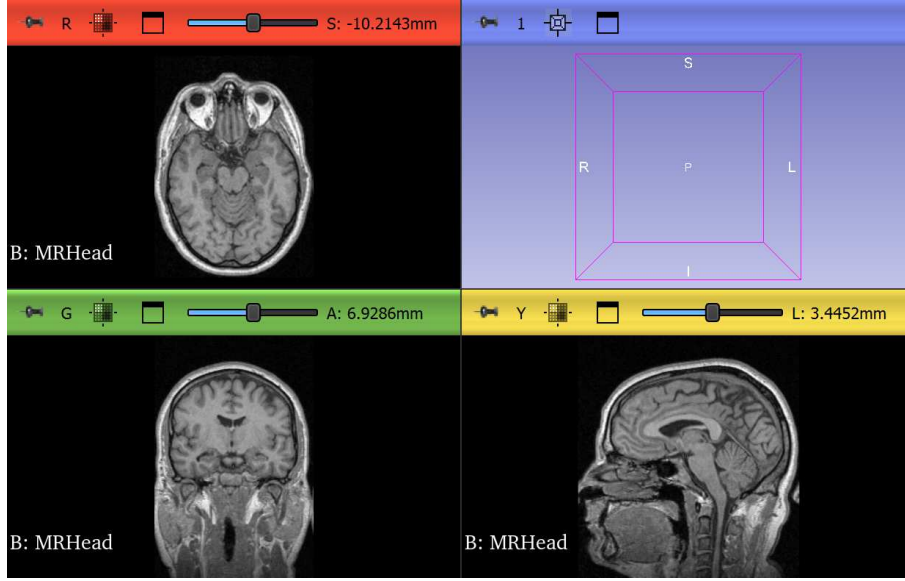


Figure 4.7: 3d Slicer environment: four panels to show the coronal, sagittal, axial view and 3d-rendering of the data loaded. The three different view of an image of example available in the software showing an MRI brain is shown.

After saving the brain map corresponding to the damage values when the disease starts from a single seed region at a specific time point in Python, the NIfTI image was loaded into 3d-Slicer. In the *Volumes* module, the color scale was set to an appropriate color-map (*viridis*), and the visualization panels for the three anatomical planes were used to scroll through the slices. In addition, a title and a color legend indicating the damage values were added to enhance the interpretability.

To include the 3d view, the *Volume Rendering* module was used. The region of interest (ROI) was set to encompass the entire brain volume, and the preset *MR-Default* was selected. The *Enable* option was activated, and *Display ROI* was checked to visualize the selected ROI (in this case, the whole brain). Finally, in the *Advanced* settings, the option *Synchronize with Volumes Module* was enabled under *Volume Properties*.

This pipeline enabled the 3d visualization of the brain maps, as well as slice-by-slice inspection through the three panels that allow scrolling across slice.

4.4 Validation of NSM results in MS

As mentioned previously, the second part of the chapter deals with the validation of the NDM results for MS. Specifically, the brain maps resulting from the simulation are compared with MS data acquired at baseline (CIS stage). A voxel-based analysis on the MS-dataset will determine differences between CIS and healthy controls in a

localized manner; similarly the comparison of regional values of quantitative metrics between CIS and healthy controls will lead to which regions are altered first in MS. The overall aim is to qualitatively assess the correspondence between NDM results and real CIS data.

In the first part, a brief overview of the main characteristics of MS will be provided, with particular emphasis on its pathophysiological mechanisms interpreted through quantitative metrics, that are selected for the voxel-wise analyses (Section 4.4.1). This is followed by a general description of the MS dataset used in this study (Section 4.4.2). The dataset not only includes DTI metrics but also measures derived from the NODDI model.

The final three sections will present the detailed pipeline, referring to the Python code used to perform voxel-wise analyses and statistical tests comparing MS patients with the control group.

The general pipeline followed to compute voxel-wise analyses through **z-score maps** is now introduced. The first step consists of data selection and includes patients at CIS stage which are then converted to MS and healthy controls. The second step involves downloading the data from XNAT. Maps are retrieved using the patient's **b-number** (unique identifier), together with the transformation matrix required to register in MNI-space, which makes them aligned and directly comparable.

The next step is the computation of voxel-wise maps for the control group. Specifically, a voxel-wise mean map and standard-deviation map, which reflect aggregated characteristics of all individual control maps, are generated. Finally, voxel-wise z-score maps are computed for each of the patient cohort. For each voxel in each patient map, the corresponding z-score is calculated. These maps can be visualized through **FSLeyes** Toolbox.

The z-score quantifies the deviation from the healthy population: higher values indicate a greater difference from the control mean. This measure helps highlight clusters of voxels, for each metric, that show stronger deviations from healthy controls, and can therefore be used to compare the model-generated maps with the empirical data.

This method provides a computational alternative to voxel-based analyses performed automatically using the SPM toolbox in MATLAB.

4.4.1 Advanced MRI parameters in MS

MS is an immune-mediated disease of the CNS whose progression reflects a shift from acute, focal inflammatory injury to more widespread inflammation and neurodegeneration, combined with a progressive failure of compensatory mechanisms such

as remyelination and neuroplasticity. Aging further increases neuronal vulnerability and reduces resilience [13]. In this framework, the biological mechanisms of MS pathology can be interpreted through quantitative advanced MRI metrics, which capture microstructural alterations in vivo.

Advanced MRI metrics vary according to the underlying tissue properties, and tissue damage produced by an MS episode can affect both the lesion core and the surrounding tissue in a widespread manner. These MRI-detectable changes reflect processes such as inflammation, neurodegeneration, and remyelination, and, more indirectly, neuroplasticity. Combining multiple quantitative metrics allows the study of microstructural alterations in a biologically meaningful way. For instance, an increase in MD accompanied by stable V_{fintra} may indicate inflammation or an early disease phase, whereas a reduction in V_{fintra} is more consistent with the onset of neurodegeneration. Variations in FA or ODI may instead reflect remyelination processes. Neuroplasticity, although relevant, is not directly captured by these structural MRI metrics and is usually assessed through functional imaging. For these reasons, this work examines metrics sensitive to microstructural changes typically associated with early-stage MS.

Underlying tissue injury in MS is mainly driven by unresolved inflammation. Focal inflammatory demyelination in WM is marked by perivenular inflammation, blood-brain barrier disruption, and a wave of demyelination that evolves over days to weeks, followed by variable degrees of tissue repair. Some inflammatory episodes resolve incompletely, leaving behind astroglial scarring, while approximately 20% of lesions does not resolve and may become chronic active lesions, including paramagnetic rim lesions. These slowly expanding lesions can extend into previously healthy perilesional tissue, causing low-grade demyelination and axonal transection. In progressive MS, additional inflammatory mechanisms, such as leptomeningeal inflammation and microglial activation, also contribute to extralesional tissue injury, although they remain less well understood.

Inflammation is tightly linked to neurodegeneration in MS. Axonal damage is evident from the earliest stages of lesion development, and neuronal loss can also begin early, becoming progressively more pronounced. Demyelination leads to abnormal redistribution of ion channels along axons, causing intra-axonal calcium accumulation and activating catabolic pathways that drive axonal degradation. Oxidative stress and mitochondrial dysfunction further contribute to glial and neuronal injury and may represent key drivers of disease progression [13].

In parallel, the CNS engages compensatory processes aimed at preserving function, chiefly remyelination and neuroplasticity. Remyelination restores thinner and shorter

myelin sheaths but helps preserve axonal conduction and protect axons from further injury. However, its efficiency varies widely across patients and lesions due to factors such as age, inflammation, lesion environment, and impaired oligodendrocyte progenitor cell differentiation. Neuroplasticity, the reorganization of structural and functional networks, may also help maintain functions but can become insufficient as damage accumulates.

Overall, disability progression in MS is thought to arise when irreversible axonal and neuronal loss outpaces remyelination and neuroplasticity. Quantitative MRI metrics offer a valuable means of capturing these early microstructural changes, providing insight into the balance between damage and repair mechanisms in MS.

4.4.2 MS dataset

The dataset used for the voxel-wise analyses includes 42 subjects, consisting of 19 patients and 23 healthy controls. Patient data were acquired within 3 months from the onset of neurological symptoms suggestive of their *first demyelination episode*. The dataset is part of a larger cohort acquired at the National Hospital of Neurology and Neurosurgery and Moorfields Eye Hospital in London, UK, which has been used in previous studies [4]. The scans were acquired using a 3T Achieva MRI scanner (Philips Medical System) equipped with a 32-channel head coil and a maximum gradient strength of 62 mT,m⁻¹. For this particular study, data from patients were selected, as previously described, in order to include individuals at the CIS stage who later converted to MS. Moreover, the cohort was chosen to include subjects with different lesion loads, in order to obtain a heterogeneous group that captures a range of possible pathological presentations. The general characteristics of the dataset are reported in Figures 4.8 and 4.9.

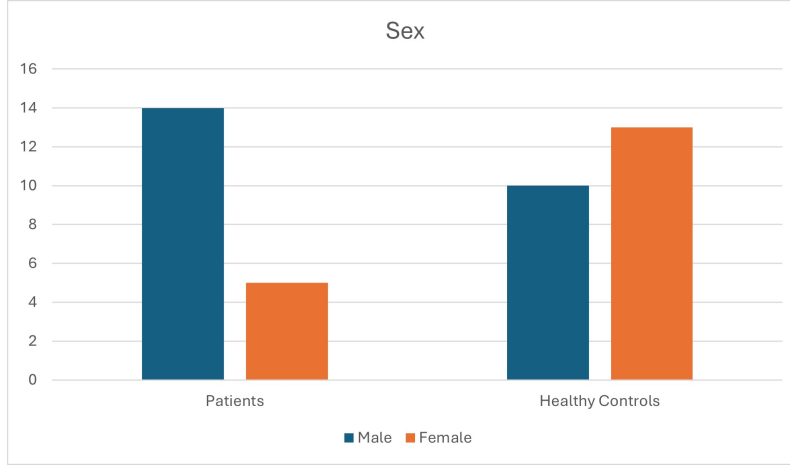


Figure 4.8: Histogram of the sex distribution in the patient (19 subjects) and healthy control (23 subjects) groups. Male subjects are represented in *blue* colour, while women in *orange*.

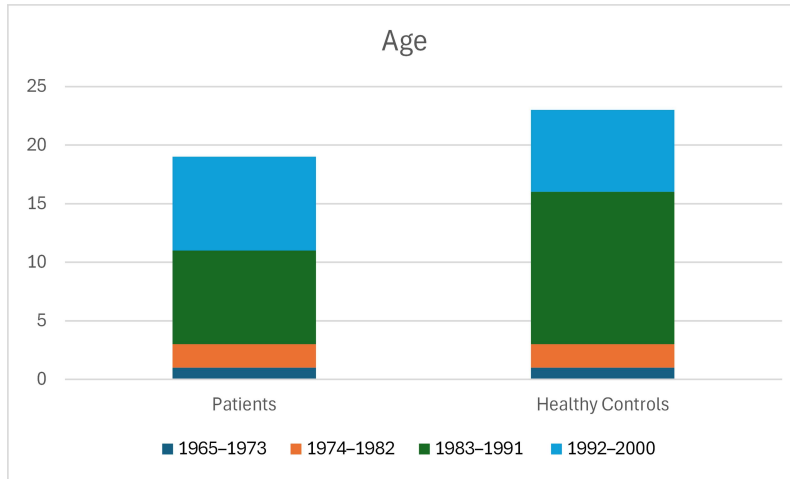


Figure 4.9: Histogram of the age distribution in the patient and healthy control groups. Five homogeneous age intervals are displayed, ranging from 1965 to 2000, each represented by a different color

For each subject, a set of specific scans was downloaded from the **XNAT** Database Environment:

- DTI metric maps: **FA** and **MD**;
- NODDI metric maps: **ODI**, $V_{\text{f intra}}$, $V_{\text{f iso}}$;
- binarized **Lesion Mask** for each patient;
- MNI-space registration transforms.

The metrics selected were chosen in order to capture alterations in tissue microstructure.

4.4.3 MNI-space registration

After downloading the entire dataset, including all metric maps for each subject, the next step consisted of registering all maps from subject space into MNI space. The MNI space is a common anatomical reference in which images are spatially aligned, allowing them to overlap and therefore be comparable across subjects.

To achieve this transformation, it was necessary to download the **transformation matrices**, which are subject-specific and stored in the **RESULTS_xnatRegT1DWI** folder in XNAT, as described in the previous section. These functions were applied to the metric maps through dedicated Bash commands, which are detailed as follows, together with significant parts of the corresponding Bash code applied to each of the metrics.

The registration in MNI space is based on two main functions available in the **NiftyReg** package, namely:

- **reg_transform**;
- **reg_resample**.

The pipeline followed to obtain the images in MNI-space, and implemented in the Bash code, consists of three steps:

- Application of the *reg_transform* command to compute the inverse transformation matrix, in order to obtain the transformation from b0 space to T1 space;
- Application of the *reg_resample* command to resample the maps in T1 space, using the default interpolation methods for all metric maps and nearest-neighbor (NN) interpolation (`-inter 0`) for the lesion mask;
- Application of the *reg_resample* command to resample the maps from T1 space in atlas space (MNI), using the same interpolation settings as above and taking as reference the standard MNI152 template with 2 mm isotropic resolution.

MNI-space registration bash code

```

1 # save files in local variables
2 T1ToB0="$trans_function1/T1ToB0_transformation.nii.gz" # T1
   ↪ to b0
3 T1ToAtlas="$trans_function1/T1ToAtlas_transformation.nii.gz"
   ↪ # T1 to Atlas

```

```

4  parcellation="/usr2/mrtools/fsl/data/standard/MNI152_T1_2mm.
    ↪ nii.gz"
5  T1="$trans_function1/T1.nii.gz"
6
7  dwi_FA="$maps1/bDWIduc-dti-fa.nii.gz" # FA
8  lesion_mask="$lesion1/lesions_mask_t1_space-bin.nii.gz" #
    ↪ Lesion mask
9
10 # MNI registration (Default interpolation)
11 reg_transform -invNrr "$T1ToB0" "$dwi_FA" "
    ↪ $trans_function1/B0ToT1_transformation.nii.gz" #
    ↪ inverse transformation
12 reg_resample -ref "$T1" -flo "$dwi_FA" -trans "
    ↪ $trans_function1/B0ToT1_transformation.nii.gz" -res
    ↪ "FA_T1_space.nii.gz" # b0 to T1
13 reg_resample -ref "$parcellation" -flo FA_T1_space.nii.gz
    ↪ -trans "$T1ToAtlas" -res "$root_pz/dti_FA/${
    ↪ subj_name}_FA_mni_space.nii.gz" # T1 to Atlas
14
15 # MNI registration lesion mask (NN interpolation)
16 reg_resample -ref "$T1" -flo "$lesion_mask" -trans "
    ↪ $trans_function1/B0ToT1_transformation.nii.gz" -res
    ↪ lesion_T1_space.nii.gz -inter 0
17 reg_resample -ref "$parcellation" -flo lesion_T1_space.
    ↪ nii.gz -trans "$T1ToAtlas" -res "$root_pz/
    ↪ lesion_mask_bin/${subj_name}_lesion_mni_space.nii.gz
    ↪ " -inter 0
    
```

4.4.4 Voxel-wise z-score maps calculation

After the registration of all metric maps in MNI-space, **voxel-wise z-score maps** are computed. Using healthy control maps as reference population, a z-score is calculated for each voxel of each map of each patient, according to the following formula:

$$z - score_{pz}(i) = \frac{voxel_{pz}(i) - mean(voxel_{hc})(i)}{std(voxel_{hc})(i)} \quad (4.2)$$

where $voxel_{pz}$ is the patient's voxel value, i is the metric considered, and $mean(voxel_{hc})$ and $std(voxel_{hc})$ represent the mean and standard deviation of the corresponding voxel in healthy controls for that metric.

First, a mean map and a standard deviation map are computed for each metric by aggregating all healthy-control maps. These maps are used as reference for the z-score calculation. The function `voxel_wise_mean_std.py` developed for this purpose is reported below. A `brain_mask` is used in order to exclude the background from the computation of both mean and standard deviation.

Then, as shown in the function `z_score_maps.py`, voxel-wise z-score maps are generated for each metric map of each patient. Within a *for loop*, the z-score formula is applied to every voxel to obtain a complete z-score map for each patient and for each metric.

Subsequently, a mean z-score map is computed across patients so that one representative z-score map is obtained for each metric. To compute this map, a voxel-wise **trimmed mean** at $p\%$ with $p = 5$ has been applied (using the `trim_mean` function in the Python code). This procedure excludes *outliers* from the voxel-wise mean calculation, ensuring that the resulting z-score values are more reliable. Outliers may arise from registration inaccuracies, particularly at the edges of the brain, where non-linear transformations tend to propagate spatial distortions. In the dataset, such outliers are indeed present: voxels with extreme values are likely caused by misalignment or edge errors introduced during registration, probably because of the two-stage registration process. Since the analyses is voxel-wise, no corrective spatial smoothing was applied, and misalignment-related artifacts can remain visible. This effect is less evident when computing the healthy-control mean maps (used as the z-score reference), because averaging naturally reduces voxel-wise noise and registration errors. For these reasons, after computing the individual voxel-wise z-score maps for each patient, the final mean metric-specific z-score maps are obtained by applying a $p\%$ trimmed mean across subjects. In particular, for each set of corresponding voxels, the trimmed mean computes the average after excluding the 5% of values lying at the extremes of the voxel-value distribution.

To facilitate cluster identification, the maps display only voxels falling outside the $|z| > 1$ range. This thresholding approach highlights areas of deviation from the mean. These maps are then overlaid onto the **lesion probability map**, which was computed by summing the lesion maps of all patients.

Generation of healthy controls mean and standard deviation maps

```

1 def voxel_wise_mean_std(
2     files: List[Union[str, Path]],
3     brain_mask,
4     ddof: int = 1):

```

```

5
6 file_list = [str(f) for f in files] #convert paths (input)
    ↪ to str
7 ref_img = nib.load(file_list[0]) # reference image
8 stack = np.stack([nib.load(f).get_fdata(dtype=np.float32)
    ↪ for f in file_list], axis=0)
9
10 # Application of the brain mask
11 brain_mask_boolean = brain_mask > 0 # TRUE=brain FALSE=
    ↪ background
12 mask_4d = np.broadcast_to(~brain_mask_boolean, stack.shape)
13 masked_stack = np.ma.array(stack, mask=mask_4d, dtype=np.
    ↪ float32, copy=False)
14
15 # Computation of mean and sd map
16 mean_map = masked_stack.mean(axis=0)
17 std_map = masked_stack.std(axis=0, ddof=ddof)
18
19 mean_img = nib.Nifti1Image(mean_map, ref_img.affine, ref_img
    ↪ .header)
20 std_img = nib.Nifti1Image(std_map, ref_img.affine, ref_img
    ↪ .header)
21
22 return mean_map, std_map

```

Generation of patients' z-score maps

```

1 def z_score_maps(
2     files: List[Union[str, Path]],
3     mean_map,
4     std_map,
5     metric):
6
7     file_list = [str(f) for f in files]
8     ref_img = nib.load(file_list[0])
9
10    low_treshold = -1 #treshold
11    high_treshold = 1
12
13    z_maps_data =[]

```

```

14  for f in file_list:
15      img = nib.load(f).get_fdata(dtype=np.float32)
16      zmap = np.zeros_like(img, dtype=np.float32) #initalize
17      np.divide(img - mean_map, std_map, out=zmap, where=
          ↪ std_map > 0) # z-score formula
18      zmap[~np.isfinite(zmap)] = 0.0
19      z_maps_data.append(zmap)
20
21  # Trimmed mean maps for each metric at p% [p = 5]
22  z_stack = np.stack(z_maps_data, axis=0).astype(np.float32)
23  z_map_mean_trimmed = trim_mean(z_stack, proportiontocut
          ↪ =0.05, axis=0) #cut 5% sx and dx
24  z_mean_img_trimmed = nib.Nifti1Image(z_map_mean_trimmed.
          ↪ astype(np.float32), ref_img.affine, ref_img.header)
25
26  # Apply treshold
27  zmap_trimmed_thr = z_map_mean_trimmed.copy()
28  mask = (zmap_trimmed_thr>low_treshold) & (zmap_trimmed_thr<
          ↪ high_treshold)
29  zmap_trimmed_thr[mask] = 0 # [-1 1] = 0
30  z_trimmed_thr = nib.Nifti1Image(zmap_trimmed_thr, ref_img.
          ↪ affine, ref_img.header)

```

Box-plot analyses

Beside the visualization of voxel-wise z-score maps, a box-plot analysis was performed to investigate differences in region-wise z-score values computed for each metric.

Since the z-score maps are voxel-wise, the first step consisted in projecting them onto the combined atlas used for the model analyses. Specifically, for each patient's voxel-wise map, the average value of z-score was calculated within each brain region of the AAL-FIRST-SUIT atlas. The resulting values were saved in a CSV file containing, for each metric, all regional averages across all patients. These box-plots then provide a visualization of the z-score distribution within selected regions of the brain in the cohort. The pipeline consists of two main steps:

- Computation of the average z-score value within each atlas region using the `seg_stats` command in Bash, followed by aggregation of all subjects' outputs in a single CSV file for each metric using `cat`;
- Generation of box-plots (`box_plot` function) for each metric, focusing on

selected ROIs, using CSV files as input.

4.4.5 Statistical tests

An additional method based on statistical approaches (hypothesis tests) was employed. Since the voxel-wise analysis did not reveal clearly visible clusters of altered voxels in GM regions, a statistical approach was introduced to further investigate potential regional patient-control differences. First, a Welch two-sample hypothesis test was applied region-wise to compare patients and controls, using the metric maps in MNI-space. To obtain region-wise measurements suitable for this analysis, the parcellation from the combined atlas was applied to each map, yielding one value per region for every subject.

After generating CSV files aggregating regional values across all subjects - separated by metric and by group (patients and controls) - the **Welch test**, combined with FDR correction, was performed to identify significant regional differences. Since this analysis did not reveal substantial group differences, a more powerful multivariate approach was subsequently adopted. For this reason, a PCA was conducted to detect broader patterns of regional variation that may distinguish patients from controls. The detailed pipeline and the python code generated to obtain the results are described below.

Region-wise measures

Using the same Bash commands adopted for generating the box-plots in the previous sections, the average value of each brain region in the combined atlas was computed for every metric map registered to MNI-space, for both patients and healthy controls. The corresponding commands are reported below. The **seg_stats** command calculates the mean metric values within each atlas region and stores them in individual CSV files, whereas the **cat** command concatenates these outputs, producing a single CSV file for each metric that contains the region-wise values of all patients or all healthy controls.

```
1 while read each; do seg_stats ${each} -A1 AAL90_SUIT32_FIRST
   ↪ -2mm.nii.gz ${each/.nii.gz/}_atlas.csv; done
2 ls *.csv | while read each; do cat "${each}"; echo ""; done >
   ↪ MNI_FA_pz.csv
```

Two samples Welch test

To compare patients and healthy controls region by region for each metric, a Welch two-sample test was selected. This choice is motivated by the fact that Welch's test does not assume equal variances or equal sample sizes between groups, and relies on the Welch-Satterthwaite equation to estimate the degrees of freedom. These properties make it more robust and reliable than the classical Student's t -test, especially in real datasets where unequal group sizes are common, as the MS-dataset. The statistical test was applied independently to each ROI, meaning that the code tested, for every metric, whether a given ROI showed a significant difference between patients and controls.

Since the atlas contains more than one hundred ROIs, performing a large number of parallel statistical tests increases the probability of obtaining false positives purely by chance. For this reason, a False Discovery Rate (FDR) correction was applied to the collection of p -values. The FDR approach controls the expected proportion of falsely declared significant results and reduces spurious findings while maintaining adequate sensitivity to detect true group differences. The Python code used for this analysis is shown below. In the script, the `ttest_ind` function applies the Welch two-sample test independently for each ROI of the patient and control groups. Subsequently, the `multipletests` function is used to recompute the p -values using the FDR correction, ensuring proper control of false discoveries across the large number of parallel statistical tests.

```

1 pz = pd.read_csv("MNI_Vfiso_pz.csv")
2 ctrl= pd.read_csv("MNI_Vfiso_hc.csv")
3 ROIS = pz.columns.intersection(ctrl.columns) #ROIs
4 results = { "ROI": [], "p_value": []}
5
6 for roi in ROIS:
7     x = pz[roi].to_numpy(dtype=float) #array
8     y = ctrl[roi].to_numpy(dtype=float)
9     t, p = ttest_ind(x, y, equal_var=False) # welch t-test
10
11     results["ROI"].append(roi)
12     results["p_value"].append(p)
13
14 out = pd.DataFrame(results)
15 reject, p_fdr, _, _ = multipletests(out["p_value"], method="
    ↪ fdr_bh") #FDR correction
16 out["p_FDR"] = p_fdr

```

```
17 out["significant_FDR"] = reject    # FALSE = ROI rejected
```

Principal Component Analyses

Since the application of the Welch t -test did not reveal any region showing significant patient-control differences - when each ROI was tested independently and given the large number of regions involved - a more robust multivariate approach was adopted: **PCA**. PCA reduces the dimensionality of the dataset (from more than one hundred ROIs to a small number of components) while retaining the majority of the variance. By constructing a new set of uncorrelated features, the Principal Components (PCs), this method increases statistical power and enables the identification of coordinated patterns of regional variation that may differentiate patients from controls.

In this framework, statistical tests (Welch t -tests) are performed on PCs rather than on individual ROIs. If a PC is found to be significantly different between groups ($p < 0.05$), its loadings are then examined to determine which ROIs contribute most strongly to the observed pattern. Positive loadings for these PC correspond to regions contributing to higher component values in patients relative to controls, whereas negative loadings correspond to regions contributing to lower component values in patients.

The complete pipeline is outlined below, together with the relevant excerpts of the Python code used to implement each step.

- Data loading and z -score normalization

The data matrix is constructed from the entire dataset, including both patients and healthy controls (processed separately for each metric), and is then standardized using the `StandardScaler` function, which compute the z -score normalization.

```
1 pz = pd.read_csv("MNI_Vfiso_pz.csv")
2 ctrl = pd.read_csv("MNI_Vfiso_hc.csv")
3 X=pd.concat([pz,ctrl], axis=0) # data matrix
4
5 scaler = StandardScaler() # z-score
6 X_std = scaler.fit_transform(X)
7 labels = np.array([0]*len(pz)+[1]*len(ctrl)) # 0 = pz, 1=
    ↪ ctrl
```

- Principal Component Analysis (PCA)

With the `fit_transform` function, the PCA model is first fitted to the data matrix and then applied to it. In practice, this step projects the original standardized data onto the new feature space defined by the PCs, yielding a transformed data matrix expressed in the PCA space.

```
1 pca = PCA(n_components=None)
2 X_pca = pca.fit_transform(X_std) #PCA
```

- Selection of the relevant PCs based on the explained variance

The selection of the PCs can be guided by the amount of **explained variance**. A common criterion consists in choosing the minimum number of PCs that together account for approximately 70/90% of the total variance. All available PCs are first computed (equal to $\min(N_{subj} - 1, N_{regions})$), the cumulative explained variance is then evaluated, and the smallest set of PCs whose cumulative contribution exceeds 80% of the total variance is selected.

```
1 explained = pca.explained_variance_ratio_
2 cumulative = np.cumsum(explained)
3 n_pc = np.argmax(cumulative >= 0.80) + 1
4 scores = X_pca[:, :n_pc]
```

- Welch two-sample *t*-test applied to the selected PCs

The selected PCs were compared using a **Welch two-sample *t*-test**. If a PC yields a *p*-value below 0.05, that component is considered significantly different between patient-control groups. In practical terms, a significant PC indicates that patients and controls differ along the multivariate pattern of ROIs represented by that component.

```
1 results = []
2 for i in range(scores.shape[1]):
3     pc_pz = scores[:len(pz), i] #pz
4     pc_ctrl = scores[len(pz):, i] #controls
5     t, p = ttest_ind(pc_pz, pc_ctrl, equal_var=False)
6     results.append({
7         "PC": f"PC{i+1}",
8         "p_value": float(p),})
```

- Inspection of the loadings for PCs showing significant patient-control differences

For each PC identified as significant, the result indicates that the corresponding multivariate combination of ROIs differs between patients and healthy controls

($p < 0.05$). To understand which regions drive this separation, the loadings of that PC are examined, as they reveal the ROIs that contribute most strongly to the component. By analyzing these loadings, specific combinations of regions that are responsible for the multivariate difference between the two groups can be identified.

```
1 loadings = pd.DataFrame( pca.components_[0:n_pc].T, index=
    ↪ X.columns, columns=[f"PC{i+1}" for i in range(n_pc)
    ↪ ])
```

The results of PCA, namely the significant PCs for each metric and their corresponding loadings, were examined through histograms generated directly within the Python code. For the final analyses, in which the model outputs are compared with the MS dataset, a subset of ROIs was selected - specifically, those showing the highest loadings for the significant PCs. These ROIs can be considered the regions that most strongly drive the separation between the two groups, and therefore represent meaningful candidates for the comparison with the model outputs.

Chapter 5

Results

In this chapter, the model results with all related maps are presented and discussed. First, a collection of model output graphs is presented (see Section 5.1), each corresponding to the results of the simulation started from a different seed region. Subsequently, selected brain maps are shown (see Section 5.2). All these results were obtained using the **NSM Toolbox**, specifically by applying the **NDM** to the healthy mean connectome.

The final section is dedicated to the **voxel-wise z-score maps** derived from the MS-dataset (see Section 5.3) and the results of the **PCA** (see Section 5.4).

5.1 Network Diffusion Model outputs

In this section, graphs illustrating the predicted spreading of the disease when it starts from each of the seed regions are shown. Each graph displays multiple curves representing the predicted spread of the disease in each region of the brain included in the combined atlas, in a range from 0 to 1, which refers to the degree to which a region is involved in the disease over a series of time points.

Model Output: Cerebellar subregions

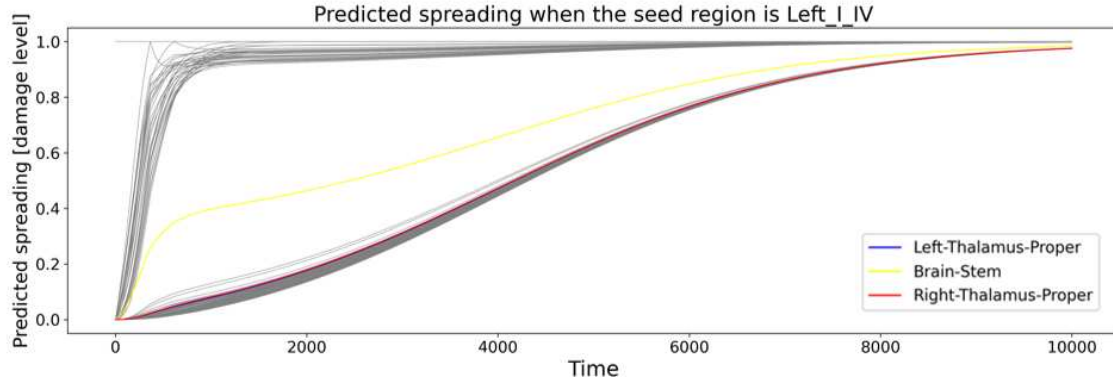


Figure 5.1: Model output showing the disease spreading across GM regions when the simulation is initiated from the cerebellar region *Left_I_IV*. Each curve represents a different brain region and its temporal involvement. The predicted spreading on the y-axis ranges from 0 to 1 and reflects how rapidly a brain region becomes involved in the disease: values close to 0 indicate no involvement, whereas values approaching 1 indicate stronger damage: the closer and sooner a region reaches 1, the earlier it is significantly damaged. The x-axis represents time points, corresponding to a fictitious temporal scale based on discrete model iterations rather than real time. When the damage originates from a cerebellar subregion, two distinct sets of curves emerge: one group shows a faster increase in spreading, corresponding to regions affected earlier due to a higher level of connection to the seed, while the second group exhibits a slower rise. Among these, the Brain-Stem (highlighted in *yellow*) appears as an isolated curve, clearly distinguishable from the others. The Thalamus (in *red* the right and in *blue* the left), although belonging to the second, slower set of curves, shows one of the fastest trajectories within this group. The first set of rapidly increasing curves corresponds to other cerebellar regions, while the second set represents the remaining brain areas. The curve that remains constantly at a value of 1 corresponds to the seed region itself.

Figure 5.1 shows the model output when the seed region is located in the cerebellum, specifically the *Left_I_IV*. The number of time points is set to 10,000, and a subset of regions is color-coded to highlight their role in the disease, as illustrated above. Specifically, the Brain-Stem curve is shown in *yellow*, while the Thalamus curve is shown in *red* and *blue*. From the figure, it can be observed that, aside from the curve located in the middle corresponding to the Brain-Stem, two distinct groups of curves can be clearly identified. The first group exhibits an early onset of damage and mainly includes cerebellar subregions, whereas the second group becomes affected at later time points. As a result, the Brain-Stem shows a temporal progression that deviates from all other spreading curves, while the left and right Thalamus display very similar temporal patterns.

Model Output: Thalamus

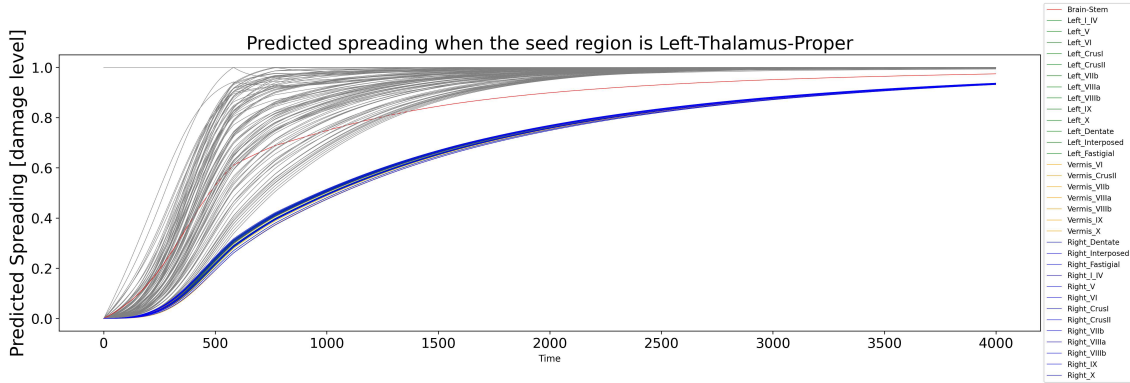


Figure 5.2: Model output showing the disease spreading across GM regions when the simulation is initiated from Left Thalamus. Each curve represents a different brain region and its temporal involvement. The Brain-Stem (highlighted in red) has a temporal evolution similar to other cortical regions, although its shape may be slightly different. It becomes isolated at a later time point. The other colored curves correspond to cerebellar subregions, which are distinguished using three different colors depending on whether they belong to the left hemisphere, the right hemisphere, or the vermis. This color coding was included to highlight how the cerebellum behaves when the damage originates from the Thalamus. Overall, this group of curves shows a slower spreading compared to the other regions.

By setting Thalamus as seed, we can observe that the Brain-Stem curve has a temporal evolution similar to other cortical regions, although its shape may be slightly different and it becomes isolated at a later time point. On the other hand, the cerebellar curves appears to have a slower spreading with a set of curves that are close to each others.

The number of time point is set to 4000.

Model Output: Motor Cortex and Visual Cortex

Lastly, the seed region was also placed in the motor cortex and in the visual cortex. In particular, as previously mentioned, two brain regions were selected as seed regions: Precentral region and Calcarine.

Also in this case, the Brain-Stem and cerebellar regions are color-coded to facilitate comparison across results.

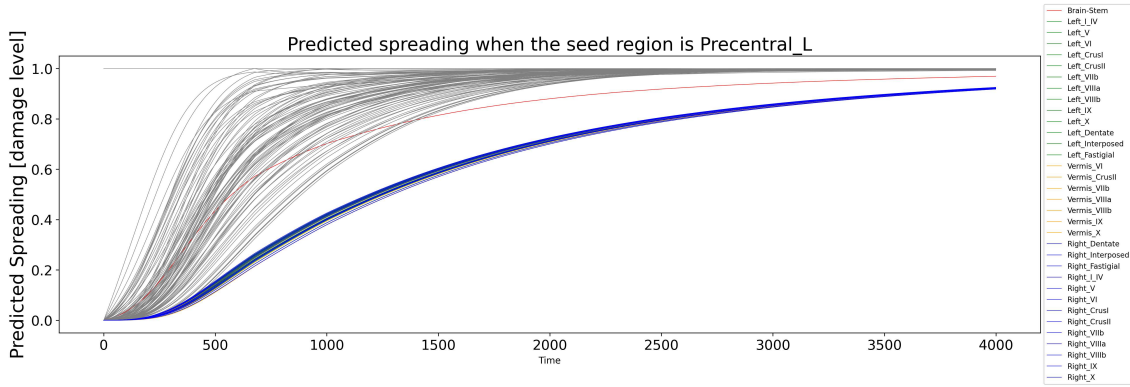


Figure 5.3: Model output showing the disease spreading across GM regions when the simulation is initiated from a motor cortex subregion: Left Precentral. Each curve represents a different brain region and its temporal involvement. The Brain-Stem (highlighted in *red*) has a temporal evolution similar to other cortical regions, although its shape may be slightly different and it becomes isolated at a later time point. The other colored curves correspond to cerebellar subregions, which are distinguished using three different colors depending on whether they belong to the left hemisphere, the right hemisphere, or the vermis. This color coding was included to highlight how the cerebellum behaves when the damage originates from the motor cortex. Overall, this group of curves shows a slower spreading compared to the other regions.

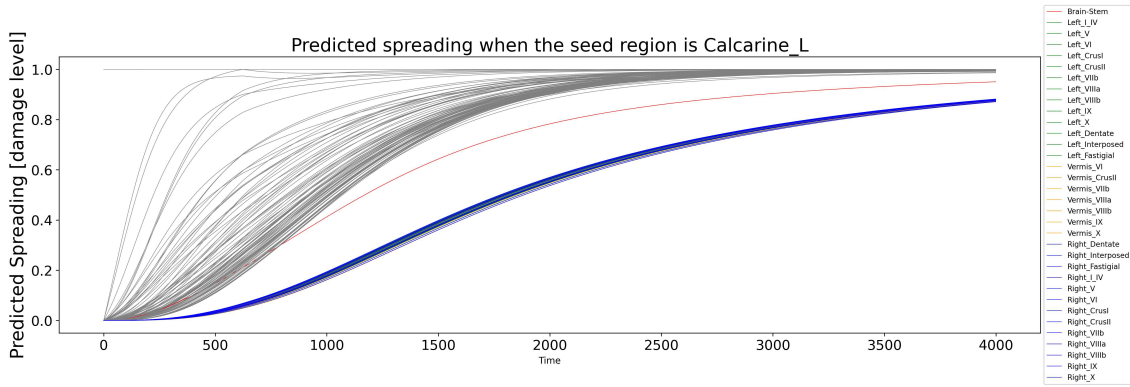


Figure 5.4: Model output showing the disease spreading across GM regions when the simulation is initiated from a visual cortex subregion: Left Calcarine. Each curve represents a different brain region and its temporal involvement. The Brain-Stem (highlighted in *red*) has a temporal evolution similar to other cortical regions, although its shape may be slightly different and it becomes isolated at a later time point. The other colored curves correspond to cerebellar subregions, which are distinguished using three different colors depending on whether they belong to the left hemisphere, the right hemisphere, or the vermis. This color coding was included to highlight how the cerebellum behaves when the damage originates from the visual cortex. Overall, this group of curves shows a slower spreading compared to the other regions.

5.2 Brain maps

5.2.1 Disease spreading maps from cerebellar subregions

Brain damage maps at a single time point with the seed placed in each cerebellar subregion



Figure 5.5: Brain damage maps (coronal, sagittal and axial views) at a single time point ($t = 500$), each representing the level of disease in every brain region at that time point, when the damage originates from each cerebellar seed region which is named above each sub-figure. The color bar shows a filtered range of values, spanning from *black* for values near 0 to *white* for higher damage levels. The filtering was applied to enhance the visualization of intensity variations across maps and to allow meaningful comparison between them. Some seed regions show lower damage propagation, such as the Fastigial nucleus, whereas others display a faster and more widespread diffusion, such as Crus I or I_IV.

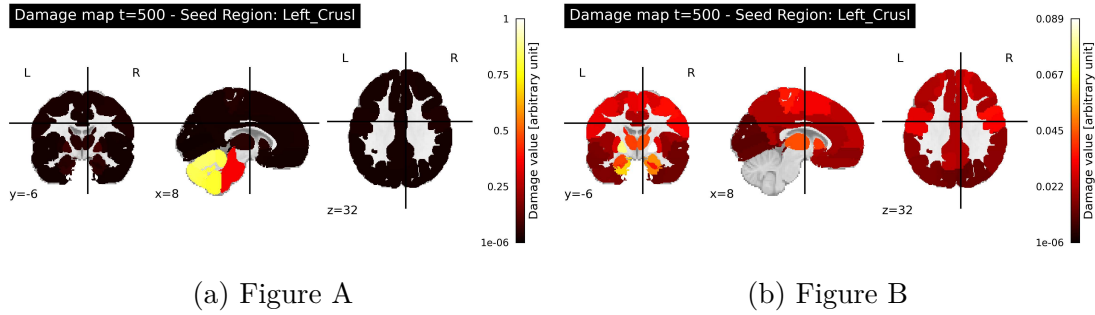
Figure 5.5 shows a comparison between brain maps for each of the three anatomical views (coronal, sagittal, and axial) when the spreading process starts from each cerebellar subregion. To better visualize the differences between the maps, the values have been filtered using an empiric threshold, with the aim of selecting a value that allows the differences between the maps to be compared as clearly as possible.

Indeed, it is clearly visible that the damage level of the same brain region depends on the chosen seed region, given the same amount of time. For this reason, some regions are not included in certain maps. For example, in most cases the Cerebellum and the Brain-Stem are not visible because they exhibit a higher damage level. However, in some maps, such as those associated with **Fastigial** and **Vermis_X**, these areas are included and clearly visible.

One of the most evident observations from the maps is that those having cerebellar regions such as *Crus_I*, *I_IV*, *V*, *VI*, *Crus_II* as seeds exhibit higher levels of damage compared to those seeded in *Fastigial* or *Vermis_X*. This difference is readily apparent from the color distributions in the maps: the first group is characterized by colors corresponding to higher damage levels, whereas the latter group generally displays darker colors across most brain regions.

These initial maps allow a qualitative visualization of the differences in spreading dynamics originating from different cerebellar subregions. However, given the large number of maps considered, the analyses and results will focus on a reduced subset of cerebellar seeds. In particular, *Left_Crus_I*, *Left_I_IV* and *Vermis_X* are selected for further investigation, as they appear to be associated with the fastest and slowest disease spreading.

Focus on Left_CrusI brain maps



(a) Figure A

(b) Figure B

Figure 5.6: Brain damage maps at a single time point ($t = 500$) with Left_CrusI as seed region; given the wide range of the predicted spreading values in the supratentorial brain and in the hind-brain, the visual representation has been split into two: A) shows the cerebellum and Brain-Stem, which are the first to be involved; B) cortical regions are damaged later, as demonstrated by the reduced colour scale of predicted spreading values [15]

In Figure 5.6.A, a brain map including all regions is illustrated; therefore, no filtering has been applied. When the spreading starts from the *Left_Crus_I*, the first regions affected are the other cerebellar subregions (shown in *yellow*), followed by the Brain-Stem (highlighted in *red*) and finally the rest of the cerebral cortex.

Figure 5.6.B instead highlights the damage values in the cortical and deep GM regions. In this case, values have been filtered. Deep GM regions appear to reach higher damage levels compared to several cortical areas, such as the visual cortex. Among the cortical regions, the motor cortex is particularly distinguishable, as it is represented by a brighter shade of red, indicating a higher level of damage compared to the surrounding areas.

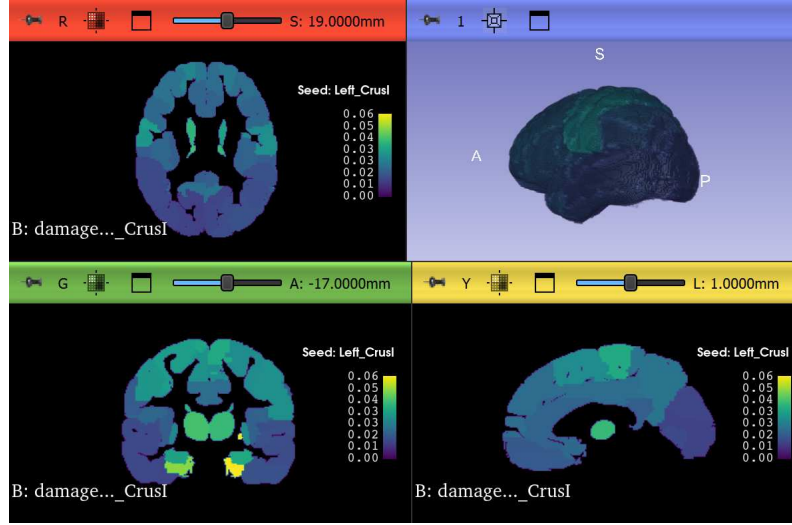


Figure 5.7: Brain damage maps filtered within a range that enhances the visualization of deep GM and cortical regions, at a selected time point ($t = 500$) with *Left_CrusI* as the seed region. The views displayed include coronal, sagittal, and axial slices, alongside the 3d rendering obtained in 3d Slicer. The color bar represents the damage values using the *viridis* color scale.

Deep gray matter maps

The maps in Figure 5.8 illustrate the temporal evolution of predicted spreading levels in deep GM regions. In particular, four time points have been selected: 300, 500, 800, and 1000.

The comparison is made among three cerebellar subregions used as seeds. The *Vermis_X* exhibits a slower progression of damage compared to both *Left_Crus_I* and *Left_I_IV*. Moreover, the **Pallidum** appears to be the first region affected, followed by the **Thalamus**, whose damage values are also reported below each map.

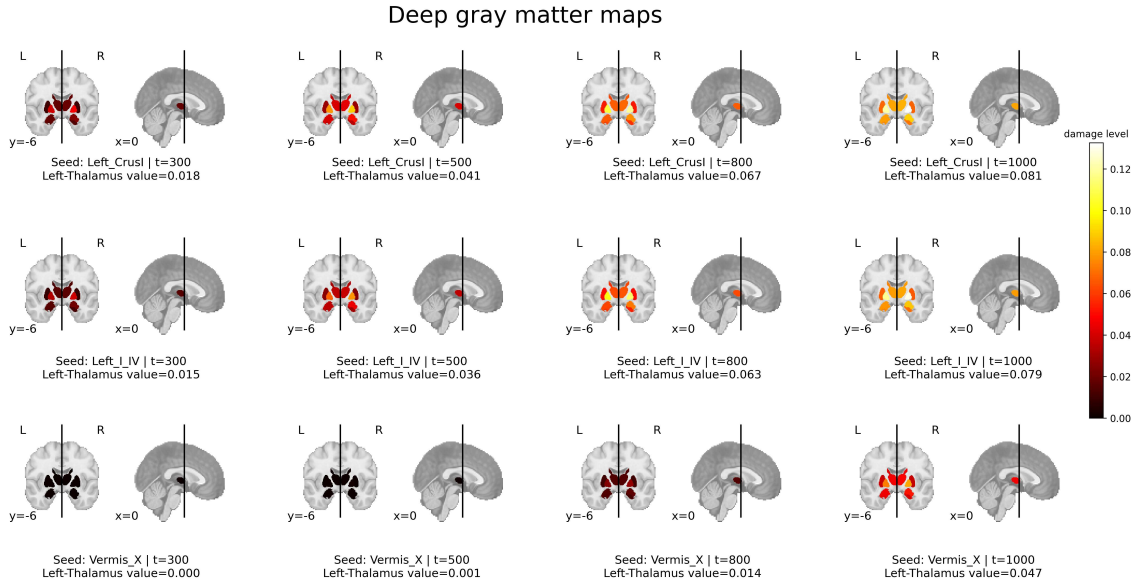


Figure 5.8: Temporal evolution brain-maps (coronal and sagittal view) at four time points ($t = 300, 500, 800,$ and 1000) showing early progression of damage in deep GM regions when the seed is placed in Left_CrusI, Left_I_IV, and Vermis_X, visualized using a hot color map. The Left Thalamus value is highlighted for reference. Among these regions, Vermis_X exhibits a slower progression compared to the other two seeds. The Pallidum emerges as the region with the fastest increase in damage, followed closely by the Thalamus. [15]

5.2.2 Disease spreading maps from the Thalamus

The brain map shown in Figure 5.9 represents the regional damage levels when the spreading starts from the *Left_Thalamus*. No filtering has been applied to the values; therefore, all regions are included and clearly distinguishable. The seed region is visible as it appears in *yellow*.

The 3d Slicer visualization enhances the interpretation of damage distributions, thanks to the possibility of scrolling through individual slices as well as exploring the data using three-dimensional renderings. In general, the left hemisphere of the cortex appears to be more affected, as indicated by brighter colors, compared to the right hemisphere, while the cerebellum exhibits a lower level of damage.

The brain maps corresponding to the cerebellar subregion shown in Figure 5.7 exhibit lower damage values compared to those associated with spreading from the Thalamus, when considering the same time point.

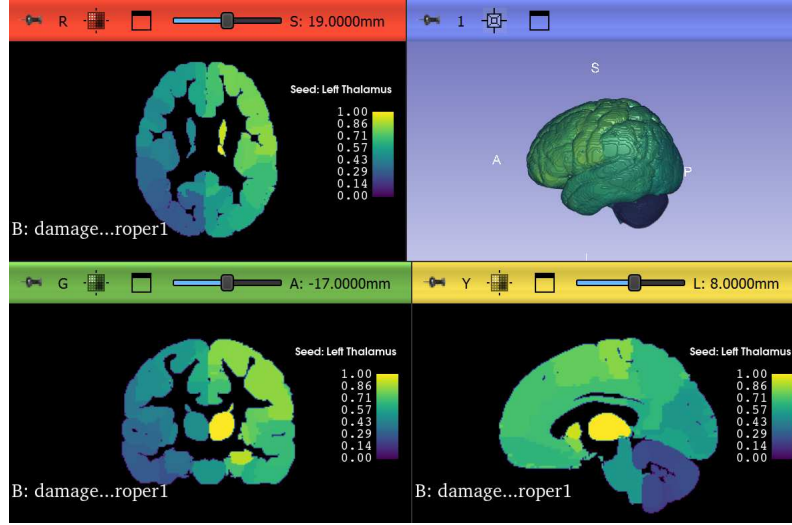


Figure 5.9: Brain damage maps at a single time point ($t = 500$) with *Left_Thalamus* as the seed region. The views displayed include coronal, sagittal, and axial slices, together with the 3d rendering obtained in 3d Slicer. The color bar represents the values using the *viridis* color scale. All regions are included in the map. The Thalamus is highlighted in yellow and is associated to the highest damage value ($=1$), as it corresponds to the selected seed region.

Visual comparison between Thalamus and Cerebellar subregions

The brain maps illustrated in Figure 5.10 and Figure 5.11 show a comparison of the temporal spread of disease originating from two cerebellar subregions (*Left_Crus_I* and *Left_I_IV*) and from the *Left_Thalamus*, in specific brain regions. In particular, Figure 5.10 highlights the **Motor Cortex** and the **Visual Cortex**, whereas Figure 5.11 focuses on the **Brain-Stem**.

Specifically, in Figure 5.10, the regions included in the maps are the following:

- Precentral_L and Paracentral_Lobule_L (*Primary Motor Cortex*), Frontal_Sup_L and Frontal_Mid_L (*Premotor Cortex*), and Supp_Motor_Area_L (*Supplementary Motor Cortex*) for the **Motor Cortex**;
- Calcarine_L, Cuneus_L, Lingual_L, Occipital_Sup_L, Occipital_Mid_L, and Occipital_Inf_L for the **Visual Cortex**.

Two different color-bars, one for the cerebellar subregions and one for the Thalamus, emphasize differences in the predicted spreading. The propagation originating from the Thalamus occurs more rapidly than that from either the *Crus_I* or *I_IV* regions. In both cases, the disease reaches the motor cortex before the visual cortex.

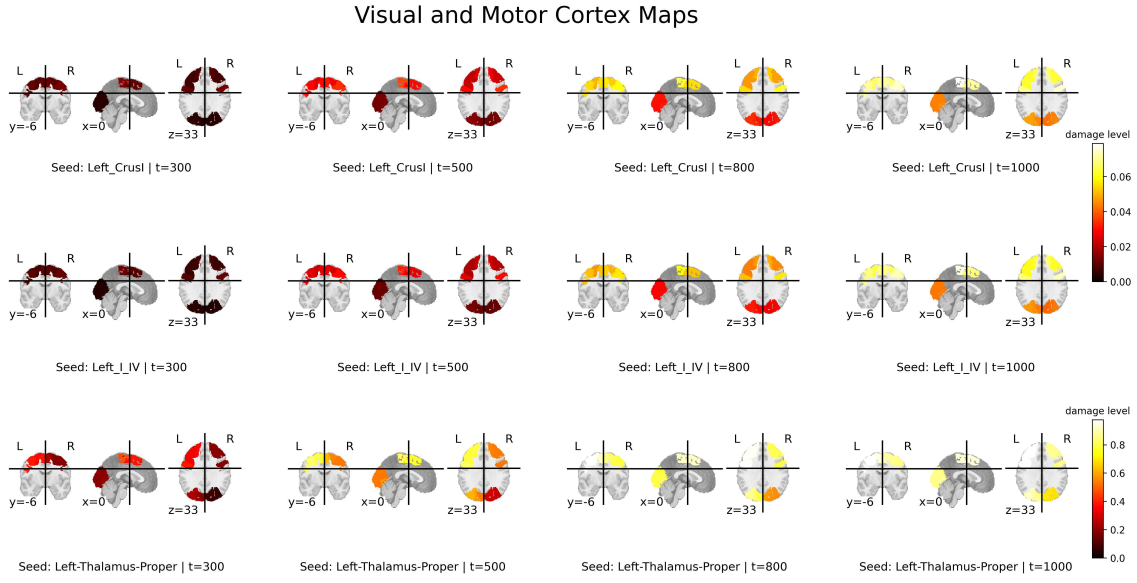


Figure 5.10: Temporal evolution brain-maps (coronal, sagittal and axial view) at four time points ($t = 300, 500, 800$, and 1000) showing the early progression of damage in the motor cortex and visual cortex when the simulation starts from two cerebellar subregions (Left_CrusI and Left_I_IV) and from the Left-Thalamus. Each row corresponds to a specific seed region, named below each map. Two separate color bars are displayed: the top one refers to the cerebellar seeds, while the bottom one refers to the thalamus. They are shown independently because the damage values increase much more rapidly when the seed is placed in the thalamus; using distinct color scales makes this difference clearly visible in the maps. [15]

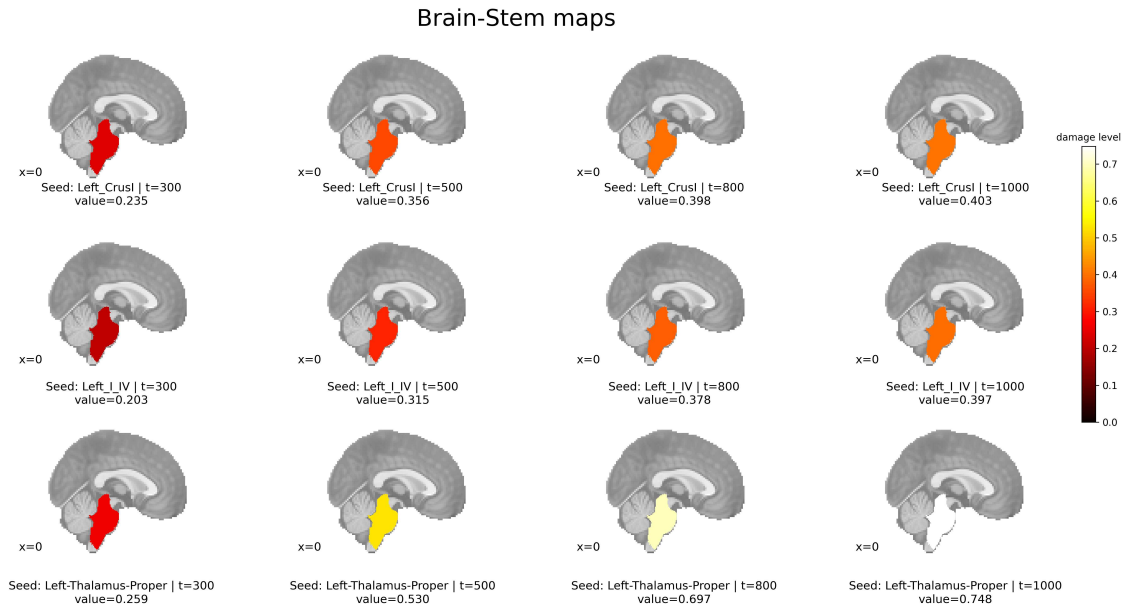


Figure 5.11: Temporal evolution brain-maps (sagittal view) at four time points ($t = 300, 500, 800$, and 1000) showing the early progression of damage in the Brain-Stem when the simulation starts from two cerebellar subregions (Left_CrusI and Left_I_IV) and from the Left-Thalamus. Each row corresponds to a specific seed region, named below each map. The spreading value in the Brain-Stem is also reported. [15]

The temporal evolution maps of the Brain-Stem are illustrated together with the corresponding damage values at each time point in Figure 5.11. The damage levels observed in the Brain-Stem are significantly higher than those reported for the Motor Cortex and the Visual Cortex. Additionally, the spreading originating from the Thalamus appears to spread more rapidly compared to the cerebellar seeds.

5.2.3 Disease spreading maps from Motor and Visual Cortices

Finally, the spread of the disease across the brain has also been investigated when it originates from other regions of interest, such as the Motor Cortex or the Visual Cortex. In particular, the selected seed regions are the **Left Precentral** and the **Left Calcarine**, as mentioned in the previous sections. The figures below show the damage level maps and the corresponding 3d renderings at time point equal to 500 for both cases. The seed regions are highlighted in *yellow*.

Motor Cortex

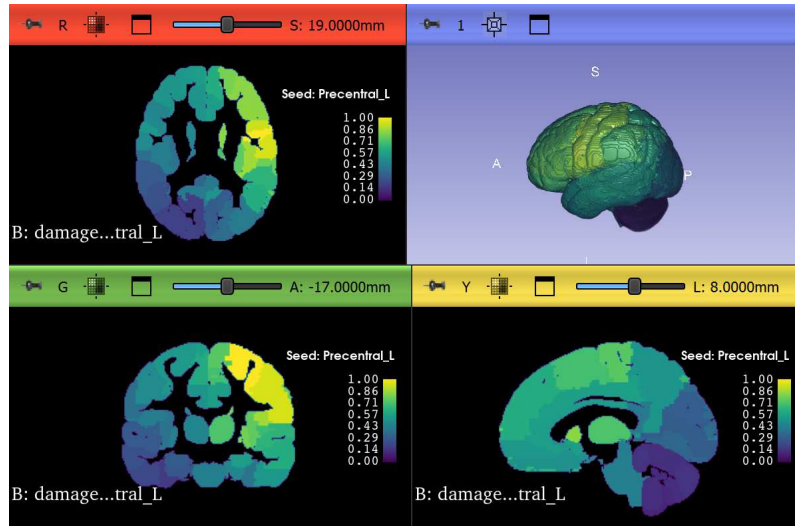


Figure 5.12: Brain damage maps at a single time point ($t = 500$) with Left_Precentral as the seed region. The views displayed include coronal, sagittal, and axial slices, together with the 3d rendering obtained in 3d Slicer. The color bar represents the values using the *viridis* color scale. All regions are included in the map. The Left Precentral is highlighted in yellow and is associated to the highest damage value ($=1$), as it corresponds to the selected seed region.

Visual Cortex

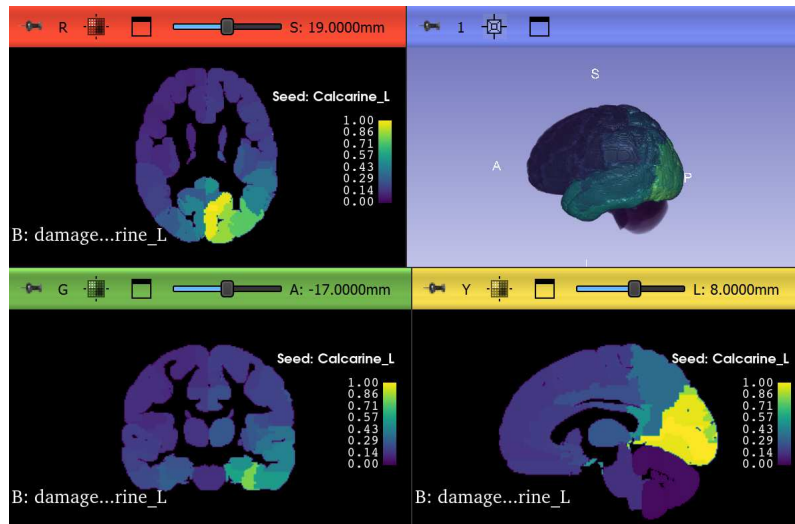


Figure 5.13: Brain damage maps at a single time point ($t = 500$) with Left_Calcarine as the seed region. The views displayed include coronal, sagittal, and axial slices, together with the 3d rendering obtained in 3D Slicer. The color bar represents the values using the *viridis* color scale. All regions are included in the map. The Calcarine is highlighted in yellow and is associated to the highest damage value ($=1$), as it corresponds to the selected seed region.

5.3 Voxel-wise z-score maps

In this section, the results of the voxel-wise analyses are presented. Specifically, the mean z-score maps in MNI-space for each metric are shown, with the **lesion probability map** and the **cortico-cerebellar tracts** overlaid. To facilitate cluster identification, the maps display only voxels below -1 and above 1. The visualization of the maps is performed in **FSLeyes**.

For each map, two separate colorbars were selected to highlight positive and negative values: the *hot* colorbar (ranging from *red* for lower values to *yellow* for higher values) for positive values and the *blue-lightblue* colorbar for negative values (ranging from *blue* for lower values to *lightblue* for higher values). The lesion probability map was overlaid using the *green* colorbar, where increasingly intense green shades indicate a higher number of patients presenting a lesion in the same region (the lesion probability map was computed as the sum of all patients' lesion maps). To ensure a clear visualization, the standard 2 mm MNI template was also included as a background image.

For each metric, a snapshot was manually selected using the software to display a representative slice in the three anatomical views (sagittal, coronal, and axial), highlighting the most relevant z-score voxel clusters.

FA voxel-wise z-score map

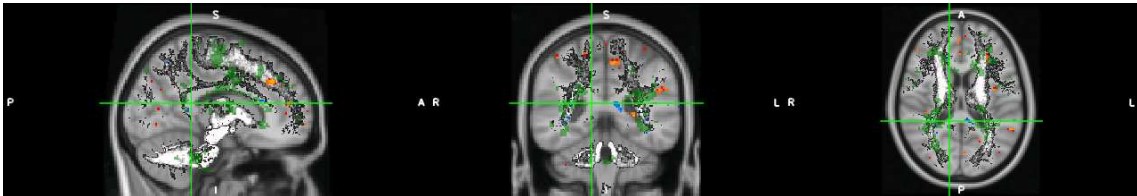


Figure 5.14: Voxel-wise FA z-score map calculated as the mean z-score map across CIS patients, visualized in FSLeyes, with the lesion probability map (in *green*) and the cortico-cerebellar tracts overlaid. The z-score map is thresholded to display values above 1 (in a *red-to-yellow* color scale) and values below -1 (in shades of *blue*). Both positive and negative clusters are present. Some of these clusters overlap with the lesion probability map or are located in adjacent tissue, while others appear as smaller or spurious clusters in different brain regions.

MD voxel-wise z-score map

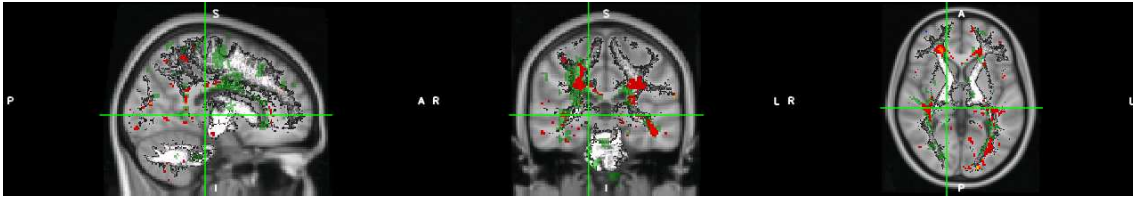


Figure 5.15: Voxel-wise MD z-score map calculated as the mean z-score map across CIS patients, visualized in FSLeyes, with the lesion probability map (in *green*) and the cortico-cerebellar tracts overlaid. Most of the voxel clusters are positive (in *red* shades) and are located in correspondence with the lesions within the WM. Some smaller clusters appears in other parts of the brain.

ODI voxel-wise z-score map

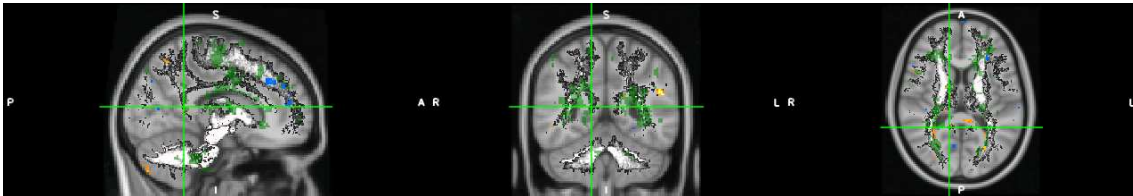


Figure 5.16: Voxel-wise ODI z-score map calculated as the mean z-score map across CIS patients, visualized in FSLeyes, with the lesion probability map (in *green*) and the cortico-cerebellar tracts overlaid. A small number of voxel clusters of limited size also appear in the map, including both positive and negative ones, some of which are located within the WM tracts.

Vfintra voxel-wise z-score map

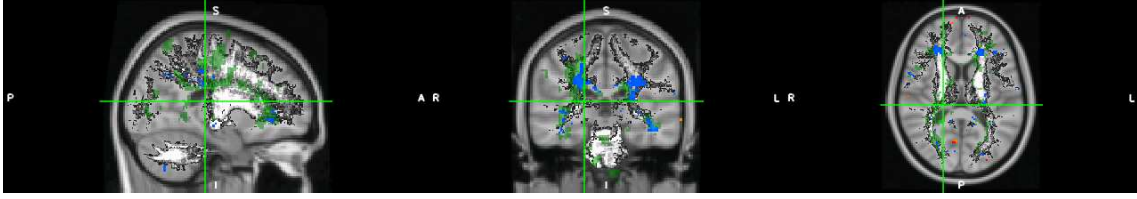


Figure 5.17: Voxel-wise Vfintra z-score map calculated as the mean z-score map across CIS patients, visualized in FSLEyes, with the lesion probability map (in *green*) and the cortico-cerebellar tracts overlaid. Most of the voxel clusters are negative (in *blue* shades) and are located in correspondence with the lesions within the WM. Some smaller clusters appears in other parts of the brain.

Vfiso voxel-wise z-score map

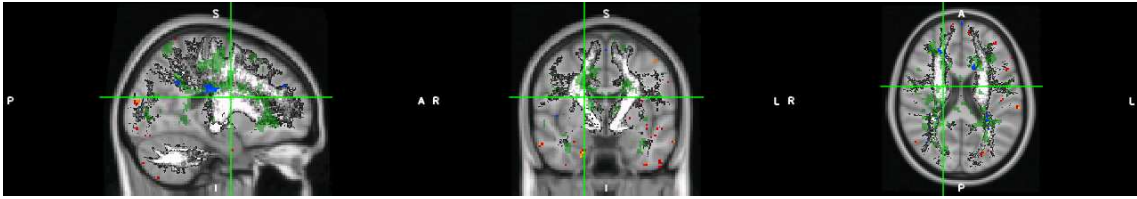


Figure 5.18: Voxel-wise Vfiso z-score map calculated as the mean z-score map across CIS patients, visualized in FSLEyes, with the lesion probability map (in *green*) and the cortico-cerebellar tracts overlaid. In this map, no large or well-defined clusters emerge; instead, only small voxel groups are present, scattered throughout the tissue. These clusters include both positive (in *red* shades) and negative (in *blue* shades) values, some of which lie in correspondence with the lesion probability map in the WM, while others appear as small or spurious clusters distributed in a more random manner.

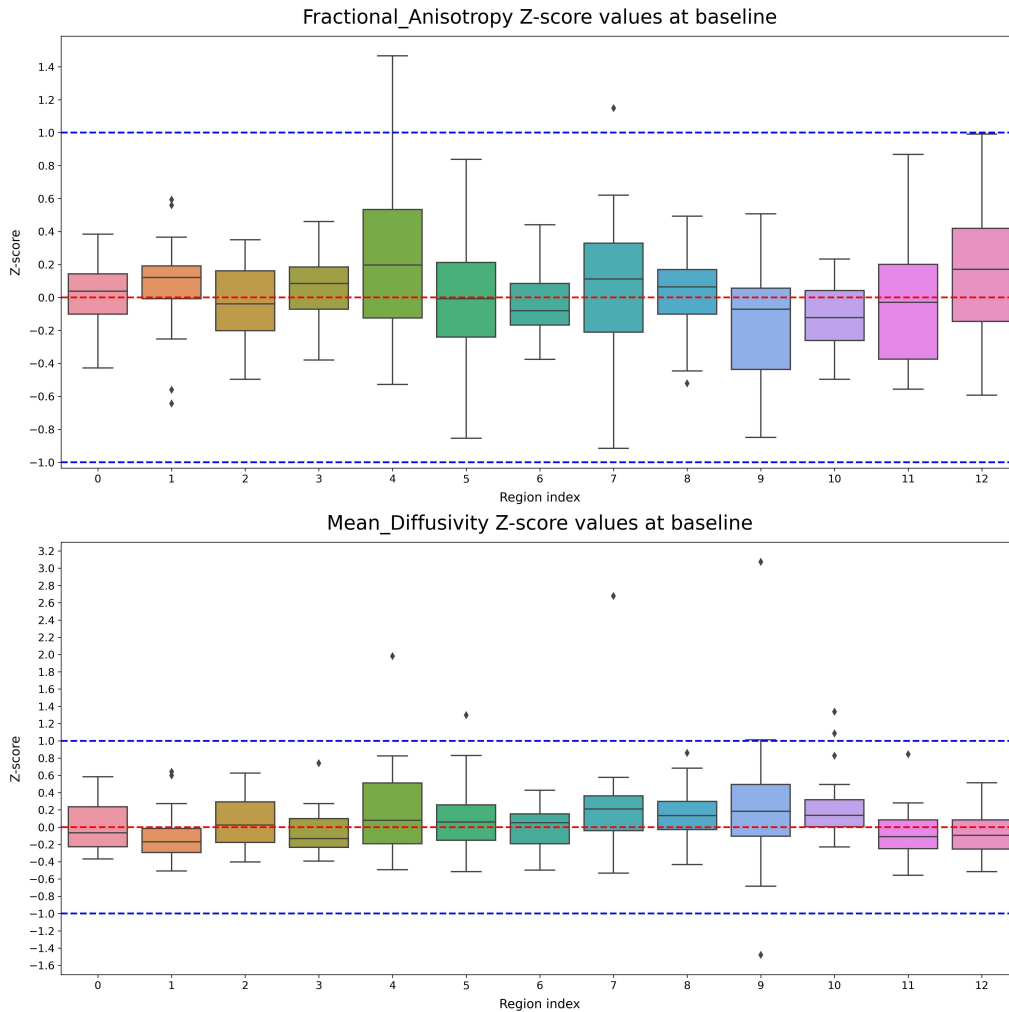
The z-score maps shown above are computed to identify clusters of voxels located within brain regions that could be used to compare the model outputs. Ideally, clusters in GM regions would be expected, since these are the areas considered by the NDM.

However, inspection of the maps reveals that, apart from a few small clusters likely attributable to noise, most voxel clusters overlap with the lesions (highlighted in *green*), predominantly within WM. In particular, the **MD map** shows a predominance of positive z-scores (in *red*), largely overlapping with WM lesions and mainly located along the cortico-cerebellar loop. In contrast, the **Vfintra map** displays mostly negative z-scores (in *blue*). This may corresponds to a general increase in MD and a general decrease in Vfintra in patients compared to controls.

The remaining three metrics exhibit more dispersed voxel patterns with smaller clusters, some overlapping with lesions and others likely reflecting noise. Notably, Vfiso appears to have more randomized values, with both positive and negative z-scores distributed across the brain.

Box-plots for selected ROIs

The z-score box-plots, as explained in the previous chapter (see Section 4.4.4), were generated to assess how the region-wise z-scores differ across patients for each metric. Only a subset of regions from the atlas was included in the box-plot (Figure 5.21) - specifically those relevant to the objectives of this thesis, such as cerebellar subregions (regions [11] and [12]), Thalamus (region [8]), Motor cortex (regions [0]) and [1]), Visual cortex (regions [2] and [3]), and the Brain-Stem (region [6]). The Pallidum (region [7]) has also been added. Other regions, such as Cingulum (region [4]), Heschl (region [5]), Putamen (region [9]), Hippocampus (region [10]), are added in the analyses for comparison.



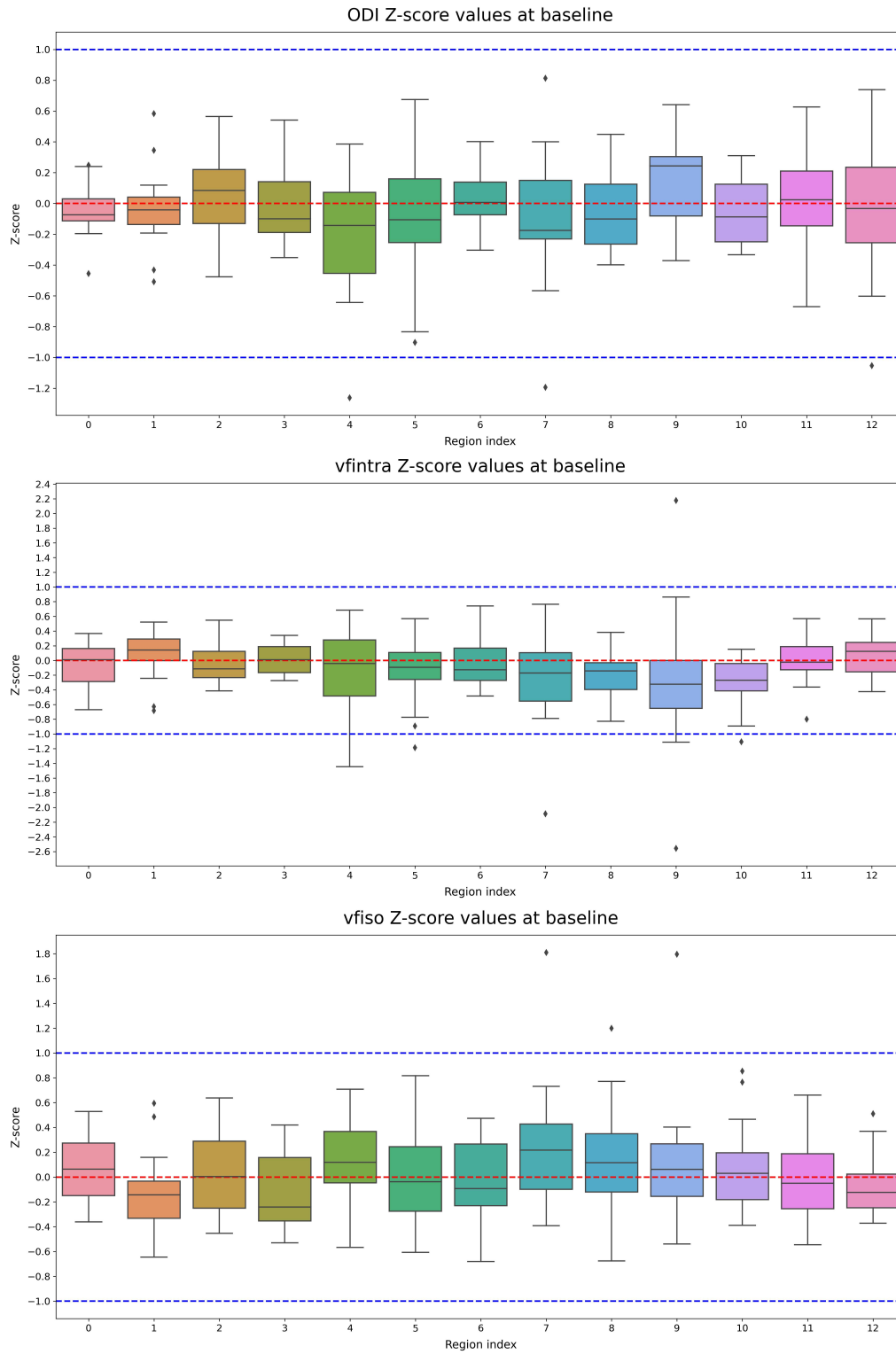


Figure 5.21: Region-wise z-score box-plots, each corresponding to a different quantitative metric indicated in the plot title; the y-axis shows the z-score values, while the x-axis displays one box per GM region, indexed numerically, with the indices corresponding to the following anatomical areas of interest: 0: Precentral_L, 1: Supp_Motor_Area_L, 2: Calcarine_L, 3: Cuneus_L, 4: Cingulum_Ant_L, 5: Heschl_L, 6: Brain-Stem, 7: Left-Pallidum, 8: Left-Thalamus-Proper, 9: Left-Putamen, 10: Left-Hippocampus, 11: Left_I_IV, 12: Left_CrusI; the z-score is computed region-wise by averaging the voxel values within each region for every subject, allowing inspection of the distribution of z-scores across patients for each region. The mean values are approximately around $z = 0$ (*red* line) and are located within the $z = [-1, 1]$ range (*blue* lines). A few outliers are also present.

The region-wise z-score box-plots do not show any statistically significant differences when specific ROIs are examined. Although some variability is visible across subjects, the mean values remain approximately around 0 (within the $[-1, +1]$ range). A few outliers are also present.

5.4 Principal Component Analyses results

This section reports the results of the statistical tests applied to the two groups (patients and controls) for each metric separately. Since the Welch test did not identify any region as significantly different between the groups, only the PCA results are presented in detail. For each metric, two different plots are generated:

- **Scree plot**, showing the cumulative explained variance as a function of the number of PCs. The selection criterion was set at the point where the cumulative variance reached approximately 80%, which determined the number of components retained.
- **Histogram** of the loadings for the first 20 ROIs of the Principal Component that resulted in a p-value < 0.05 .

From the PCA analyses, the only metrics for which at least one Principal Component resulted significantly different between the two groups were **FA**, **ODI** and **Vfintra**. Therefore, results are only reported for these metrics. The p-values taken into consideration for the following analyses are the ones without the correction for multiple comparison. Only the first ten ROIs with the highest positive and negative loadings are considered for comparison with the model outputs.

The p-values are then corrected using the FDR procedure. After applying the correction, none of the metrics showed a PC with an FDR-adjusted p-value below

the significance threshold of 0.05. However, for both **FA** and **Vfintra**, the p-value of the same component that was initially significant before correction increased to approximately 0.065. This suggests that these components may still carry a meaningful signal and that further investigation is warranted - ideally using additional data or a larger dataset.

Fractional Anisotropy

Figure 5.22 shows the selection of PCs based on the scree plot. In this case, the number of components required to explain at least 80% of the total variance (indicated by the *red* line in the figure) is 15. Therefore, the first 15 PCs were selected for the Welch tests. The PCs are colored in *blue*.

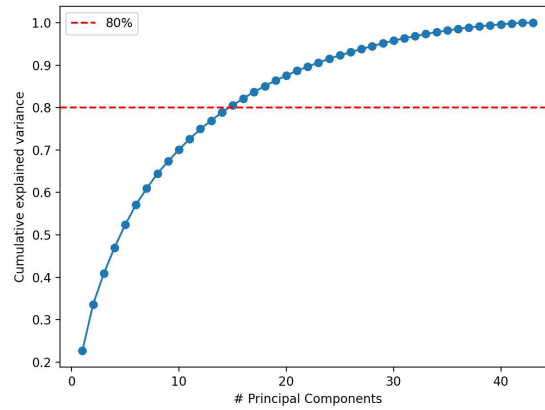


Figure 5.22: Fractional Anisotropy scree plot that can be used to choose the number of PCA components that explain at least 80% of the total variance.

The histogram in Figure 5.23 displays the 20 largest positive and negative loadings for the PC that remained significant after the Welch test and was identified as differing between the two groups (with a p-value of 0.004). Although the interpretation of a PC requires considering the full spatial pattern of all regional contributions, only the top 20 loadings are shown to highlight the regions with the strongest influence on this component. For FA, regions such as the **Supramarginal gyrus**, several **cerebellar subregions**, and **orbital areas** contribute positively, whereas **Temporal regions**, the **Amygdala**, and the **Brain-Stem** contribute negatively.

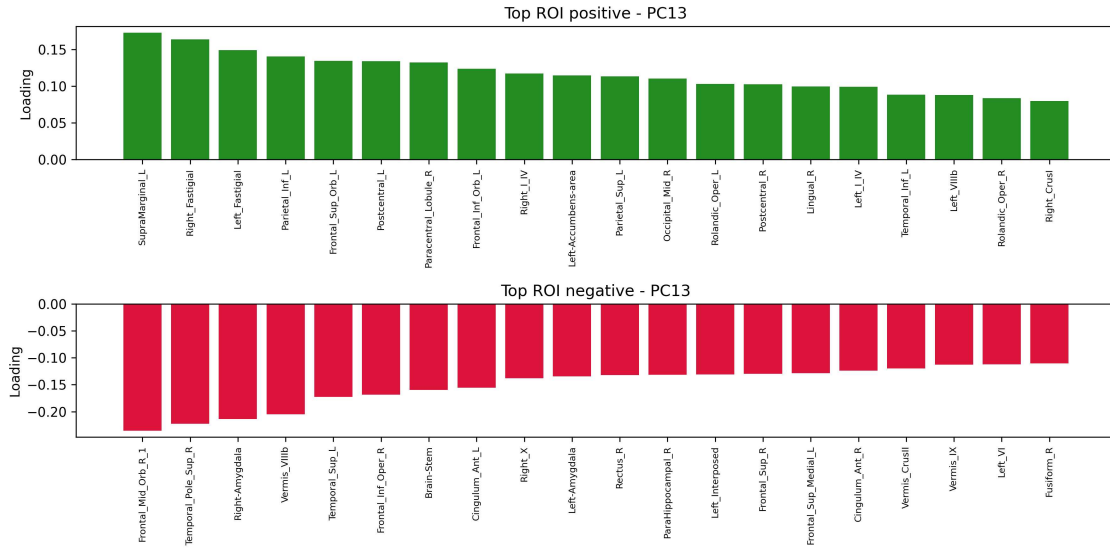


Figure 5.23: FA: Histogram of the top 20 positive and negative loadings for the Principal Component that showed a p-value < 0.05 in the Welch test.

Orientation dispersion index

For ODI, 16 Principal Components were required to explain 80% of the total variance, and after applying the Welch test, only one component resulted significant (p-value = 0.03). As shown in Figure 5.25, regions belonging to the **Frontal** and **Motor cortex** exhibit the strongest positive contributions, whereas cerebellar subregions such as the **Vermis** show the largest negative contributions.

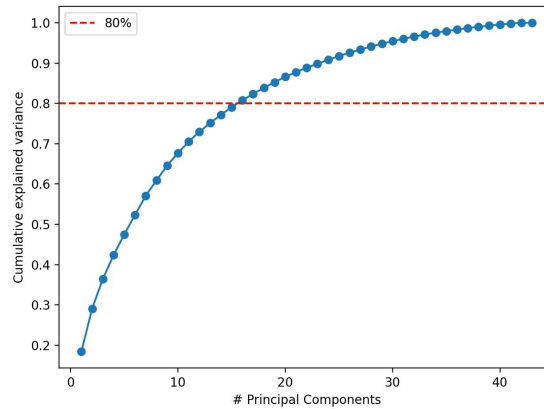


Figure 5.24: ODI scree plot that can be used to choose the number of PCA components that explain at least 80% of the total variance.

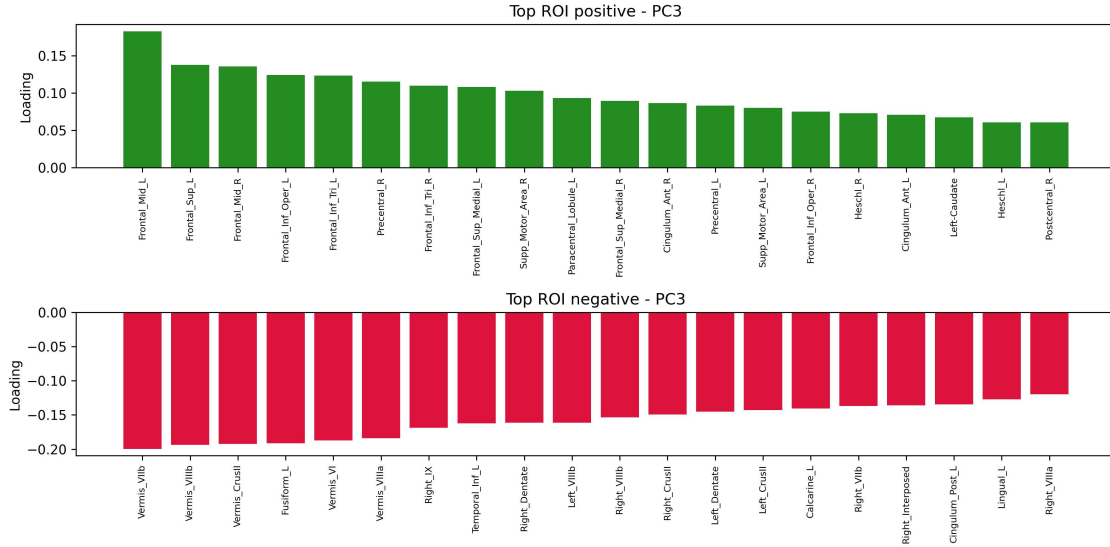


Figure 5.25: ODI: Histogram of the top 20 positive and negative loadings for the Principal Component that showed a p -value < 0.05 in the Welch test.

Intracellular volume fraction

In the case of **Intracellular Volume Fraction**, the number of components required to explain at least 80% of the total variance (indicated by the *red* line in the figure) is 11 (*blue* line).

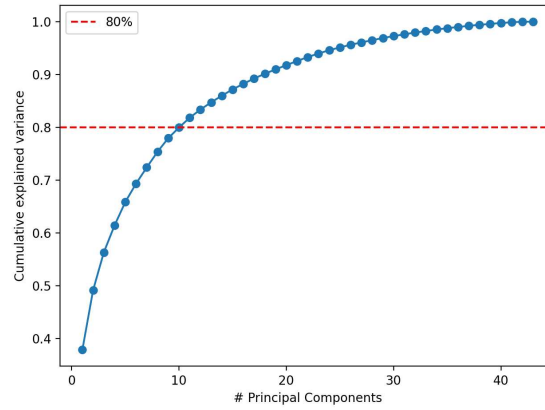


Figure 5.26: Vfintra scree plot that can be used to choose the number of PCA components that explain at least 80% of the total variance.

The Welch test identified one Principal Component as significant (p -value = 0.006), and its loadings are reported below. In this case, the **Thalamus**, several **cerebellar subregions**, and the **Cingulum** show the strongest positive contributions, whereas the **Motor area**, the **Temporal area**, and additional **cerebellar subregions** exhibit the largest negative contributions.

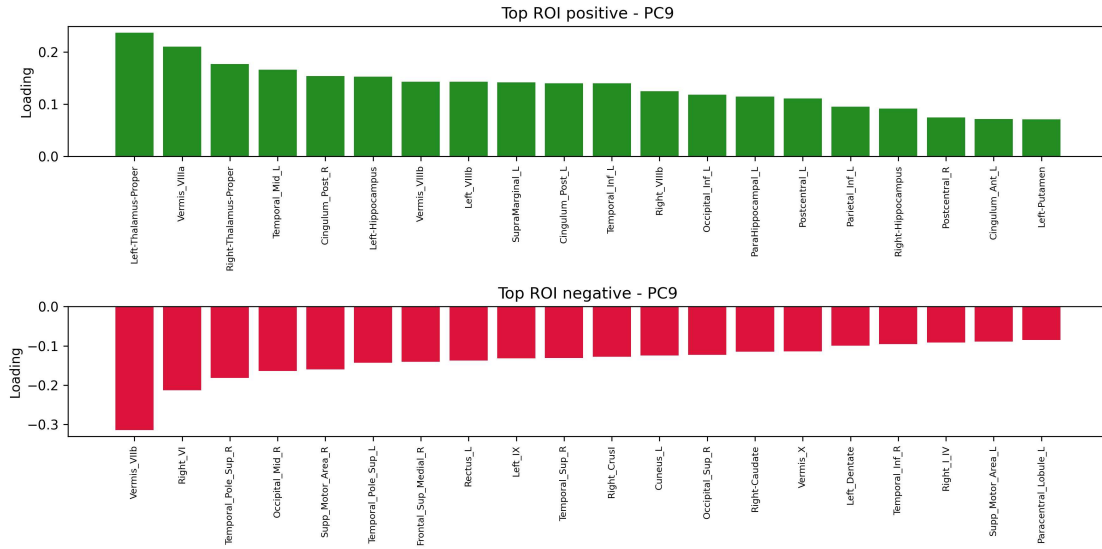


Figure 5.27: Vfintra: Histogram of the top 20 positive and negative loadings for the Principal Component that showed a p-value < 0.05 in the Welch test.

Summary of PCA analyses

For each metric, the ROIs with the highest absolute loadings from the PCs that survived the Welch t-test were selected. These ROIs represent the main contributors to the **multivariate pattern** captured by the significant PCs, that are the dimensions along which patients and healthy controls differ the most in the PCA space. The ROIs identified for each metric were then aggregated into a single table (see Table 5.28) with the aim of highlighting the regions that contribute most consistently to the observed multivariate group differences. It is important to note that this is an *exploratory multivariate analysis*: the loadings do not indicate univariate differences of individual ROIs, but rather their relative contribution to the overall multivariate pattern distinguishing patients from controls.

The regions in the table are grouped according to their anatomical classification: cerebellum, deep gray matter, and the remaining cortical regions included in the AAL parcellation. These GM regions are used as comparison with the model outputs, as shown below.

Cerebellum	Deep Gray Matter	Cortical regions (AAL)
I_IX	Accumbens Area	Brain-Stem
I_IV	Amygdala	Cingulum_Ant
VIIIb	Hippocampus	Cingulum_Post
X	Thalamus	Frontal_Inf_Oper
Vermis VIIb		Frontal_Inf_Orb
Vermis CrusII		Frontal_Inf_Tri
Vermis VIIIb		Frontal_Mid_Orb
Vermis VI		Frontal_Sup_Med
Vermis VIIa		Frontal_Sup_Orb
Vermis VIIa		Fusiform
Fastigial		Occipital_Mid
Dentate		Paracentral_Lobule
		Parietal_Inf
		Postcentral
		Precentral
		Rectus
		Supp_Motor_area
		SupraMarginal
		Temporal_Inf
		Temporal_Mid
		Temporal_Pole_Mid
		Temporal_Pole_Sup
		Temporal_Sup

Figure 5.28: Table summarizing the ten regions with the highest PCA loadings (10 positive and 10 negative) for each of the 3 metrics, collected together into a single table and grouped according to their anatomical classification with different color each - cerebellum (*purple*), deep GM (*red*), and the remaining cortical regions included in the AAL parcellation (*green*). These regions are used as comparison with the model results.

Several cerebellar subregions and the thalamus appear in the list of regions, whereas the Pallidum does not. Other relevant regions include the Brain-stem. Finally, in terms of the cerebral cortex, the simulations focus on the motor and visual cortices: some motor cortical areas, such as the Precentral gyrus and the Supplementary Motor Area, are present, while the primary visual cortex, beside the (Occipital_Mid), is not included.

Comparison with the model outputs

Finally, the results of the simulations are compared with the MS dataset at the CIS stage, using as reference the table shown above, which reports the results of the PCA analyses. The model outputs are grouped in a table (Table 5.29), which shows, for each of the four seed regions tested in the thesis (Crus I, Thalamus, Precentral, and Calcarine), the first ten brain areas affected by the spreading of the disease (time point = 500).

The comparison is carried out using colors: specifically, each region in the list that

also appears in the PCA summary table is color-coded according to its anatomical area (i.e. purple = cerebellum, red = deep GM and green = cortical regions), following the same color scheme used in Table 5.28.

From the table, it is evident that many of the regions most affected when the damage originates from these seeds also appear in the PCA results. Cerebellar regions listed in the Crus I column refer collectively to all cerebellar areas. For this reason, Crus I is the seed with the highest number of affected regions that overlap, including cerebellar regions, deep GM structures, and several cortical areas. The Thalamus also shows a wide range of affected regions; such as deep GM regions and cortical regions, but not the cerebellum. In contrast, the Motor Cortex and Visual Cortex seeds share fewer regions with the PCA results, and these overlaps are restricted to cortical areas only.

Left CrusI	Left Thalamus	Left Precentral (Motor Cortex)	Left Calcarine (Visual Cortex)
Cerebellar regions	Caudate	Postcentral	Lingual
Brain-Stem	Hippocampus	Frontal_Inf_Oper	Cuneus
Pallidum	Postcentral	Rolandic_Oper	Occipital_Sup
ParaHippocampal	Precentral	Caudate	Occipital_Inf
Thalamus	Pallidum	Frontal_Mid	Fusiform
Amygdala	Paracentral Lobule	Pallidum	Occipital_Mid
Caudate	Amygdala	Frontal_Sup	ParaHippocampal
Paracentral Lobule	Frontal_Inf_Orb	Putamen	Temporal_Pole_Mid
Hippocampus	Frontal_Mid_Orb	Parietal_Inf	Temporal_Inf
Precentral	Accumbens Area	SupraMarginal	Cingulum_Post

Figure 5.29: Table summarizing the model outputs for each of the four seed regions analyzed: *Left_CrusI*, *Left_Thalamus*, *Left_Precentral*, and *Left_Calcarine*. Each column of the table reports the first ten regions that become affected by the disease when the pathology originates from the corresponding seed region. The regions are listed in order of damage onset, from the earliest to the tenth affected. The region names highlighted with colors refer to regions that also appeared in the previous table: purple for cerebellar regions, red for deep GM regions, and green for all remaining regions. These color-coded regions correspond to the areas identified in the PCA analyses. Notably, several regions in the present list also appear in the PCA results, meaning that the regions contributing most to the multivariate CIS-HC group differences are also among those predicted to be early affected in the simulated disease spreading model. Cerebellar regions refer collectively to all cerebellar areas. For this reason, the *CrusI* column show a higher number of regions involved in the disease that also appear in the PCA results, in contrast to the others.

Chapter 6

Discussion and conclusions

6.1 Discussion

This thesis originated from the study of a specific class of models, known as **Network Spreading Models**, which have already been used to investigate the propagation of neurodegenerative diseases such as Alzheimer’s and Parkinson’s disease. The aim was to explore whether these models could also be applied to analyze the spread of neuroinflammatory diseases, such as **Multiple Sclerosis**, which is a chronic autoimmune disorder of the CNS. The strength of this thesis lies in the fact that this is the first time that these specific spreading models, and in particular the **Network Diffusion Model**, have been applied to study the propagation of pathology across the brain in MS. These models use structural connectivity as a constraint on pathology diffusion, relying on the brain network derived from the neuronal pathways that link different regions of the brain. Furthermore, since these models can reveal spatial patterns of pathological diffusion across GM regions, this thesis also employs a whole-brain atlas that includes the cerebellum, a region typically excluded from structural connectivity maps due to the complexity of the tracts connecting it to the cerebrum. These tracts contain a high number of crossing fibers, which are not resolved by standard tractography methods, but can be reconstructed using advanced diffusion techniques [18]. This approach enables the reconstruction of cortico-cerebellar pathways and therefore allows the inclusion, and consequently the study, of the **cerebellum**. In fact, the cerebellum, which contains more than 100 billion neurons, has only recently begun to be investigated and is gaining increasing importance, particularly regarding its involvement in MS. The cerebellum is commonly affected in MS, including both early-onset and progressive forms, and cerebellar symptoms are associated with increased disability and poorer prognosis (see Section 1.1.3). Thanks to that, the application of the model allowed the study

of the disease spreading across the entire brain, including the cerebellum.

The main goal of this thesis is to study how the disease spreads across the entire brain when an initially affected region is set as seed, with the aim of validating the model for studying MS. This is done by comparing the spatial pattern produced by the model with a real dataset of MS patients acquired at University College London, UK, in the CIS stage of the disease, hence at the time of the first demyelinating event. Although the progression of MS has been widely investigated in its progressive phases, its origin and the mechanisms underlying disease onset are still unclear.

Given the involvement of the cerebellum in the disease, the focus is the application of the model in order to simulate the spread of the disease when it originates primarily from the cerebellum, thus assuming that the cerebellum acts as a pathological trigger - the main hypothesis explored in this thesis. Therefore, the combination of NDM, which simulate the propagation of pathology across brain regions, the possibility of selecting any region as the starting point - allowing the examination of how the resulting patterns differ - and the validation of model outputs using a real CIS-stage MS dataset, offers a framework to improve the understanding of MS, particularly at its early stages, and to shed light on the role of the cerebellum in the onset of the disease.

NDM simulates the spreading of disease across GM regions, and in the case of MS, this corresponds to the diffusion of inflammation throughout the entire brain via neuronal pathways, either intra or extracellularly. Indeed, neuroactive substances, including pro-inflammatory cytokines, chemokines, and reactive oxygen species, travel within the extracellular space by diffusion, creating a "volume transmission" that allows inflammation to spread beyond the immediate site of injury. Although demyelination occurs in both WM and GM, it has been demonstrated that the extent of GM demyelination is greater (in terms of extension) than that of WM in several brain regions, including the cerebellum (which shows the largest discrepancy between GM and WM demyelination), the thalamus, and the motor cortex [8].

Beside the cerebellum, other brain regions were selected as seed to study the spreading of the disease: the Thalamus, the Motor Cortex, and the Visual Cortex. The Thalamus is another region known to be particularly vulnerable in the early stages of MS. One possible explanation is its proximity to the ventricular CSF, which may expose it to inflammatory and neurotoxic soluble factors originating from the CSF. Furthermore, the Thalamus is directly connected to the cerebellum through the contralateral CTC pathway, which links the cerebellum to the thalamus and supports the hypothesis explored in this thesis. The Motor Cortex and Visual Cortex were also selected because they are regions frequently involved in the symptomatic

presentation of MS. Indeed, MS is often diagnosed following motor impairments such as gait difficulties or visual disturbances such as optic neuritis, which are among the most common initial symptoms reported by patients.

The NDM, as mentioned earlier, takes a healthy structural connectivity matrix as input and simulates the spreading of damage when it originates from a selected seed region within the model. By applying the model to a mean healthy whole-brain structural connectivity matrix (computed as the average of 10 healthy subjects) and selecting the seed regions justified above, the results show that the spatial pattern of GM regions involved in the simulated disease spread depends on the chosen seed. This demonstrates that the model is capable of simulating disease propagation along neuronal pathways in a way that varies according to the location from which the pathology is assumed to originate.

What we expect from the simulation is that brain areas known to be particularly vulnerable in MS show earlier damage compared to other regions. In line with this expectation, when analyzing spatial spreading pattern from a cerebellar seed region (Figure 5.6), the first areas to be affected are other cerebellar subregions that are more strongly connected to the seed. This finding is consistent with previous studies showing, through a data-driven event-based model, that the cerebellum is among the earliest structures to exhibit atrophy in CIS/relapse-onset MS [7]. Immediately after the cerebellum, the Brain-Stem emerges as one of the most affected regions. In the temporal evolution graphs of the model output, the Brain-Stem curve is clearly distinguishable, appearing in the central portion of the plot and isolated from the others (Figure 5.1). This behavior is reasonable, since, even though it has been less frequently examined in imaging studies, it is known to be involved in MS, with a crucial role in disability progression. In addition, it is in close continuity with the spinal cord, whose atrophy is observed in the early stages of the disease, independently of cortical or deep GM involvement. This aligns with evidence showing a significantly reduced cross-sectional area of the Brain-Stem in relapsing-remitting MS patients [3]. Furthermore, it has also been shown that the Brain-Stem exhibits early atrophy together with the Thalamus [7]. This is consistent with our model results: the thalamus appears among the most affected deep GM structures, immediately after the Brain-Stem and before cortical areas. This supports the expected involvement of the cortico-cerebellar loop, where strong connectivity between the cerebellum and Thalamus is well established.

Regarding the cortical regions, the first area to be affected (according to the model) is the motor cortex. This also aligns with our expectations, given the prominent role of the motor-system involvement in MS symptomatology.

From the model output, another deep GM region stands out from the others: the **Pallidum** appears as the most affected structure within this group of regions, showing an earlier high level of damage than the Thalamus. Evidence for the involvement of the Globus Pallidus in MS has been highlighted by a multicenter study reporting increased T1/T2 ratio values -an imaging measure sensitive to demyelination, neurodegeneration, and iron accumulation, each occurring at different stages of the disease and therefore reflecting clinically relevant microstructural abnormalities. An increased T1/T2 ratio has been associated with more severe disability and longer disease duration. In this study, the Thalamus did not show any relevant alteration in T1/T2 ratio values, regardless of clinical phenotype or disability level [16]. As part of the basal ganglia network, the Pallidum is not only involved in gait control but also plays a role in regulating cognitive and emotional functions. Another study focusing on MS-related alterations in functional connectivity reported increased connectivity between the internal part of the Globus Pallidus and the Thalamus, along with reduced connectivity between the external part of the Globus Pallidus and cortical regions. These findings have been interpreted as a compensatory mechanism aimed at counterbalancing the reduced efficiency of communication between the basal ganglia and other functional networks of the brain, a dysfunction likely driven by focal lesions and neurodegeneration [25]. This supports the hypothesis that Globus Pallidus may have a more central role within these pathways, given its contribution to the communication between the direct and indirect circuits and its extensive connections with multiple cortical areas. All these regions also emerged as critical regions in the PCA analysis when differentiating MS-patients from HC as discussed further below.

When the simulation is initiated from the Motor Cortex or the Visual Cortex, the predicted pattern of disease diffusion does not follow the expected hypothesis (Figure 5.12 5.13). Specifically, the Cerebellum, Brain-Stem, and Thalamus do not appear to be affected in the same order of disease involvement. This behavior is coherent with the fact that these cortical regions are more closely associated with the symptomatic manifestations of MS rather than with the underlying pathophysiological mechanisms that drive early structural damage.

When seeding from the thalamus (Figure 5.9), the regions are affected similarly to when seeding the cerebellum, apart from the fact that the cerebellum is affected last. This makes this model unsuitable given that there are several papers showing an initial involvement of the cerebellum. This could be an effect of the connectivity of the thalamus which is asymmetric with the cerebellum, reflecting anatomical truth. The thalamus could therefore be a secondary trigger of the disease if first affected

from the cerebellum.

The second part of the thesis focuses on validating the simulation results using a real MS dataset. The dataset includes 19 patients at the CIS stage and 23 healthy controls. For each subject, five quantitative brain maps were considered for comparison: FA, MD, ODI, V_{fintra} and V_{fiso}. The primary validation approach employed is a **voxel-wise analysis**, performed by computing voxel-wise z-scores for each patient, using the healthy control group as reference for the standardization. This computation was performed to identify clusters of voxels in the z-score maps that differ between the two groups. These clusters serve as a basis for comparison, allowing us to assess potential similarities in the spatial patterns of the model outputs.

Analyzing the z-score maps for each metric, clusters of voxels deviating from the healthy controls can be identified; however, these clusters are predominantly overlaid with WM lesions. Indeed, the maps are visualized together with the lesion probability map and the cortico-cerebellar loop.

For example, the MD z-score map (Figure 5.15) shows clusters with positive values, indicating an overall increase of the metric in patients. Except for a few scattered small clusters, these alterations are mainly located in correspondence with WM lesions and their surrounding tissue. An increase in MD may reflect early signs of increased water diffusivity, potentially resulting from inflammation and demyelinating events. The lesion probability map highlights numerous lesions overlapping with the cortico-cerebellar loop, which supports our hypothesis that cerebellar involvement may trigger the spread of damage along connected tracts and adjacent tissues.

Regarding the NODDI metrics, the V_{fintra} (Figure 5.17) exhibits a general decrease, with clusters of voxels showing negative z-scores that may reflect axonal loss. As in the other maps, these voxels capture microstructural changes within WM, and most are spatially aligned with the lesion probability map. The ODI (Figure 5.16) shows both positive and negative clusters, which is reasonable given that this metric reflects the distribution of fiber orientations. Thus, a more variable distribution may indicate increasing fiber disorganization. In the ODI maps, a few clusters are present outside lesion areas, suggesting subtle tissue changes that might not be directly lesion-driven.

Overall, the maps reflect microstructural alterations, as evidenced by the substantial overlap between abnormal clusters and the lesion probability map. However, they primarily capture WM changes and do not show clear, spatially localized alterations within GM regions. It is important to note that the absence of visible, well-defined GM clusters does not necessarily imply the absence of GM damage. This may be due to several factors, including the limited sample size or the possibility that GM

alterations are predominantly metabolic rather than structural.

Since the z-score map results primarily reflect WM lesions, PCA-based hypothesis tests were performed to assess whether GM regions might show greater differences between patients and controls. From the PCA analyses, for three of the metrics (FA, ODI and Vfintra) a Principal Component was found to differ significantly between patients and healthy controls. Since this investigation is an exploratory multivariate analysis, the uncorrected p-value was considered for the interpretation of the results. For each of these significant PCs, the loadings were analyzed to identify the ROIs with the highest absolute loading values, as these regions contribute the most to the multivariate pattern captured by the component that differentiates the two groups. In other words, these ROIs are those that drive the largest portion of the variation distinguishing patients from controls within the multivariate space defined by the PCA. By grouping together these GM ROIs (see Table 5.28) showing the strongest contributions, the set of regions emerging from PCA are then compared with the spatial patterns obtained from the model output (Table 5.29). In general, the comparison show some similarities between GM regions.

As highlighted in Table 5.28, several **cerebellar subregions** appear in the list of regions of PCA, consistent with the model in which these areas are set as seed regions for the hypothesis. The **Thalamus** is also present: in the model output, it emerges as particularly affected when the simulated damage originates from the cerebellum. In fact, it was the second most damaged deep GM structure after the Pallidum. Notably, the **Pallidum** does not appear in the table, but its damage may be linked to other factors such as iron deposition (not captured by diffusion-based metrics) as mentioned above. Future work should also analyze for example T1/T2 maps to assess other contributions to very early MS pathology.

Other relevant regions include the **Brain-Stem**, which the model identifies as the area most affected after the cerebellum. Finally, regarding the cerebral cortex, the model specifically focused on the motor and visual cortices. When damage originates in the cerebellum, the motor cortex becomes distinguishable as it appears more affected than other cortical areas, including the visual cortex. In the list of regions, some motor cortical areas such as the **Precentral** gyrus appear, while the primary visual cortex, apart from a single region, does not.

From this analysis, similarities emerge between the main regions identified by the simulations and the GM regions highlighted by the PCA. From Table 5.29, several regions in the list also appear in the PCA results, meaning that the regions contributing most to the multivariate CIS-HC group differences are also among those predicted to be early affected in the simulated disease spreading. It is therefore

evident that damage originating from a cerebellar subregion results in a higher number of early-affected regions that also appear in the PCA table, compared with the other seed locations. This suggests that, overall, the model that best captures the pattern observed in the MS dataset at the CIS stage is the simulation that starts from the cerebellum, which support the hypothesis of this thesis.

In the PCA analyses, the p-values were not corrected for multiple comparisons, and when the FDR correction was applied, they were increased to just above the significance threshold. This outcome is reasonable given the exploratory nature of the study, although it indicates that further investigations are needed - potentially by increasing the sample size or by validating the findings using an independent dataset.

In this thesis, PCA was used - specifically focusing on ROI loadings - to investigate multivariate patterns and identify differences between patients and controls. However, PCA is not the only possible approach for analyzing MRI-derived features. An alternative method that could have been employed, and which is often considered easier to interpret, is the Random Forest algorithm. Machine-learning analyses of MRI features have shown that such techniques can help delineate characteristic patterns of tissue alteration and improve the understanding of the underlying disease. Random Forest offers several advantages, including the ability to model nonlinear relationships, robustness to noise and outliers, and minimal assumptions about data distribution. It also provides estimates of feature importance, which can aid interpretability. It performs well even when handling high-dimensional feature spaces, making it suitable for MRI datasets. In a previous study, a Random Forest classifier was trained to discriminate between healthy controls and CIS patients using different types of MRI-derived features, and feature-importance scores were used to identify the most informative metrics for the classification [24]. Interestingly, microstructural features emerged as the strongest contributors to the model's predictions. An alternative approach in the present work could therefore have been to apply a Random Forest classifier to the entire dataset - including all five metrics - to obtain a feature-ranking analysis across the full feature set.

6.2 Considerations and future work

The **Network Diffusion Model** was applied to simulate disease spreading through a diffusive mechanism giving as input a structural connectome and propagating across GM regions. The model performed as expected and revealed distinct spatial patterns depending on the initial seed region manually selected within the model.

Although the model has already been used to study neurodegenerative diseases - also in more advanced and integrated forms - this thesis represents, to the best of current knowledge, the first application of the NDM to investigate MS.

The implementation used here corresponds to the basic formulation of the model; however, some extensions have been proposed in the literature to adapt the framework to the study of neurodegenerative disorders, resulting in more complex versions that incorporate additional components and biological mechanisms. In particular, in a study [23] the NDM was integrated into a combined aggregation-network-diffusion framework (the AND model), which incorporates all the hallmark processes of misfolded-protein progression: the initial seeding and production, subsequent aggregation, and the network-mediated spreading across brain regions. Additionally, this extended model has the capability to predict patients' CSF profiles of phosphorylated misfolded proteins as a function of disease progression, providing a more comprehensive characterization of pathological dynamics.

These developments highlight the potential for future work to integrate similar enhancements in order to better characterize the progression of MS.

In the application of the model to MS, it is important to note that the model simulates propagation along neuronal pathways through a diffusive mechanism, but it does not incorporate the impact of WM lesions or alterations in WM integrity on the diffusion of pathology. In its current form, the model simulates damage only within GM regions and assumes that pathological spread occurs through the brain network via slow and local mechanisms. In addition, it does not account for alternative routes of propagation, such as the spread of pathology through the CSF or the bloodstream. These represent important limitations and future work should consider integrating these mechanisms to make the model more specific and biologically appropriate for MS.

Future work could also explore the use of different types of structural connectivity maps. In this thesis, connectivity was weighted by fiber density, but other approaches that are more directly related to myelin content, such as conduction-velocity-based connectivity, could be considered. In addition, only patients with CIS-stage were included in the present analyses. Future research could benefit from incorporating longitudinal data to investigate how the spatial pattern of pathology evolves over time and to construct dynamic, patient-specific maps of disease progression. Furthermore, other metrics sensitive to myelin content could have been used to analyze the model outputs, such as qT1 or Radial Diffusivity. These qMRI metrics have been associated with demyelination and, in a recent study, were shown to have the highest ability to discriminate between patients and controls [27].

6.3 Conclusions

To summarize and conclude the thesis, the NDM simulates disease spreading through a diffusive mechanism which propagates across GM regions from a seed. The model revealed distinct spatial patterns that depend on the manually selected seed region. The current formulation does not incorporate the impact of WM lesions or tract integrity on pathological propagation, nor does it account for alternative routes such as CSF or blood-borne spread. Future extensions should therefore integrate WM integrity measures to better align simulations with the biological features of MS.

A deeper understanding of MS pathophysiology is also required to determine whether inflammation originates primarily from GM alterations, WM damage, or metabolic dysfunction. Incorporating these mechanisms will be crucial to develop a framework that more accurately reflects the complexity of MS and its early pathological triggers, and could be integrated into more comprehensive platforms such as The Virtual Brain, which can simulate whole-brain dynamics through macroscale modeling of brain activity: NSM could identify networks particularly vulnerable to pathological changes and TVB could model the dynamics of such networks to capture functional characteristics otherwise hidden to the macroscopic changes of the brain. Integrating this network model could help quantify the contribution of specific circuits to large-scale brain dynamics and, in turn, improve our understanding of MS.

Bibliography

- [1] Wilkins Alastair. “Cerebellar dysfunction in multiple sclerosis”. In: *Frontiers in Neurology* 8 (June 2017). DOI: 10.3389/fneur.2017.00312.
- [2] Peter J. Basser and Carlo Pierpaoli. “Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI”. In: *Journal of Magnetic Resonance, Series B* 111.3 (June 1996), pp. 209–219. DOI: 10.1006/jmrb.1996.0086.
- [3] Tomos R. Chivers, Cris S. Constantinescu, and Christopher R. Tench. “MRI-based measurement of Brain Stem cross-sectional area in relapsing-remitting multiple sclerosis”. In: *Journal of Neuroimaging* 25.6 (Apr. 2015), pp. 1002–1006. DOI: 10.1111/jon.12244.
- [4] Sara Collorone et al. “Brain microstructural and metabolic alterations detected in vivo at onset of the first demyelinating event”. In: *Brain* 144.5 (Apr. 2021), pp. 1409–1421. DOI: 10.1093/brain/awab043.
- [5] Michael Deppe et al. “Evidence for early, non-lesional cerebellar damage in patients with multiple sclerosis: DTI measures correlate with disability, atrophy, and disease duration”. In: *Multiple Sclerosis Journal* 22.1 (Apr. 2015), pp. 73–84. DOI: 10.1177/1352458515579439.
- [6] Arman Eshaghi et al. “Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data”. In: *Nature Communications* 12.1 (Apr. 2021). DOI: 10.1038/s41467-021-22265-2.
- [7] Arman Eshaghi et al. “Progression of regional grey matter atrophy in multiple sclerosis”. In: *Brain* 141.6 (May 2018), pp. 1665–1677. DOI: 10.1093/brain/awy088.
- [8] C P Gilmore et al. “Regional variations in the extent and pattern of grey matter demyelination in multiple sclerosis: A comparison between the cerebral cortex, cerebellar cortex, deep grey matter nuclei and the spinal cord”. In: *Journal of Neurology, Neurosurgery & Psychiatry* 80.2 (Feb. 2009), pp. 182–187. DOI: 10.1136/jnnp.2008.148767.

- [9] Maha Haki et al. “Review of Multiple Sclerosis: Epidemiology, Etiology, Pathophysiology, and treatment”. In: *Medicine* 103.8 (Feb. 2024). DOI: 10.1097/md.00000000000037297.
- [10] Anhar Hassan. “Cerebellum”. In: *Mayo Clinic Neurology Board Review*. Ed. by Kelly D. Flemming. 2nd ed. Oxford University Press, 2021, pp. 147–157. ISBN: 9780197512166. DOI: 10.1093/med/9780197512166.003.0018. URL: <https://doi.org/10.1093/med/9780197512166.003.0018>.
- [11] Tiantian He et al. “A coupled-mechanisms modelling framework for neurodegeneration”. In: *Lecture Notes in Computer Science* (2023), pp. 459–469. DOI: 10.1007/978-3-031-43993-3_45.
- [12] Christopher C. Hemond and Rohit Bakshi. “Magnetic Resonance Imaging in multiple sclerosis”. In: *Cold Spring Harbor Perspectives in Medicine* 8.5 (Jan. 2018). DOI: 10.1101/cshperspect.a028969.
- [13] Tanja Kuhlmann et al. “Multiple sclerosis progression: Time for a new mechanism-driven framework”. In: *The Lancet Neurology* 22.1 (Jan. 2023), pp. 78–88. DOI: 10.1016/s1474-4422(22)00289-7.
- [14] Fred D. Lublin et al. “Defining the clinical course of multiple sclerosis”. In: *Neurology* 83.3 (July 2014), pp. 278–286. DOI: 10.1212/wnl.0000000000000560.
- [15] Arianna Mancini et al. “Investigating Network Spreading Models of disease spreading from a whole-brain healthy structural connectome: the role of the cerebellum”. In: *School of Brain Cells and Circuits “Camillo Golgi”, Ettore Majorana Foundation and Centre for Scientific Culture , Erice (Italy)* (2025).
- [16] Monica Margoni et al. “Exploring in-vivo brain microstructural damage in multiple sclerosis through T1W/T2W-ratio mapping: A multicenter study (P14-4.009)”. In: *Neurology* 98.18 supplement (May 2022). DOI: 10.1212/wnl.98.18supplement.1297.
- [17] D.W. McRobbie. “Fundamentals of MR Imaging”. In: *Comprehensive Biomedical Physics*. Ed. by Anders Brahme. Vol. 3. Elsevier, 2014, pp. 1–19.
- [18] Fulvia Palesi et al. “Contralateral cerebello-thalamo-cortical pathways with prominent involvement of associative areas in humans in vivo”. In: *Brain Structure and Function* 220.6 (Aug. 2014), pp. 3369–3384. DOI: 10.1007/s00429-014-0861-2.

-
- [19] Fulvia Palesi et al. “Contralateral cortico-ponto-cerebellar pathways reconstruction in humans in vivo: Implications for reciprocal cerebro-cerebellar structural connectivity in motor and non-motor areas”. In: *Scientific Reports* 7.1 (Oct. 2017). DOI: 10.1038/s41598-017-13079-8.
 - [20] Fulvia Palesi et al. “The importance of cerebellar connectivity on simulated brain dynamics”. In: *Frontiers in Cellular Neuroscience* 14 (July 2020). DOI: 10.3389/fncel.2020.00240.
 - [21] Katrin Parmar et al. “The role of the cerebellum in multiple sclerosis—150 years after Charcot”. In: *Neuroscience & Biobehavioral Reviews* 89 (June 2018), pp. 85–98. DOI: 10.1016/j.neubiorev.2018.02.012.
 - [22] Ashish Raj, Amy Kuceyeski, and Michael Weiner. “A network diffusion model of disease progression in dementia”. In: *Neuron* 73.6 (Mar. 2012), pp. 1204–1215. DOI: 10.1016/j.neuron.2011.12.040.
 - [23] Ashish Raj et al. “Combined model of aggregation and network diffusion recapitulates alzheimer’s regional tau-positron emission tomography”. In: *Brain Connectivity* 11.8 (Oct. 2021), pp. 624–638. DOI: 10.1089/brain.2020.0841.
 - [24] Antonio Ricciardi et al. “Patterns of inflammation, microstructural alterations, and sodium accumulation define multiple sclerosis subtypes after 15 years from onset”. In: *Frontiers in Neuroinformatics* 17 (Mar. 2023). DOI: 10.3389/fninf.2023.1060511.
 - [25] Francesco Romanò et al. “Structural and functional correlates of disability, motor and cognitive performances in multiple sclerosis: Focus on the Globus pallidus”. In: *Multiple Sclerosis and Related Disorders* 86 (June 2024), p. 105576. DOI: 10.1016/j.msard.2024.105576.
 - [26] Elinor Thompson et al. “Demonstration of an open-source toolbox for network spreading models: Regional amyloid burden promotes tau production in alzheimer’s disease”. In: *Alzheimer’s & Dementia* 20.S9 (Dec. 2024). DOI: 10.1002/alz.093791.
 - [27] Carmen Tur et al. “What contributes to disability in Progressive MS? A brain and cervical cord-matched quantitative MRI study”. In: *Multiple Sclerosis Journal* 30.4–5 (Feb. 2024), pp. 516–534. DOI: 10.1177/13524585241229969.
 - [28] David C. Van Essen et al. “The Wu-minn human connectome project: An overview”. In: *NeuroImage* 80 (Oct. 2013), pp. 62–79. DOI: 10.1016/j.neuroimage.2013.05.041.

- [29] Johannes Weickenmeier et al. “A physics-based model explains the prion-like features of neurodegeneration in alzheimer’s disease, parkinson’s disease, and amyotrophic lateral sclerosis”. In: *Journal of the Mechanics and Physics of Solids* 124 (Mar. 2019), pp. 264–281. DOI: 10.1016/j.jmps.2018.10.013.
- [30] Alexandra Young et al. “Data-driven modelling of neurodegenerative disease progression: Thinking outside the Black Box”. In: *Nature Reviews Neuroscience* 25.2 (Jan. 2024), pp. 111–130. DOI: 10.1038/s41583-023-00779-6.
- [31] Hui Zhang et al. “Noddi: Practical in vivo neurite orientation dispersion and density imaging of the human brain”. In: *NeuroImage* 61.4 (July 2012), pp. 1000–1016. DOI: 10.1016/j.neuroimage.2012.03.072.