Open University of Catalonia (UOC) Master's Degree in Data Science

MASTER'S THESIS preliminary draft

Area: Medical Area (TFM-Med)

From brain disconnection to atrophy

Evaluating multi-modal brain network connectivity and regional grey matter volumes in Machine Learning-based classification tasks of Multiple Sclerosis Disease

Author: Salvador Sanchis Beneseit

Tutor: Eloy Martínez de las Heras

Professor: Laia Subirats

Barcelona, January 7, 2024

Credits/Copyright

FINAL PROJECT RECORD

Title of the project:	From brain disconnection to atrophy: Evaluating multimodal brain network connectivity and regional grey matter volumes in Machine Learning-based classification tasks of Multiple Sclerosis Disease
Author's name:	Salvador Sanchis Beneseit
Collaborating teacher's name:	Eloy Martínez de las Heras
PRA's name:	Laia Subirats
Delivery date (mm/yyyy):	01/2024
Degree or program:	Master's Degree in Data Science
Final Project area:	Medical Area (TFM-Med)
Language of the project:	English
Keywords	Brain Network Connectivity, Brain Atrophy, SVM Classification

Dedication/Quote

To my family, Georgia, Mirtó and Aris, for their support, patience and understanding during all the hours spent behind a closed door.

Acknowledgements

To my tutor Eloy, who set the bar real high, but offered invaluable support and guidance.

Abstract

Departing from sets of brain connectivity data obtained from 147 multiple sclerosis (MS) patients and 18 healthy volunteer control subjects, we first describe the demographic characteristics of the sample, and then we implement a pre-processing pipeline in order to, subsequently, determine statistically significant differences between patients and controls in three different types of measures: 1) Structural and functional connectivity measures in the form of adjacency matrices of 76x76 nodes (76 brain regions), 2) Graph connectivity measures, both global and nodal, derived from the same adjacency matrices, and 3) Individual brain volume measures of the same 76 brain areas. For each of these types of data, the subset of information that has shown statistically significant differences between groups (patients vs. controls) is fed to a classification task using a classical Support Vector Machine algorithm (SVM), thus obtaining different levels of precision in the classification task for each type of data. Results will show that direct classification using structural FA-weighted connectivity (Fractional Anisotropy) yields the best precision. However, some graph connectivity measures derived from the same data attain high levels of precision as well, especially when these are centered on specific brain regions that seem to be particularly affected by the pathological processes of the disease. Finally, brain volumes, a measure that portrays the process of brain atrophy, affords the possibility to discern patients from controls too, albeit with slightly less precision compared to brain connectivity. We will discuss the potential and eventual shortcomings of this approach to draw an integrated image of the various processes involved in the pathophysiogical course of multiple sclerosis.

Keywords: brain connectivity, graph theory, brain atrophy, fractional anisotropy (FA), support vector machines (SVM), data augmentation.

Resum

Partint d'un conjunt de dades de connectivitat cerebral obtingudes de 147 pacients amb esclerosi múltiple i 18 subjectes de control sans, aquest projecte descriu primerament les característiques demogràfiques de les dades i, a continuació, implementa una seqüència de pre-processament per a poder determinar posteriorment diferències estadísticament significatives entre pacients i controls en tres tipus diferents de mesures: 1) Mesures de connectivitat estructural i funcional en forma de matrius d'adjacència de 76 nodes (regions cerebrals), 2) Mesures de grafs, tant globals com nodals, derivades de les mateixes matrius d'adjacència, i 3) Mesures individuals de volum cerebral de les mateixes 76 àrees cerebrals. Per a cada tipus de dades, el subconjunt d'informació que ha mostrat diferències estadísticament significatives entre els grups (pacients vs. controls) es sotmet a una tasca de classificació utilitzant un algorisme clàssic de Màquines de Suport Vectorial (SVM), obtenint així diferents nivells de precisió en la tasca de classificació per a cada tipus de dades. Els resultats mostraran que la classificació directa utilitzant la connectivitat ponderada de FA estructural (Anisotropia Fraccional) ofereix la millor precisió. No obstant això, algunes mesures de connectivitat gràfica derivades de les mateixes dades aconsegueixen també nivells alts de precisió, especialment quan es centren en les regions cerebrals que semblen ser particularment afectades pels processos patològics de la malaltia. Finalment, els volums cerebrals, una mesura que fa patent el procés d'atrofia cerebral, també permet la possibilitat de discernir entre pacients i controls, tot i que amb una precisió lleugerament menor en comparació amb la connectivitat cerebral. Discutirem el potencial i les eventuals limitacions d'aquest enfocament per dibuixar una imatge integrada dels diversos processos involucrats en el curs patofisiològic de l'esclerosi múltiple.

Keywords: connectivitat cerebral, teoria de grafs, atròfia cerebral, anisotropia fraccional (FA), màquines de suport vectorial (SVM), augmentació de dades.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Context and motivation

Multiple sclerosis (MS) affects over 2 million people worldwide and has no known cure [20]. Its most common type is relapsing-remitting MS (RRMS), characterized by sustained cognitive and motor deterioration through periods of relapse and recovery. Other courses of the disease include secondary progressive MS (SPMS), often a continuation of RRMS wherein deterioration progresses independently of remission or relapse, and primary progressive MS (PPMS), in cases where continued deterioration is present from the very onset of the disease [19]. Cognitive impairment affects up to 65% of MS patients and includes deficits in attention, memory, executive functions and information processing [3]. On the other hand, the disease also affects sensorimotor processes [13], resulting in impaired motor control and balance, as well as weakness and spasticity. These cognitive and motor impairments have been correlated to a number of well-described processes typical of the course of the disease, which include demyelination, focal White Matter (WM) lesions, neuroinflammatory processes and both general and regional atrophy. Graph Theory based analysis techniques have unveiled specific patterns of structural and functional network connectivity that correlate with visible symptoms of cognitive and motor impairment. However, the relations between all these (physical and network connectivity) factors and the resulting symptoms remains complex and elusive, as impairment can appear and evolve as a result of different combinations of the aforementioned processes, and is mediated by demographic variables that result in notable interpersonal variety in the ongoing course of the disease.

Despite of the complex interrelation of processes underlying MS, there is general agreement over the fact that atrophy of GM is associated with the most severe cognitive and motor impairment symptoms seen in later stages of the disease, even if the temporal patterns of

atrophy spreading are still poorly understood [9].

Studies have posed the possibility that focal lesions and demyelination in WM structures might not be the sole reason for the unfolding of atrophy and that, instead, there might be "hidden processes at a microstructural level" that underlie and precede visible atrophy [8]. Graph Theory based analysis of brain connectivity patterns proves to be an ideal tool to analyze these "less visible" processes of brain disfunction. In recent years, an array of statistical measures of both structural and functional brain connectivity have been developed that unveil patterns of adaptation, resilience and eventual breakdown of adaptive neural communication between different parts of the brain. These patterns are obtained from measures of different kinds, such as Fractional Anisotropy (FA), for structural connectivity, and functional Magnetic Resonance (fMRI), both being processed in the form of adjacency matrices, like the one we show in fig.1.1, that map the presence and intensity of connections (edges) between different areas (nodes) of the brain. One way in which these connectivity measures are revealing of underlying patterns is, for instance, the discovery of resilience and reorganization mechanisms of certain areas of the brain that counter the advance of the pathological processes in MS, but that reach a certain threshold, referred to as *network collapse*, from which point on the symptoms worsen and accelerate[19].

Despite of all these advances in our understanding of the workings and adaptations of brain connectivity, the temporal and mechanic relation between WM pathology and GM atrophy is still poorly understood[28]. Our initial general objective for this project was to use the demographic, functional/structural brain connectivity, and regional brain volume data available from a sample of multiple sclerosis patients in order to explore possible correlations between connectivity deterioration and atrophy processes through the course of the disease, bearing in mind that our data was not longitudinal, but that we could compare patients with healthy control volunteers across a diversity of variables at a single point in time.

Finally, as we will show in the following chapters, the most time and effort invested in this project have been dedicated to reproducing analysis methods of previous research on brain connectivity in multiple sclerosis[4, 19] with the data available to us. This involved preprocessing the data, testing for statistically significant differences between groups, and implementing machine learning classification tasks on various subsets of the data, making along the way all the necessary adaptations that our specific sample required, which involved selecting different graph measures and different brain regions compared to the ones used in the aforementioned studies. Once this was accomplished, we included brain volume measures

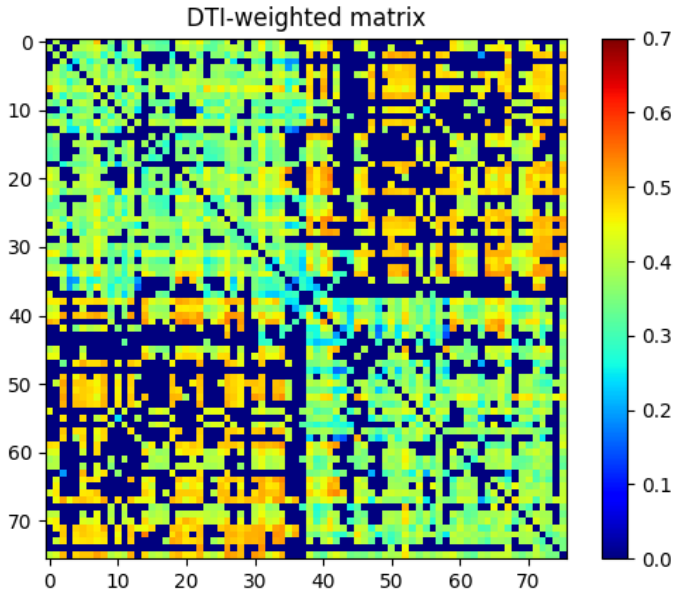

Figure 1.1: Adjacency matrix.

in the same preprocess-test-classification pipeline in order to explore the correlation between brain connectivity deterioration and brain atrophy.

1.2 Goals

1. To understand the relevance of brain connectivity data in relation to Multiple Sclerosis.
2. To deepen our knowledge on the application of graph theory to the study of brain connecivity, and its potential to characterize pathological brain states.
3. To implement the necessary preprocessing pipelines on the available raw data that will enable and support a correct analysis.
4. To define and correctly implement the necessary statistical tests on the data that can validate and filter which amount of data is used for classification tasks, and in which way.
5. To learn to correctly implement Machine Learning models to classify real subjects starting from different types of brain imaging data.
6. To test the appropriateness and efectiveness of different types of data (FA, functional, graph derived and volumetric) in Machine Learning classification tasks.

7. To use our data analysis to assess the relation between brain connectivity, brain atrophy and motor/cognitive impairment.

8. To discuss the relevance of the results in regard to the specific challenges in Multiple Sclerosis research, considering previous findings and future challenges.

1.3 Sustainability, diversity and ethical/social challenges

1.3.1 Sustainability

This project has been design and implemented with the most basic and strictly necessary research resources. Data storage as well as the implementation of machine learning classification tasks have taken place on a local machine (linux OS laptop), with no use of remote processing or access to data centers. Thus, there has been no traceable resource consumption, above the regular use of a laptop, that may conflict with any of the following Sustainable Development Goals (SDG): SDG 7 - Affordable and clean energy; SDG 9 - Industry, innovation, and infrastructure; SDG 11 - Sustainable cities and communities; SDG 12 - Responsible consumption and production; SDG 13 - Climate action; SDG 14 - Life below water; SDG 15 - Life on land. In fact, the use of a classical machine learning algorithm, such as SVM, for the classification task, in comparison to more sophisticated algorithms that require access to remote GPU resources but that are not necessarily better suited to the nature of the data in this project and therefore do not necessarily yield better results, may be considered an ecologically more responsible choice.

1.3.2 Ethical Behaviour and social responsibility

This project analyzes data from multiple sclerosis patients and healthy control subjects. The data is obtained from the ImagineEM research group at the Hospital Clinic Barcelona, under a data confidentiality agreement signed with the tutor of this thesis, and member of the ImagineEM group, Eloy Martínez de las Heras. Data transfer consists of adjacency matrices derived from pseudonymised images, and demographic information devoid of personal details such as names, income, residence or ethnicity. Further than the scientific and medical advance that this project may contribute to, and which could hopefully benefit the condition of future patients, we consider that the outcome of this project has no impact on ethical or social aspects, nor on any law or regulation that we have knowledge of. We also do not consider it affects in any way any of the Sustainable Development Goals (SDG) that involve ethical and social responsibility: SDG 1 - No poverty; SDG 2 - Zero hunger; SDG 6 - Clean water and sanitation;

SDG 8 - Decent work and economic growth; SDG 16 - Peace, justice, and strong institutions.

1.3.3 Diversity, gender and human rights

The data analyzed does not contain any information concerning diversity of race, religion, ethnicity, sexual orientation, functional, or of ideology. The gender variable was included in the demographic study of our sample, but this pursues the sole objective of describing the sample, and does not affect our results or their interpretation. In the preprocessing phase, the gender variable was considered again during the data harmonization process, together with the age variable, but finally these variables did not influence the harmonization results and thus were not included as parameters. Therefore, we consider that there was no conflict in terms of the Sustainable Development Goals (SDG) concerned with diversity, gender and human rights: SDG 5 - Gender equality; SDG 10 - Reduced inequalities.

1.4 Approach and methodology

After the initial background research and the setting of the preliminary objectives, the core of the project involved, first, the state of the art research, and second, the implementation.

The state of the art research consisted of an extensive review of specialized literature on the subject, departing from a handful of articles that set the immediate precedent to our project, with the tutor of this thesis as a co-author[4, 19, 26]. We find it important to mention that the state of the art review was written previous to having access to our data, and before being faced with the technical challenge of the preprocessing-testing-classification work. Obviously, our goals were adapted once we realized which operations our data afforded us, and when we discovered necessary procedures we couldn't anticipate. However, as we nevertheless tried to maintain our initial objectives throughout the whole project, we have found it important to include in this document the state of the art review in its entirety, as it was originally written, only with minor adaptations to facilitate its contextual reading.

In the implementation phase we obtained our data, which included the following:

1. Demographic information of the sample of subjects, including age, gender, multiple sclerosis disease subtype (Relapsing-remitting/Primary progressive/Secondary progressive), disease duration, and disability score.

2. Adjacency matrices for all subjects, representing their structural brain connectivity network as a measure of weighted Fractional Anisotropy (FA), built from a diffusion tensor imaging

model (DTI). These were 76x76 symmetric matrices, representing 76 brain areas of interest (referred to as nodes) that include cortical as well as sub-cortical areas of the brain.

3. Adjacency matrices representing functional brain connectivity, obtained from resting-state functional magnetic resonance (rs-fMRI). These had the same dimensions as the structural connectivity matrices.

4. Adjacency matrices representing morphological gray matter network (finally not used in our project)

5. A list of the names of the 76 nodes (brain regions of interest).

6. A list of regional brain volumes of the 76 regions of interest for all the patients and control subjects. On a longitudinal study, these would be used to quantify brain atrophy.

After an initial demographic study of the sample, the implementation phase continued with the preprocessing of the data (normalization, outlier detection, harmonization and thresholding). Once the data was preprocessed, different graph measures were derived from it, both at a global (whole brain) and at a nodal level (specific brain regions). Compared to the structural and functional connectivity matrices, these measures apply graph theory based network connectivity principles to portray different features of brain connectivity, such as clustering or centrality tendencies, and can themselves be used to describe or classify brain connectivity profiles.

At this point, we had three different types of data available for analysis: 1) structural and functional adjacency matrices, 2) global and nodal graph connectivity measures, and 3) regional brain volumes. Our next step was to check for statistically significant differences between groups for each of these 3 types of measures. For this task we used Student's t-test of independence (or U Mann-Whitney as a non-parametric alternative) to determine on which measures patients and controls differed from each other. We also checked for statistical differences between different groups of patients according to their disease type, but we couldn't find any.

Next, we used only those measures where we obtained a statistically significant difference to implement classification tasks using a support vector machine algorithm. This way we were able to compare the precision of machine learning classification tasks based on structural connectivity, global/nodal graph measures and brain volumes, respectively. Having completed the classification tasks, we are ready to discuss the results in relation to the objectives set in the state of the art research phase.

All the steps of the implementation phase, which have been summarized here, will be described in further detail in the 3rd chapter of this document.

1.5 Summary of the outputs of the project

The output of this project consists of three elements:

1. This document.
2. A video presentation (in process)(link...)
3. A GitHub repository containing the totality of the code written for each phase of the implementation of the project. This repository includes a ReadMe page with instructions on how to navigate the material included in it. For confidentiality reasons, the patients' demographic information and the actual data have been removed from the public version of the repository, leaving visible the implementation code, the results of the statistic tests and the machine learning classification tasks, and the visualizations appearing in the current document. The repository can be found at:

   https://github.com/ssanchisb/tfm

   The repository includes the following documents:

   (a) Demographic study of the sample (R notebook).

   (b) Data loading and preprocessing (set of Python scripts)

   (c) Statistic tests, graph measures, machine learning classification tasks and visualizations (Python notebooks)

1.6 Brief description of the remaining chapters of the report

Chapter 2: State of the art. In this chapter we present the state of the art review, as it was written before the implementation phase and before having access to our data. While the main objectives and conclusions presented in it have been left untouched, minor syntactic corrections have been made to provide for a contextualized reading within the overall contents of this thesis.

Chapter 3: Implementation. Description of all the processes involved in the analysis of the data, presented separately in two sections: methods and results.

Chapter 4: Discussion. Interpretation of the results, followed by a presentation of attained objectives and limitations of the study.

Chapter 2

State of the art

2.1 Introduction

As mentioned in section 1.1, the main pathological phenomena currently under scrutiny in the study of Multiple Sclerosis(MS) are White Matter(WM) lesions, Grey Matter(GM) atrophy, cognitive impairment, structural and functional connectivity, and heterogeneity of disease courses due to individual factors (age, etc.). We have already described how these different elements are all present in MS, but how their (temporal/causal) relations are complex and poorly understood. Within this complexity, we are particularly interested in the central role of atrophy because of its irreversibility, as a sign of severe impairment, and as a potential predecessor of all other factors. Our proposal for a main research objective is to study the correlations between atrophy and brain connectivity measures, both structural and functional.

With this state of the art review we aim, on the one hand, to expand our understanding of the several elements at play in the pathological dynamics of MS, and on the other hand, to look into recent literature for examples of studied correlations between atrophy and (graph theory based) measures of brain connectivity.

The results of our search confirm the intricacy and complexity of the relations between atrophy, cortical lesions, cognitive impairment and brain connectivity through the testimony of recent research that offers clear paths for future inquiry but is yet inconclusive. We will show examples of recent research that study atrophy in MS, showing its importance in the dynamics of the disease and therefore supporting the relevance of our research departure point; but we will also point at the fact that we couldn't find studies that approach the subject through the correlation between atrophy and graph connectivity measures, at least not in the context of MS, which supports the potential originality of our proposal.

The relation between atrophy, cognitive impairment and brain connectivity in the context MS is an extremely active field of research. We would like to point out that from the 29 articles that we include in this state of the art review, 17 have been published in 2020 or later, and 13 have been published in 2023.

We want to conclude this introduction with the following quote:

"Neurodegenerative diseases appear to progress by spreading via brain connections"[2]

The apparent simplicity of this sentence *purposefully* obviates the fact that atrophy is not like a virus, physically crawling its way through brain connections. This quote implicitly points at the complexity of the subject while reminding us of the evidence that brain connections must play a central role in the spread and course of the disease. If this is so, network connectivity might just be the right tool for a better understanding of the problem.

2.2 MS and network connectivity

In this section, we will succinctly enumerate the few articles that we will keep as a main reference and base to depart from, especially in order to develop and implement methods of classification and techniques of graph theory network analysis. First and foremost, the articles of Martínez-Heras et al., and Casas-Roma et al., both use similar data to that which will be available to our research, and both offer to us concrete tools and guidelines to approach a descriptive and classification task by using brain connectivity measures [19, 4]. While the former sets the base for efficient methods of classification, pointing out the superior accuracy of Fractional Anisotropy (FA) measures to model structural networks, describing the importance of thalamic connections, which is an idea that will become important to us and we will describe further in the coming sections, and offers a testimony on both the difference and continuity between different MS phenotypes, the latter we will refer to as a guide when it comes to the more technical aspects of measuring connectivity, specifically in terms of setting up a multi-modal approach that integrates functional and structural network measures.

Further, we will also refer to Solana et al. as a reference of how Support Vector Machines (SVM) with k-fold cross-validation can be used as a classification method in the concrete case of MS to differentiate cognitively preserved (CP) and impaired (CI) patients, a classification

that will be instrumental to our main objective [26]. We also include in our reference list a previous article of Solana et al. that, without referring to atrophy as such, does *"investigate the association between the loss of diffusion-based structural connectivity, measured with graph theory metrics, and magnetic resonance (MR) markers of microstructural damage"*, showing that *"patients with multiple sclerosis (MS) display reduced structural connectivity among brain regions"* and describing how *"the accumulation of neuroaxonal pathological burden seems to magnify the risk of global network collapse"*[27]. We consider that these are central issues to take into account when using structural connectivity to understand pathophysiological processes.

Finally, we include as our main go-to references the short [25] and extensive [10] summaries of most used brain network connectivity measures.

2.3 Atrophy in MS

In this section we will delve into the research that has been done on the specific issue of atrophy in MS. This will include i.a. the ideas of atrophy as a predictor of impairment, the relative independence and temporal dynamics of WM lesions and GM volume loss, especially at early stages of the disease, the search for potential biomarkers of atrophy, and the central role of the thalamus in the spread of atrophy.

Conti et al. remind us of a generally accepted premise, which is that "the mechanisms underlying the loss of GM in MS remain largely undetermined"[7]. These researchers proposed a *"machine learning approach to investigate whether cortical tissue loss in MS is mainly dependent on local pathological processes or disconnection from distant WM"* and showed in their results that *"the contribution of cortical lesions to cortical neurodegeneration increases with the duration of the disease"* [7]. This connection between cortical lesions and neurodegeneration at later stages of the disease has as its counterpart the fact that *"decreased volumes of the corpus callosum and thalamus in the relatively early stage of MS may predict the development of Brain Volume Loss(BVL)"*[12]. Also, lower baseline WM integrity seems to be related to a more severe later spread of cortical atrophy in RRMS subjects, which shows that *"WM damage seems to drive atrophy more than conversely"* [32]. Other baseline factors that can predict future CI include *"education, disease severity, lesion burden, and volume of limbic structures"*. These factors *"may be helpful when identifying at-risk patients"*[17].

Thalamic atrophy has also been pinpointed as a cause for cognitive impairment in early

stages of the disease [22] in patients with relapsing-remitting MS (RRMS). Early onset dynamics in the relation between WM lesions, mild cognitive impairment (CI) and atrophy are central to our research objectives, and their study is both relevant and challenging because it seems that atrophy often precedes more visible symptoms and may even take place before a physician's diagnose. In these early stages, *"atrophy measures targeted to specific brain regions may provide improved markers of neurodegeneration"*, which leads to the observation that *"atrophy appears to be a better predictor of clinical disability and deterioration than WM lesions"* [21]. In the same line, other studies confirm that *"the rate of brain atrophy during the RRMS disease course seems to correlate better with future clinical disability than disability measures at that time"*[13].

In earlier stages of the disease, atrophy appears regionally segregated in a clearer way, and it is then relatively easier to see that *"the effect of inflammatory activity is more prominent in the thalamus than in the cortical gray matter"*, and this confirms that *"the thalamus is one of the first atrophic regions in MS, later followed by regions of cortical gray matter"* [23]. Of main importance to our study, evidence shows that *"sensitivity of the thalamus to early volume changes in active MS may be explained by its profile as a connectivity hub"*[23]. In the effort of describing the regional progression of the spread of atrophy, studies show that *"the posterior cingulate cortex, precuneus, and thalamus are among the earliest regions to become atrophic in both relapse-onset phenotypes and primary-progressive multiple sclerosis"*, and thus we find again that research confirms that *"atrophy is associated with disease duration and with disability accumulation over time"* [9].

Concerning the relation between WM lesions and atrophy, this same focus on early stages of the disease reveals that *"in SPMS patients cortical GM atrophy and WM damage are (at least partly) independent disease processes"*[28]. In terms of this independence, the role of thalamic atrophy as a distinct, if not single, trigger for neurodegenerative processes becomes inconspicuous. In that sense, some studies describe how *"silent (not visible lesion-generating) destructive processes already alter the thalamic microstructure and cause measurable atrophy"* at early stages of the disease, while at the same time reminding us that *"however, in the very early phase of MS, acute (nonlesional) inflammatory processes in the thalamus can cause temporary edema that will lead to a specific increase of longitudinal diffusivity and thus to a FA increase"* [8]. In other words, there are compensatory processes that augment structural resilience before atrophy sets in. This is an important fact to take into account when interpreting FA measures at early stages of the disease.

Further, research confirms that together with cortical lesion volume and annualized relapse

rate, *"BVL is one of the components of clinical disability worsening"*[6]. More specifically, an increase in the Expanded Disability Status Scale score, which is one of the measures that will be available to us in our study, can be *"significantly correlated with the volume decrease in right nucleus accumbens in both RRMS and SPMS"*, which leads to the consideration that some *"subcortex structures (especially the nucleus accumbens) could be considered as marker for the transition of RRMS to SPMS"*.[30]

Complementary to these lines of research, and confirming the central role of atrophy as precursor and predictor of the disease spread, efforts are made to identify markers for atrophy. In that sense, serum neurofilament light chain (Nfl) appears as a strong candidate for a neuroaxonal injury biomarker. [24]. Although it is not the aim of our study to evaluate Nfl as a neuromarker (our data will not include such measurements), this study seems relevant to us inasmuch as it evaluates the influence of neuromarkers on the connectome, using *"connectivity-based approaches incorporating the distribution and magnitude of the extended brain network aberrations caused by lesions"*, which *"may offer higher sensitivity for axonal damage in patients with multiple sclerosis (MS) than conventional lesion characteristics"*[24].

To conclude this section, we want to give special attention to a study by Cen et al. that we find sums up several of the issues discussed so far while adding novel perspectives and implementing groundbreaking state-of-the-art techniques. These researchers point at the fact that brain atrophy is *"spreading about 3 times faster in MS patients than in healthy controls"*, and that *"the thalamus is particularly susceptible to neurodegeneration and has been shown to be one of the earliest regions impacted by atrophy"*[5]. The authors remind us that brain atrophy occurs in healthy individuals as part of normal aging, and that this natural process has been described extensively, to the point that healthy brain normative charts exist and can be compared to brain atrophy maps of individual MS patients [1]. Cen et al. leverage these charts to create (healthy) digital twins (HDT) of MS patients, and subsequently *"develop statistical learning models to estimate when the thalamic atrophy trajectory of an MS patient deviated from their expected thalamic atrophy trajectory based on their corresponding HDT"*[5]. The most interesting result of this strategy is that they find that *"progressive brain tissue loss"* on thalamic areas *"precedes clinical disease onset in MS by a mean of 5.1 to 3.8 years and a median of 6 years"*[5]. This constitutes a strong case not only for the independence between WM lesions and atrophy, but for atrophy as a silent trigger that hides in itself *"a 'true' biological onset of MS that so far remains unknown"*[5]. Another aspect of this research that we will refer to in our study is the fact that it implements complex data processing techniques, namely, Multivariate Adaptive Regression Splines (MARS), to augment cross-sectional data into longitudinal data.

While these techniques will probably fall out of reach of our technical and data availability possibilities in our research, the fact that the longitudinal augmentation has been validated by contrast with actual longitudinal data tells us of the relevance and predictive power of measures of (thalamic) atrophy at a single point in time.

In this section we have pointed at the centrality of thalamic atrophy processes as predecessors of all other visible factors of the disease, at the independence between WM lesions and atrophy, and at the degree of brain volume loss as indicator of the severity of cognitive impairment. Although references have been made to affectations of the connectivity network, we haven't found a study that explicitly explored a direct correlation between atrophy measures and brain connectivity measures. In order to find such an approach, we have looked beyond the case of MS, and we have found studies on other neurodegenerative pathologies where a connection is made between atrophy and network connectivity. We are describing some of these in the following section.

2.4 Atrophy and network measures

The studies we review in this section are using Alzheimer's Disease (AD), Temporal Lobe Epilepsy (TLE), Multiple System Atrophy (MSA), Primary Lateral Sclerosis (PLS), dementia, and aphasia as case studies to relate neuro-degenerative physiopathological processes to network connectivity measures. Despite the specific characteristics of each of these diseases, we consider that the approach of the studies we will review is highly informative to our objectives and is comparable to the aim of describing correlations between atrophy and network connectivity in the case of MS.

In the case of MSA, Lyu et al., depart from previous evidence that shows *"reduced gray matter volume and cortical hypometabolism of the INS and temporal gyrus, which were associated with cognitive decline in patients with MSA"*, and they subsequently find that *"patients with MSA exhibited significant FC aberrances between the CAN and brain areas of sensorimotor control, limbic network, putamen, and cerebellum"*[18]. Other studies have also found that *"structural and functional changes in the subcortical limbic structures and disrupted cerebello-cerebral networks may be associated with early cognitive decline in MSA"*[14]. These are just 2 examples of studies that describe correlations between cognitive impairment, the decrease of functional connectivity and GM atrophy in patients with MSA.

We find similar approaches in the study of PLS, where results show a significant degree of *"fractional anisotropy reduction"*, and correlations are found between *"cortical thickness alterations, structural and functional connectivity changes (...) in the right hemisphere"*[29].

As to AD, we find a rather straightforward but inspiring approach to the study of atrophy in relation to connectivity in the research conducted by Wang et al., who propose *"a validation framework that proves that AD atrophy propagates along the main structural connections in the network"*[31]. By simply looking at the immediate neighbor nodes of the areas with atrophy, they find that *"node atrophy is significantly related to neighbor atrophy"*, a fact that they consider proves *"that AD atrophy is closely related to connectivity"*[31]. Although this may sound intuitively evident, we find it important to have evidence that the spread of atrophy doesn't follow random topological patterns, and that because of that, structural connectivity maps can be informative of the possible regional spread course of neurodegeneration.

Another example comes from the study of Dravet syndrome, a phenotype of seizure epilepsy with cognitive and motor impairment, where a relation has been found between structural connectivity and atrophy patterns. The structural changes have been found to originate at *"a region of the hippocampal formation"*, which has been defined as the *"epicenter"* of the changes [16]. The method to define this epicenter consists of a connectivity analysis pipeline defined by Larivière et al. in their study of Temporal Lobe Epilepsy (TLE)[15]. TLE is often studied as a network disorder, and Larivière et al. set up a cross-sectional mega-analysis where they *"integrate neuroimaging and connectome analysis to identify network associations with atrophy patterns"*[15]. They depart from the idea that hub regions are especially vulnerable to pathological processes because of their high centrality. This so-called *"nodal stress hypothesis"* finds empiric confirmation in the fact that *"hubs typically show greater atrophy than locally connected peripheral nodes"* [15]. By using centrality measures, the researchers find hubs that are more (and earlier) vulnerable to processes of atrophy, both for groups as in individual patients, and they call those nodes "epicenters". They subsequently *"detect syndrome-specific disease epicenters, by comparing every region's functional and structural connectivity profiles to whole-brain patterns of atrophy"* and are able to define syndrome-specific "atrophy maps" that can also become patient-tailored, and which they defend can *"predict spatial patterns of disease duration and age-related effects"*[15]. We find that this analysis pipeline, which we describe here in an extremely succinct manner and obviating technical details, is a valuable referent inasmuch as it attempts to prove that *"brain network architecture governs whole-brain atrophy"*[15]. One last example where the same "epicenter-defining" techniques are used is the research by Brown et al. applied to cases of frontotemporal dementia and progressive

aphasia. The authors derive 2 specific graph-theoretical measures to predict future atrophy: "shortest path length to the epicenter", and "nodal hazard", which is *"the cumulative atrophy of a region's first-degree neighbors"*. Using these predictors and baseline atrophy, and validating their predictions with longitudinal data, the authors are able to demonstrate that *"a patient-tailored model using network-based predictors can accurately estimate the spatial pattern of subsequent atrophy"*. Importantly as well to our objectives, the authors tackle the issue of inter-individual heterogeneity of atrophy spread by *"using principal component analysis on the atrophy maps for all patients at all time points to determine the atrophy latent space"*, and find that *"regional atrophy severity depends on the shortest path length to the epicenter (SPE)"*[2].

2.5 Last considerations in relation to the implementation phase

In this review, we have put our focus almost exclusively on the most recent recent improvements in the study of MS, and also in the use of brain network connectivity approaches applied to the study of neurodegenerative pathologies. In order to tackle the next phase of our project, we are aware that both classical and generative AI models for correlation and classification will have to be implemented to obtain significant results out of our data. However, we have considered that a general state of the art review of classification algorithms was beyond the scope and objective of this assignment, and maybe not relevant at this point, as we consider that understanding the field of application is the necessary first step before deciding on the analysis techniques that will be implemented.

2.6 Summary and list of research objectives

As a result of this state of the art review we want to add the following points of attention to our initial goals for the project:

1. The central importance of thalamic (and other subcortical) regions in the spread of atrophy.
2. The independence between GM atrophy and WM cortical lesions.
3. The fact that atrophy may precede, by years, the onset of the disease as it is visibly manifested in the eyes of physicians.

4. Atrophy as an independent indicator of cognitive impairment.

5. Hub centrality as a measure of vulnerability in respect to atrophy spread (nodal stress hypothesis).

In our first proposal for this project we had conceptualized the importance of atrophy as a cause for the most severe impairment symptoms of MS, and in relation to its temporality, under the rationale that functional connectivity changes happen at a very short time-scale, structural connectivity changes happen on a mid-term scale, and atrophy progresses later and on a longer term scale. This state of the art review has modified and complexified this view, as we have discovered research that proves that atrophy appears earlier than any cortical lesion and potentially long before any symptom, and we have also seen that brain volume loss alone (both at baseline and at different stages of the disease) can be a reliable indicator of impairment and of the future evolution of the disease. Seeing what the state of the art offers, we are now more convinced of the importance of atrophy as a potential prognostic tool, and we are also confident that atrophy can be studied through structural connectivity measures.

We aim to include these elements in our study of the possible correlations between structural and functional connectivity measures, regional and whole-brain BVL atrophy measures and cognitive impairment.

Chapter 3

Implementation

3.1 Methods

3.1.1 Participants

This is a cross-sectional study involving 178 subjects. The data was obtained, via the tutor of this thesis, from the database of the ImaginEM research group at the Hospital Clinic Barcelona. It consisted of 5 different types of measurements: demographic variables of the sample of subjects, structural, functional and morphological adjacency matrices, and brain volumetric measures. Due to inconsistencies between the number of files for each type of information, we reduced the sample down to those subjects from which we could have complete information, ending up with a final sample of 165 subjects (147 MS patients and 18 healthy control volunteers).

3.1.2 Data preprocessing

Different types of connectivity data have previously been used to construct multilayer networks, wherein morphological network data acted as an interlink between functional and structural connectivity data, and the result is a multi-dimensional network[4]. This type of data requires specific processing in order to derive graph measures from it, and the construction of the multidimensional matrices is not trivial. Considering our objectives and the time available, we decided early on in the project to work only with structural (FA weighted) and functional (obtained from rsFMRI) networks, discarding morphological networks. Eventually, we did generate a combination of functional and structural networks through a simple weighted sum of values, which produced matrices of the same dimensions of the original functional and structural networks, but combining both types of information. So, for the preprocessing phase, we tested and corrected values of the 165 adjacency matrices representing structural, functional and

combined networks, respectively.

Using a linear regression model, we obtained a p value > 0.9 when testing the variability of the sample due to gender or age, and thus no adaptation was made to correct for those variables. Next, we harmonized all data using the Combat model[11]. This was necessary due to the fact that original measurements were obtained from 2 different types of scanners. The Combat model, implemented with the neuroCombat python library, includes the possibility of testing for the influence of secondary variables in the data variability. Using this feature, we again tested for the influence of gender and age, finding no relevant results and therefore not including those parameters in the final harmonization step. Since structural and functional connectivity matrices differ in their diagonal values (zeros for the former, ones for the latter), we implemented the harmonization steps separately, and we turned diagonal values of functional matrices to zero. Concerning the values in functional connectivity matrices, as they range from -1 to 1, turning them to absolute values simplifies the information and avoids issues in several graph measures[4]. We went one step further in terms of simplicity, and turned all negative values to zero, so that only positive correlation would be considered in the interpretation of the results. Next, by filtering out 0.001 of the data of both structural and functional we could discard a number of outlier values that were affecting the centrality measures of both sets of matrices, and we normalized all values from 0 to 1. Finally, we removed all values (weighted edges) smaller than 0.1 in the structural matrices, as well as those edges that were not present in at least 60% of the healthy control subjects. With this last operation, 1128 edges were removed, leaving matrices with 1798 non-zero edge connections. We removed the equivalent edges on the functional matrices, so that functional connectivity would only be taken into account on those edges we had selected from the structural connectivity measures.

The code for all steps described in this section can be found in the code section of the GitHub repository.

3.1.3 Graph network measures

Our data consists of adjacency matrices of 76x76 nodes. The matrices are symmetric, which means that the connections are undirected (value of connection from node 1 to node 2 is equal to value from node 2 to node 1), and their values are representing the connectivity weight of 2850 edges. Different graph measures were derived from structural, functional and combined matrices. Combined matrices were generated through a weighted sum of values, giving 75% weight to structural connectivity. For each measure, we tested for statistically significant differences between groups, comparing on the one hand patients vs. controls, and

on the other hand patients with RRMS vs. all other patients (SPMS and PPMS). Each graph measure reveals a different feature of network connectivity; even though we derived up to 11 different network metrics from our data, we summarize here a brief theoretical description only of the measures that yielded significant difference between patients and controls:

1. **Node connectivity measures:** These are the most basic measures, with which we can obtain a summary of the connectivity of each node in a network[10]. Along with **node degree**, which measures the amount of edges each node is connected to, **strength** measures the sum of weights of all edges connected to a node. This metric can be obtained at a nodal level, or globally, as an average of strengths across the network.

2. **Centrality measures:** Centrality describes the differential influence in the network of a node, compared to other nodes. What this influence concretely means depends on the specific centrality measure. **Degree centrality** is the most basic measure of centrality, and it is equivalent to node degree, that is, it is the sum of edges of each node. This measure counts the quantity of connections, but does not consider their quality[10]. In contrast, **eigenvector centrality** takes into account both the degree of a node and the degree of its neighbors, and thus, according to this measure, a node is more central when it connects to other central nodes. A different type of centrality measure altogether is obtained with the **closeness centrality** metric, which takes into account what is known as the average shortest path length, that is, the average of shortest paths from a node to any other node in the network. According closeness centrality, a node is considered more central if it can reach any other node more efficiently (with less hops over other nodes). Finally, **betweenness centrality** measures the capacity of nodes to act as links, according to their placement along shortest paths and independently of their degree. A node with high betweeness centrality might not necessarily have a high number of connections, but will be placed on a high traffic path, acting as a central passage.

3. **Rich Clubs:** The **Rich Club** coefficient measures the importance of nodes taking into account subgraph areas within the overall network. For each sub-zone, a level of richness is calculated according a specific threshold in degree or strength, thus defining zones of different varying richness. The rich club coefficient will indicate whether or not 'rich areas' of a network have a tendency to connect preferentially to each other.

4. **Clustering measures:** Clustering describes the extent to which a network presents differentiated subgroups of nodes, opposite to consisting of a random or regular web of connections. It is a particularly relevant metric in the context of brain connectivity, since modularity and differentiated functionality are defining features of nervous systems.

Concretely, the **clustering coefficient** measures the probability of finding a connection between any two neighbors of a given node. The clustering coefficient can be averaged into a global metric.

3.1.4 Brain volumes

A matrix with the normalized brain volumes of the 76 nodes of all subjects was available to us. We created three separate sets that we subsequently used in statistical tests and classification tasks. The first set included all volumes for the 76 nodes. Following the objectives stated on our state of the art review, the second set included only the volumes for right and left thalamus. Finally, we calculated the differences in average volume between patients and controls for all 76 nodes, and we created a third set that included only the 5 nodes with the biggest difference between groups. The values of these three sets were normalized from 0 to 1 prior to statistical testing and classification.

3.1.5 Statistics

For our demographic study, we used a Wilcoxon signed-rank test to compare patients and controls on the age variable, and Dunn's tests with Bonferroni correction to compare the 3 different phenotypes (RRMS, SPMS, PPMS) on the age, disease duration and EDSS score variables. For the gender comparisons we performed a Chi-squared test.

Concerning the network graph measures, we compared patients vs. controls, as well as phenotype groups, but the latter were joined into 2 distinct groups (RRMS vs. PPMS/SPMS), taking into account the low number of subjects with PPMS or SPMS in our sample. We constructed functions in python that would, for each graph measure derived from functional, structural or combined networks, perform a Shapiro-Wilk test to determine whether comparisons between groups should be parametric or non-parametric, performing then accordingly a Student's t-test of independence, or a Mann-Whitney U test, respectively. Due to the large number of pairwise comparisons made in each test, and in order to ensure the relevance of the chosen metrics, we decided that only results obtaining a $p<0.001$ significance level would be used further in the classification tasks. Statistical comparisons were performed both for global graph measures and for nodal measures, so that we were able to obtain sublists of nodes with differences between groups for every graph metric.

The same statistical procedures were applied to the volumetric data, implementing a t-test of independence between patients and controls, and between phenotypes, for the complete set

of 76 brain volumes, the set of 5 most different nodes between patients and controls, and using only left and right thalamus volumes.

3.1.6 SVM classification

Support Vector Machines (SVM) have proven to be a powerful method to classify brain connectivity data, including data from multiple sclerosis patients [26, 4, 19]. Here, our purpose was to use SVM to compare the accuracy of classification tasks that use full adjacency matrices (structural, functional or combined), graph derived measures (global or nodal), and brain volume data. For each of these types of data, we built functions that would perform a grid search cross validation in order to select the optimal C and *gamma* parameters in each task. When splitting our sample into train and test, although we increased our test size to 0.25, we found that we had only 5 control subjects in our test group, which pushed the algorithm to classify all subjects as positive in all tasks, obtaining therefore always the same accuracy (over 0.9). In order to avoid this overfitting effect, we decided to increase the size of our control group using SMOTE, an oversampling technique for imbalanced classification. After that, the classification was further validated using a stratified k-fold iteration (n splits = 5), from which the best accuracy, precision, recall and F1 scores were kept. We used this same procedure to perform the different classification tasks of patients vs. controls: using connectivity based matrices (structural, functional and combined), global graph measures of strength (based on structural networks) and Rich Club coefficient (based on functional connectivity), and brain volumes (all volumes, thalamic, and 5 most different). Concerning nodal graph measures, departing from the results summarized in table 3.2, we proposed a classification task in which each subject was characterized by a vector of values, starting from the 54 structural node strength values with statistically significant differences, continuing with the 3 structural node degree centrality values... and so on, until completing a vector of 75 values representing diverse nodal graph measures obtained from structural, functional and combined connectivity matrices. We call this a 'nodal embedding' of graph measures.

3.2 Results

3.2.1 Demographic analysis

Our n=147 sample of patients was composed of 143 women and 43 men, with an average age of 47.33 years, and an average disease duration of 15.96 years. The average EDSS score was 2.66, and the maximum score was 7.50 (scale from 0, no symptoms, to 10, death). The sample of healthy controls (n=18) had an average age of 36.25 years, and was composed of 15 women and 3 men. Table 3.1 shows a summary of the characteristics of our sample, including p value comparisons of patients vs. controls, as well as different MS phenotypes. Figure 3.1 shows the relative unbalance between gender, with more women in the sample, and phenotypical expression of the disease, with RRMS being the most prevalent form. Further, figure 3.2 shows quite clearly how higher scores in the Expanded Disability Scale (EDSS) correlate with PPMS and SPMS, and also how we can differentiate groups according to their disease duration, seeing that PPMS have had the disease for shorter time, while we have patients with RRMS or SPMS along the whole range of disease duration. Concerning brain volumetric data, we explored the correlation between atrophy and impairment. Figure 3.3 shows what seems to be a tendency for patients with higher scores in EDSS to have diminished volumes of left and right thalamus. We verified this tendency with an ANOVA test, obtaining a p=0.0025 significance level.

Table 3.1: *Characteristics of the sample.*

	Age	Disease Duration	EDSS	Female(%)
HCs	36.25 (28.64 - 43.18)	-	-	15 (83%)
RRMS	45.28 (38.21 - 51.00)	12.79 (8.66 - 18.78)	2.00 (1.50 - 2.50)	90 (72%)
SPMS	53.80 (50.59 - 65.27)	23.62 (16.45 - 33.01)	6.00 (5.88 - 6.12)	10 (62%)
PPMS	56.81 (55.90 - 57.91)	10.65 (8.23 - 15.68)	5.50 (5.00 - 6.00)	4 (67%)
p value HCs vs. MS	p<0.001*	-	-	0.398***
p value MS types	p<0.001**	p<0.001**	p<0.001**	0.716***

Note: The data represent the absolute numbers and the proportions of the qualitative data, or the median and IQR for the quantitative data. *Wilcoxon signed-rank test. **Dunn's test. Across the different variables, the tests showed significant difference between groups, except for the pairwise comparison between PPMS and SPMS. ***Chi-squared test.

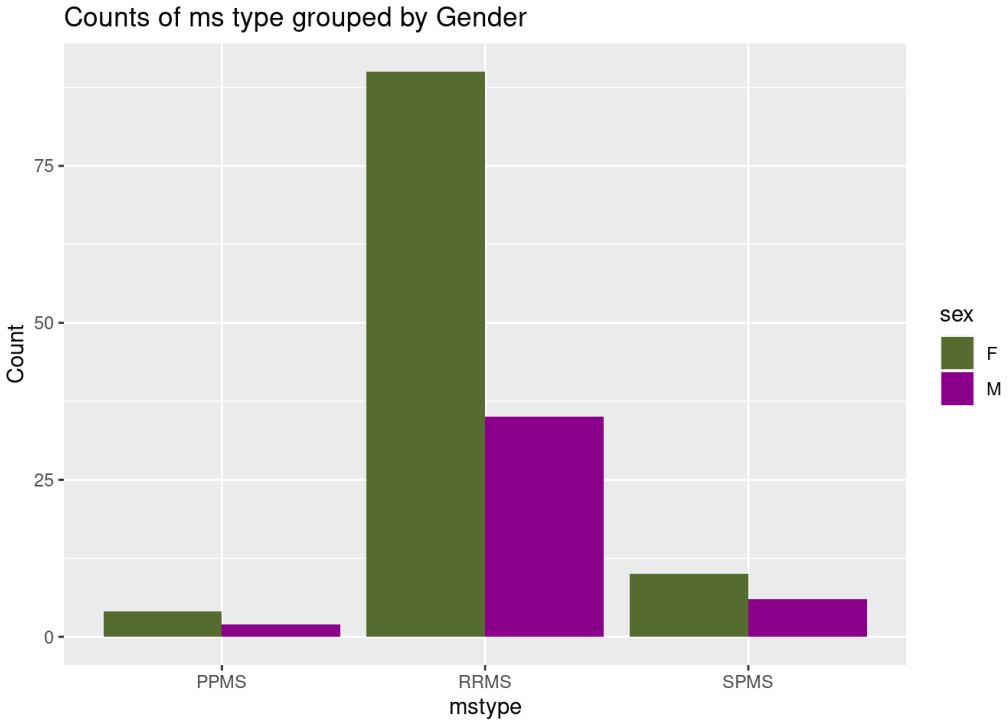

Figure 3.1: Comparison of gender and phenotype.

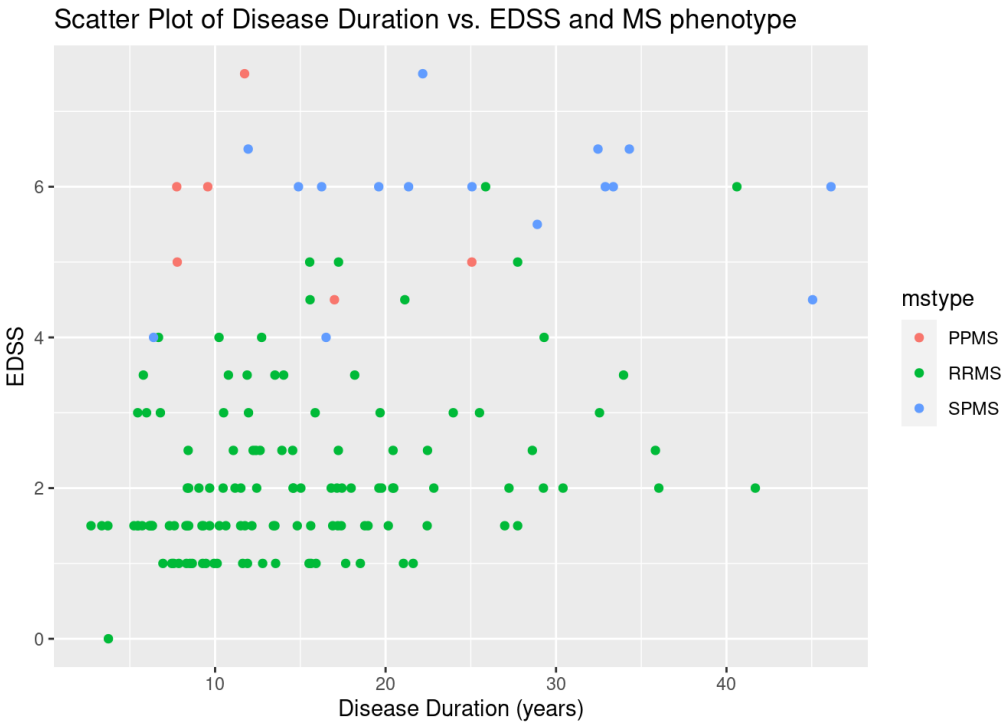

Figure 3.2: Comparison of EDSS, disease duration and MS phenotype.

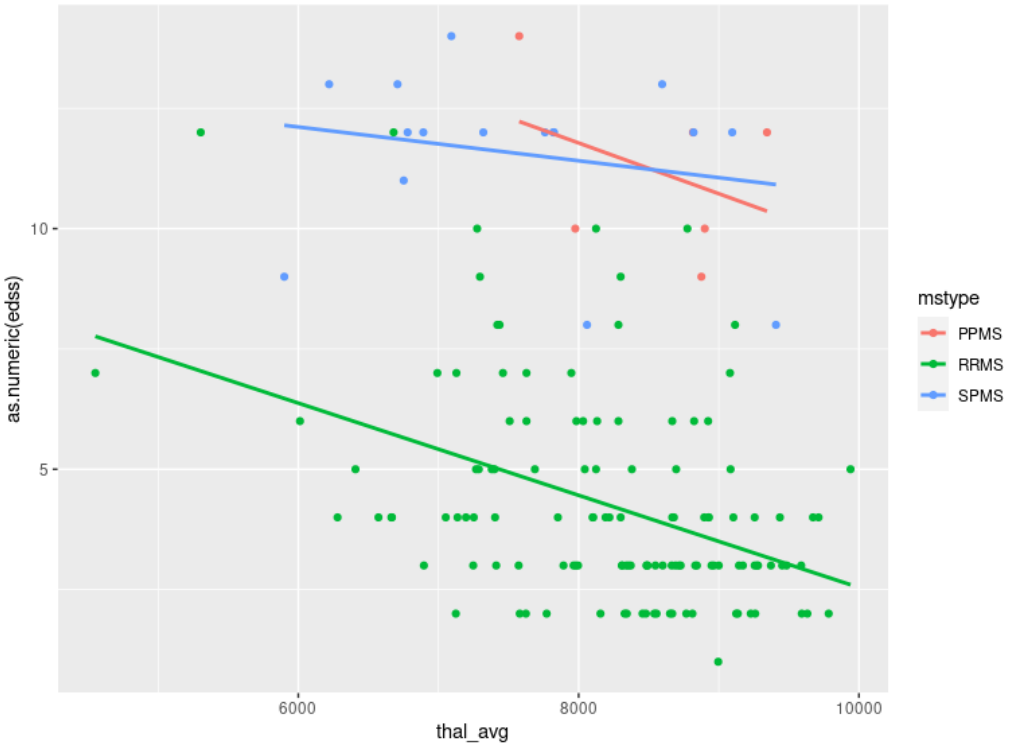

Figure 3.3: Comparison of EDSS and thalamic volume.

3.2.2 Statistics and SVM classification

As we stated in the methods section, we used a threshold value of $p<0.001$ to determine significant difference between groups. With this threshold value, we did not obtain a significant difference between RRMS patients vs. PPMS/SPMS. Therefore, all results from here on concern comparisons between patients and controls.

Global graph measures of strength, based on the structural networks, and of Rich Club coefficient, based on the functional networks, obtain both a $p<0.001$ value. As to nodal measures, different measures obtained statistically significant differences between patients and controls on different number of nodes. The nodal graph results are summarized in table 3.2. Among all measures, strength is the one where we see most significant nodal differences between patients and controls, with 54 of the 76 nodes.

Table 3.2: *Nodal graph comparison of patients vs. controls.*

	structural	functional	combined
strength	54/76	n.a.	5/76
degree centrality	3/76	1/76	n.a.
closeness centrality	3/76	1/76	n.a.
betweenness centrality	1/76	2/76	1/76
eigenvector centrality	5/76	n.a.	n.a.
clustering coefficient	n.a.	1/76	n.a.

Amount of nodes with significant differences between patients and controls.

Figure 3.4 shows the location of the 18 nodes with the highest difference in strength between patients and controls, making visible the affectation of subcortical areas, in this case left and right hippocampus. In order to explore further the relevance of the fact that so many nodes show a difference in strength between patients and controls, we generated some connectogram plots showing structural connectivity FA weights of only those 54 nodes where we found significant differences in strength. Figure 3.5 shows all the 54 nodes of interest, with FA weights difference between patients and controls represented across a color scale (darker blue indicating bigger difference in weight, controls having higher weight of FA connection relative to patients in all cases). Further, figure 3.6 highlights the 2 pericalcarine cortex areas, showing the important decrease of weight in the connection between these cortex areas and several subcortical structures, including the thalamus, the hippocampus and the accumbens area.

Figure 3.4: Location of nodes with biggest strength difference.

Image shows the 18 nodes with the biggest structural strength difference between controls and patients, value 1.0 representing maximum relative difference in strength, taking into account differences in strength in all 76 nodes of the define connectivity network.

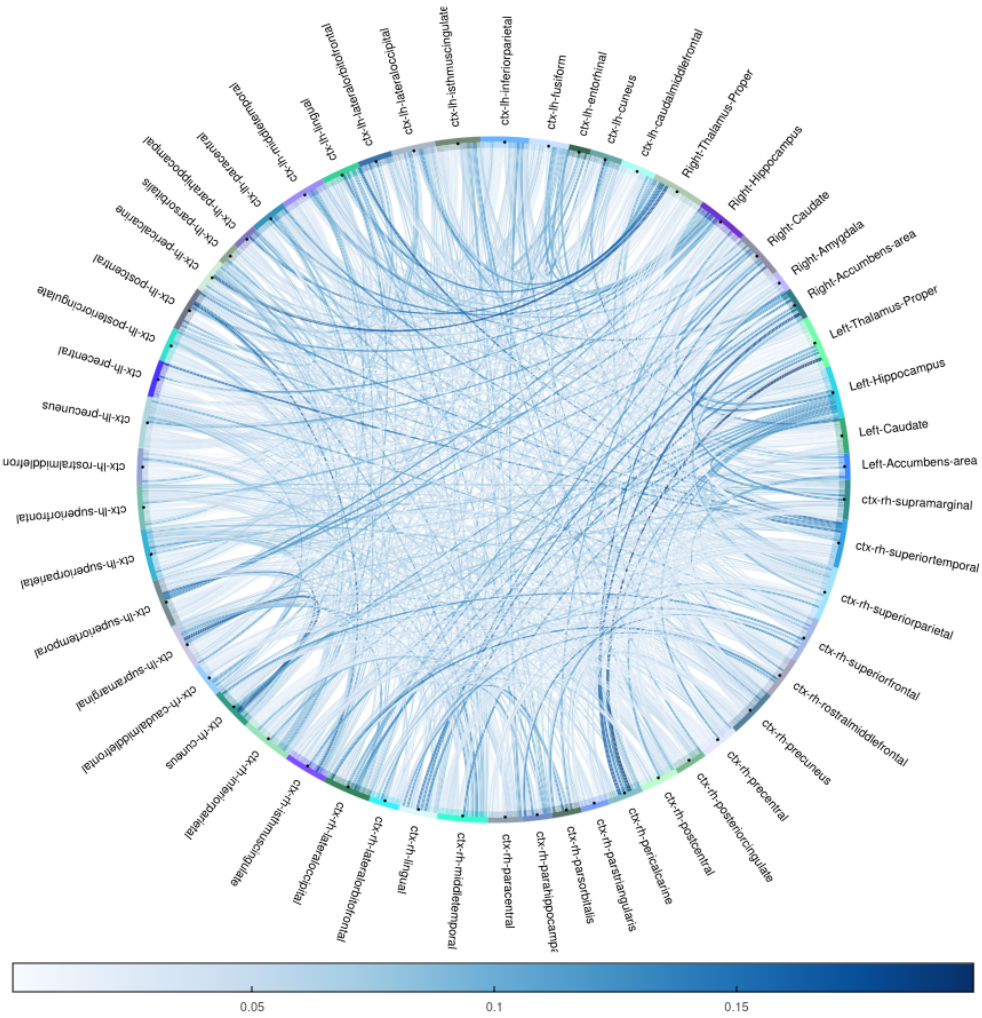

Figure 3.5: Full Connectogram.

The plot shows a selection the 54 nodes (out of a total of 76) with the biggest difference in graph strength measure. The edges are representing FA weight difference between controls and patients, with controls having higher structural connectivity weights in all cases. Differences are represented along the color scale, where darker blue showing bigger magnitude in difference. Even if we discarded 22 nodes for this plot, the result is still a rather cluttered image that makes visible the complexity of network connections.

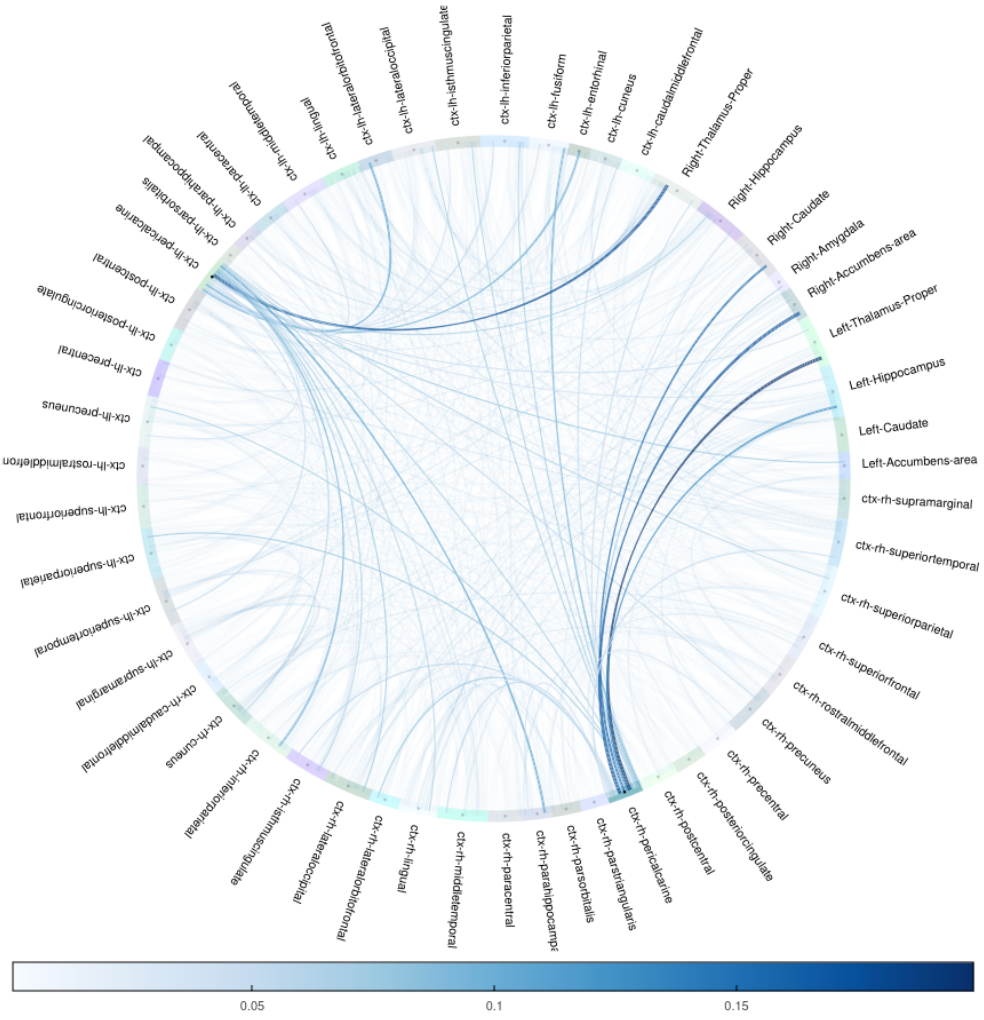

Figure 3.6: Plot of selected nodes.

Through simple visual inspection of the previous full connectogram plot, and despite its complexity, we can see that the precalcarine cortex areas receive several edge connections with important weight differences (darker blue lines). In this plot, we highlight the 2 precalcarine areas (left and right), to better show the areas to which they are connected. We find therefore that most differences of FA weight connectivity between controls and patients happen in the connection between the precalcarine cortices and several subcortical areas, such as the thalamus, the hippocampus and the accumbens area.

Table 3.3: *SVM classification results.*

Data type	measure	Metrics			
		accuracy	precision	recall	F1
Connectivity	structural	1.0	1.0	1.0	1.0
	functional	0.94	0.94	1.0	0.97
	combined	0.97	0.97	1.0	0.98
Graph	structural strength	0.97	0.75	1.0	0.86
	functional Rich Club	0.94	0.67	0.67	0.67
	node embedding	0.97	0.75	1.0	0.86
Brain Volumes	all	0.94	1.0	0.50	0.67
	thalamic	0.67	0.10	0.33	0.15
	5 most different	0.76	0.14	0.33	0.20

The SVM classification task results are summarized in table 3.3. In these results, we witness the superior efficiency, across all metrics, of classification with structural connectivity matrices (accuracy=1.0, precision=1.0), followed by functional (accuracy=0.94, precision=0.94) and combined (accuracy=0.97, precision=0.97). However, we note as well the very high efficiency of classification using the strength graph measure based on structural networks (accuracy=0.97, precision=0.75), as well the good results on the classification using embedded vectors of nodal graph measures (accuracy=0.97, precision=0.75). Finally, we can see how the classification tasks based on brain volumes also attain acceptable levels of accuracy, but with more irregular values on other metrics (lower precision, recall or F1 values).

Chapter 4

Discussion

Our results show that statistical tests confirm significant differences between patients and controls when using connectivity matrices, graph measures and brain volume measures. We haven't been able to find significant differences between phenotype groups. Previous research has been able to find these differences by using the same procedures on similar data[19]. We assume that the reason we couldn't differentiate between phenotype groups is the reduced size of our sample. With 147 patients, we end up having only 6 subjects in the PPMS group (and 16 in the SPMS group).

As we proceeded to the classification of patients vs. controls, we obtained the highest accuracy when using structural FA connectivity, though functional and combined matrices also yielded high accuracy. Concerning global graph measures, we only obtain statistically significant differences between groups using network global strength of structural matrices. Strength is maybe the simplest of all graph measures, reflecting in an almost direct way the importance of weights in the connectivity network, and therefore it is no surprise that the metric of graph strength behaves and offers similar results to using FA connectivity directly. More remarkably, rich club measure of functional connectivity also yields significant difference between groups, though the precision it delivers in SVM classification tasks is sensibly lower compared to strength. We do not deem it is straightforward to propose an interpretation of the possible relevance of the Rich Club coefficient in functional networks, when the same coefficient does not yield significant differences in structural networks for the same sample. We take into consideration that global graph measures produce a single value for each matrix, and this brings up again the limitation given by the size of the sample (165 matrices = 165 metrics); thus, we have to interpret with caution the fact that no other global graph measure delivers significant differences between groups.

On the other hand, nodal graph measures offer more complexity and information in their result: we obtain 76 node measures for each subject, and we can compare node to node across the whole sample. Thanks to this, we are able to find significant differences on different amounts of nodes across different types of measures despite the limited size of the sample. We propose this fact as a first conclusion of our study; when the size of the sample is limited, graph and connectivity measures allow us to *multiply* the amount of information generated by each subject in it. In terms of efficiency, there lies an interest in producing nodal graph measures as well. Compared to the full FA structural or fMRI functional adjacency matrices, classification via nodal graph measures is computationally less expensive, even for a non-deep model such as SVM. Next to that, there is the added detail in the information we obtain from nodal graph measures. Not only we reveal the different features associated with each graph metric, but we can associate those features to specific brain regions too, helping us understand the regional and connectivity dynamics of the disease. In this sense, and keeping in mind that we exclude from this report a number of nodal graph measures that did not yield statistical significance, our results are rather clear in pointing out that, beyond connectivity strength (54 nodes affected out of 76), the affectations we observe when we compare controls and patients have to do mainly with different aspects of node centrality. The fact that eigenvector centrality is affected (5 nodes affected out of 76) portrays how the quality, and not just the quantity, of connections is affected by multiple sclerosis. On the other hand, closeness centrality being affected (3 nodes out of 76) talks about the efficiency in the routing of information, with patients having to resort to relaying information through longer connectivity paths. Again, this reveals more information than just quantitative. We believe that a larger sample would yield significant differences across a larger array of nodes, but even the relatively limited findings in this study make an argument for the advantages that graph connectivity metrics offer over an analysis exclusively based on the full adjacency matrices, in terms of gaining qualitative information without loosing much power in classification tasks. In this sense, it is worth to point out the high levels of accuracy achieved in the classification task that uses embedded vectors of varied nodal measures; although the information obtained across all the different metrics involved in the construction of the vectors seems initially sparse and unconnected, the fact that it characterizes the subjects correctly enough to deliver a 0.97 level of accuracy in a classification task validates this specific approach. Even when we only obtain significant difference on a single node of the whole 76-node network, as it is our case for the measure of betweenness centrality, this single value can become relevant in the construction the embedded vector, thus minimizing loss of information and gaining precision and detail in the characterization of the subjects.

In our state of the art review we introduced the objective of relating connectivity affectations

and gray matter atrophy through the analysis of brain volumetric data. Here, we must take into account the fact that, strictly speaking, atrophy can only be quantified using intra-individual longitudinal data. Therefore, any correlation between connectivity and atrophy we could establish with our cross-sectional data sample would be at best, and by definition, indirect and hypothetical. Given these limitations, we argue that our study nevertheless manages to include a focus on the deterioration of gray matter subcortical structures to complete and extend the findings based on graph metrics of connectivity. On the one hand, the results presented in figure 3.4 clearly show how subcortical structures of MS patients undergo a substantial transformation in their connectivity, with great difference in strength compared to controls. However, since this concrete visualization relies on nodal measures of strength, and does not take into account volumetric data, it completes only half of our objective: it supports the argument of the importance of subcortical areas, but it does not establish a correlation with atrophy. The other half of our objective is tackled by the statistical tests of independence between groups based on brain volumes, and the subsequent classification tasks. Brain volumetric data has proven efficient in the classification through SVM (0.94 accuracy) despite comparatively lower values on precision, recall and F1 values. Concerning the role of the thalamus in the (early) development of the disease, which is another of the main ideas included in our state of the art review, the results we obtain with our sample are less definite (0.67 accuracy and down to 0.10 precision), but we did find statistical significance between average brain volumes and EDSS score (**??**). Our most inconclusive findings come from the subset of the 5 nodes with the highest difference in volume between controls and patients. On the one hand, these 5 areas are all cortical areas; this does not mean there aren't subcortical areas affected by atrophy, but it does diminish their comparative importance. On the other hand though, classification tasks using these same 5 nodes yield slightly better results than classification using thalamic regions. Brought together, the results obtained from the inclusion of brain volumetric data in our study adds partial evidence but does not clarify the role of subcortical atrophy in MS. Prior to establishing a correlation between, on the one hand, connectivity patterns analyzed through graph metrics, and on the other hand, atrophy represented by inter-individual comparison of brain volumes, the robustness of both methods has to be established. We consider that our study firmly validates the power of graph measures to characterize and correctly classify pathological cases vs. controls, and does validate as well the importance of volumetric data by showing significant differences of average brain volume between patients and controls. Having reached this point, we would now be ready to conduct a study that focuses solely on establishing a correlation between the two aspects; for this we would need more time and, ideally, longitudinal data on a larger sample of subjects.

Bibliography

[1] R. A.I. Bethlehem, J. Seidlitz, S. R. White, J. W. Vogel, K. M. Anderson, C. Adamson, S. Adler, G. S. Alexopoulos, E. Anagnostou, A. Areces-Gonzalez, D. E. Astle, B. Auyeung, M. Ayub, J. Bae, G. Ball, S. Baron-Cohen, R. Beare, S. A. Bedford, V. Benegal, F. Beyer, J. Blangero, M. Blesa Cábez, J. P. Boardman, M. Borzage, J. F. Bosch-Bayard, N. Bourke, V. D. Calhoun, M. M. Chakravarty, C. Chen, C. Chertavian, G. Chetelat, Y. S. Chong, J. H. Cole, A. Corvin, M. Costantino, E. Courchesne, F. Crivello, V. L. Cropley, J. Crosbie, N. Crossley, M. Delarue, R. Delorme, S. Desrivieres, G. A. Devenyi, M. A. Di Biase, R. Dolan, K. A. Donald, G. Donohoe, K. Dunlop, A. D. Edwards, J. T. Elison, C. T. Ellis, J. A. Elman, L. Eyler, D. A. Fair, E. Feczko, P. C. Fletcher, P. Fonagy, C. E. Franz, L. Galan-Garcia, A. Gholipour, J. Giedd, J. H. Gilmore, D. C. Glahn, I. M. Goodyer, P. E. Grant, N. A. Groenewold, F. M. Gunning, R. E. Gur, R. C. Gur, C. F. Hammill, O. Hansson, T. Hedden, A. Heinz, R. N. Henson, K. Heuer, J. Hoare, B. Holla, A. J. Holmes, R. Holt, H. Huang, K. Im, J. Ipser, C. R. Jack, A. P. Jackowski, T. Jia, K. A. Johnson, P. B. Jones, D. T. Jones, R. S. Kahn, H. Karlsson, L. Karlsson, R. Kawashima, E. A. Kelley, S. Kern, K. W. Kim, M. G. Kitzbichler, W. S. Kremen, F. Lalonde, B. Landeau, S. Lee, J. Lerch, J. D. Lewis, J. Li, W. Liao, C. Liston, M. V. Lombardo, J. Lv, C. Lynch, T. T. Mallard, M. Marcelis, R. D. Markello, S. R. Mathias, B. Mazoyer, P. McGuire, M. J. Meaney, A. Mechelli, N. Medic, B. Misic, S. E. Morgan, D. Mothersill, J. Nigg, M. Q.W. Ong, C. Ortinau, R. Ossenkoppele, M. Ouyang, L. Palaniyappan, L. Paly, P. M. Pan, C. Pantelis, M. M. Park, T. Paus, Z. Pausova, D. Paz-Linares, A. Pichet Binette, K. Pierce, X. Qian, J. Qiu, A. Qiu, A. Raznahan, T. Rittman, A. Rodrigue, C. K. Rollins, R. Romero-Garcia, L. Ronan, M. D. Rosenberg, D. H. Rowitch, G. A. Salum, T. D. Satterthwaite, H. L. Schaare, R. J. Schachar, A. P. Schultz, G. Schumann, M. Schöll, D. Sharp, R. T. Shinohara, I. Skoog, C. D. Smyser, R. A. Sperling, D. J. Stein, A. Stolicyn, J. Suckling, G. Sullivan, Y. Taki, B. Thyreau, R. Toro, N. Traut, K. A. Tsvetanov, N. B. Turk-Browne, J. J. Tuulari, C. Tzourio, Vachon-Presseau, M. J. Valdes-Sosa, P. A. Valdes-Sosa, S. L. Valk, T. van Amelsvoort, S. N. Vandekar, L. Vasung, L. W. Victoria, S. Villeneuve, A. Villringer, P. E. Vértes, K. Wagstyl, Y. S. Wang, S. K. Warfield,

V. Warrier, E. Westman, M. L. Westwater, H. C. Whalley, A. V. Witte, N. Yang, B. Yeo, H. Yun, A. Zalesky, H. J. Zar, A. Zettergren, J. H. Zhou, H. Ziauddeen, A. Zugman, X. N. Zuo, C. Rowe, G. B. Frisoni, A. Pichet Binette, E. T. Bullmore, and A. F. Alexander-Bloch. Brain charts for the human lifespan. *Nature 2022 604:7906*, 604:525–533, 4 2022.

[2] Jesse A. Brown, Jersey Deng, John Neuhaus, Isabel J. Sible, Ana C. Sias, Suzee E. Lee, John Kornak, Gabe A. Marx, Anna M. Karydas, Salvatore Spina, Lea T. Grinberg, Giovanni Coppola, Dan H. Geschwind, Joel H. Kramer, Maria Luisa Gorno-Tempini, Bruce L. Miller, Howard J. Rosen, and William W. Seeley. Patient-tailored, connectivity-based forecasts of spreading brain atrophy. *Neuron*, 104:856–868.e5, 12 2019.

[3] Massimiliano Calabrese, Federica Agosta, Francesca Rinaldi, Irene Mattisi, Paola Grossi, Alice Favaretto, Matteo Atzori, Valentina Bernardi, Luigi Barachino, Luciano Rinaldi, Paola Perini, Paolo Gallo, and Massimo Filippi. Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis. *Archives of neurology*, 66:1144–1150, 9 2009. old. cortical lesions and GM atrophy as independent predictors of cognitive impairment.

[4] Jordi Casas-Roma, Eloy Martinez-Heras, Albert Solé-Ribalta, Elisabeth Solana, Elisabet Lopez-Soley, Francesc Vivó, Marcos Diaz-Hurtado, Salut Alba-Arbalat, Maria Sepulveda, Yolanda Blanco, Albert Saiz, Javier Borge-Holthoefer, Sara Llufriu, and Ferran Prados. Applying multilayer analysis to morphological, structural, and functional brain networks to identify relevant dysfunction patterns. *Network Neuroscience*, 6:916–933, 7 2022.

[5] Steven Cen, Mulugeta Gebregziabher, Saeed Moazami, Christina J. Azevedo, and Daniel Pelletier. Toward precision medicine using a "digital twin" approach: modeling the onset of disease-specific brain atrophy in individuals with multiple sclerosis. *Scientific Reports 2023 13:1*, 13:1–14, 9 2023.

[6] Magdalena Chylińska, Jakub Komendziński, Adam Wyszomirski, and Bartosz Karaszewski. Brain atrophy as an outcome of disease-modifying therapy for remitting-relapsing multiple sclerosis. *Multiple Sclerosis International*, 2023:1–11, 8 2023.

[7] Allegra Conti, Constantina Andrada Treaba, Ambica Mehndiratta, Valeria Teresa Barletta, Caterina Mainero, and Nicola Toschi. An interpretable machine learning model to predict cortical atrophy in multiple sclerosis. *Brain Sciences*, 13(2), 2023.

[8] Michael Deppe, Julia Krämer, Jan Gerd Tenberge, Jasmin Marinell, Wolfram Schwindt, Katja Deppe, Sergiu Groppa, Heinz Wiendl, and Sven G. Meuth. Early silent

microstructural degeneration and atrophy of the thalamocortical network in multiple sclerosis. *Human Brain Mapping*, 37:1866, 5 2016.

[9] Arman Eshaghi, Razvan V. Marinescu, Alexandra L. Young, Nicholas C. Firth, Ferran Prados, M. Jorge Cardoso, Carmen Tur, Floriana De Angelis, Niamh Cawley, Wallace J. Brownlee, Nicola De Stefano, M. Laura Stromillo, Marco Battaglini, Serena Ruggieri, Claudio Gasperini, Massimo Filippi, Maria A. Rocca, Alex Rovira, Jaume Sastre-Garriga, Jeroen J.G. Geurts, Hugo Vrenken, Viktor Wottschel, Cyra E. Leurs, Bernard Uitdehaag, Lukas Pirpamer, Christian Enzinger, Sebastien Ourselin, Claudia A. Gandini Wheeler-Kingshott, Declan Chard, Alan J. Thompson, Frederik Barkhof, Daniel C. Alexander, and Olga Ciccarelli. Progression of regional grey matter atrophy in multiple sclerosis. *Brain*, 141:1665–1677, 6 2018.

[10] Alex Fornito, Andrew Zalesky, and Edward T. Bullmore. *Fundamentals of Brain Network Analysis*. Elsevier, 1 2016.

[11] Jean-Philippe Fortin, Nicholas Cullen, Yvette I. Sheline, Warren D. Taylor, Irem Aselcioglu, Philip A. Cook, Phil Adams, Crystal Cooper, Maurizio Fava, Patrick J. McGrath, Melvin McInnis, Mary L. Phillips, Madhukar H. Trivedi, Myrna M. Weissman, and Russell T. Shinohara. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, 167:104–120, 2018.

[12] Juichi Fujimori and Ichiro Nakashima. Early-stage volume losses in the corpus callosum and thalamus predict the progression of brain atrophy in patients with multiple sclerosis(prepint version available at research square https://doi.org/10.21203/rs.3.rs-2626591/v]). 3 2023.

[13] Louis Hofstetter, Yvonne Naegelin, Lukas Filli, Pascal Kuster, Stefan Traud, Renata Smieskova, Nicole Mueller-Lenke, Ludwig Kappos, Achim Gass, Till Sprenger, Iris Katharina Penner, Thomas E. Nichols, Hugo Vrenken, Frederik Barkhof, Chris Polman, Ernst Wilhelm Radue, Stefan J. Borgwardt, and Kerstin Bendfeldt. Progression in disability and regional gray matter atrophy in relapsing-remitting multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England)*, 20:202, 2 2014.

[14] Kurt A. Jellinger. Mild cognitive impairment in multiple system atrophy: a brain network disorder. *Journal of Neural Transmission*, 130:1231–1240, 10 2023.

[15] Sara Larivière, Raúl Rodríguez-Cruces, Jessica Royer, Maria Eugenia Caligiuri, Antonio Gambardella, Luis Concha, Simon S. Keller, Fernando Cendes, Clarissa Yasuda, Leonardo Bonilha, Ezequiel Gleichgerrcht, Niels K. Focke, Martin Domin, Felix Von Podewills,

Soenke Langner, Christian Rummel, Roland Wiest, Pascal Martin, Raviteja Kotikalapudi, Terence J. O'Brien, Benjamin Sinclair, Lucy Vivash, Patricia M. Desmond, Saud Alhusaini, Colin P. Doherty, Gianpiero L. Cavalleri, Norman Delanty, Reetta Kälviaïnen, Graeme D. Jackson, Magdalena Kowalczyk, Mario Mascalchi, Mira Semmelroch, Rhys H. Thomas, Hamid Soltanian-Zadeh, Esmaeil Davoodi-Bojd, Junsong Zhang, Matteo Lenge, Renzo Guerrini, Emanuele Bartolini, Khalid Hamandi, Sonya Foley, Bernd Weber, Chantal Depondt, Julie Absil, Sarah J.A. Carr, Eugenio Abela, Mark P. Richardson, Orrin Devinsky, Mariasavina Severino, Pasquale Striano, Domenico Tortora, Sean N. Hatton, Sjoerd B. Vos, John S. Duncan, Christopher D. Whelan, Paul M. Thompson, Sanjay M. Sisodiya, Andrea Bernasconi, Angelo Labate, Carrie R. McDonald, Neda Bernasconi, and Boris C. Bernhardt. Network-based atrophy modeling in the common epilepsies: A worldwide enigma study. *Science advances*, 6, 11 2020.

[16] Matteo Lenge, Simona Balestrini, Davide Mei, Letizia Macconi, Maria Eugenia Caligiuri, Valeria Cuccarini, Domenico Aquino, Federica Mazzi, Ludovico D'Incerti, Francesca Darra, Bernardo Dalla Bernardina, and Renzo Guerrini. Morphometry and network-based atrophy patterns in scn1a-related dravet syndrome. *Cerebral Cortex*, 33:9532–9541, 8 2023.

[17] Elisabet Lopez-Soley, Eloy Martinez-Heras, Magi Andorra, Aleix Solanes, Joaquim Radua, Carmen Montejo, Salut Alba-Arbalat, Nuria Sola-Valls, Irene Pulido-Valdeolivas, Maria Sepulveda, Lucia Romero-Pinel, Elvira Munteis, Jose E Martínez-Rodríguez, Yolanda Blanco, Elena H Martinez-Lapiscina, Pablo Villoslada, Albert Saiz, Elisabeth Solana, Sara Llufriu, and Cristina M Ramo-Tello. Personalized medicine dynamics and predictors of cognitive impairment along the disease course in multiple sclerosis. *J. Pers. Med*, 11:1107, 2021.

[18] Haiying Lyu, Xue Zhu, Naying He, Qing Li, Qianyi Yin, Yufei Huang, Fuhua Yan, Jun Liu, and Yong Lu. Alterations in resting-state mr functional connectivity of the central autonomic network in multiple system atrophy and relationship with disease severity. *Journal of Magnetic Resonance Imaging*, 58(5):1472–1487, 2023.

[19] Eloy Martinez-Heras, Elisabeth Solana, Francesc Vivó, Elisabet Lopez-Soley, Alberto Calvi, Salut Alba-Arbalat, Menno M. Schoonheim, Eva M. Strijbis, Hugo Vrenken, Frederik Barkhof, Maria A. Rocca, Massimo Filippi, Elisabetta Pagani, Sergiu Groppa, Vinzenz Fleischer, Robert A. Dineen, Barbara Bellenberg, Carsten Lukas, Deborah Pareto, Alex Rovira, Jaume Sastre-Garriga, Sara Collorone, Ferran Prados, Ahmed Toosy, Olga Ciccarelli, Albert Saiz, Yolanda Blanco, and Sara Llufriu. Diffusion-based structural

connectivity patterns of multiple sclerosis phenotypes. *Journal of neurology, neurosurgery, and psychiatry*, 2023.

[20] Rozanna Meijboom, Elizabeth N York, Agniete Kampaite, Mathew A Harris, Nicole White, Maria del C Valdés Hernández, Michael J Thrippleton, Peter Connick, David Hunt, Siddharthan Chandran, Adam D Waldman, and on behalf of the FutureMS Consortium. Patterns of brain degeneration in early-stage relapsing-remitting multiple sclerosis. *medRxiv*, 2022.

[21] Rozanna Meijboom, Elizabeth N. York, Agniete Kampaite, Mathew A. Harris, Nicole White, Maria del C. Valdés Hernández, Michael J. Thrippleton, N. J.J. MacDougall, Peter Connick, David P.J. Hunt, Siddharthan Chandran, and Adam D. Waldman. Patterns of brain atrophy in recently-diagnosed relapsing-remitting multiple sclerosis. *PLOS ONE*, 18:e0288967, 7 2023.

[22] Saba Naghavi, Fereshteh Ashtari, Iman Adibi, Vahid Shaygannejad, Neda Ramezani, Ahmad Pourmohammadi, Fariba Davanian, Zahra Karimi, Seyed-Mahdi Khaligh-Razavi, and Mehdi Sanayei. Effect of deep gray matter atrophy on information processing speed in early relapsing-remitting multiple sclerosis. *Multiple Sclerosis and Related Disorders*, 71:104560, 2023.

[23] Irene Pulido-Valdeolivas, Magí Andorrà, David Gómez-Andrés, Kunio Nakamura, Salut Alba-Arbalat, Erika J. Lampert, Irati Zubizarreta, Sara Llufriu, Eloy Martinez-Heras, Elisabeth Solana, Nuria Sola-Valls, María Sepulveda, Ana Tercero-Uribe, Yolanda Blanco, Anna Camos-Carreras, Bernardo Sanchez-Dalmau, Pablo Villoslada, Albert Saiz, and Elena H. Martinez-Lapiscina. Retinal and brain damage during multiple sclerosis course: inflammatory activity is a key factor in the first 5 years. *Scientific Reports 2020 10:1*, 10:1–11, 8 2020.

[24] Henning H. Rise, Synne Brune, Claudia Chien, Tone Berge, Steffan D. Bos, Magí Andorrà, Irene Pulido Valdeolivas, Mona K. Beyer, Piotr Sowa, Michael Scheel, Alexander U. Brandt, Susanna Asseyer, Kaj Blennow, Mads L. Pedersen, Henrik Zetterberg, Michel Thiebaut de Schotten, Maria Cellerino, Antonio Uccelli, Friedemann Paul, Pablo Villoslada, Hanne F. Harbo, Lars T. Westlye, and Einar A. Høgestøl. Brain disconnectome mapping derived from white matter lesions and serum neurofilament light levels in multiple sclerosis: A longitudinal multicenter study. *NeuroImage : Clinical*, 35:103099, 1 2022.

[25] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52:1059–1069, 9 2010.

[26] Elisabeth Solana, Eloy Martinez-Heras, Jordi Casas-Roma, Laura Calvet, Elisabet Lopez-Soley, Maria Sepulveda, Nuria Sola-Valls, Carmen Montejo, Yolanda Blanco, Irene Pulido-Valdeolivas, Magi Andorra, Albert Saiz, Ferran Prados, and Sara Llufriu. Modified connectivity of vulnerable brain nodes in multiple sclerosis, their impact on cognition and their discriminative value. *Scientific Reports 2019 9:1*, 9:1–8, 12 2019.

[27] Elisabeth Solana, Eloy Martinez-Heras, Elena H. Martinez-Lapiscina, Maria Sepulveda, Nuria Sola-Valls, Nuria Bargalló, Joan Berenguer, Yolanda Blanco, Magi Andorra, Irene Pulido-Valdeolivas, Irati Zubizarreta, Albert Saiz, and Sara Llufriu. Magnetic resonance markers of tissue damage related to connectivity disruption in multiple sclerosis. *NeuroImage: Clinical*, 20:161–168, 1 2018.

[28] Martijn D. Steenwijk, Marita Daams, Petra J.W. Pouwels, Lisanne J. Balk, Prejaas K. Tewarie, Jeroen J.G. Geurts, Frederik Barkhof, and Hugo Vrenken. Unraveling the relationship between regional gray matter atrophy and pathology in connected white matter tracts in long-standing multiple sclerosis. *Human brain mapping*, 36:1796–1807, 5 2015.

[29] Ee Ling Tan, Marlene Tahedl, Jasmin Lope, Jennifer C. Hengeveld, Mark A. Doherty, Russell L. McLaughlin, Orla Hardiman, Kai Ming Chang, Eoin Finegan, and Peter Bede. Language deficits in primary lateral sclerosis: cortical atrophy, white matter degeneration and functional disconnection between cerebral regions. *Journal of Neurology 2023*, 1:1–15, 9 2023.

[30] Artem Trufanov, Alexander Krasichkov, Alexey Polushin, Dmitry Skulyabin, Aleksandr Efimtsev, Igor Litvinenko, Evgeniya Kuznetsova, Dmitrii Medvedev, and Gennady Bisaga. Basal ganglia atrophy as a marker of multiple sclerosis progression. *Biomarkers in Neuropsychiatry*, 9:100073, 12 2023.

[31] Tong Wang, Bin Wang, Zhifeng Li, and Yunxiao Ma. The relationship between brain atrophy and brain network connection in alzheimer's disease. *2022 16th ICME International Conference on Complex Medical Engineering, CME 2022*, pages 131–134, 2022.

[32] Merlin M. Weeda, D. R. van Nederpelt, J. W.R. Twisk, I. Brouwer, J. P.A. Kuijer, M. van Dam, H. E. Hulst, J. Killestein, F. Barkhof, H. Vrenken, and P. J.W. Pouwels. Multimodal mri study on the relation between wm integrity and connected gm atrophy and its effect on disability in early multiple sclerosis. *Journal of Neurology*, 1:1–19, 9 2023.