Multimodal Assessment of Neurodegenerative Diseases

by

Ava Sheikhi Shoshtari

B.A., The University of British Columbia, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Biomedical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

June 2016

© Ava Sheikhi Shoshtari, 2016
Abstract

There is growing recognition that accurate assessment of brain function includes activity at multiple temporal and spatial scales. In this thesis, we explored ways to combine clinically-relevant imaging information derived from subjects with neurodegenerative disease.

In the first work, we investigated a two-step framework to determine both joint and unique biomarkers from structural and functional MRI in 18 healthy control (HC) and 12 Parkinson’s disease (PD) subjects. Three matrices (structural, functional, and structural/functional interactions) were derived from a subset of features in both modalities that were likely candidates for discrimination between PD and HC subjects. Finally, Least Absolute Shrinkage and Selection Operator (LASSO) regression was performed to determine if subjects’ clinical characteristics such as gender, smoking history, smell performance, Hoehn and Yahr Scale (H&Y Stage), and Unified Parkinson’s Disease Rating Scale (UPDRS) values, could be accurately predicted based on the imaging features. The results revealed that complementary biomarkers were most informative in predicting clinical scores in both groups.

In the second work, for analyzing imaging data from subjects with Multiple Sclerosis (MS), we employed a joint Multimodal Statistical Analysis Framework, a data fusion approach that used Latent Variables (LV). We studied fusion of information from seven different imaging modalities: Myelin Water Imaging (MWI), Diffusion Tensor Imaging (DTI), resting state functional MRI (rsfMRI), cortical thickness of the right and left hemisphere, MS lesion load, and normalized brain volume from 47 subjects with MS. Decomposed common and unique information in each modality were acquired and their relationships with disease duration (DD), the Expanded Disability Status Scale (EDSS), and age, were analyzed through LASSO
regression. We noted that common components of the seven modalities were the most accurate in predicting clinical indices. Results further revealed the regional importance of each modality by indicating a unique pattern of degeneration in MS and an asymmetry between the cortical thickness components in the two hemispheres.

Our results demonstrate the power of utilizing multimodal imaging biomarkers in neurodegenerative diseases. Since structural imaging data is acquired along with functional data, we propose that fusion of information from both types of data should become part of routine analysis.
Preface

This thesis is primarily based on a few manuscripts, resulting from collaboration between multiple researchers. In all publications, the contribution of the author was in developing, adapting, implementing, and evaluating the methods. All co-authors contributed in data acquisition, or evaluating and editing the manuscript.

The study described in Chapter 3 is based on work conducted by:


The contribution of the author was in developing, adapting, implementing, and evaluating the method. Sue-Jin Lin contributed in editing the manuscript. Profs. McKeown and Wang helped with their valuable suggestions in improving the methodology.

The study described in Chapter 4 is based on work conducted by:


The contribution of the author was in developing, adapting, implementing, and evaluating the method. Profs. McKeown and Wang helped with their valuable suggestions in improving the methodology. Profs. Riddehough, Tam, and Vavasour helped with data acquisition.

A version of the study described in Chapter 4 will be published in:

The contribution of the author was in developing, adapting, implementing, and evaluating the method. Sue-Jin Lin contributed in editing the manuscript. Profs. McKeown and Wang helped with their valuable suggestions in improving the methodology.
Table of Contents

Abstract ...................................................................................................................... ii
Preface ................................................................................................................... iv
Table of Contents .................................................................................................... vi
List of Tables ......................................................................................................... viii
List of Figures ....................................................................................................... ix
List of Abbreviations ............................................................................................ x
Acknowledgements .............................................................................................. xii
Dedication ................................................................................................................ xiii

Chapter 1: Introduction ............................................................................................ 1
  1.1 Data Analysis of Brain Imaging .................................................................... 1
  1.2 Multimodal Imaging Analysis ....................................................................... 3
  1.3 Proposed Framework .................................................................................... 3
  1.4 Structure of Thesis ....................................................................................... 6

Chapter 2: Datasets .................................................................................................. 9
  2.1 Parkinson’s Disease Dataset ....................................................................... 9
  2.2 Multiple Sclerosis Dataset ......................................................................... 9

Chapter 3: A Two-Step Approach Assessing Complementary Information from Structural and Functional Magnetic Resonance Imaging in Parkinson’s Disease ..................... 11
  3.1 Introduction ................................................................................................. 11
  3.2 Methods ..................................................................................................... 13
  3.3 Results ....................................................................................................... 21
  3.4 Discussion ................................................................................................. 27

Chapter 4: Multimodal Imaging Assessment of Multiple Sclerosis ...................... 29
  4.1 Introduction ................................................................................................. 29
  4.2 Methods ..................................................................................................... 33
  4.3 Results ....................................................................................................... 44
  4.4 Discussion ................................................................................................. 49

Chapter 5: Conclusion and Future Work .............................................................. 52
  5.1 Conclusion .................................................................................................. 52
  5.2 Future Work ............................................................................................... 53
List of Tables

Table 3.1. Demographics and clinical indices ........................................................................................................... 14
Table 3.2. PD cohort: depicted ROIs, to-from connections, and a combination of both that are highly correlated with clinical scores........................................................................................................ 23
Table 3.3. HC cohort: depicted ROIs, to-from connections, and a combination of both that are highly correlated with clinical scores........................................................................................................ 24
Table 4.1. MS Subject Demographics ..................................................................................................................... 34
Table 4.2. Bilateral regions of interest and their abbreviations ....................................................................................... 35
Table 4.3. Bilateral cortical thickness regions of interest ............................................................................................ 40
List of Figures

Figure 3.1. Proposed framework for the two-step assessment of complementary information from both structural and functional magnetic resonance imaging in Parkinson’s disease. ........................................... 13
Figure 3.2. Preprocessing pipeline............................................................................................................. 16
Figure 3.3. To- From rsfMRI Connectivity Map. L and R denote left and right hemisphere’s of the brain respectively. The red illustrates the desired unique connections. ......................................................... 18
Figure 3.4. PD cohort: LASSO prediction results for all three datasets (sMRI, rsfMRI connectivity, and interaction term). .................................................................................................................. 25
Figure 3.5. HC cohort: LASSO prediction results for all three datasets (sMRI, rsfMRI connectivity, and interaction term). .................................................................................................................. 26
Figure 4.1. FA label extraction process .................................................................................................... 36
Figure 4.2. MWF label extraction process ................................................................................................. 38
Figure 4.3. Regression for predicting clinical indices based on lesion load and normalized brain volume ............................................ .................................................................................................................................................................. 45
Figure 4.4. Regression results based on the combined lesion load and normalized brain volume ............ 45
Figure 4.5 LASSO components predicting DD ......................................................................................... 46
Figure 4.6 LASSO components predicting EDSS .................................................................................... 47
Figure 4.7 LASSO components predicting age.......................................................................................... 48
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADLs</td>
<td>Activities of Daily Life</td>
</tr>
<tr>
<td>BOLD</td>
<td>Blood Oxygen Level Dependent</td>
</tr>
<tr>
<td>BPF</td>
<td>Brain Parenchymal Fraction</td>
</tr>
<tr>
<td>CBG</td>
<td>Cortico-Basal Ganglia</td>
</tr>
<tr>
<td>CCA</td>
<td>Canonical Correlation Analysis</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>DD</td>
<td>Disease Duration</td>
</tr>
<tr>
<td>DTI</td>
<td>Diffusion Tensor Imaging</td>
</tr>
<tr>
<td>EDSS</td>
<td>Expanded Disability Status Scale</td>
</tr>
<tr>
<td>FA</td>
<td>Fractional Anisotropy</td>
</tr>
<tr>
<td>FLAIR</td>
<td>Fluid-Attenuated Inversion-Recovery</td>
</tr>
<tr>
<td>FLIRT</td>
<td>FMRIB’s Linear Image Registration Tool</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional MRI</td>
</tr>
<tr>
<td>GLM</td>
<td>Generalized Linear Model</td>
</tr>
<tr>
<td>HC</td>
<td>Healthy Control</td>
</tr>
<tr>
<td>HMAT</td>
<td>Human Motor Area Template</td>
</tr>
<tr>
<td>H &amp;Y Stage</td>
<td>Hoehn and Yahr Scale</td>
</tr>
<tr>
<td>IndFeat</td>
<td>Independent Significance Feature Test</td>
</tr>
<tr>
<td>LASSO</td>
<td>Least Absolute Shrinkage and Selection Operator</td>
</tr>
<tr>
<td>LV</td>
<td>Latent Variable</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini Mental State Examination</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>MS</td>
<td>Multiple Sclerosis</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>Multi-LV</td>
<td>Multidirectional LV</td>
</tr>
<tr>
<td>MWF</td>
<td>Myelin Water Fraction</td>
</tr>
<tr>
<td>MWI</td>
<td>Myelin Water Imaging</td>
</tr>
<tr>
<td>OSC</td>
<td>Orthogonal Signal Correction</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s Disease</td>
</tr>
<tr>
<td>PLS</td>
<td>Partial Least Square</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
</tr>
<tr>
<td>RRMS</td>
<td>Relapsing-Remitting MS</td>
</tr>
<tr>
<td>rsfMRI</td>
<td>Resting State Functional MRI</td>
</tr>
<tr>
<td>sMRI</td>
<td>Structural MRI</td>
</tr>
<tr>
<td>SNPc</td>
<td>Substantia Nigra pars compacta</td>
</tr>
<tr>
<td>SPM</td>
<td>Statistical Parametric Mapping toolbox</td>
</tr>
<tr>
<td>SPMS</td>
<td>Secondary-Progressive MS</td>
</tr>
<tr>
<td>subLV</td>
<td>Sub-LV</td>
</tr>
<tr>
<td>supLV</td>
<td>Super LV</td>
</tr>
<tr>
<td>UPDRS</td>
<td>Unified Parkinson’s Disease Rating Scale</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to thank the faculty, staff, and my fellow students and lab mates at the University of British Columbia (UBC), who have supported my work in this field. I owe particular thanks to my supervisor, Dr. Martin J. McKeown for his guidance, scientific as well as personal support, and dedication to research. He consistently allowed this thesis to be my own work, but steered me in the right direction whenever he thought I needed it. He truly inspires me in every aspect of life. I also would like to thank Prof. Wang, my co-supervisor, whose office was always open whenever I ran into a trouble spot or had a question about my research.

Furthermore, I like to state my appreciation for people who were involved in the validation of this research project: Dr. Andrew Riddehough, Dr. Roger Tam, Dr. Irene Vavasour, and Sue-Jin Lin. Without their passionate participation and input, this thesis could not have been successfully conducted.

Finally, I must express my very profound gratitude to my parents and my sister for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.
Dedication

I would like to dedicate my thesis to my beloved parents, *Banafsheh* and *Farhad*, and my lovely sister *Tina*. 
Chapter 1: Introduction

1.1 Data Analysis of Brain Imaging

With the growing accessibility to advance brain imaging technologies, it is becoming apparent that numerous neurological disorders, including Parkinson’s disease (PD) and Multiple Sclerosis (MS), are associated with alteration in both brain structure and the pattern of brain regional activity. These changes can be observed using both functional and structural Magnetic Resonance Imaging (fMRI, sMRI). In the past decades, studies have investigated brain functions associated to emotion, cognition, language, memory, and responses to many other external stimuli, along with resting-state brain function [1].

Magnetic Resonance Imaging (MRI) is the standard clinical imaging method that detects the relaxation properties of hydrogen atoms in strong magnetic fields in order to visualize three-dimensional views of the brain [2]. Based on the region of the brain that is being imaged, MR signal differs according to tissue type [1]. This enables the construction of structural images that distinguish grey matter, white matter, and cerebral spinal fluid [1].

Functional MRI (fMRI) is a variation of MRI that identifies fluctuations in local blood flow and hemoglobin oxygenation in response to different neuronal activity [2]. Studies show that when different areas of one’s brain are being used to perform a task, there is a related localized increase in oxygen intake and as a result oxygen-rich blood flow to the active brain areas [1]. The MR scanner archives the blood oxygenation level–dependent (BOLD) signals that differ with the degree of oxygenation [1]. As a result, fMRI allows the mapping of the functional centers of the human brain [2].
Our brain is a network consisting of spatially distributed, but functionally connected regions that constantly share information with each other [3]. Functional connectivity is defined as the chronological reliance of neuronal activation patterns of anatomically separated brain regions [4]. Recently, neuroimaging studies have started to focus on investigating functional connectivity by measuring the level of co-activation of resting-state fMRI time-series between brain regions [3]. Studies have shown interesting new findings about the functional connections of specific brain regions and local networks, in addition to significant new understandings in the overall organization of functional communication in the brain network [3].

Neuroimaging has developed influential noninvasive techniques in order to capture properties of the human brain in vivo [1]. Imaging studies disclose understandings about normal brain function and structure, neural processing and neuroanatomic manifestations of psychiatric and neurological disorders, and neural processing changes connected to treatment response [1]. Currently, with the vast volumes of data available for each patient and the complex temporal and spatial dependencies present in the collected data, there exists numerous challenges for linking and fusion of multimodal data [1].

Incorporating multimodal data not only gives important insights into brain processes and structures, but also provides spatiotemporal resolution complementarity [5]. Some merits of using multimodal datasets include a complete physiological view on brain processes and structures, quantification, generalization and normalization, along with availability of biomarkers [5].
1.2 Multimodal Imaging Analysis

In order to better understand complex neurodegenerative diseases such as PD and MS that affect structural, functional, and other aspects of the brain, it is important to determine features that are complementary to different modalities. In the past few decades, multi-modal approaches have been used in order to determine joint biomarkers combining information across different imaging modalities [6]. With a multimodal approach, complementary information about various tissue types with different characteristics as well as information about different temporal and spatial properties of the brain regions can be obtained. Multimodal analysis is majorly divided into two groups of data-integration and data fusion methods [7].

Data integration techniques utilize information from one imaging modality in order to improve the results from another modality. For example, there are a number of studies on EEG-fMRI integration for the study of human brain function that take advantage of the complementarity between the two signals [8].

Data fusion models on the other hand, take advantage of the combined information between different modalities used. There are a number of data-fusion multivariate methods that establish relationships between two types of data sets. In recent years, methods that take advantage of latent variable (LV) approaches have become popular [9] [10].

1.3 Proposed Framework

Multimodal analysis of brain imaging data is an effective approach to better understand brain diseases. The approach allows the identification of relationships among multiple data types and recognizes whether certain disease risk factors are common or discrete across multiple modalities. Here, we propose two multimodal approaches applied to two different
neurodegenerative diseases (PD, MS). The first approach uses data integration, in order to assess complementary information from sMRI and fMRI in PD. It aims to determine whether one type of imaging is more clinically informative than the other, or whether a combination of both can be significantly correlated with clinical indices. The second approach utilizes data fusion methods and proposes a joint Multimodal Statistical Analysis Framework. Using this data driven, multimodal, latent variable (LV) approach, common and unique information in seven different modalities (diffusion tensor imaging (DTI), myelin water imaging (MWI), rsfMRI, right and left hemisphere’s cortical thickness, normalized brain volume, and lesion load) were acquired, and subsequently fused to demonstrate the power of multimodal imaging markers.

1.3.1 Objective

The global objective of this thesis is to propose frameworks that form joint imaging biomarkers that are able to predict clinical scores. To this end, we proposed to use multimodal analysis on brain imaging data of two different neurodegenerative diseases, PD and MS, in order to take advantage of the complementary information that exist within the imaging modalities. This thesis develops new methods and optimizes and employs existing methods that are suitable for joint analysis of multimodal brain imaging data.

1.3.2 Contributions

This study investigates multimodal analysis techniques in the course of achieving complementary biomarkers that are more informative in predicting clinical scores of neurodegenerative diseases. The main contributions of the two studies are as follows:

- In the first study:
We have proposed a new two-step analysis to assess complementary information from structural and functional data in both PD and HC subjects. This method determines biomarkers from structural and functional MRI jointly, as well as biomarkers that are unique to each imaging modality.

With this study we demonstrate that complementary biomarkers are more informative in predicting clinical scores in both PD and HC cohorts.

Since structural data is routinely obtained when obtaining fMRI data, we propose that the combination of the two modalities may be more informative than analyzing each modality separately.

In the second study:

We have introduced a new way to combine brain information such as diffusion tensor imaging (DTI), myelin water imaging (MWI), rsfMRI, cortical thickness, normalized brain volume and lesion load within the framework. This application is novel in a sense that no previous study has employed the methods in this study in order to analyze multimodal imaging data in MS.

We have proposed a two-step registration analysis to further extract regions of interest from Myelin Water Fraction (MWF) and Fractional Anisotropy (FA) maps.

With this study we demonstrate that compared with the current regression approaches that are used in studies for combining two modalities, the methods used in this study for combining seven modalities result in stronger prediction of the clinical indices. Furthermore, results indicate a unique pattern of degeneration in MS and an asymmetry between the cortical thickness LASSO components in
left and right hemispheres. Observations reveal left cortical thickness components having a stronger effect on predicting the clinical scores. Our results show the power of multimodal imaging markers in MS.

1.4 Structure of Thesis

This thesis is further subdivided into the following 4 chapters.

Chapter 2: Data Cohorts

The findings of this thesis are based on data from two different diseased cohorts. Chapter Three investigates the PD cohort that also consists of age matched healthy controls (HCs) and Chapter Four analyzes the MS cohort that consists of only MS subjects, without controls. The imaging modalities utilized in both studies have been widely used in many studies in order to reflect properties of either brain structure or function [1]. In the next two sub-sections the dataset used for the two applications defined in the framework is described.

Chapter 3: A Two-Step Approach Assessing Complementary Information from Structural and Functional Magnetic Resonance Imaging in Parkinson’s Disease

Structural and functional magnetic resonance imaging (MRI) data are usually assessed independently in Parkinson’s disease (PD). We explored the complementary clinically-relevant information produced by both modalities by examining the relations between structural and functional imaging features and clinical indices. We propose a two-step framework that determines biomarkers from structural and functional MRI jointly, as well as biomarkers that are unique to each imaging modality. Structural and functional MRI data were recorded for 12 PD patients and 18 healthy control (HC) subjects as well as patient characteristics such as gender, smoking history, smell performance and Unified Parkinson’s Disease Rating Scale (UPDRS)
values. Volumetric values within Regions of Interest (ROIs) were obtained from the structural data, and instantaneous and time-lagged connectivity was extracted from resting state fMRI. As a preprocessing step, both types of MRI data subsequently underwent a preliminary feature selection algorithm, based on the ability to discriminate between PD and HC subjects. This subset of features was then used to form three matrices with three (structural, functional, and structural/functional interactions) groups of features. Finally, Least Absolute Shrinkage and Selection Operator (LASSO) Regression was performed to determine if clinical scores can be accurately predicted based on the imaging features. Resting state fMRI (rsfMRI) connectivity features alone accurately predicted UPDRS III and gender scores in the PD cohort, and smoking, gender, and smell differentiation in the HC cohort. Structural data alone predicted Hoehn and Yahr Scale (H&Y stage) and smell detection scores in the PD cohort but no indices in the HC cohort. The interaction between structural and functional features predicted UPDRS III, UPDRS, disease duration, H&Y stage, smell detection, and gender scores in the PD cohort, and gender, smoking, and smell differentiation scores in the HC cohort. With this study we concluded that complementary biomarkers are more informative in predicting clinical scores in both PD and HC cohorts. Since structural data is routinely obtained when obtaining fMRI data, the two modalities may be more informative than analyzing each modality separately.

Chapter 4: Multimodal Imaging Assessment of Multiple Sclerosis

Magnetic Resonance Imaging (MRI) biomarkers of multiple sclerosis (MS), particularly fluid-attenuated inversion-recovery (FLAIR) sequences, are considered standard of clinical care. However, advanced analytical methods, capable of fusing results from different MRI modalities, could potentially be more informative. Here we attempt to predict disease duration (DD), Kurtzke expanded disability status scale (EDSS), and age in MS subjects (n=47) based on fusion
of information from seven imaging modalities: Myelin Water Imaging (MWI), Diffusion Tensor Imaging (DTI), resting state functional MRI (rsfMRI), right & left hemisphere’s cortical thickness, lesion load, and normalized brain volume, by adapting a joint Multimodal Statistical Analysis Framework. Using this data driven, multimodal, latent variable (LV) approach, common and unique information in each dataset was acquired and their relationship with DD, EDSS, and age is analyzed through the Least Absolute Shrinkage and Selection Operator (LASSO) regression. The common components between the seven modalities, but not the unique components of each modality, accurately predicted DD and demonstrated a strong linear relationship with EDSS and age. To further investigate the regional importance of each modality, LASSO was performed on decomposed unique components of each modality. Results indicate a unique pattern of degeneration in MS and an asymmetry between the cortical thickness LASSO components in left and right hemispheres. Observations reveal left cortical thickness components having a stronger effect on predicting the clinical scores. Our results show the power of multimodal imaging markers in MS.

Chapter 5: Conclusion and Future Work

This chapter consists of an overall conclusive summary followed by suggestions for future work.
Chapter 2: Datasets

Methods proposed by this thesis were utilized for the analysis of two cohorts with neurodegenerative disease, in order to analyze the importance and strength of joint biomarkers and multimodal imaging. The following two sub-sections briefly describe the two datasets used.

2.1 Parkinson’s Disease Dataset

This dataset consists of 12 PD and 18 HC. All subjects in the cohorts are aged matched. Each subject’s dataset consists of their sMRI, rsfMRI, and clinical scores. Each PD patient’s clinical set consists of their age, Unified Parkinson’s Disease Rating Scale (UPDRS) values, duration of being diagnosed with the disease, Hoehn and Yahr Scale (H&Y Stage), the Mini Mental State Examination (MMSE) score, smoking history, and smell performance scores. All patients were right handed. Out of the 12 patients, 5 were females. Two of the male patients were smokers and two of the male patients with one of the female patients show initial symptoms of bradykinesia while the rest had predominately tremor. Similar to PD clinical scores, the chosen age matched control cohort was also right handed and 3 of the male subject were smokers. In this set, there were 10 males and 8 females. The collected clinical score of the control set consists of each subject’s age, MMSE score, smoking history, and smell performance scores.

2.2 Multiple Sclerosis Dataset

The MS dataset used in this thesis consisted of 47 MS subjects 36 of which are females and 11 subjects that are males. 43 of the patients had relapsing-remitting MS (RRMS) and 4 had secondary-progressive MS (SPMS). The set included each patient’s anatomical and functional MRI scans as well as demographic scores. The multimodal data used in this cohort included:
diffusion tensor imaging (DTI), myelin water imaging (MWI), rsfMRI, cortical thickness in both left and right hemispheres of the brain, normalized brain volume and lesion load.
Chapter 3: A Two-Step Approach Assessing Complementary Information from Structural and Functional Magnetic Resonance Imaging in Parkinson’s Disease

3.1 Introduction

Parkinson’s disease (PD) is a progressive neurodegenerative disorder that predominately affects movement, but also impacts cognition [11]. Pathologically, there is significant degeneration of dopamine-producing neurons in the Substantia Nigra pars compacta (SNpc), due to unknown causes [12]. When dopamine levels decrease, a number of signs and symptoms emerge, including tremor, bradykinesia (slowness of movement), rigidity (stiffness of muscles), impaired posture and balance, decreased ability to perform automatic/unconscious movements, and speech and writing difficulties [11].

Neurodegeneration within the central nervous system (CNS) is typically associated with cell loss, microglial proliferation, astroglial reaction, or increased age-related deposition of iron [13] – all features that typically happen at the microscopic level and may not be obvious in structural MRI (sMRI) scans during visual inspection, particularly in early disease stages. Nevertheless, with careful quantifiable methods, sMRI is able to detect and provide details on local changes in tissue volume and shape, offering the option of determining macroscopic cerebral lesions in a quantifiable approach [13]. sMRI allows clinicians and researchers to assess and visualize in vivo alterations of CNS structures with some regional distribution patterns distinctive to PD [13], and this may allow sMRI to be used as an early biomarker even before the onset of clinical
symptoms [14]. As the disease progresses, and people age, volumetric and geometrical changes of the structure in the brain may correlate with, e.g. axial features [15].

Resting state fMRI (rsfMRI) is a technique used to measure the correlation between spontaneous activation patterns of discrete brain regions [16]. Low-frequency (<0.1 Hz) blood oxygen level dependent (BOLD) fluctuations often display strong correlations at rest in grey matter regions [16]. Functional connections of specific brain regions and local networks can foster new understandings in the overall organization of functional communication in the brain network [16], [17]. A number of studies have looked at the relationship of connectivity in the brain with clinical scores using rsfMRI. For example, alterations of the cortico-basal ganglia (CBG) motor network in patients with PD off dopaminergic medication results in impaired information transfer within the CBG motor pathway [17].

While both sMRI and rsfMRI have been assessed in PD, they are not typically assessed jointly. Here we examine structural and functional MRI data in order to determine whether one type of imaging is more informative or whether the combination of both techniques are more predictive of clinical disease status. The study consisted of data collected from 12 PD and 18 healthy control (HC) subjects. The proposed novel framework (Fig. 3.1), details a two-step assessment of complementary information from both sMRI and fMRI in PD. The datasets form cohorts which undergo different structured analyses, and are pre-processed for further analyses. As part of the first step, the method identifies candidate discriminative features between the two cohorts. In the second step, with the features selected from before, an interaction term encapsulating both sMRI and rsfMRI was defined. Finally, the relationship of all three groups of features (sMRI features, rsfMRI features, and interaction terms) for both PD and HC cohorts was measured in order to
determine whether complementary multimodal imaging biomarkers were more beneficial in predicting clinical indices.

Figure 3.1. Proposed framework for the two-step assessment of complementary information from both structural and functional magnetic resonance imaging in Parkinson’s disease.

3.2 Methods

3.2.1 Demographics and Clinical Data

The dataset consisted of 30 age-matched subjects, 18 of whom were healthy controls without neurological disease. Similar to PD group, the age-matched control cohort were right-handed and
three of the male subjects were smokers. In this set, there were 10 males and 8 females. The clinical score of the control set consisted of each subject’s age, MMSE score, smell detection, and their smell differentiation scores.

The ages of PD and HC subjects were not significantly different (unpaired two samples t-test, p > 0.05) Clinical indices (Table 3.1), included age, UPDRS, UPDRS III, disease duration, H&Y Stage, MMSE score, smell detection and smell differentiation scores. All patients were right handed. Out of the 12 patients, 5 are females. Two of the male patients were smokers. Two of the male patients and one of the female patients showed initial symptoms of bradykinesia while the rest had predominantly tremor.

<table>
<thead>
<tr>
<th>Group</th>
<th>Sex M-F</th>
<th>Smoking</th>
<th>Initial Symptoms B-Bradykinesia T-Tremor</th>
<th>Age (Years) (Mean ± STD)</th>
<th>UPDRS (Mean ± STD)</th>
<th>UPDRS III (Mean ± STD)</th>
<th>Duration (Years) (Mean ± STD)</th>
<th>H&amp;Y Stage (Mean ± STD)</th>
<th>MMSE (Mean ± STD)</th>
<th>Smell Detection (Mean ± STD)</th>
<th>Smell Differentiation (Mean ± STD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>5 F 7 M</td>
<td>2 M</td>
<td>3 B (2 M, 1 F) 9 T (5 M, 4 F)</td>
<td>58.33 ± 10.30</td>
<td>48.66 ± 15.00</td>
<td>30.29 ± 9.88</td>
<td>8.33 ± 7.54</td>
<td>1.75 ± 0.72</td>
<td>29.33 ± 1.23</td>
<td>0.79 ± 0.83</td>
<td>2.48 ± 0.65</td>
</tr>
<tr>
<td>Controls</td>
<td>8 F 10 M</td>
<td>3 M</td>
<td>NA</td>
<td>58.17 ± 8.47</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>27.89 ± 2.40</td>
<td>-0.56 ± 0.87</td>
<td>1.18 ± 0.53</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1. Demographics and clinical indices.

We used a two-step procedure to assess the complementary information from the imaging data (sMRI + fMRI). The first step included pre-processing and feature extraction. The second step consisted of defining a complementary interaction term between the two sMRI and fMRI selected features and determining the relationship with clinical scores via LASSO regression.

3.2.2 Step 1: Preprocessing and Feature Selection

3.2.2.1 Voxel-Based Ratio Analysis

As seen in Figure 3.2, the corrected output label file from the open source program, FreeSurfer [19] was used to determine the number of voxels in each Region of Interest (ROI). 171 ROIs
were robustly segmented in PD and HC. Once the common labels were extracted from the
corrected output label file of FreeSurfer, the following ratio with n=171 labels, was determined
for further analysis:

$$\frac{v_i}{\sum_{k=1}^{n} v_k}, \text{where } i = 1, 2, \ldots, n$$ (3.1)

and $v_i$ is the number of voxels in the $i^{th}$ ROI.

Taking the ratio of each label for all 30 subjects, a matrix of 30 subjects x 171 label files was
prepared. Since the range of values was between 0 and 1, a logarithmic transform of: $\log(1 + \text{Ratio}_i)$ was performed.
Figure 3.2. Preprocessing pipeline.
3.2.2.2 Resting State fMRI (rsfMRI) Analysis

rsfMRI scans were acquired with scan resolution of 80x80 mm, 32 slices, 180 volumes and TR 2000 ms, for a total of 6 min. For each subject’s data, functional images were corrected for differences in slice acquisition times using the Statistical Parametric Mapping toolbox (SPM) [20] and each volume was re-sliced to generate isotropic voxels. Secondly, using the SPM Realign module, movement artifacts in the rsfMRI time-series were corrected. In parallel, registration label files were created for each subject’s anatomical files using the open source program FreeSurfer (Harvard, MA). Cortical labels derived from high resolution T1 weighted images were registered to mean functional image using FMRIB’s Linear Image Registration Tool (FLIRT) in FSL [21]. A transformation matrix of structural T1 to mean rsfMRI was calculated to correct and improve FreeSurfer registration label file estimations. The corrected labels along with labels from the Human Motor Area Template (HMAT) [22] were registered to the structural files and the time courses of previously identified ROIs were selected to be further analyzed (Fig. 3.2). A connectivity matrix was then derived based on the following method. In order to make a fair comparison between structural and functional features, we restricted ourselves to 56 ROIs that had been previously found to be important for connectivity in PD [23], [24].

3.2.2.2.1 Connectivity Matrix

Connectivity features were derived using the covariance between the time series obtained for different ROIs on a subject-by-subject basis. The mean time course (after linear detrending) from all voxels from each ROI was extracted to create a 180 time points x 56 ROIs matrix, $X$. This was then circularly shifted to create a lag-1 matrix (ignoring edge effects), $X_{-1}$. We then calculated the covariance matrix of the augmented matrix, $[X, X_{-1}]$. The unique covariance
matrix elements (see figure 3.3), were then collected together as a subject-specific row vector. This resulted in a functional connectivity matrix of 30 subjects x 4676 covariance elements.

Figure 3.3. To- From rsfMRI Connectivity Map. L and R denote left and right hemisphere’s of the brain respectively. The red illustrates the desired unique connections.

3.2.2.3 Feature Selection: Independent Significance Feature Test (IndFeat)

Since the size of the collected data from both structural data (30 subjects x 171 regions) and the functional data (30 subjects x 4676 regions) was large, we first eliminated features in both datasets that were unlikely to be informative. Using the Independent Significance Feature Test (IndFeat), a hypothesis test was carried out on each candidate predictor variable in each data set to gauge whether that predictor was likely to be informative [25]. Note that this method does not precisely select features to be used, rather it discards features that seem obviously useless [25]. This method assumes that the data are uncorrelated, the target variable is categorical, and that it can only be a two-class variable, which is satisfied in our case (PD vs Control).
IndFeat was performed on structural and functional data separately. The IndFeat algorithm uses two inputs: a matrix of predictor variables (examples in rows, variables in columns), and a target vector. The two matrices of predictor variables (structural data and functional data) were inputted into the IndFeat algorithm separately. The target vector was a binary vector of size 30 x 1, consisting of 18 controls with a binary value of 1 and 12 PDs with a binary value of 0. IndFeat uses

\[ s = \frac{\mu(X_{PD}) - \mu(X_N)}{\sqrt{\frac{\text{var}(X_{PD})}{n_{PD}} + \frac{\text{var}(X_N)}{n_N}}} \]  

(3.2)

where IndFeat returns a significance value (s) for each predictor [25]. A bigger value of s demonstrates more significance. Previous heuristics have suggested that features be retained only if they are at a value of 2.0 or higher [25].

In the structural dataset out of the 171 regions, 12 regions were found to be informative. For the functional dataset, 502 connections were found to be potentially useful.

3.2.3 Step Two: Interaction Term and Clinical Relations

In this step, an interaction term between both groups of extracted features (structural and functional) was defined, as we aimed to determine if one modality of imaging was more informative than the other, or that the combination of both was highly correlated with clinical indices. Least Absolute Shrinkage and Selection Operator (LASSO) Regression was performed on the three (structural, functional, and interaction) datasets to see correlations with the clinical scores of both cohorts, in order to further investigate the hypothesis-driven analysis.
3.2.3.1 Interaction Term

Selected features of structural (12 features) and functional (502 features) data were derived in the previous step. An interaction term based on the Kronecker product of each feature of structural data and the features of functional data was established [26]. Assume $S_i$ is the $i$th structural feature ($S$ is an $m \times p$ matrix) and $F_i$ is the $i$th functional feature ($F$ is a $m \times q$ matrix). The defined interaction matrix is of size $m \times nq$ and in the order of the pairwise interaction of $(S_1, F_1)$, $(S_1, F_2)$, ..., $(S_1, F_q)$, $(S_2, F_1)$, ..., $(S_p, F_q)$, where $m=30$ subjects, $n=12$ structural features, and $q=502$ functional features.

3.2.3.2 Least Absolute Shrinkage and Selection Operator (LASSO) Regression

LASSO is a variable selection technique for regression that minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant [28]. This method aids to improve the prediction accuracy when dealing with multicollinearity data while carrying interpretability and numerical stability of the data [27].

In a linear regression model with data $(x_{i1}, x_{i2}, \ldots, x_{ip}, y_i), i = 1, \ldots, n$, where $x_{ij}$ is a linear regressor, $y_i$ is the response variable of the $i$th observation, and $N$ is the number of observations, "LASSO" minimizes the residual sum of squares [28]:

$$\hat{\beta}_L = \arg \min \left\{ \frac{1}{2N} \sum_{i=1}^{N} (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{3.3}$$

subject to the sum of absolute values of the coefficient being less than a constant:

$$\sum_{j=1}^{p} |\hat{\beta}_j^L| \leq t \ (Constant) \tag{3.4}$$

LASSO sacrifices some bias into the model, in order to reduce the variance of the predicted values [28]. In equation 3.3, $\lambda$ is a nonnegative regularization parameter such that as $\lambda$ increases,
the number of nonzero components of $\beta$ decreases [28]. This will improve the overall prediction accuracy and induce sparsity to the coefficients so many of them are equal to zero.

In this paper, LASSO was applied to all three different groups of features (structural, functional, and interaction) to analyze the significance of each feature group in predicting clinical scores (response variable) for the PD and control groups. The MATLAB `lasso` function was used with 10-fold cross-validation in order to prevent overfitting. The value of lambda associated with the lowest mean square error (MSE) with 10-fold cross-validation was used. We then compared actual clinical scores with the predicted values:

$$C_{pred} = \text{ Intercept}(j) + X * \beta(i,j)$$

(3.5)

where $X$ is the input data set, once inputted with HC patients and once with PD separately, $N$ is the length of $X$, and $i= 1,2,\ldots,N$.

For the binary data set of PD and controls, the logistic LASSO model was used since the clinical response variable of these sets contained only ordinal data.

Significant components extracted via LASSO were then used as inputs of the Generalized Linear Model (GLM), in order to determine the significance value of each component in the two established groups.

### 3.3 Results

The LASSO results demonstrated that the interaction term defined between the two datasets better predicted clinical scores in both cohorts (figures 3.4 and 3.5). The selected LASSO components for each clinical score in the two cohorts for all 3 datasets were then placed into a GLM system and the significance of each component was determined. Components with a p-value $<0.05$ were determined to be of significance and as a results were chosen as predicting
biomarkers for each clinical score. The components were then mapped back onto each dataset in order to identify the ROIs, connections, or combination of ROIs and connections that they represent. These regions/connections acting as a biomarker are representative of the area of the brain that is related to each clinical score. Tables 3.2 and 3.3 provide the p-values of each depicted component in both cohorts for the three datasets. The results show that overall, the interaction term provided a better joint biomarker for predicting a combination of clinical scores in PD and HC cohorts (Fig. 3.4 and Fig. 3.5). The results indicate that connectivity from rsfMRI predicts UPDRS III and gender scores in the PD cohort and smoking, gender, and smell differentiation in the HC cohort. Structural data on the other hand only shows correlations with the H&Y stage and smell detection scores in the PD cohort and cannot predict scores in the HC cohort. However, the features of significance from the interaction data are correlated with UPDRS III, UPDRS, disease duration, H&Y stage, and gender scores in the PD cohort, and gender, smoking, and smell differentiation scores in the HC cohort.
<table>
<thead>
<tr>
<th>Clinical Score</th>
<th>sMRI Components (ROI)</th>
<th>p-Value</th>
<th>rsfMRI Components (Connections)</th>
<th>p-Value</th>
<th>Interaction Term Components (ROI +Connections)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPDRS III</td>
<td>Lag 1: Left Amygdala &amp; Right Pre-SMA*&lt;sup&gt;(1)&lt;/sup&gt;</td>
<td>0.0011</td>
<td>Ctx_lh_S_oc_sup_and_transversal &amp; Lag 1: Left Amygdala &amp; Right Pre-SMA</td>
<td>1.3898&lt;sup&gt;4&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ctx_lh_S_oc_sup_and_transversal &amp; Lag 1: Left Amygdala &amp; Right Pre-SMA</td>
<td>4.9107&lt;sup&gt;4&lt;/sup&gt;</td>
</tr>
<tr>
<td>Duration</td>
<td></td>
<td></td>
<td>Ctx_lhG_Inslg_and_scent_ins &amp; Lag 1: Left Posteriorcingulate cortex &amp; Right Cerebellum cortex Ctx_lh_S_oc_sup_and_transversal &amp; Lag 1: Left M1&lt;sup&gt;(2)&lt;/sup&gt; &amp; Right Occipital Parietal Visual Association Area</td>
<td>0.0064</td>
<td>Ctx_lhG_Inslg_and_scent_ins &amp; Lag 1: Left Posteriorcingulate cortex &amp; Right Cerebellum cortex Ctx_lh_S_oc_sup_and_transversal &amp; Lag 1: Left M1&lt;sup&gt;(2)&lt;/sup&gt; &amp; Right Occipital Parietal Visual Association Area</td>
<td>0.0118</td>
</tr>
<tr>
<td>H&amp;Y Stage</td>
<td>Ctx_lhG_precuneus Ctx_lh_S_oc_sup_and_transversal Ctx_rh_S_collat_transv_ant</td>
<td>0.0108 5.5426&lt;sup&gt;5&lt;/sup&gt; 0.0202</td>
<td>Ctx_lhG_precuneus &amp; Lag 1: Posteriorcingulate cortex &amp; Right Cerebellum cortex Ctx_lh_S_oc_sup_and_transversal &amp; Lag 1: Right Caudalanteriorcingulate and Right vmPFC&lt;sup&gt;(3)&lt;/sup&gt;</td>
<td>0.0302 0.0118</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lag 1: Right Hippocampups &amp; Left VLPFC&lt;sup&gt;(4)&lt;/sup&gt; Lag 1: Left Pre-SMA &amp; Left lateraloccipital cortex Lag 1: Right VLPFC &amp; Left Occipital Temporal Gyrus</td>
<td>0.0356 0.0242 0.0107</td>
<td>Ctx_lhG_precuneus &amp; Lag 1: Right Pre-Motor Cortex &amp; Right Pre-SMA</td>
<td>0.0428</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Lag 1: Right Hippocampups &amp; Left VLPFC&lt;sup&gt;(4)&lt;/sup&gt; Lag 1: Left Pre-SMA &amp; Left lateraloccipital cortex Lag 1: Right VLPFC &amp; Left Occipital Temporal Gyrus</td>
<td>0.0356 0.0242 0.0107</td>
<td>Ctx_lhG_precuneus &amp; Lag 1: Right Pre-Motor Cortex &amp; Right Pre-SMA</td>
<td>0.0428</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smell Detection</td>
<td>Ctx_lh_S_oc_sup_and_transversal</td>
<td>0.0222</td>
<td>Ctx_lh_S_oc_sup_and_transversal &amp; Lag 0: Left Superior Frontal Gyrus &amp; Left Frontal Middle Gyrus</td>
<td>0.0434</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2. PD cohort: depicted ROIs, to-from connections, and a combination of both that are highly correlated with clinical scores.

*<sup>(1)</sup> Pre-SMA: Pre- Supplementary Motor Area
*<sup>(2)</sup> M1: The sphenoidal segment covered by the middle cerebral artery in the cerebrum
*<sup>(3)</sup>: vmPFC: Ventromedial Prefrontal Cortex
Table 3.3. HC cohort: depicted ROIs, to-from connections, and a combination of both that are highly correlated with clinical scores.

* VLPFC: Ventrolateral Prefrontal Cortex

<table>
<thead>
<tr>
<th>Clinical Score</th>
<th>sMRI Components (ROI)</th>
<th>rsfMRI Components (Connections)</th>
<th>Interaction Term</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Lag 0: Left Hippocampus and Right Parahippocampal cortex</td>
<td>9.347e^-4</td>
<td>Ctx_lh_S_oc_sup_and_transversal &amp; Lag 1: Left Amygdala &amp; Right Pre-SMA</td>
<td>1.3898e^-4</td>
</tr>
<tr>
<td>Smoking</td>
<td>Lag 1: Right Pre-Motor Cortex &amp; Right Precuneus Cortex</td>
<td>8.5579e^-4</td>
<td>Ctx_lh_G_precuneus &amp; Lag 1: Left Pallidum &amp; Right Caudalanterior-cingulate Cortex</td>
<td>0.0038</td>
</tr>
<tr>
<td>Smell Differentiation</td>
<td>Lag 0: Left Occipital Parietal Visual Association Area &amp; Left Putamen</td>
<td>0.0189</td>
<td>Ctx_lh_G_ins_lg_and_S_cent_ins &amp; Lag 1: Right Superior Parietal Cortex &amp; Right Caudalanterior-cingulate Cortex</td>
<td>0.0067</td>
</tr>
</tbody>
</table>

* (4): VLPFC: Ventrolateral Prefrontal Cortex

* (5): S1: Primary Somatosensory Cortex
Figure 3.4. PD cohort: LASSO prediction results for all three datasets (sMRI, rsfMRI connectivity, and interaction term).
Figure 3.5. HC cohort: LASSO prediction results for all three datasets (sMRI, rsfMRI connectivity, and interaction term).
3.4 Discussion

Consistent with our results, a number of prior studies of neurodegenerative diseases have demonstrated the value of joint biomarkers in predicting clinical scores. For example, joint evaluation of brain structure and function via MRI yields higher classification accuracy of prodromal Alzheimer’s disease than either alone [29]. We found that neither sMRI nor fMRI could be used to predict UPDRS and disease duration in the PD cohort; however, the interaction of both modalities was superior (Fig. 3.4).

In healthy controls, our results are consistent with a number of prior studies. We found that rsFMRI and the interaction between structural and rsfMRI were roughly equivalent in detecting gender, smoking and smell differentiation. rsFMRI differences have been previously demonstrated with respect to gender [52], and smoking [51]. Most olfactory studies have emphasized localization of olfactory cortex [50] rather than connectivity changes, but connectivity changes have also been demonstrated [49].

In PD subjects the interaction term was able to predict H&Y stage, disease duration, UPDRS and UPDRS-III, and gender, while rsFMRI predicted UPDRS III and gender scores. A recent study found gender differences in $^{[123]I}$FP-CIT single-photon emission computed tomography binding in the caudate in PD and so altered connectivity differences appear plausible [30]. The UPDRS contains information regarding mentation, behavior, and mood, activities of daily life (ADLs) in addition to clinician-assessed motor performance (i.e., UPDRS-III) which may explain why both structural and fMRI results are needed for its accurate prediction.

The prominent clinically-relevant structural changes we observed that included occipital regions are perhaps surprising. Previous studies have suggested prominent posterior atrophy in
Parkinson’s disease with dementia [35], but none of our subjects had frank dementia (as it was an exclusion criterion). This suggests that the interaction of early structural posterior atrophy and altered functional connectivity may represent an early and sensitive biomarker for disease.

One of the limitations of this study is that subjects were recruited from a single site in order to prevent multi-site scanner effects; however, in further studies subjects could be recruited from multiple sites in order to increase statistical power. Furthermore, future studies may employ a larger sample size as the samples used in this study is relatively small and the results may represent an unusual sample of subjects. In order to mitigate sample size effects, cross-validation is used throughout the framework.
Chapter 4: Multimodal Imaging Assessment of Multiple Sclerosis

4.1 Introduction

Multiple sclerosis (MS) is an inflammatory, demyelinating condition of the central nervous system (CNS), generally considered to be autoimmune in nature [33]. This condition affects the functionality of the brain, spinal cord, and the optic nerves in eyes, causing problems with vision, balance, muscle control, and other brain functions [34]. In MS, the immune system attacks a fatty material (myelin), which acts as a sheath around nerve axons. Without this outer shell, nerve axons are more likely to be damaged and the neuronal signals may not be propagated appropriately [34].

MRI has become a powerful tool for detecting demyelination of the central nervous system [36]. Different MRI sequences, such as fluid-attenuated inversion-recovery (FLAIR), have been proposed for routine use, as they are particularly sensitive for detecting demyelination. Accurate assessment of MS is important, as clinical symptoms can be vague, and MRI may provide a quantitative metric of disease severity and likely duration. Disease modifying therapy is usually more effective if given early in the disease [37]. As a result, predicting the duration of the disease is of importance.

Other MRI methods, such as Diffusion Tensor Imaging (DTI), have also been studied to assess MS. The principal application of DTI is in the imaging of white matter where the location, orientation, and anisotropy of tracts can be measured [39]. In fibrous tissues including white matter, water diffusion is greatest in the direction parallel to the fiber orientation; however, water
diffusion is highly restricted and hindered in the directions perpendicular to the fibers [40]. DTI has been shown to be an effective measure of demyelination and axonal loss [41], and can detect alterations in white matter fibers before being visible in conventional MRI [42], making it useful for prognostic measures [41].

Over the past few years, researchers have consistently found an increased diffusivity, and decreased Fractional Anisotropy (FA) compared with normal appearing white matter, in lesions visible in T2-weighted MRI [43], [44], [45]. Several studies have explored the relations between MRI measures and disease duration. Fractional Anisotropy may correlate with disease duration in early MS patients [46], [47], while another study found that FA reduction is similar in patients with early and moderate disease duration periods [47]. Studies have also explored the relationship between FA & EDSS and have found significant associations between FA of the corticospinal tracts and EDSS [48].

MRI-visible brain lesions that become apparent with progressive disease and lesion load measured early in the disease are associated with future relapses in patients with relapsing-remitting MS (RRMS) [54], [55], [56] disability accumulation [54], [57] and cognitive deficits [58], [54].

Resting state fMRI (rsfMRI) is a more recent technique used to obtain a better understanding of functional impairments of MS subjects. Studies generally use rsfMRI to measure the correlation between spontaneous activation patterns of brain regions [3]. Low-frequency (<0.1 Hz) Blood Oxygen Level Dependence (BOLD) fluctuations often display strong correlations at rest even in distant grey matter regions [59]. Because it is noninvasive and does not require active patient engagement, rsfMRI may be particularly useful in patients who are not able to undergo currently
available methods for lesion localization [61]. In MS, rsfMRI has illustrated a negative correlation between disease duration and inter-thalamic connectivity [38]. Another fMRI study revealed a lack of activation in frontal regions was correlated with longer disease duration in MS [39]. A link between impaired regional integration in the cerebellum and increasing global disability (EDSS) has been suggested [62].

Another emerging imaging technique that is used in MS studies is Myelin Water Fraction (MWF), which attempts to quantitate the amount of water within myelin versus the total amount of water in the tissue scanned [63]. MWF is defined as the fraction of a T2 MR signal below 30 milliseconds divided by the total signal. This quantitative technique has many practical applications for the in vivo monitoring of demyelination and remyelination in a variety of disorders of myelin such as MS [64]. This modality has also shown significant correlation results with disease duration [63], [65], and EDSS [66].

Cortical thickness can be examined either in local terms or as a global average for the entire brain [67]. Cortical thickness is determined on the basis of the grey matter set in neuroimaging data, usually from the local or average distance between the white matter surface and the pial surface [67]. Studies show that this measure, changes minimally with brain size both within and across species [67]. The typical values in adult humans are between 1.5 and 3 mm [68]. Researchers have determined that during aging, a cortical thinning which is associated with a decrease on the order of about 10 μm per year is observed in healthy subjects [68]. Furthermore, cortical thinning is observed in multiple cortical regions in MS subjects [69]. In a comprehensive analysis of the global and regional values of cortical thickness based on 3D MRI, a modest age- dependent correlation was observed between cortical thickness and clinical measures that included the disease duration, and EDSS [69].
New MRI practices and advances in software development present an opportunity to include brain atrophy measurements in the regular management of MS patients [70]. Measurements of alterations in normalized brain volume and brain parenchymal fraction (BPF) over time are among the best-studied techniques for quantifying neurodegeneration in MS [71]. Research shows that brain atrophy takes place more rapidly in patients with MS than in healthy controls [71]. Findings demonstrate shrinkage of 0.5–1% per year in brain volume for patients with MS and 0.1–0.3% in healthy individuals which seems to correlate with measures of disability [71].

Despite that numerous studies have investigated the relationships of different MRI modalities with different demographical clinical scores, to our knowledge, there has not been a study that combines multiple modalities (e.g. 7 modalities in this case) and analyzes the multimodal data jointly. While there are a number of data fusion multivariate methods in the literature that establish relationships between two types of data sets, extending this beyond pairwise comparisons is less common. Partial least square (PLS) and canonical correlation analysis (CCA) methods, both latent variable (LV) approaches, have been proposed for pair-wise data fusion [72], [73]. Since real-world raw data are often high dimensional and collinear, applying CCA directly to the data set can result in numerical instability. As a result, PLS has been used more frequently than CCA [74]. PLS attains the co-variation between the predictor and the response variables, finding a set of LVs that are maximally related to each other [72]. There are, however, two major concerns with PLS and its extended applications. First, the framework is modeled so that it is only able to process two data sets jointly [74]. Secondly, there is information mixing concern with PLS, so that during the modeling process information outside the space may be introduced that contaminates the model [74], [75]. PLS regresses one data set (X₁) on the other (X₂) in unidirectional fashion. As a result, the variations are modeled in the X₁ space and then
used to predict $X_2$ by passing the information contributed by $X_2$ to $X_1$ [74]. However, in the deflation procedure of PLS, only the information contained by $X_1$ is used and once $X_1$ is deflated, the updated data still stay in the original space, while the deflated $X_2$ does not automatically lie in the original space [74].

In order to address the aforementioned concerns, here we employ a technique similar to that proposed by Chen et al [74]. The method relates to multiple data sets by modeling multiple data spaces simultaneously in a multidirectional fashion, unlike conventional approaches [74]. The algorithm consists of a two-step modeling strategy: In the first step, a multidirectional LV extraction solution (multi-LV extraction) is employed through an optimization problem (see methods below) [74]. In the second step, a joint post-processing is utilized on the extracted LVs in order to acquire common and specific information in each data space [74].

The relationship between DD, EDSS, and age and the decomposed common and unique information of each modality for all subjects can then be analyzed through LASSO and informative features selected.

### 4.2 Methods

The dataset used in this paper consists of 47 MS subjects (36 females and 11 males) with either relapsing-remitting MS (RRMS) or secondary-progressive MS (SPMS). The set consists of each patient’s anatomical and functional MRI scans as well as demographic scores (Table 4.1).
<table>
<thead>
<tr>
<th>Number of Subjects</th>
<th>Age (Years) (Mean ± STD)</th>
<th>Gender (M=Male, F= Female)</th>
<th>Disease Duration (Months) (Mean ± STD)</th>
<th>MS Type (RRMS, SPMS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>42.83 ± 10.76</td>
<td>36 F</td>
<td>136.98 ± 103.76</td>
<td>43 RRMS 4 SPMS</td>
</tr>
</tbody>
</table>

Table 4.1. MS Subject Demographics

4.2.1 Resting-State fMRI (rsfMRI)

Eyes-closed resting state fMRI scans were acquired with resolution 80x80 mm, 36 slices, 240 volumes and TR 2000 ms (8 min). For each subject’s data, primarily the functional images were corrected for differences in slice acquisition times using the Statistical Parametric Mapping toolbox (SPM) [20] and later each volume was re-sliced to generate isotropic voxels. Secondly, using the SPM Realign module, movement artifacts in the resting-state fMRI time-series were corrected. In parallel, registration label files were created for each subject’s anatomical files using the open source program FreeSurfer (Harvard, MA). Cortical labels derived from high resolution T1 weighted images were registered to mean functional image using FMRIB’s Linear Image Registration Tool (FLIRT) in FSL [21]. A transformation matrix of structural T1 to mean rsfMRI was calculated to correct and improve FreeSurfer registration label files estimations. The corrected labels along with labels from the Human Motor Area Template (HMAT) [22] were registered to the structural files and the time courses of previously identified regions of interest (ROIs) were selected to be further analyzed. Selected ROIs are listed in table 4.2.

Connectivity features were derived using covariance between different ROIs on a subject-by-subject basis (see section 3.2.2.2.1). From this step 1926 unique connections were obtained for each subject.
### Bilateral regions-of-interest

<table>
<thead>
<tr>
<th>Superior frontal gyrus (SFG)</th>
<th>Occipital- parietal visual area (OPV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medial frontal gyrus (MFG)</td>
<td>Lateral occipital lobe (LOL)</td>
</tr>
<tr>
<td>Inferior prefrontal cortex (IPC)</td>
<td>Anterior cingulate cortex (ACC)</td>
</tr>
<tr>
<td>Temporal Pole, insula cortex and amygdala (TIA)</td>
<td>Posterior cingulate cortex (PCC)</td>
</tr>
<tr>
<td>Superior temporal cortex (STC)</td>
<td>Precuneus (P)</td>
</tr>
<tr>
<td>Posterior parietal cortex (PPC)</td>
<td>Medial orbitofrontal cortex (MOC)</td>
</tr>
<tr>
<td>Post central cortex (PC)</td>
<td>Lateral orbitofrontal cortext(LOC)</td>
</tr>
<tr>
<td>Supramarginal region (SR)</td>
<td>Fusiform gyrus (FG)</td>
</tr>
<tr>
<td>Medial temporal lobe, hippocampus and parahippocampal gyrus (MHPG)</td>
<td>Superior parietal cortex (SPC)</td>
</tr>
</tbody>
</table>

Table 4.2. Bilateral regions of interest and their abbreviations.

#### 4.2.2 Diffusion Tensor Imaging (DTI)

The FSL toolbox was used to pre-process the data and extract the FA maps to deal with distortion from a combination of eddy currents, subject motion, and magnetic field inhomogeneities.

4.2.2.1 FA-Label Extraction

In order to extract Regions of Interest (ROIs), the extracted FA maps were processed through a two steps registration process using FSL’s FLIRT function (Fig. 4.1). Primarily, a standard FA map (input) was registered to the extracted FA map of each subject (reference) in order to obtain a transformation matrix. In the second step, white matter labels (JHU-ICBM), using nearest
neighbor interpolation, were registered to the extracted FA maps of each subject (reference) while taking into account the weight of the transformation matrix obtained in the previous step. In this step an affine transformation was used to correct for geometric distortion or deformation.

48 common ROIs between all subjects were selected. The mean of the FA map at each ROI for every subject was then calculated and used in further analysis.

Figure 4.1. FA label extraction process.
4.2.3 Myelin Water Fraction (MWF)

4.2.3.1 MWF-Label Extraction

ROIs were extracted from the MWF maps in the same manner as the FA maps. Using FSL’s FLIRT function, primarily, the JHU-ICBM-T2 white matter standard (input) was registered to the T2 weighted extracted first echo (reference) in order to obtain a transformation matrix. In the second step, same as before, white matter labels (JHU-ICBM), using nearest neighbor interpolation, were registered to the extracted MWF maps of each subject (reference) while taking into account the weight of the transformation matrix obtained in the first step (Fig. 4.2). In addition, in this step an affine transformation was used to correct for geometric distortion or deformation and 35 common ROIs between all subjects were selected. Similar to FA, the mean of the MWF map at each ROI, for every subject was then calculated and used in further analysis.
Figure 4.2. MWF-Label extraction process.
4.2.4 Lesion Load & Normalized Brain Volume & Cortical Thickness

A preprocessing was done on FLAIR images in order to reduce MR inhomogeneity using a modified version of N3 [76] and structure-preserving noise removal using FSL SUSAN [77]. Secondly, marking of each lesion with one or more seed points was done by a radiologist and an automatic segmentation of all seeded lesions using Parzen windows, with heuristics to reduce false positives was established [78]. Finally the images were visually reviewed for quality control.

Whole brain atrophy was measured with respect to brain surface detection and brain volume calculation. Based on the established segmentation, a normalized measure of atrophy, the brain parenchymal fraction (BPF), was calculated as the ratio of brain parenchymal tissue volume to the total volume contained within the brain surface contour [79], [60].

Furthermore, cortical thickness was obtained from the structural data for both left and right hemisphere of each patient’s brain (see Table 4.3 for bilateral ROIs).
### Bilateral Cortical Thickness ROIs

<table>
<thead>
<tr>
<th>Region</th>
<th>Left Hemisphere</th>
<th>Right Hemisphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>bankssts</td>
<td>G_roc-temp.latifus</td>
<td>S_circular_insula_inf</td>
</tr>
<tr>
<td>caudalanteriorcingulate</td>
<td>G_roc-temp.med-Lingual</td>
<td>S_circular_insula_sup</td>
</tr>
<tr>
<td>cingulummedialfrontal</td>
<td>G_roc-temp.med-Parahip</td>
<td>S_collat_transv_ant</td>
</tr>
<tr>
<td>cuneus</td>
<td>G_orbital</td>
<td>S_collat_transv_post</td>
</tr>
<tr>
<td>entorhinal</td>
<td>G_pariet_inf-Angular</td>
<td>S_front_inf</td>
</tr>
<tr>
<td>fusiform</td>
<td>G_pariet_inf-Supramar</td>
<td>S_front_middle</td>
</tr>
<tr>
<td>inferiorparietal</td>
<td>G_parietal_sup</td>
<td>S_front_sup</td>
</tr>
<tr>
<td>inferiortemporal</td>
<td>G_postcentral</td>
<td>S_interm_prim-Jensen</td>
</tr>
<tr>
<td>isthmuscingulate</td>
<td>G_parcenral</td>
<td>S_intrapariet_and_P_trans</td>
</tr>
<tr>
<td>laterooccipital</td>
<td>G_precuneus</td>
<td>S_occ_middle_and_Lunatus</td>
</tr>
<tr>
<td>lateralorbitofrontal</td>
<td>G_rectus</td>
<td>S_occ_sup_and_transversal</td>
</tr>
<tr>
<td>lingual</td>
<td>G_subcallosal</td>
<td>S_occipital_ant</td>
</tr>
<tr>
<td>medialorbitofrontal</td>
<td>G_temp_sup-G_T_transv</td>
<td>S_occ-temp_lat</td>
</tr>
<tr>
<td>middletemporal</td>
<td>G_temp_sup-Lateral</td>
<td>S_occ-temp_med_and_Lingual</td>
</tr>
<tr>
<td>parahippocampal</td>
<td>G_temp_sup-Plan_polar</td>
<td>S_orbital_lateral</td>
</tr>
<tr>
<td>paracentral</td>
<td>G_temp_sup-Plan_tempo</td>
<td>S_orbital_med-olfact</td>
</tr>
<tr>
<td>parasphericus</td>
<td>G_temporal_inf</td>
<td>S_orbital_H_Shaped</td>
</tr>
<tr>
<td>parstriangular</td>
<td>G_temporal_middle</td>
<td>S_parieto_occipital</td>
</tr>
<tr>
<td>pericalcarine</td>
<td>Lat_Fis-ant-Horizont</td>
<td>S_pericallosal</td>
</tr>
<tr>
<td>postcentral</td>
<td>Lat_Fis-post</td>
<td>S_postcentral</td>
</tr>
<tr>
<td>posteriorcingulate</td>
<td>Pole_occipital</td>
<td>S_precentral-inf-part</td>
</tr>
<tr>
<td>precuneus</td>
<td>Pole_temporal</td>
<td>S_precentral-sup-part</td>
</tr>
<tr>
<td>rostralanteriorcingulate</td>
<td>S_calcarine</td>
<td>S_suborbital</td>
</tr>
<tr>
<td>rostralmiddlefrontal</td>
<td>S_cingul-Marginalis</td>
<td>S_subparietal</td>
</tr>
<tr>
<td>superiorfrontal</td>
<td>S_circular_insula_ant</td>
<td>S_temporal_inf</td>
</tr>
</tbody>
</table>

*Table 4.3. Bilateral cortical thickness regions of interest.*

**4.2.5 Multidirectional Latent Variable Extraction (Multi-LV)**

In this analysis, we propose to employ a Multidirectional LV (Multi-LV) analysis [74] in order to extract eight sets (47 subjects x 32 LV ~90% of the explained variance by the data set) of latent variables: 1 common set among the seven modalities and 7 unique to each modality. LASSO Regression is then performed on the highly correlated decomposed extracted components from the Multi-LV method in order to determine whether the common or the specific information in each data space provides clinically relevant information.

For the purpose of this analysis, we define a space where both the common systematic variations for all datasets and the unique systematic variations in each data space are represented [74]. For this purpose, this method defines sub-latent variables (subLVs) in each modality inputted, that
are linear combinations of the original variables. For instance, presume there are \( m \) (in this analysis 3) modalities \( \mathbf{X}_1 \) with size \( N \times P_1 \), \( \mathbf{X}_2 \) with size \( N \times P_2 \), \ldots, and \( \mathbf{X}_m \) with size \( N \times P_m \), then the subLVs will be \( \mathbf{X}_1 \mathbf{w}_1 \), \( \mathbf{X}_2 \mathbf{w}_2 \), \ldots, \( \mathbf{X}_m \mathbf{w}_m \). This algorithm also characterizes a variable \( \mathbf{t}_g \) that is a super latent variable (supLV) and represents a relationship between the established subLVs, relating them simultaneously. \( \mathbf{t}_g \) is obtained through the following optimization problem:

\[
\begin{align*}
\max_{\mathbf{t}_g, \mathbf{w}_i} & \sum_{i=1}^{m} \left( \mathbf{t}_g^T \mathbf{X}_i \omega_i \right)^2, \\
\text{s.t.} & \quad \mathbf{t}_g^T \mathbf{t}_g = 1, \quad \omega_i^T \omega_i = 1, \forall i = 1, 2, \ldots, m.
\end{align*}
\]

(4.1)

In the above equation, all columns \( \mathbf{X}_i \) for \( m \) data matrices are normalized to unit variance in advance and are assumed to be zero-mean [74]. According to the above optimization problem it is expected that the extracted subLVs, \( \mathbf{X}_i \mathbf{w}_i \), carry as many variations as possible in each modality and at the same time be correlated with each other as closely as possible [74]. In addition, \( (\mathbf{t}_g^T \mathbf{X}_i \mathbf{w}_i)^2 \) models the covariance information that is between each subLV and the supLV, \( \mathbf{t}_g \) [74]. The subLVs in this algorithm are characterized through the following equation:

\[
t_i = \sqrt{\frac{1}{\mathbf{t}_g^T \mathbf{X}_i \mathbf{X}_i^T \mathbf{t}_g}} \mathbf{X}_i \mathbf{X}_i^T \mathbf{t}_g, \\
\forall i = 1, 2, \ldots, m.
\]

(4.2)

The number of subLVs is determined by the ratio of explained variance, in this paper 90%, of all 47 subjects for the seven modalities [74]. The algorithm also modifies the subLVs so that each modality is orthogonal to the other ones and is ordered according to the descending average correlation [74]. This also ensures that \( \mathbf{t}_g \) is ordered accordingly.
4.2.6 Common and Unique Components

Since t_g plays a role as a ‘link bridge’ between the seven modalities, it stands for the similarity among the multiple data sets [74]. In order to extract the unique components of each data space, the following decomposition is performed:

\[
\begin{align*}
    P_{ic}^T &= (t_g^T t_g)^{-1} t_g^T X_i \\
    X_i &= X_{ic} + X_{io} = t_g P_{ic}^T + X_{ie} \\
    \forall i &= 1,2,\ldots,m
\end{align*}
\]  

(4.3)

where \( P_{ic} \) is the loading matrix and \( X_{ie} \) is the residual information within the \( i \)th data space. \( X_{ie} \) may encapsulate specific information, unique to \( X_i \) besides noise [74]. In order to extract these components, the method of orthogonal signal correction (OSC) is employed, in order to extract the orthogonal components from the residuals [74], [81].

\[
\begin{align*}
    P_{io}^T &= (T_{io}^T T_{io})^{-1} T_{io}^T T_i \\
    X_{ie} &= X_{io} + X_{ir} = T_{io} P_{io}^T + X_{ir} \\
    \forall i &= 1,2,\ldots,m
\end{align*}
\]  

(4.4)

where \( T_{io} \) is the orthogonal component revealing the unique information associated with each modality, \( P_{io} \) is the loading matrix and \( X_{ir} \) is the final residual within each data space.

4.2.7 Least Absolute Shrinkage and Selection Operator (LASSO) Regression

LASSO is an advanced variable selection technique for regression [27]. LASSO minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant [28]. This method aids to improve the prediction accuracy when dealing with multicolinearity data while carrying interpretability and numerical stability of the data [27].
In a linear regression model with data \((x_{i1}, x_{i2}, \ldots, x_{ip}, y_i), i = 1, \ldots, n\), where \(x_{ij}\) is a linear regressor, \(y_i\) is the response variable of the \(i\)th observation, and \(N\) is the number of observations, “LASSO” minimizes the residual sum of squares [28]:

\[
\hat{\beta}^L = \arg\min_{\beta} \left\{ \frac{1}{2N} \sum_{i=1}^{N} (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\}
\]

subject to the sum of absolute values of the coefficient being less than a constant:

\[
\sum_{j=1}^{p} |\hat{\beta}^L_j| \leq t \ (\text{Constant}).
\] (4.6)

LASSO sacrifices some bias into the model, in order to reduce the variance of the predicted values [27]. In equation 4.5, \(\lambda\) is a nonnegative regularization parameter corresponding to one value of Lambda [27]. As \(\lambda\) increases, the number of nonzero components of \(\beta\) decreases. This will improve the overall prediction accuracy and as a result some of the produced coefficients by this method are equal to zero.

In this paper, LASSO was applied to all eight components (1 common and 7 unique) separately, in order to analyze the significance of each component in predicting the disease. The MATLAB ‘lasso’ function was used with 10-fold cross-validation in order to prevent overfitting. The value of lambda associated with the lowest mean square error (MSE) with K-fold cross-validation was used.

Next the scalar lambda value with minimum MSE was selected and its index \((j)\) was used to determine the LASSO predicted value as follows:

\[
C_{pred} = \text{Intercept}(j) + X * \beta(i,j)
\]

(4.7)
where $X$ is the input data set, $N$ is the length of $X$, and $i= 1,2,\ldots,N$.

4.3 Results

Lesion load and normalized brain volume are vastly used in MS studies [54], [55], [56], [57], [58], [71]. In order to demonstrate the significance of combining 7 imaging modalities with the proposed methods for the purpose of predicting clinical indices, we have primarily combined lesion load and normalized brain volume. Through regression (Fig. 4.3), the strength of the prediction based on the two combined modalities for DD, EDSS, & age was analyzed (Fig. 4.4). Next, seven data sets from 7 different imaging modalities (FA, MWF, rsfMRI, lesion load, BFS, and left and right hemisphere’s cortical thickness), were analyzed by using the proposed methods in section 4.2. Each data set was pre-processed and 7 subspaces, mean MWF, $X_1$ (47 subjects x 35 ROIs), mean FA, $X_2$ (47 subjects x 48 ROIs), and rsfMRI, $X_3$ connectivity matrix (47 subject x 1926 connections), lesion load, $X_4$ (47 subject x 1 value), BFS, $X_5$ (47 subject x 1 value), and left, $X_6$, and right hemisphere’s cortical thickness, $X_7$ (47 subjects x 108 ROIs, for each hemisphere) were defined. Through the Multi-LV extraction processes 32 subLVs were defined representing 90% of the ratio of the explained variance. The supLV $t_g$, defining the common components within the 7 data spaces, has a size of 47 subjects by 32 components. Subsequently, each unique component’s set (47 Subjects x 32 components) was established through the method described in section 4.2.6. LASSO was then applied on each dataset of the 8 datasets (1 common and 7 unique to each modality) separately. Figures 4.5, 4.6, and 4.7 show the relationship between each dataset and the acquired clinical scores (DD, EDSS, & age).
Figure 4.3. Regression for predicting clinical indices based on lesion load and normalized brain volume. Y is the target variable (DD or EDSS or Age). X is a matrix of predictors. β is a vector of regressors and ε is a vector of error terms.

Figure 4.4. Regression results based on the combined lesion load and normalized brain volume.
Figure 4.5. LASSO components predicting DD. The $y$-axis consists of the predicted LASSO values for the clinical score on the $x$-axis. 

a) LASSO of all latent components with DD. LASSO depicted 14 significant components, $Y=0.25X + 105$ and $R^2=0.51$. 

b) LASSO of decomposed common components among all seven modalities with DD. LASSO depicted 20 significant components. $Y=0.74X + 35.8$ and $R^2=0.90$. 

c) LASSO of decomposed unique BFS components with DD. LASSO depicted 1 significant component. $Y=0.27X + 99.6$ and $R^2=0.27$. 

d) LASSO of decomposed unique mean FA components with DD. LASSO depicted 1 significant component. 

e) LASSO of decomposed unique mean MWF components with DD. No significant components were detected. 

f) LASSO of decomposed unique rsfMRI connectivity components with DD. No significant components were detected. $Y=0.05X + 131$ and $R^2=0.08$. 

g) LASSO of decomposed unique right hemisphere’s cortical thickness components with DD. LASSO depicted 8 significant components. $Y=0.26X + 102$ and $R^2=0.27$. 

h) LASSO of decomposed unique left hemisphere’s cortical thickness components with DD. LASSO depicted 7 significant components $Y=0.06X + 129$ and $R^2=0.08$. 

i) LASSO of decomposed lesion load components with DD. LASSO depicted 1 significant component.
Figure 4.6. LASSO components predicting EDSS. The y-axis consists of the predicted LASSO values for the clinical score on the x-axis. a) LASSO of all latent components with EDSS. LASSO depicted 14 significant components. $Y=0.38X + 1.53$ and $R^2=0.76$. b) LASSO of decomposed common components among all seven modalities with EDSS. LASSO depicted 17 significant components. $Y=0.38X + 1.54$ and $R^2=0.65$. c) LASSO of decomposed unique BFS components with EDSS. LASSO depicted 2 significant components. $Y=0.04X + 2.39$ and $R^2=0.05$. d) LASSO of decomposed unique mean FA components with EDSS. LASSO depicted 7 significant components. $Y=0.26X + 1.83$ and $R^2=0.52$. e) LASSO of decomposed unique mean MWF components with EDSS. LASSO depicted 5 significant components. $Y=0.25X + 1.86$ and $R^2=0.48$. f) LASSO of decomposed unique rsfMRI connectivity components with EDSS. No significant components were detected. g) LASSO of decomposed unique right hemisphere’s cortical thickness components with EDSS. LASSO depicted 11 significant components. $Y=0.04X + 2.37$ and $R^2=0.07$. h) LASSO of decomposed unique left hemisphere’s cortical thickness components with EDSS. LASSO depicted 15 significant components. $Y=0.05X + 2.36$ and $R^2=0.07$. i) LASSO of decomposed lesion load components with EDSS. LASSO depicted 8 significant components. $Y=0.04X + 2.38$ and $R^2=0.07$. 

47
Figure 4.7. LASSO components predicting age. The y-axis consists of the predicted LASSO values for the clinical score on the x-axis. a) LASSO of all latent components with age. LASSO depicted 18 significant components. $Y=0.30X + 29.9$ and $R^2=0.64$. b) LASSO of decomposed common components among all seven modalities with age. LASSO depicted 4 significant components. $Y=0.15X + 36.4$ and $R^2=0.35$. c) LASSO of decomposed unique BFS components with age. LASSO depicted 5 significant components. $Y=0.20X + 34.1$ and $R^2=0.20$. d) LASSO of decomposed unique mean FA components with age. LASSO depicted 3 significant components. $Y=0.12X + 37.7$ and $R^2=0.32$. e) LASSO of decomposed unique mean MWF components with age. No significant components were detected. f) LASSO of decomposed unique rsfMRI connectivity components with age. No significant components were detected. g) LASSO of decomposed unique right hemisphere’s cortical thickness components with age. LASSO depicted 8 significant components. $Y=0.02X + 42.0$ and $R^2=0.03$. h) LASSO of decomposed unique left hemisphere’s cortical thickness components with age. LASSO depicted 13 significant components. $Y=0.19X + 34.6$ and $R^2=0.20$. i) LASSO of decomposed lesion load components with age. LASSO depicted 3 significant components. $Y=0.02X + 42.1$ and $R^2=0.03$. 
4.4 Discussion

Although there have been a number of studies that investigate joint biomarkers in different disease, most are limited to 2 modalities. For instance, one study demonstrated that joint evaluation of brain structure and function via MRI can yield higher classification accuracy of prodromal Alzheimer’s disease than either alone [29]. Our results validate the power of joint biomarkers by combing seven different imaging modalities.

In figures 4.5a, 4.6a, and 4.7a, the latent variable components consist of the common component between the seven modalities, the unique components to each modality, and an error/residual term. Comparing the results of the regression of the two modalities (lesion load and normalized brain volume) for each clinical score in figure 4.4, with the employed methods on the seven modalities (Fig. 4.5, 4.6, and 4.7), we observe that both the latent variables and the common components, in sub graphs a and b respectively, demonstrate a higher correlation with the clinical score. As a result we can conclude that the combination of the seven modalities with the methods of this study result in more precise predictions of the clinical indices.

LASSO results demonstrated that the common components ($t_a$) could accurately predict disease duration (Fig. 4.5b) with 20 significant components. Furthermore, a relationship between DD and the unique sets of components for each data space is seen in figure 4.5. As demonstrated in the graphs, only BFS (Fig. 4.5c), mean FA (Fig. 4.5d), and lesion load (Fig. 4.5i) with 1 significant LASSO component each and right and left hemisphere’s cortical thickness connections with 8 (Fig. 4.5g), and 7 (Fig. 4.5h) significant components respectively were able to illustrate weak predictions of DD. The results illustrate MS following a unique pattern of degeneration.
In addition, LASSO components also demonstrate a strong relationship between age and EDSS (Fig. 4.6). As seen in the figures 4.5g and 4.5h, 4.6g and 4.6h, and 4.7g and 4.7h, LASSO components also illustrate an asymmetry between left and right hemisphere’s cortical thickness components. Unique components of left hemisphere’s cortical thickness are able to construct a better linear relationship between the modality and the clinical scores. Studies have previously observed structural asymmetries, including cortical thickness differences, in the two hemispheres [82]. Findings show an increase in thickness with aging in the left hemisphere in healthy controls [83]. This phenomenon can also be observed in MS patients (Fig. 4.6g, Fig 4.6h). In addition, studies suggest left hemisphere to be more affected than the right hemisphere in both MS patients and healthy volunteers [84], [18]. This study confirms the previous findings and shows that the left hemisphere, which is considered to be the dominant hemisphere, is not only affected more than the right hemisphere, but also can better predict pattern in degeneration.

In general, unique components of MWF and rsfMRI were not able to establish a relationship between the modality and the clinical score. Unique components of mean MWF with 5 significant LASSO components were only able to establish a weak linear relationship between MWF and EDSS (Fig 4.6e). Common components were more strongly correlated with clinical scores. Furthermore, as can be observed in figure 4.6, except for unique MWF components and unique rsfMRI components, all other unique structural data were able to demonstrate a relationship with age through LASSO.

For future work, we propose also investigating other feature selection methods (e.g., based on mutual information) for detecting which features of the seven modalities are more relevant. The number of latent variables desirable to acquire the best generalization for the prediction of new observations can be attained by cross-validation methods such as bootstrapping [53].
In future, we intend to improve our registration and interpolation techniques. We also plan to look into using non-linear registrations.
Chapter 5: Conclusion and Future Work

5.1 Conclusion

In this thesis we proposed two frameworks that can analyze multimodal data in neurodegenerative disease in order to distinguish biomarkers that are predictable of clinical scores. This work emphasizes the potential utility of combining imaging modalities into a single biomarker. To this end, in the third chapter we proposed to use a data integration approach, utilizing both sMRI and fMRI in order to take advantage of the complementary information in the two imaging modalities. We used IndFeat’s, feature selection algorithm as a preprocessing step in order to extract a subset of features from the two MRI modalities separately. Using these features we then formed an interaction event matrix based on Kronecker products of the two modalities and performed LASSO Regression on the three (structural, functional, and interaction) datasets separately. We showed that complementary biomarkers are more informative in predicting clinical scores in both PD and HC cohorts. Since structural data is routinely obtained when obtaining fMRI data, the two modalities may be more informative than analyzing each modality separately.

In chapter four, we proposed a data fusion method based on latent variables in order to assess multimodal imaging in MS. Multimodal data from mean MWF, mean FA, rsfMIR connections, left and right hemisphere’s cortical thickness, lesion load, and normalized brain volume, were fused together through a joint Multimodal Statistical Analysis Framework. We then further decomposed the extracted LVs into 8 components, 1 of which contains the common information among the seven modalities and the other 7 carry the unique information in each modality.
Through LASSO, we observed that the common components between the seven modalities, accurately predicted disease duration and demonstrated a strong linear relationship with EDSS and age. To further investigate the regional importance of each modality, LASSO was performed on decomposed unique components of each modality. Results indicate a unique pattern of degeneration in MS and an asymmetry between the cortical thickness LASSO components in left and right hemispheres. Observations demonstrate left cortical thickness components having a stronger effect on predicting the clinical scores. Findings show the importance of multimodal imaging markers in MS.

In conclusion, this thesis demonstrates the power of utilizing multimodal imaging biomarkers in neurodegenerative diseases. Since structural imaging data is acquired along with functional data, we propose that fusion of information from both types of data should become part of routine analysis.

5.2 Future Work

This work demonstrates the potential of combining imaging modalities into a joint biomarker. In practice, all of the studied image sequences can be obtained with a typical MRI scanner, albeit with slightly longer acquisition times.

Subjects in both the PD & MS study were recruited from a single site in order to prevent the multi-site scanner effects; however, in further studies subjects could be recruited from multiple sites in order to increase the statistical power needed to clarify effects. It is noteworthy to mention that the main effects of site for each modality must then be removed in order to lessen the multi- scanner issue. Furthermore, future studies may employ a larger sample size as the samples used in both studies are relatively small and the results may represent an unusual sample
of subjects. In order to mitigate sample size effects, cross-validation is used throughout the frameworks.

In addition, in the established studies, the utilized fMRI may explore functional connectivity that can be observed between regions where no structural connectivity exists. For instance, the functional connection between the right temporal pole region to the left caudal medial frontal gyrus has no known structural connections. This functional connection is mainly due to some polysynaptic connections at the temporal resolution between the two regions and can appear to be co-activating instantaneously. Prior studies have proposed that there may exist a mutual influence by a third region [31], [32]. By examining lagged connectivity, this may partly circumvent this consideration.

Furthermore, FA is highly sensitive to microstructural variations but it is not precisely explicit to the kind of changes [38] order to maximize the specificity, it is suggested that future studies utilize several diffusion tensor measures (e.g. Mean Diffusivity (MD) and FA) to improve the characterization of the tissue microstructure [38].

For the study in chapter four the number of latent variables desirable to acquire the best generalization for the prediction of new observations can be attained by cross-validation methods such as bootstrapping [53]. In addition, we intend to investigate other feature selection methods (e.g., based on mutual information) for detecting which features of the seven modalities are more relevant. We also intend to further improve our registration and interpolation techniques. We also plan to look into using non-linear registrations.

In future, we intend to apply the multimodal frameworks used in this thesis in order to improve the clinical diagnosis of brain disorders using multimodal imaging.
We note that structural imaging is routinely obtained with functional MRI, often just for visualization purposes. Given the results of this thesis, it would seem reasonable to routinely consider joint analyses of these data.
Bibliography


[54]A. Lavery, L. Verhey and A. Waldman, "Outcome Measures in Relapsing-Remitting Multiple Sclerosis: Capturing Disability and Disease Progression in Clinical Trials", 60


[78] J. McAusland, R. Tam, E. Wong, A. Riddehough and D. Li, "Optimizing the Use of


