Classification of Resting-State fMRI using Evolutionary Algorithms: Towards a Brain Imaging Biomarker for Parkinson's Disease

Amir Dehsarvi

Doctor of Philosophy

University of York

Electronic Engineering

January 2018

To my Jennifer

&

our soon to be born baby

و به پدر و مادر نازنینم با عشق...

Abstract

It is commonly accepted that accurate early diagnosis and monitoring of neurodegenerative conditions is essential for effective disease management and delivery of medication and treatment. This research develops automatic methods for detecting brain imaging preclinical biomarkers for Parkinson's disease (PD) by considering the novel application of evolutionary algorithms. An additional novel element of this work is the use of evolutionary algorithms to both map and predict the functional connectivity in patients using rs-fMRI data. Specifically, Cartesian Genetic Programming was used to classify dynamic causal modelling data as well as timeseries data. The findings were validated using two other commonly used classification methods (Artificial Neural Networks and Support Vector Machines) and by employing *k*-fold cross-validation. Across dynamic causal modelling and timeseries analyses, findings revealed maximum accuracies of 75.21% for early stage (prodromal) PD patients in which patients reveal no motor symptoms versus healthy controls, 85.87% for PD patients versus prodromal PD patients, and 92.09% for PD patients versus healthy controls. Prodromal PD patients were classified from healthy controls with high accuracy – this is notable and represents the key finding since current methods of diagnosing prodromal PD have low reliability and low accuracy. Furthermore, Cartesian Genetic Programming provided comparable performance accuracy relative to Artificial Neural Networks and Support Vector Machines. Nevertheless, evolutionary algorithms enable us to decode the classifier in terms of understanding the data inputs that are used, more easily than in Artificial Neural Networks and Support Vector Machines. Hence, these findings underscore the relevance of both dynamic causal modelling analyses for classification and Cartesian Genetic Programming as a novel classification tool for brain imaging data with medical implications for disease diagnosis, particularly in early stages 5-20 years prior to motor symptoms.

Keywords:

Evolutionary Algorithms; Cartesian Genetic Programming; Classification; Parkinson's Disease; Prodromal Parkinson's Disease; Resting-state fMRI; Dynamic Causal Modelling.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Thank you to my beautiful Jennifer and to my family, for their tireless love and support. Without them all, completing my studies at both degree and masters level would not have been possible and it is through them that I was able to achieve my goal of working towards a PhD.

A special thanks to Professor Stephen Smith for providing incredible supervision and guidance. I would, further, like to thank my thesis advisory panel, Dr. David Halliday, for his valuable comments and input. Without them, this work simply would not have been possible.

Author's Declaration

This thesis was written by myself and represents my original work, with supervision from Professor Stephen Smith. This thesis has not been submitted for any other award at this or any other institution. All sources are acknowledged as List of Abbreviations

ACO	Ant Colony Optimisation
AD	Alzheimer's Disease
ADASYN	Adaptive Synthetic Sampling
AI	Artificial Intelligence
AIS	Artificial Immune System
ANN	Artificial Neural Networks
Bagging	Bootstrap Aggregating
BOLD	Blood-Oxygen Level Dependent
BRAIN	Bradykinesia Akinesia Incoordination Test
CAT	Computed Axial Tomography
CBD	Corticobasal Degeneration
CDR	Clinical Dementia Rating
CGP	Cartesian Genetic Programming
CI	Computational Intelligence
CONN	Functional Connectivity Toolbox
CSF	Cerebrospinal Fluid
CT	See CAT
CV	Cross-validation
DAT	Dopamine Transporter Imaging
DCM	Dynamic Causal Modelling
DMN	Default Mode Network
EAs	Evolutionary Algorithms
EEG	Electroencephalography
EPI	Echo Planar Imaging

EU	European Union
FA	Flip Angle
FFT	Fast Fourier Transform
FID	Free-Induction Decay
Flair	Fluid Attenuated Inversion Recovery
fMRI	Functional Magnetic Resonance Imaging
FoV	Field of View
FPC	Fronto Parietal Control
FSL	FMRIB Software Library
FWHM	Full Width at Half-Maximum
GA	Genetic Algorithms
GP	Genetic Programming
GWAS	Genome Wide Association Surveys
HD	Huntington's Disease
kNN	k-Nearest Neighbours
LB	Lewy Body
LID	Levodopa-Induced Dyskinesia
LIPC	Left Inferior Parietal Cortex
MBRS	Modified Bradykinesia Rating Scale
MDS-UPDRS	Movement Disorder Society sponsored revision of Unified Parkinson's Disease Rating Scale
MEG	Magneto Encephalography
MoCA	Montreal Cognitive Assessment
mPFC	Medial Prefrontal Cortex
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
MSA	Multiple System Atrophy
NDDs	Neurodegenerative Diseases
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NINDS	National Institute of Neurological Disorders and Stroke
NIRS	Near Infrared Spectroscopy
PCC	Posterior Cingulate Cortex
PD	Parkinson's Disease
PET	Positron Emission Tomography
PPMI	Parkinson's Progression Markers Initiative
PSO	Particle Swarm Optimisation
PSP	Progressive Supranuclear Palsy
RBD	Rapid Eye Movement Sleep Behaviour Disorder
RCGP	Recurrent CGP
RF	Radio Frequency

RIPC	Right Inferior Parietal Cortex
rs-fMRI	Resting State fMRI
RSN	Resting State Neural Network
SD	Standard Deviation
SEM	Structural Equation Modelling
SIGN	Scottish Intercollegiate Guidelines Network
SPECT	Single Photon Emission Computed Tomography
SMOTE	Synthetic Sampling with Data Generation
SNR	Signal to Noise Ratio
SPM	Statistical Parametric Mapping
SVM	Support Vector Machines
TE	Echo Time
TR	Repetition Time
UHDRS	Unified Huntington's Disease Rating Scale
UK	United Kingdom
UPDRS	Unified Parkinson's Disease Rating Scale
UPSIT	University of Pennsylvania Smell Identification Test
USA	United States of America
WM	White Matter

References.

Portions of the data analyses from Chapter 6 were presented at the Parkinson's UK Research Conference, 2016, as follows:

A. Dehsarvi, and S. Smith, "Evolutionary Algorithms Meet fMRI: Diagnosis of Parkinson's Disease.," *Parkinson's UK Research Conference.*, poster presentation, 2016.

Chapter 1. Introduction

By 2050, 22% of the global population will be over 60 years of age, which is double the current proportion (estimated at 11%, 2000). During early 20th Century, life expectancy in developed nations was an average 50 years of age, whereas in 2014, life expectancy was estimated at an average 80 years of age. This explosion in life expectancy is attributed largely to advances regarding sanitation, medicine, and living conditions [1]–[3]. An ageing population is not without challenges as the biological and cognitive decline associated with ageing is the leading cause of non-communicable diseases, such as cancer, stroke, type-two diabetes, Parkinson's disease (PD) [4], [5] and Alzheimer's disease (AD), amongst others. As such, the World Health Organization's "Draft Twelfth General Programme of Work" (April 13, 2013) lists "Addressing the challenge of non-communicable diseases" [6] as a salient issue. The current research develops a new method of diagnosing PD, a neurodegenerative disease (a class of non-communicable diseases), using evolutionary algorithms (EAs).

This thesis firstly outlines why it is important to examine learning algorithms in the context of the classification for PD, prodromal PD (early stages of PD), and healthy age-matched controls. Secondly, computational intelligence methods are briefly reviewed, including EAs, Cartesian Genetic Programming (CGP), and classification. Furthermore, this thesis includes a discussion of methods for working with class-imbalanced data, given the limitations in testing and recruiting patient and healthy matched samples. For instance, the research presented uses a dataset with 102 PD patients, 18 prodromal PD patients, and eight healthy controls. Thirdly, functional magnetic resonance imaging (fMRI) is reviewed, including an overview of resting state fMRI (rs-fMRI) as the findings reported in this thesis examine classification using rs-fMRI data. Finally, an analysis of clinical rs-fMRI data is presented using EAs with the overarching objective of detecting a brain imaging biomarker for PD and developing novel tools for accurate monitoring of treatment of PD.

1.1. Research Aims

New tools for early diagnosis and accurate monitoring of PD are required as the prevalence of PD is expected to increase with an ageing population. Early diagnosis and accurate

monitoring of PD facilitates drug dosages to be patient-tailored, which boosts therapeutic benefits and limits side effects and functional disability, whilst improving patient quality of life. Difficulties providing a differential diagnosis of PD abound, for example, community studies indicate that PD diagnoses have 15-26% false positive rates [7], [8], which is further exacerbated in the early stages of PD. Therefore, quick and non-invasive methods for diagnosis and monitoring with good accessibility are necessary.

The research presented in this thesis proposes a novel brain imaging biomarker for PD. There are two overarching research questions.

Research question 1:

> *Can accurate monitoring of PD be achieved using EAs on rs-fMRI data for patients prescribed Modafinil?*

Research question 2:

> *Can early stage PD be diagnosed using EAs on rs-fMRI data?*

Research question 1 focuses on developing novel clinical monitoring tools using data from a controlled experiment in which participants were administered the drug (versus a control group) Modafinil (typically prescribed for PD patients to relieve physical fatigue). Research question 2 applies the techniques developed to the diagnosis of participants with PD, early stage PD, and healthy age-matched controls. These research questions are both novel and exploratory, examining a timely issue with numerous medical implications, including concerning the transferability of medical research based on limited and unequal sample sizes.

Developing a tool that can differentiate between the various stages of disease severity is a fundamental aim of the work presented, with therapeutic consequences in terms of tailoring medication dosage and monitoring medication in accordance with symptoms exhibited and overall PD stage. The research presented analyses open data taken from the Parkinson's Progression Markers Initiative database (PPMI; http://www.ppmi-

info.org/data), a longitudinal study where participants underwent a comprehensive longitudinal follow-up schedule of clinical, imaging and bio-specimen assessments.

This research develops automatic procedures for identifying PD brain imaging preclinical biomarkers, which enhances the confidence of methods involved in early PD diagnosis. A core research aim is to identify the applicability of Cartesian Genetic Programming (CGP) and recurrent CGP (RCGP) classification for both timeseries and Dynamic Causal Modelling (DCM) analyses. The timeseries values and DCM values from the rs-fMRI data are subjected to supervised classification and the findings are validated with two other commonly used classification methods (Artificial Neural Networks, referred to as ANN, and Support Vector Machines, referred to as SVM). EAs, such as CGP and RCGP, have not previously been applied to brain imaging data. A crucial advantage of EAs, specifically CGP, is that they offer a *white box solution* providing more information on the inputs used and better understanding of the final solution obtained in classification, relative to ANN and SVM. Moreover, research on the classification of rs-fMRI data has typically used statistical-based classifiers (e.g., independent components analysis and multivariate pattern analysis [9]–[13]; for an example of independent components analysis in rs-fMRI for PD data, see [14]). This thesis examines an additional novel question: is DCM analysis useful for classification? Previous research has not examined the applicability of DCM values in classification and little research has applied DCM to PD data [15]–[20].

A common limitation with medical data involves recruiting low numbers of patients, which can result in class-imbalanced data (e.g., high numbers of controls versus patients). This research examines the applicability of classification methods to two datasets with heavily class-imbalanced data, which mimics the conditions prolific in medical research, enabling the research findings presented in this thesis to be more easily generalised to clinical settings.

1.2. Key Novel Aspects Examined in this Thesis

The work presented in this thesis aims to develop novel automatic methods for identifying PD brain imaging preclinical biomarkers, which can aid clinical diagnosis, monitoring and

investigation of PD. Currently prodromal PD diagnosis (5-20 years before motor symptoms are apparent) is in its infancy with typically low accuracy rates and high levels of misdiagnosis with other Parkinsonian conditions. Hence, prodromal PD diagnosis is highly novel and relevant, ensuring access to early treatment for patients 5-20 years prior to motor symptoms occurring, and providing overall better disease management, thus, increasing patient quality of life.

Research questions 1 and 2 respectively develop novel clinical monitoring and diagnosis tools using open rs-fMRI data, with a data-driven exploratory approach. Additional novelties presented in this work include the examination of CGP and RCGP as applied to both timeseries and DCM analyses, with findings being validated by ANN and SVM. EAs have not previously been applied to brain imaging data (which is typically analysed using statistical-based classifiers) and contain numerous benefits, including providing a *white box solution* that presents researchers with detailed information concerning the derived outcomes. In addition, DCM analysis is advantageous in that it provides information on the causal connectivity between different brain regions, yet, this technique has not previously been subjected to classification and, moreover, there is only limited research applying DCM to PD clinical data [15]–[20].

A final core novelty of the work explored in this thesis involves the use of clinical class-imbalanced samples. Medical data typically includes heavily class-imbalanced data (i.e., unequal numbers of patients relative to controls). Hence, this research examines the previous crucial research points whilst using heavily class-imbalanced datasets that mimic the conditions that clinical research typically encounter, applying techniques to minimise the impact of class-imbalanced data on classification accuracy. Outcomes are highly relevant for the generalisation of clinical findings derived from limited samples.

1.3. Outline of Thesis Chapters

The research presented in this thesis examines the application of computational intelligence methods to classification of rs-fMRI data with the aim to identify brain imaging PD biomarkers for disease diagnosis and monitoring. Chapter 5 and Chapter 6

apply EAs to brain imaging data derived from open databases. Chapter 5 investigates disease monitoring via the classification of participants administered a fatigue-reducing drug (Modafinil) commonly prescribed to PD patients, using a double-blind paradigm. Chapter 6 examines the diagnosis of prodromal PD patients, PD patients, and healthy controls. The classification of prodromal PD represents a crucial advance in the field given that prodromal PD occurs 5-20 years before motor symptoms are present and the findings could facilitate early diagnosis and better disease management. Both Chapter 5 and Chapter 6 apply CGP classification to rs-fMRI timeseries and DCM analyses and findings are validated using ANN and SVM classifiers. Furthermore, ADASYN is applied due to both datasets being highly class imbalanced.

Chapter 1 outlines the research aims. Chapter 2 presents a review of PD, including epidemiology, diagnosis, treatment, and monitoring. Chapter 3 reviews MRI and fMRI approaches including resting state and default mode network, and Chapter 4 outlines computational intelligence techniques, emphasising CGP and RCGP for timeseries analysis and classification of dynamic signals. Chapter 5 presents Study 1: classification of participants administered Modafinil versus controls. Chapter 6 presents Study 2: classification of PD patients versus prodromal PD patients versus healthy controls. Finally, Chapter 7 includes conclusions and directions for future research.

Chapter 2. Parkinson's Disease

Neurodegenerative Diseases (NDDs) globally are the leading cause of disability, with patients facing on average over 20 years of chronic disability. NDDs cause neural degeneration, resulting in diverse behavioural and cognitive deficits. Over 50% of all patients with dementia go undiagnosed by their primary care provider [21]–[24]. The common symptomologies between different NDDs make a differential diagnosis challenging, especially during the early phases. As each NDD requires its own specific treatment plan, more efficient and reliable methods to differentially diagnose NDDs are required, particularly focusing on the initial stages of the disease [25]. The research considered here examines a novel approach: the use of EAs to diagnose early stage PD, a class of NDDs.

PD has a 15-26% rate of misdiagnosis by general neurologists [7], [8]. Based on similar physical and cognitive symptoms in early PD, patients are often mistakenly diagnosed with another frequently occurring NDD (e.g., AD and Huntington's Disease, referred to as HD) or even diagnosed with a less common NDDs (e.g., progressive supranuclear palsy[1], multiple system atrophy[2], and normal essential tremor) [7]. These difficulties in reaching a differential diagnosis, particularly in early PD, reiterate the need for an objective diagnostic aid such as is developed in the research presented in this thesis. This section briefly outlines two common NDDs (AD and HD) before outlining the epidemiology, pathology, clinical phenotype, and assessment strategies involved in PD, amongst other relevant information. The literature reviewed outlines the context in which PD is diagnosed and highlights the relevance of future work in applying EAs to other NDDs.

2.1. Alzheimer's Disease (AD)

AD is both the most frequently occurring NDD and the most common type of dementia [26]. Like PD, AD is caused by numerous interacting factors including age, environment, genetic factors, and lifestyle [27], with age being an important trigger of AD. For example,

[1] Also known as PSP. This is an uncommon brain disorder that affects movement, control of walking (gait) and balance, speech, swallowing, vision, mood and behaviour, and thinking. The disease results from damage to nerve cells in the brain.
[2] Also known as MSA. This is a rare NDD characterised by symptoms such as tremors, slow movement, muscle rigidity, and postural instability balance difficulties (collectively known as Parkinsonism) due to dysfunction of the autonomic nervous system and ataxia.

individuals aged 65-80 years are twice as likely to develop AD relative to younger adults and, in North America, the incidence per age group doubles every 5 years, from 0.17% at 65, to 2.92% at 85 [28]. AD symptoms involve memory loss, reduction of eyesight, visuospatial deficits, delusions, and obsessive behaviour.

AD is diagnosed and monitored via physical and cognitive assessments, yet AD patients show greater cognitive loss relative to PD patients (yet these typically reveal greater motor loss). For instance, cognitive symptoms are evaluated with a range of cognitive assessment tests, one of which is the Rey-Osterrieth Complex Figure [29], [30], in which patients copy a drawing of connected shapes and, following a brief period, patients are asked to replicate this same drawing from memory; hence, visuospatial skills and memory are both evaluated. Medication can slow disease development and temporarily reduce the cognitive symptoms but it only has a mild effect on disease outcomes [31]. There is yet no cure for AD. As AD advances, comprehensive care is required for catering to patients' needs, with financial and emotional implications.

2.2. Huntington's Disease (HD)

HD is hereditary and individuals whose parents are a gene carrier have a 50% likelihood of developing HD [32]. HD has numerous cognitive, motor, and psychiatric symptoms. Motor symptoms can be similar to PD (rigidity, and postural problems) but also include uncontrollable jerky movements (labelled chorea). Cognitive impairment is typically greater in HD relative to PD, and includes memory loss (sometimes resulting in dementia), depression, anxiety, aggression, and obsessive behaviour [33]. The Unified Huntington's Disease Rating Scale (UHDRS) is used to diagnose HD [34]. This scale evaluates motor, cognitive, behavioural, and functional abilities. Psychiatric medication is used for many cognitive symptoms (e.g., depression, irritability, and mood swings). In addition, PD medication can treat chorea and other motor symptoms [35]. Nevertheless, these drugs only treat HD symptoms and no cure is available.

2.3. Parkinson's Disease (PD)

In the UK, one in 500 people have PD and it is the 2nd most prevalent NDD globally, trailing only Alzheimer's disease (AD) [36], [37]. PD occurs in 1.3% European adults and increases to almost 3% in adults over 80 years of age [38]. PD is a progressive and neurodegenerative condition, in which disordered movement is a key symptom. Motor symptoms include slowness (bradykinesia), stiffness (rigidity), shaking (tremor) and impaired balance (postural instability) [39], [40]. Neurological markers include a reduction of dopaminergic neurons and Lewy body (LB) deposition, the cause of which is not known [41], [42]. There is currently no cure for PD and medication helps temporarily with controlling some of the symptoms [43].

2.3.1. Historical Overview

PD-related symptoms are first described in Ancient Indian Sanskrit texts from 3000 BC[3]. Reports of PD-related symptoms are also described in ancient Chinese, Greek and Biblical texts[4]. In 1817, James Parkinson provided the first a detailed clinical account of PD in 1817, within his *An Essay on The Shaking Palsy*, focusing on six patients, three of which were his patients and the remaining three individuals he observed in London [44]. The early neurologists Charcot, Trousseau, Gowers and Erb subsequently developed the clinical phenotype of PD. PD was initially labelled *paralysis agitans,* although in 1877, to recognise Parkinson's discoveries, Charcot renamed the disease to *maladie de Parkinson* – Parkinson's Disease [45]. Parkinson described that a common symptom involves "*lessened muscular power*" [46] that was reiterated in Charcot's description of non-tremulous type of PD, in which Charcot specifies that slow movements need to be differentiated from weakness or "*lessened muscular power*".

The subsequent century marked the gradual investigation of PD and broadening of its pathological underpinnings. In 1912, Friedrich Lewy revealed that bodies were a pathological marker of PD. In 1919, Konstantin Tretiakoff revealed the deterioration of

[3] In Ayurveda, the Indian system of medicine, there are descriptions of a condition called 'kampavata' comprising tremors, stiffness, depression and movement depletion.
[4] Ecclesiastes 12:3 "*In the day when the keepers of the house shall tremble…*"

pigmented substantia nigra cells in the brainstem of encephalitic PD patients. In 1957, Carlsson et al. revealed that dopamine[5] is a putative neurotransmitter [47]. Furthermore, in 1960, Ehringer and Hornykiewicz revealed that PD patients have significantly lower dopamine concentrations in the striatum [47], [48], leading to the start of levodopa trials for PD patients in 1961 [49]. In 1961, findings revealed that injected levodopa improves akinesia in PD patients and, by 1970, oral administration of levodopa was developed [50], [51], resulting in Carlsson winning the Nobel Prize in Medicine in 2000. To date, PD research has revealed that genetic mutations, abnormal handling of misfolded proteins by the ubiquitin–proteasome and the autophagy-lysosomal systems, increased oxidative stress, mitochondrial dysfunction, inflammation and other pathogenic mechanisms are contributing components in the death of dopaminergic and non-dopaminergic cells in the brains of PD patients [52], [53]. Dopamine replacement therapy is currently recognised as the most effective treatment for PD patients [54].

2.3.2. Epidemiology

PD occurs in 0.1-0.2% of the population, in 1-2 % of individuals over 60 years old [55], and it is more common in men [56]. Mean age of onset is 60 years, yet 5% of PD patients are diagnosed under 40 years old. Age is the largest risk factor for PD [5], [57], and increases prevalence by over 400%. Indeed, in the USA, 0.02% individuals who died aged 45 – 54 years were diagnosed with PD, relative to almost 9% individuals who died aged over 85 years. Age further increases PD severity to a greater extent than disease duration [4], [57]. PD occurs in all ethnic groups and countries, yet, it is more frequent in Caucasians relative to Asians and Africans [56].

In the United Kingdom (UK), research using data acquired from the National Health Service (NHS) General Practice Research Database[6] (2009) estimated PD prevalence at 0.3% (0.3% for men and 0.2% for women), which corresponds to 126,893 PD patients[7].

[5] Dopamine is a neurotransmitter that is located primarily in the neurons of substantia nigra and ventral tegmental area and is imperative in movement control and reward behaviour.
[6] World's largest database of anonymised longitudinal medical records, comprising approximately 3.4 million people's records
[7] Approximate number of cases per country: 108,000 in England, 10,000 in Scotland, 5,900 in Wales and 3,000 in Northern Ireland.

Given a globally ageing population, UK prevalence is estimated to raise 28% by 2020. Nevertheless, this data focuses exclusively on patients diagnosed PD, hence, actual prevalence is likely to be higher as many individuals either are misdiagnosed or do not obtain medical advice. Indeed, door-to-door community studies reveal that 25% PD patients are undiagnosed [58].

2.3.3. Pathology

PD is produced by damage to dopaminergic neurons in the substantia nigra, causing reduced dopamine within the striatum, diminished dopamine transporters, and LB deposition. Figure 2.1 and Figure 2.2 present the pathology of PD and some other NDDs; the full neuropathological spectrum of these disorders is much more complex than depicted in Figure 2.2. Figure 2.1 shows the histology of substantia nigra using haematoxylin and eosin stain with $\times 200$ magnification. Figure 2.1 (a) depicts normal substantia nigra pathology and (b) demonstrates neuronal loss and LB (see arrow) deposition in a patient with PD. Substantia nigra deterioration is considered a cause of PD motor symptoms.

LB are inclusions of alpha-synuclein, ubiquitin and neuro-filament proteins contained within the cytoplasm of neurons. PD patients reveal a distinct dispersion of LB impacting on the autonomic nervous system and medium to large monoaminergic and cholinergic neurons. Braak et al. [59] described six stages of LB deposition and related cell death: stages 1-2 focus on pathology in the dorsal motor nucleus of the vagus nerve and anterior olfactory structures, stages 3-4 focus on spread of PD to the midbrain and basal ganglia, and stages 5-6 involve spread to cortical regions.

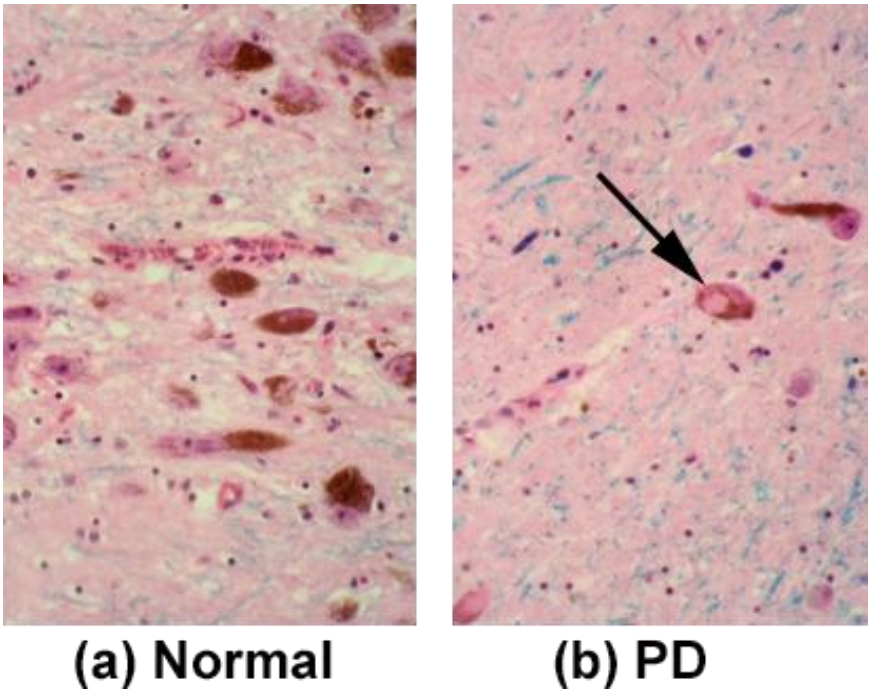

Figure 2.1 - Parkinson's disease pathology; Images are provided by Dr. Ismail, Consultant Neuro-pathologist at Leeds Teaching Hospitals NHS Trust.

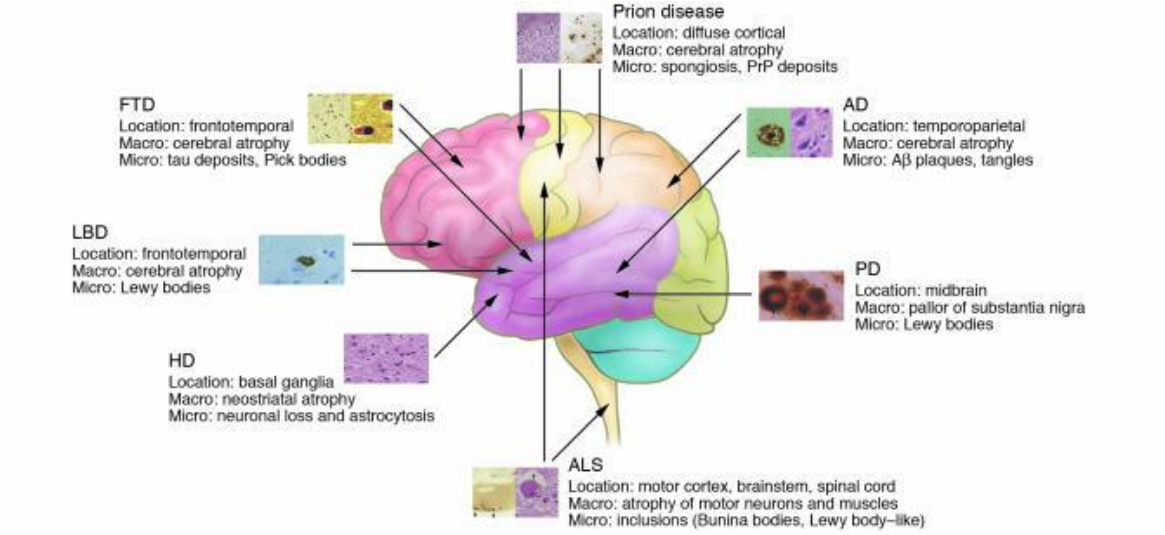

Figure 2.2 - Overview of the anatomical location of and macroscopic and microscopic changes characteristic of the NDDs discussed in [37].

PD is typically sporadic although first-degree relatives are twice as likely to develop PD [60]. There are multiple theories regarding the cause of PD, such as ageing, genetic factors [61], and environmental factors. It is likely that a combination of genetic and environmental factors result in the alpha-synuclein becoming a toxic protein that amasses

within the neurons. Neurons react by creating LB round the alpha-synuclein to shield themselves, although often this protective manoeuvre is not successful and the neuron subsequently dies. Research has revealed that long-term exposure to neurotoxins (such as pesticides) can lead to PD-related symptoms [62] and other researchers have suggested that pesticides can result in oxidative stress that hastens the natural ageing process [63].

Some research suggests that PD is just fast-tracked normal ageing since LB and reduced substantia nigra cells can also be present in healthy senior adults who have no PD symptoms. Nevertheless, in PD patients the reduction in substantia nigra cells with age occurs significantly faster, to levels of approximately 30% of age-matched controls [64]. There is greater loss in the lateral ventral substantia nigra of PD patients, whereas healthy controls experience greater loss in the medial ventral and dorsal regions [64]. Other research has identified that PD is not only related to dopaminergic deficits but also to changes in relation to other neurotransmitters (including acetylcholine and serotonin).

2.3.4. Financial Costs

The UK spent £600 million yearly on economic costs associated with 100,000 PD patients in 2003 (corresponding to £5993 per patient) [65]. Current expenditure is likely to be exacerbated by an ageing UK population with now almost 130,000 PD patients. Costs are increased with certain PD symptomology, ageing, and PD severity [66]–[68]. For PD patients with motor fluctuations and dyskinesia costs double [69] and dementia triples the costs [70], relative to PD patients without these symptoms. Further, older PD patients often live in care homes and are more likely to have falls and non-motor symptoms (e.g., dementia, hallucinations) [71].

Beyond the financial and time costs required to employ specialised clinicians, there are economic costs associated with training and using the diagnostic scale. The Unified Parkinson's Disease Rating Scale (UPDRS) scale is free for clinical purposes, yet there are costs incurred for its use in research: $1000 (approximately £750) charge for funded research and £20,000 for industry-funded research. Moreover, the online training programme costs $1,000-1,500 (approximately £750-1,125) per clinician for use in clinical

research and $250-500 (approximately £188-375) for non-members of the Movement Disorders Society to use the scale for personal reasons or non-profit research [72].

Indirect financial costs associated with PD include loss of earnings via early retirement of the PD patient or a family member who cares for them. Until a cure for PD is developed, early diagnosis and treatment are essential in improving quality of life for PD patients and their family, which will help to reduce any financial burdens linked to PD on both the patients' family and the NHS.

2.3.5. Clinical Phenotype

PD typically occurs in individuals over 60 years old and is characterised by slowness of movement (bradykinesia) and the presence of one additional symptom: extrapyramidal rigidity, rest tremor, and postural instability in the absence of exclusion criteria. Rest tremor, rigidity, and bradykinesia are considered to be key PD symptoms [8]. Listed as follows in Table 2.1 are the PD symptoms taken from the NHS guidelines [73]. In the last 20 years, research has revealed that PD is not solely a movement disorder as was thought for over 200 years. Non-motor PD symptoms can include hyposmia[8], depression, anxiety, hallucinations, sleep disturbance, autonomic dysfunction, and cognitive impairment, bowel problems, pain, and communication problems; mental health, delusions, depression, and impulsive behaviour [59], [74].

[8] Hyposmia is a reduced sense of smell and may occur many years before the onset of motor signs.

Table 2.1 - Parkinson's disease symptoms

Main symptoms

1. Tremor – shaking, which usually begins in the hand or arm and is more likely to occur when the limb is relaxed and resting
2. Slowness of movement (bradykinesia) – where physical movements are much slower than normal, which can make everyday tasks difficult and can result in a distinctive slow, shuffling walk with very small steps
3. Muscle stiffness (rigidity) – stiffness and tension in the muscles, which can make it difficult to move around and make facial expressions, and can result in painful muscle cramps (dystonia)

Physical symptoms

1. Balance problems – these can make someone with the condition more likely to have a fall and injure themselves
2. Loss of sense of smell (anosmia) – sometimes occurs several years before other symptoms develop
3. Nerve pain – can cause unpleasant sensations, such as burning, coldness or numbness
4. Problems with urination – such as having to get up frequently during the night to urinate or unintentionally passing urine (urinary incontinence)
5. Constipation
6. Inability to obtain or sustain an erection (erectile dysfunction) in men
7. Difficulty becoming sexually aroused and achieving an orgasm (sexual dysfunction) in women
8. Dizziness, blurred vision, or fainting when moving from a sitting or lying position to a standing one – caused by a sudden drop in blood pressure
9. Excessive sweating (hyperhidrosis)
10. Swallowing difficulties (dysphagia) – this can lead to malnutrition and dehydration
11. Excessive production of saliva (drooling)
12. Sleeping (insomnia) – this can result in excessive sleepiness during the day

Cognitive and psychiatric symptoms

1. Depression and anxiety
2. Mild cognitive impairment – slight memory problems and problems with activities that require planning and organisation
3. Dementia – a group of symptoms, including more severe memory problems, personality changes, seeing things that are not there (visual hallucinations) and believing things that are not true (delusions)

On average, 10% of PD patients develop severe dementia [75]. LB Disorder is diagnosed when cognitive dysfunction precedes motor symptoms, whereas Parkinson's with Dementia is diagnosed when the reverse occurs. Hence, some researchers characterise PD as a neuro-psychiatric disorder.

PD clinical sub-types are diagnosed according to age of onset or key symptoms. The presence of bradykinesia is obligatory for PD diagnosis. Bradykinesia is defined as minor, dysrhythmic and progressively decelerating movements that become lesser with repetition [76]. Consequences include slow gait with little shuffling steps, reduced dexterity, absence of gesticulation, decreased blink frequency and facial expression, low voice, and problems swallowing. Given the multi-faceted elements involved in bradykinesia, clinicians can find it problematic to diagnose – a difficulty which is exacerbated in early PD stages – with consequences on PD diagnosis accuracy [77]. Approximately 70% of PD patients experience muscular tremor (regular rhythmic oscillating movement with 4-6 Hz frequency), which is common in the hands but can be present in the arms, legs, jaw and tongue. Tremors are often conspicuous when patients are at rest but are lessened when patients engage in deliberate movement [78]. 89-99% of PD patients experience rigidity, which causes individuals to have a bent, stiff, posture [7], [78]. When the clinician moves a patient's body part, rigidity can be perceived as constant resistance, and this resistance is heightened when patient moves a contralateral limb. When both tremor and rigidity occur in the same body part, this is labelled as cogwheel rigidity since there is an irregular recurring resistance to passive movement.

Other early onset clinical symptoms include handwriting changes progressing towards micrographia, decreased facial expressiveness, no arm swinging on one side of the body, and a decreased sense of smell [79]. The glabellar tap reflex[9] is neither sensitive nor specific to PD. PD onset is slow; hence, people may initially feel these symptoms are due to normal ageing. Motor symptoms tend to occur only (or mainly) in one limb and this asymmetry continues as the disease progresses with one side of the body typically being more symptomatic. Indeed, the initially unaffected limb becomes affected 2-3 years after detection of the first symptoms in the contralateral limb [78]. Clinical symptoms characterising advanced PD include hypophonia, drooling of saliva due to decreased swallowing, and postural reflexes may, additionally, worsen. Furthermore, depression is present in 40% PD patients.

[9] Known also as the glabellar tap sign, this is a primitive reflex elicited by repetitive tapping on the forehead. Participants blink in response to the first few taps, although, if blinking persists this is abnormal (labelled as Myerson's sign) and is a symptom of PD.

2.3.6. Prodromal PD

Early PD diagnosis is essential for the administration of neuroprotective drugs given that treatment in the latter phases of the disease is not as effective due to the widespread decline of dopaminergic neurons. PD motor symptoms are apparent in the early phases of the disease, typically occurring following a loss of 60% of dopaminergic neurons [80], [81]. The premotor or prodromal PD phase is the period of time from the start of neurodegeneration and detection of clinical motor symptoms, between 5 and 20 years [82]. In the prodromal phase, individuals typically have non-motor symptoms (rapid eye movement, sleep behaviour disorder and olfactory dysfunction) [83]. Braak et al.'s staging system provides a pathological account for the occurrence of prodromal non-motor symptoms as Lewy pathology impacts on the lower brainstem before affecting the substantia nigra pars compacta in the midbrain [59]. For instance, it is likely that anosmia and rapid eye movement sleep behaviour disorder (two common prodromal non-motor symptoms) occur due to LB respectively penetrating the olfactory bulb and the pontine subcoeruleus nucleus [59]. Prodromal symptoms can be used in combination with biomarkers (cerebrospinal fluid, referred to as CSF, measurements and dopamine transporter imaging, referred to as DAT) to detect people with enhanced probability of developing PD [82]–[84].

Prodromal PD is typically misdiagnosed as another Parkinsonian disorder. Hence, a key step in ensuring accurate diagnosis involves early identification of prodromal PD patients, with consequences for patient treatment and disease management. The research presented in this thesis examines a novel diagnostic tool to identify prodromal PD patients using a brain imaging biomarker.

2.3.7. Assessment

PD progression is measured via the Unified Parkinson's Disease Rating Scale (UPDRS) [85] – a widely used and validated scale. The Movement Disorder Society sponsored revision of Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [86], [87] included rephrasing some questions to enhance clarity, adding new questions, and rephrasing other

questions to detect symptoms which may differ among patients in early PD. The UPDRS contains four sections: three sections focus on motor symptoms and the patient's ability to undergo routine daily activities and one section involves a motor exam (assessing speed, amplitude, and rhythm); the latter is essential in examining the various elements involved in bradykinesia. The motor exam involves patients completing 18 tasks and performance on each is evaluated on a 0-4 Likert scale, with ratings of 4 corresponding to more severe motor symptoms. Like the UPDRS, the MDS-UPDRS also evaluates motor symptoms via an aggregated score, which removes relevant information regarding a given patient's motor symptoms.

Much research reveals low reliability for bradykinesia assessments (relative to assessments of other PD symptoms) when using the MDS-UPDRS items [88]–[91] due in part to insufficient medical training regarding delivering the MDS-UPDRS [92]. Indeed, clinicians administering the MDS-UPDRS motor examination typically emphasise movement amplitude [93], rather than other features such as acceleration and velocity. Kishore et al. developed the Modified Bradykinesia Rating Scale (MBRS), which is more sensitive than the MDS-UPDRS regarding evaluating the multifaceted PD motor symptoms, even though it does not provide a completely objective and quantitative evaluation [94]. Correctly evaluating movement symptoms is essential with direct implications on the administration of medication. For instance, Levodopa improves movement speed better than movement amplitude or rhythm [95].

Cognitive PD symptoms are evaluated using the Montreal Cognitive Assessment (MoCA). This measurement tool was designed to detect mild cognitive impairment [96], although research has revealed that it can be suitably applied to PD screening [97]–[99] and it takes only 10 minutes to administer. MoCA provides an in-depth evaluation of cognitive abilities (visuospatial awareness, memory, attention, and language). The Clinical Dementia Rating (CDR) is another measure of cognitive impairment that, like the MDS-UPDRS focuses on cognitive implications on daily life.

Another relevant issue within the UK is that, sadly, there are not enough PD experts to meet demand and the services available differ significantly by region [100], with

implications on patient care and their quality of life. The limited support available underscores the need for developing automised methods of detecting PD biomarkers, as per the research presented in this thesis.

2.3.8. Diagnosis

UK guidelines suggest that individuals with Parkinsonian-like symptoms are referred to a specialist for diagnosis [101], [102]. The diagnostic criteria for PD are well established and are presented in Table 2.2 [7]. These criteria are specified by the UK Parkinson's Disease Society Brain Bank and National Institute for Health and Care Excellence (NICE) guidelines.

Table 2.2 - PD – UK Parkinson's Disease Society Brain Bank diagnostic criteria; taken from [103]

Step 1: Diagnosis of a Parkinsonian syndrome Bradykinesia (slowness of initiation of voluntary movement with progressive reduction in speed and amplitude of repetitive actions) and at least one of the following: (i) muscular rigidity, (ii) 4-6 Hz rest tremor and (iii) postural instability not caused by primary visual, vestibular, cerebellar or proprioceptive dysfunction.
Step 2: Exclusion criteria for PD 1. History of repeated strokes with stepwise progression of Parkinsonian features 2. History of repeated head injury 3. History of definite encephalitis 4. Oculogyric crises 5. Neuroleptic treatment at the onset of symptoms 6. More than one affected relative 7. Sustained remission 8. Strictly unilateral features after 3 years 9. Supranuclear gaze palsy 10. Cerebellar signs 11. Early severe autonomic involvement 12. Early severe dementia with disturbances of memory, language and praxis 13. Babinski's sign 14. Presence of cerebral tumour or communicating hydrocephalus on CAT scan 15. Negative response to large doses of levodopa (if malabsorption excluded)
Step 3: Supportive criteria for PD (three or more required for diagnosis of definite PD) 1. Unilateral onset 2. Rest tremor present 3. Progressive disorder 4. Persistent asymmetry affecting side of onset most 5. Excellent response (70 – 100%) to levodopa 6. Severe levodopa-induced chorea 7. Levodopa response for 5 years or more 8. Clinical course of 10 years or more

2.3.8.1. Misdiagnosis

Rates of accurate PD diagnoses are particularly low in both primary care by general practitioners [104] and secondary care by specialists [77], resulting in patients obtaining inadequate treatment. A challenge involves differentiating PD from alternative diseases that present similar symptomologies; a difficulty that is compounded in the early phases of PD. PD has rates of misdiagnosis of 25% as it is often mistakenly labelled as another

NDD, for example, PSP [77], [105]. Research reveals that 10% of PD cases are misdiagnosed as atypical Parkinsonism or a Parkinson's plus syndrome [106]. Further, post mortem research has revealed that up to 15-26% of PD cases were misdiagnosed by general neurologists [7], [8], with only 8-15% of misdiagnosis when diagnosed by expert movement disorder clinics [77], [107], [108]. Yet, due to limited financial and time-related resources, it is not feasible for movement disorder experts to diagnose each case of potential NDDs.

There are various common disorders that have Parkinsonian symptoms and which are frequently misdiagnosed as PD; these are presented in Table 2.3 [103]. These disorders typically do not respond well to levodopa and are frequently more symmetrical from the beginning, with minimal or no rest tremor. For example, essential tremor may be a more appropriate diagnosis for patients with a symmetrical limb tremor, sometimes including head and voice tremor, and sometimes due to autosomal dominant inheritance. Regarding essential tremor, there is no rigidity or Bradykinesia present and tremors are suppressed by alcohol. Asymmetrical rest tremor can, also, occur in patients with adult onset dystonia and these patients are sometimes misdiagnosed as *benign tremulous PD*, as their scans reveal no dopaminergic deficit. Furthermore, postural instability is characteristic of PD although this does not typically occur within the first year or two of PD. Postural instability is also characteristic of PSP, particularly when this occurs in the early stages and is accompanied by a history of falls. Vascular PD and drug induced PD are among other conditions that have Parkinsonian symptoms [103].

Table 2.3 - Differentiating commoner causes of Parkinsonism; taken from [103]

Condition	History	Clinical Features	Investigations	Management
Drug-induced Parkinsonism	Previous exposure to drugs mainly neuroleptic treatment and anti-emetics	May be associated with akathisia and oro-mandibular dystonia	Based on history	Discontinue offending drug. Anticholinergic drugs may be helpful for tremor
MSA	Parkinsonism and or gait unsteadiness with or without autonomic dysfunction	Orthostatic hypotension, absence of tremor, symmetrical signs, cerebellar features, erectile dysfunction, poor response to levodopa	MRI brain, sphincter EMG	Levodopa trial, amantidine measures to control postural hypotension, e.g., fludrocortisone
PSP	Early falls backwards, cognitive or behavioural changes	Gaze palsy (down more than up), axial rigidity, frontal and pyramidal signs, poor response to levodopa	MRI brain	Levodopa trial
Normal-pressure hydrocephalus	Urinary incontinence, ataxia, cognitive impairment	Dementia festinating gait	CAT or MRI brain, therapeutic lumbar puncture	Evaluate for ventriculoperitoneal shunt
Multiple lacunar strokes	Stepwise neurological impairment	Focal findings, sensory or motor loss	CAT or MRI Brain	Antiplatelet treatment, control of risk factors (e.g., diabetes, hypertension, increased cholesterol)
Cortico basal degeneration	Associated cognitive impairment	Marked asymmetry of clinical findings, dyspraxia, cortical sensory loss, myoclonus, dystonia, alien limb phenomenon, absence of response to levodopa	EEG, Psychometry	
Dementia with LB	Dementia occurring before or concurrently with Parkinsonism	Visual hallucinations	MRI brain, Psychometry	Consider cholinesterasae inhibitor

The research presented in this thesis identifies an imaging biomarker for PD using functional brain imaging (fMRI), an ideal tool that can increase diagnostic accuracy dramatically given that it is both non-invasive and not reliant on diagnostic tests that may involve subjective evaluations. Objective evaluations can be challenging, specifically given the complexity of certain symptoms, such as bradykinesia.

2.3.9. Treatment

Management of PD is achieved via a multidisciplinary team [109], in which a neurologist, geriatrician, or PD nurse specialist coordinates a team and ensures specialist consultations if and when appropriate. Healthcare professions necessary for PD treatment often include general practitioners, physiotherapists, occupational therapists, psychiatrists, psychologists, speech and language therapists, dieticians and palliative care physicians.

There is no cure for PD, hence, treatments are symptomatic and do not provide clinical neuro-protective effects. PD treatments focus primarily on medication to temporarily manage symptoms and, therefore, improve patient quality of life. Even though there is medication for PD, it is very difficult to confirm the diagnosis. Early diagnosis results in better effect of the medication for patients. Medication dosage is monitored to enhance time periods when patients exhibit few symptoms (time *on*) and to reduce periods of time when motor symptoms are prominent (time *off*). The *therapeutic window* is defined as the varying dosages of medication that result in *on* time. Outside the therapeutic window, patients will experience either sub-therapeutic levels PD symptoms (time *off*, see bottom section of Figure 2.3) or side effects (time *on*, see top section of Figure 2.3). The therapeutic window decreases as PD progresses over the years. In advanced PD, clinicians may administer less medication (increasing motor symptoms) to manage side effects associated with medication. Clinicians do not aim to treat all PD symptoms given that higher drug doses can result in greater short and long-term side effects.

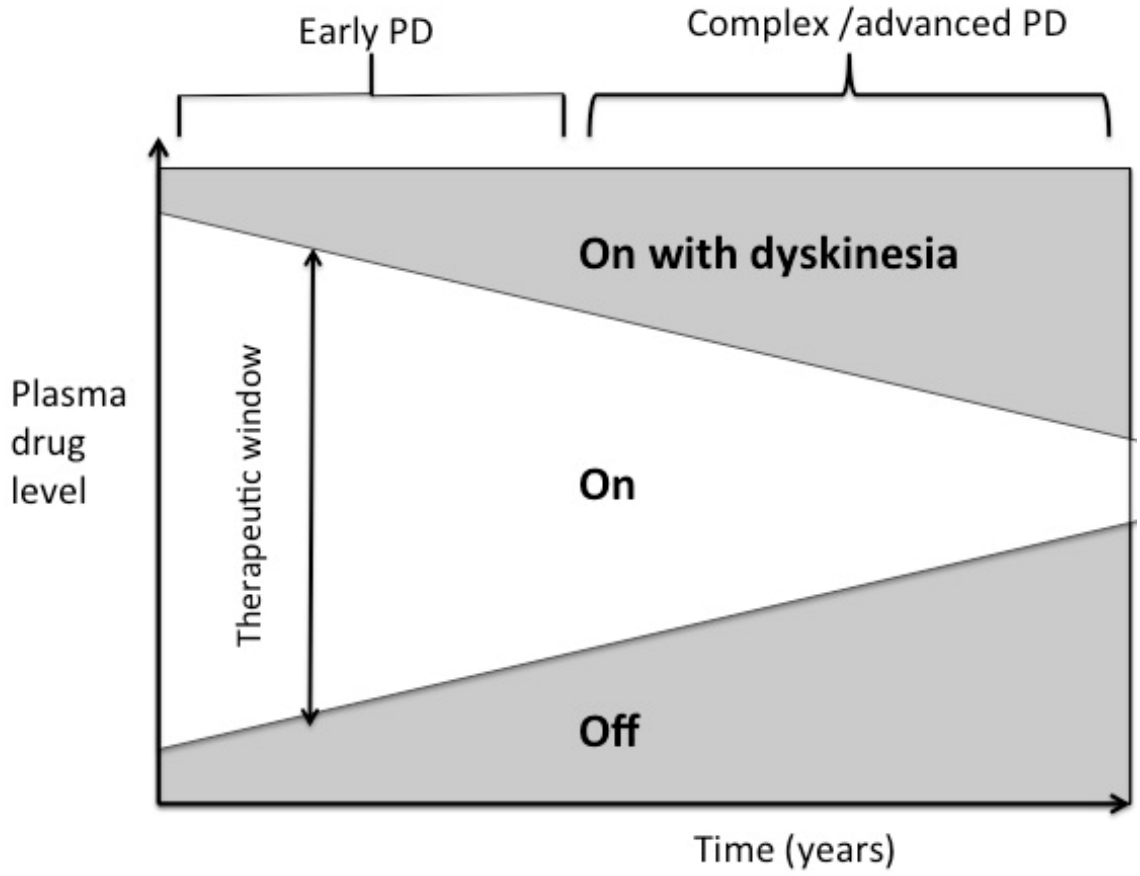

Figure 2.3 - Therapeutic window in PD

There is a variety of drugs used to treat PD, each with a specific potency and associated side effects. Treatment is tailored to each patient given the unique profile of PD symptoms exhibited, which changes according to the stage of PD. Indeed, NICE guideline specifies the following:

> *"There is no single drug of choice in the initial pharmacotherapy of early PD."*
>
> *"The choice of drug should take into account clinical and lifestyle characteristics of the patients and their preference once they have been informed of the short and long-term benefits and drawback of the drug classes."*

Hence, specialist monitoring of disease progression and drug side effects is essential for maintaining patient quality of life. Non-motor PD symptoms include disorders of mood, sleep, cognition and autonomic function and are often treated using another class of medication. The dopaminergic drugs to treat motor symptoms can themselves cause non-motor side effects similar to these PD non-motor symptoms. For advanced PD, it is sometimes recommended to administer medication through enteric or subcutaneous pathways or even for patients to obtain deep brain stimulation surgery.

2.3.10. Medication

Levodopa (this medicine is labelled co-beneldopa or co-careldopa) is prescribed for treating symptoms that impact on day-to-day movement and balance, which enables the brain to convert it into dopamine. When movement and balance symptoms are mild then other drugs are available as options. Table 2.4 compares three commonly prescribed drugs for early PD, including the benefits (e.g., minimising tremor) and costs (e.g., side effects) of each [110].

Schrag and Quinn conducted a review, revealing that 70% of PD patients are administered levodopa [43]. Indeed, levodopa is seen typically as the best treatment option available to PD patients [111], [112]. Side effects associated with levodopa can be severe, in some cases irreversible. These side effects include adverse motor complications, which are divided into two categories of motor fluctuations and dyskinesia (movement disorders similar to bradykinesia). These are labelled levodopa-induced dyskinesia (LID) [113]–[115] and their frequency increases with time spent using levodopa [116]. Hence, there is often a trade-off between PD symptoms and LID symptoms, typically managed by manipulating levodopa dosage or using alternative medication.

Table 2.4 - Comparing drug options for early PD; taken from [110]

Effects and side effects	Names of medicines		
	Levodopa	Dopamine agonists	MAO-B (monoamine oxidase B) inhibitors
Can it help me with movement problems such as slowness and tremor?	☺☺☺ More improvement	☺ Less improvement	☺ Less improvement
Can it help me with day-to-day activities?	☺☺☺ More improvement	☺ Less improvement	☺ Less improvement
Might it cause dyskinesia (uncontrolled movements)?	☹☹☹ More likely	☹ Less likely	☹ Less likely
Might I get other side effects such as sleep problems, impulse control disorder or psychotic symptoms?	☹ Fewer side effects	☹☹☹ More side effects	☹ Fewer side effects

Other PD symptoms include drooling, memory loss, and depression, amongst others. Drooling can be reduced with speech and language therapy or via mediation (e.g., anticholinergic medicines such as glycopyrronium bromide or atropine). Memory loss and confusion can occur as a result of PD, termed Parkinson's disease dementia, and can be treated with medication (cholinesterase inhibitors or memantine). Depression is frequent in PD patients and can be treated via therapy (e.g., cognitive behavioural therapy).

Common side effects from medication include abnormal daytime sleepiness, psychotic symptoms (hallucinations and delusions), orthostatic hypotension (light-headedness, faintness or dizziness when patients stand up rapidly), and impulse control disorders (e.g., compulsive gambling, binge eating, and obsessive shopping, amongst others). These can be treated by changing PD medication type/dose, taking additional medication (e.g., quetiapine or clozapine for psychotic symptoms [117]–[119], midodrine or fludrocortisone for orthostatic hypotension [120]–[122], and Modafinil for certain sleep-related conditions [123]–[126]), and/or cognitive behavioural therapy.

2.3.10.1. Modafinil

Modafinil (Provigil) is a multipurpose drug used in the treatment of sleep disorders, such as narcolepsy [127], [128], and it improves cognitive function in psychiatric disorders, for instance schizophrenia [129]–[131]. Approximately 50% of PD patients experience physical fatigue [124], for which Modafinil is a common medication that clinicians prescribe to PD patients [123]–[126]. Research has linked Modafinil to enhanced attention and memory [132]–[134] and it is well tolerated by patients and has no effect on PD movement symptoms [135]. More recently, Modafinil has been used as a "smart drug", with approximately 25% of students at Oxford, Newcastle, and Leeds having tried this drug [136]. In a 2013 interview with the Telegraph, Professor Sahakian of clinical neuropsychology (University of Cambridge) mentioned that the number of students taking Modafinil was increasing alarmingly [137]. There is much controversy regarding the use of psychostimulants or "smart drugs" to heighten cognitive abilities [138]. Research suggests that Modafinil results in less side effects and lower addiction risks, relative to stimulants such as methylphenidate [139] and amphetamine [140], with benefits similar to those provided by these drugs. Yet, Modafinil has a large impact on the dopaminergic system, indicating possibly stronger addiction risks than previously estimated [141].

Even though Modafinil is regularly prescribed to PD patients, evidence supporting its effectiveness as a fatigue-reducing drug is mixed. For instance, Lou et al. conducted a longitudinal study testing participants multiple times over two months and revealed that Modafinil can diminish physical fatigue in PD patients, as noted by the Epworth Sleepiness Scale and by a finger-tapping task following two months of treatment [125]. Other research has revealed that Modafinil relieves fatigue, as measured by the Global Clinical Impression Scale for Fatigue, and reduces excessive daytime somnolence [126]. Nevertheless, this same research revealed no relief from fatigue as measured by the Fatigue Severity Scale. A crucial point is that, regardless of the effectiveness of Modafinil in reducing fatigue, it is commonly prescribed to PD patients. Hence, the research presented in this thesis examines the classification of brain imaging data for participants administered Modafinil (versus a control group; see Chapter 5) to improve the validity of

treatment, to develop novel monitoring methods, and to further develop techniques for automatic PD biomarker identification (see Chapter 6).

2.3.11. Monitoring

Regular monitoring every six months of motor and non-motor symptoms via outpatient clinic appointments is necessary to ensure patients are administered appropriate levels of medication, given their changing symptomology (and medication side effects). Unfortunately, it is common for drugs administered to treat one symptom to also aggravate another symptom. For instance, D1 dopamine agonists reduce bradykinesia but worsen hallucinations and levodopa reduces tremor but worsens dyskinesia. Therefore, clinicians need to make a trade-off between the benefits (reducing PD symptoms) and costs (side effects) when prescribing medication.

2.3.12. Diagnostic Tools

PD diagnosis is based on clinical assessment although an autopsy is necessary for confirmation of diagnosis with, to date, no validated diagnostic PD biomarker. Key motor symptoms (Parkinsonism) are also present in other disorders, compounding difficulties in reaching a differential diagnosis. PD symptoms include motor disorders as well as non-motor disorders (olfactory and autonomic dysfunction, and sleep and cognitive symptoms) that can be apparent before motor symptoms occur [82], [142]. Hence, non-motor symptoms can be a prodromal/pre-clinical PD marker [143]–[145]. For instance, 90% of PD patients have olfactory dysfunction. Given that olfactory dysfunction is less frequent in other Parkinsonian diseases, this symptom can be used to differentially diagnose PD from other Parkinsonism diseases (e.g., the University of Pennsylvania Smell Identification Test and Sniffing Sticks reveal moderate sensitivity, 0.77, and specificity, 0.85 [146]).

Current diagnostic methods focus primarily on clinical assessment with limitations in correctly diagnosing early PD stages and differentially diagnosing PD from other disorders with Parkinsonian symptoms. Disease progression and treatment are also based on clinical assessment of symptoms (e.g., using the UPDRS scale), with a key focus on motor

dysfunction. Nevertheless, the presence of non-motor symptoms that can appear before motor symptoms has highlighted the need to re-conceptualise PD to provide a greater focus on non-motor clinical signs, such as anosmia. This can be used in addition to the current clinical criteria to provide a more nuanced account of early PD symptoms and to improve diagnostic accuracy. Still, some proposals for broadening PD assessment have been criticised for not mentioning dopamine deficit or false positives/negatives [147].

There exist many Parkinsonian conditions and tremulous movement disorders whose symptoms (particularly in the early stages) can overlap with those of PD. Hence, accurately diagnosing PD when using only clinical tools remains a challenge. Diagnostic tools for PD have primarily been tested in patients with confirmed PD diagnosis through longitudinal follow-up and/or blinded specialist assessment and, as such, their diagnostic accuracy has not been examined in early PD patients. Note that the majority of PD patients used in research on diagnostic tools did not undergo an autopsy to confirm their PD condition. Some clinicians adopt a *watchful waiting* strategy in which patient symptoms are observed over 6-12 months before a definitive diagnosis is issued. Alternatively, ancillary tests can be used for diagnosis in addition to clinical assessment. NICE guidelines directed at patients in England and Wales specify the following:

> *"PD should be diagnosed clinically and based on the UK Parkinson's Disease Society Brain Bank Criteria".*

Concerning ancillary tests, NICE guidelines specify that:

> *"123 I-FP-CIT SPECT should be considered for people with tremor where essential tremor cannot be clinically differentiated from Parkinsonism."*
>
> *"Structural MRI may be considered for the differential diagnosis of Parkinsonian syndromes".*

Up to 15-26% of patients with early stages of PD are misdiagnosed by clinicians who have limited PD expertise [7], [8], which underscores the need to identify a PD biomarker, as

per the work presented in this thesis. Beyond improving diagnostic accuracy, identifying an early stage PD biomarker would facilitate the development and assessment of treatment drugs with putative neuroprotective agents that could be administered to patients with early stages of PD, before long-term neural damage has taken place.

2.3.13. PD Biomarkers

A biomarker is "*a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention*" (Biomarkers Definitions Working Group [148]). Patients with early phases of PD are typically misdiagnosed and currently PD diagnosis can only be confirmed following an autopsy. Clear diagnostic PD criteria are challenging given the range of other disorders that share Parkinsonism symptoms, particularly in the early stages of the disease, yet this is essential before a biomarker can be identified and validated. To date, no PD biomarker has been identified to detect early stages of PD or to confirm PD diagnosis, which is the focus of the research presented in this thesis.

Given limitations in accurately diagnosing PD, research examining the identification of reliable, sensitive, reproducible, cost-effective, non-invasive and well validated biomarkers is required [149]. Biomarkers can be identified using clinical, imaging, genetic, proteomic, and/or biochemical data [82]. For instance, some research has already tested potential biomarkers using olfactory testing, tissue and fluid analysis, functional neuroimaging, and genetic risk factors. Indeed, future PD diagnosis is likely to rely on numerous interacting data sources, such as clinical, laboratory, imaging and genetics data [150]–[152].

Based on different stages of PD, three key PD biomarkers are required (Figure 2.4) [153], [154]:

1. Prodromal, preclinical, or premotor biomarkers: these are diagnostic with the objective of categorising early stage PD patients to administer neuroprotective therapies before significant neuronal loss occurs. Premotor symptoms may be caused by pathogenic processes with etiological implications [155].

2. Risk or susceptibility biomarkers: identify individuals particularly likely to develop PD.
3. Motor stage biomarkers: to chart PD progression and to examine treatment efficacy, occurring when motor dysfunction is apparent.

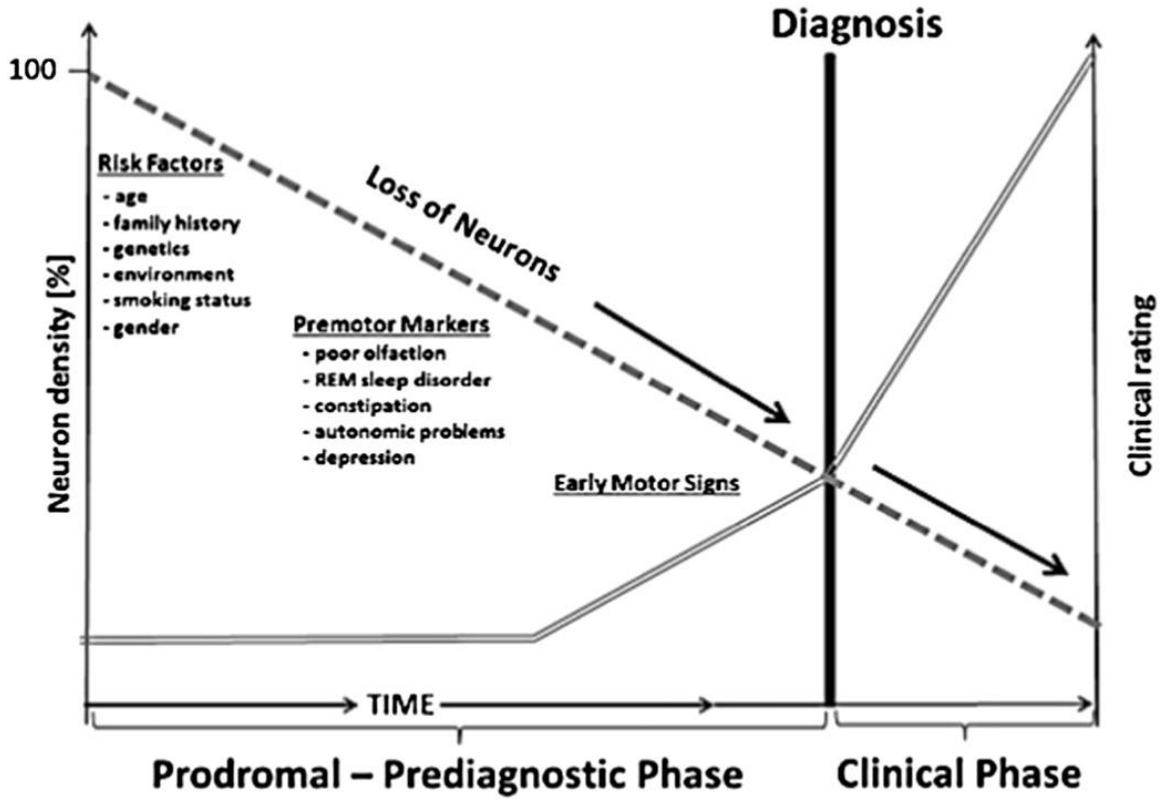

Figure 2.4 - Factors and premotor markers associated with loss of neurons (----) prior to onset of motor signs and clinical diagnosis (====) of PD; taken from [42]

Much research has been conducted to identify prodromal, motor stage and other biomarkers, with limited success. Possible reasons for this limited success involves the participant sample characteristics (PD research typically focuses on patients with a confirmed diagnosis and in the late PD stages), diagnostic criteria, and poor storage and collection of data [156], [157]. The Parkinson's Progression Markers Initiative (PPMI) focuses on eliminating many of these key limitations by aiming to recruit equal numbers of early stage (pre-medication) PD patients and healthy age-matched controls [158]. PPMI contains 20 centres across the EU and the USA and these follow standardised procedures for repeated bio-sampling (blood, CSF, urine), clinical assessments, and imaging as well as

rigorous standards for data storage and analysis. The NINDS has adopted a similar approach for its Parkinson's Disease Biomarkers Program, which involves research collaboration among consortium members, a data repository and a Data Management Resource.

2.3.13.1. Genetic and Molecular Biomarkers

Genetics is now a known PD risk factor, yet, research examining links between single nucleotide polymorphisms in various candidate genes (e.g. ones related to detoxification like Cytochrome P450s gene family or dopamine) has revealed either weak or no relationship with PD risk [159]–[161]. Genome wide association surveys (GWAS) have explored 16 PD loci although these account for a small percentage of heritability. Nevertheless, rare early onset familial PD due to mutations in a single gene has underscored the relevance of genetic factors. For instance, point mutations, duplications, and triplication in the SNCA[10] gene are related high risk of PD and some single nucleotide polymorphisms[11] of the SNCA gene are also linked to PD [153], [162]. As such, research on candidate genetic biomarkers has focused on mitochondrial dysfunction and mutations in mitochondrial genes (e.g., SNCA) and gene products (alpha-synuclein). Indeed, certain genetic mutations can result in familial Parkinsonism. Tests for these genetic mutations are costly and typically restricted to research labs. Exceptions include recessive Parkin gene mutation tests directed at young onset PD patients that is positive in 5% patients under 40 years old and LRRK2[12] tests in patients with autosomal dominant pedigrees that is positive in 5-6% these patients [163].

PD involves damage to the pigmented dopamine neurons in the substantia nigra, leading to motor dysfunction (bradykinesia, tremor at rest, rigidity, and postural instability). Post mortem brain tissue has highlighted key molecular pathways and genes, enabling targeted therapies, the advance of animal PD models, including novel drug delivery mechanism [153], [164]–[167]. Post-mortem examination has focused on the molecular

[10] The SNCA gene provides instructions for making a small protein called alpha-synuclein.
[11] The most common type of genetic variation in humans
[12] The LRRK2 (Leucine-rich repeat kinase 2) gene provides instructions for making a protein called dardarin.

neuropathology [168] in an effort to identify both a prodromal biomarker and potential treatment to stop/reverse disease progression.

Post mortem research has primarily explored two areas: firstly, the degeneration and death of the melanin containing neurons of the substantia nigra and, secondly, Lewy pathology (intra-cytoplasmic LB with inclusions containing alpha-synuclein and ubiquitin). Cell death is a key research focus given that this occurs in the dopamine containing neurons of the substantia nigra. Findings have revealed that PD brains are characterised by differences in mitochondrial function, greater oxidative stress, lysosomal dysfunction, protein aggregation and reduced degradation, deposition of iron, including inflammation and glial activation. Current methods for identifying potential biomarkers focus on global non-targeted procedures, for instance, omics (e.g., genomics, proteomics, metabolomics, etc.), comparing tissue samples from PD patients to healthy age-matched controls [145], [150], [169]–[171]. These findings are relevant in identifying a possible PD biomarker (e.g., alpha-synuclein) and to advance animal models of PD [149], [167], [168], [172]–[174].

PD neuroprotection clinical trials have explored these identified mechanisms (e.g., neuronal loss such as apoptosis and mitochrondrial dysfunction), yet, no neuroprotective or PD modification treatments have been successfully developed. It is likely that multiple mechanisms are involved in disease progression and these would need targeting in drug trials. Hence, further research is required to identify validated PD biomarkers across different disease phases to gain a more complete understanding of the disease. The research presented in this thesis focuses on identifying a brain imaging biomarker for both early stage PD as well as for the clinical phase of PD.

2.3.13.2. Functional/Behavioural Biomarkers

Early stage PD non-motor symptoms (e.g., dysfunction in olfaction, sleep, visuospatial abilities, cognition, and executive function) are due to damage to the extra nigral areas, occurring prior to degeneration of dopamine nigral neurons [175]. Functional tests may signal whether individuals are at risk of developing PD. Benefits of these tests include that they are non-invasive, cost effective, often simple to administer, and may contain some

tests that can be completed by the patient at home and/or online. Functional measures include olfaction acuity tests (e.g., the University of Pennsylvania Smell Identification Test – the UPSIT), the rapid eye movement sleep behaviour disorder (RBD)[13] [84], [176] screening questionnaire, a keyboard tapping test, the bradykinesia akinesia incoordination test (BRAIN), and accelerometer based exams. These measures are currently used in research (e.g., PREDICT-PD) and can potentially be developed for identifying at risk individuals. Data from functional tests can be used in combination with other risk factors (e.g., family history) and used as inputs to algorithms to automatically calculate a given patient's risk factor [177]–[180], although there are ethical issues regarding false positives/negatives [152], [155], [180]. The research presented in this thesis examines the use of EAs to automatically detect PD patients from healthy age-matched controls.

2.3.13.3. Brain Imaging Biomarkers

Brain imaging techniques enable clinicians and researchers to determine brain activation and neurological problems, without invasive neurosurgery. These methods can provide relevant information regarding brain structure and function, in addition to standard clinical assessments. The most commonly used brain imaging tools are described as follows.

2.3.13.3.1. Computed Axial Tomography (CAT)

CAT scanning (also known as computed tomography, CT, or computed aided tomography) uses the differential absorption of x-rays to create a brain image. During a CAT scan, an individual lies on a table that moves inside and outside of a hollow, cylindrical apparatus. Figure 2.5 illustrates a CAT scanner with a participant. An x-ray source is positioned on a ring around the inner section of the tube, with its beam directed at the individual's head. Once the individual's head moves past the x-ray beam, the beam is sampled by one of the various detectors that are positioned on the apparatus' circumference. The images generated are influenced by the absorption of the x-ray beam concerning the body tissue that moves through this beam. Bone and hard tissue have good absorption of x-rays,

[13] Approximately 50% of people who are diagnosed with REM-sleep behaviour disorder in adulthood will develop neurodegenerative Parkinsonism with mean latency of 12 years (61)

although air and water have poor absorption, and soft tissue has mediocre absorption. Hence, CAT scans reveal the structural brain elements but cannot provide information regarding specific brain functioning during cognitive tasks. Adverse effects of CAT scans include tissue damage resulting from prolonged or repeated exposure to ionising radiation. As such, like x-rays, CAT scans are used cautiously.

Figure 2.5 - CAT Scan; taken from [181]

CAT brain scans are not typically used for PD patients as these tend to reveal normal images in PD patients or indicate incidental basal ganglia calcification or age-related changes, for instance, generalised atrophy. CAT scans are useful when other neurological

symptoms that might indicate normal pressure hydrocephalus[14] or space occupying lesion in the brain are present.

2.3.13.3.2. Positron Emission Tomography (PET)

PET uses trace levels of short-lived radioactive material to map functional processes in the brain, similar to the action that glucose or oxygen perform in the brain. During radioactive decay, positrons are emitted and these are detected. Figure 2.6 depicts a PET scanner. Regions with high radioactivity are correlated with brain activity. Disadvantages of PET scans include that they are relatively expensive, inaccessible, and have poor temporal (40 seconds) and spatial (1 cm) resolution.

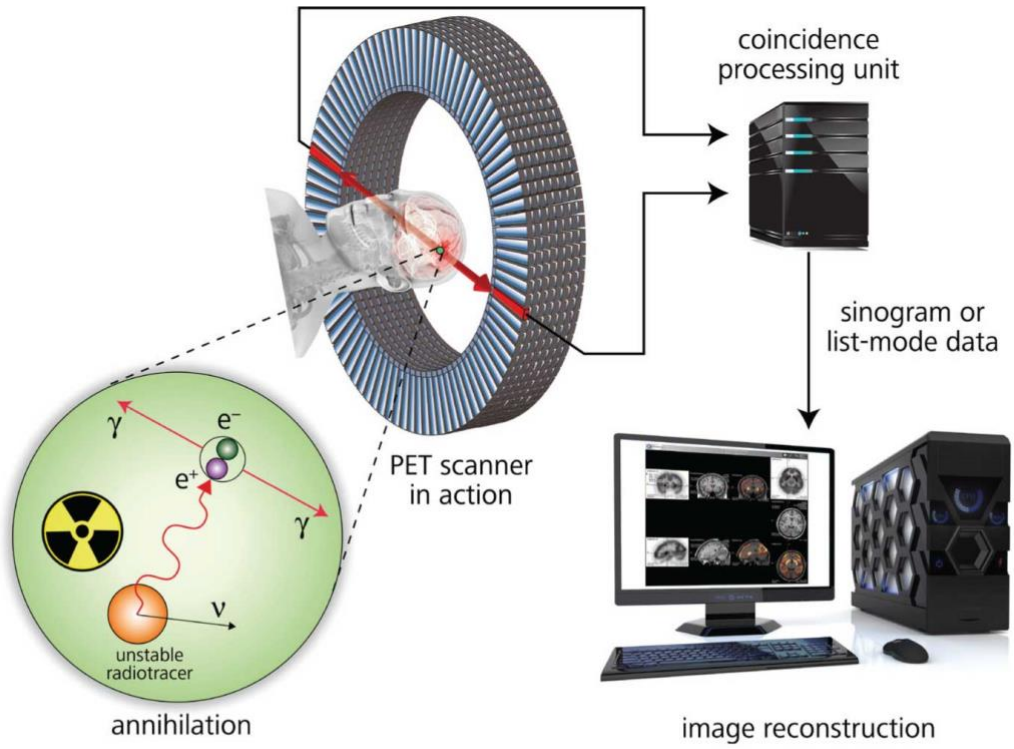

Figure 2.6 - PET Scanner; taken from [182]

Research has explored the use of PET as a diagnostic tool for PD, and typically involves a fluorine-18-labelled-dopa isotope; although PET is not currently advocated by NICE. PET

[14] Normal pressure hydrocephalus occurs when excess CSF accumulates in the brain's ventricles, which are hollow fluid-filled chambers.

can differentially diagnose PD, MSA, PSP, and corticobasal degeneration[15] from each other with good sensitivity [183]. In addition, findings reveal diminished asymmetric pre-synaptic striatal uptake in PD patients and pre-clinical nigrostriatal impairment in first degree relatives of PD patients (note that some relatives later developed PD in follow-up tests) [184]. Limitations of PET include cost, radiation exposure to patients, and poor scan accessibility for clinicians.

2.3.13.3.3. Single Photon Emission Computed Tomography (SPECT)

SPECT is a nuclear imaging method, which integrates CAT and a radioactive material (tracer), that allows the clinicians to see how blood flows to tissues and organs. The process involves first injecting radionuclide into a participant and, after 3-6 hours, a SPECT head scan is then performed to capture the uptake in the nigrostriatal (presynaptic) nerve endings. The computer collects the information emitted by gamma rays and translates them into two-dimensional cross-sections. These cross-sections can be added back together to form a 3D image of a given organ (e.g., the brain). The radioisotopes typically used in SPECT to label tracers are iodine-123-beta-CIT labelled ioflupane (FP-CIT)[16], technetium-99m, xenon-133, thallium-201, and fluorine-18. These radioactive forms of natural elements pass safely through the body and can be detected by the scanner. To safeguard patients from iatrogenic hypothyroidism (reducing thyroid uptake of radio-iodine) patients are administered potassium bromide two hours prior to the radionuclide injection and 24 hours following the injection. Further, patients are advised to refrain from consuming medication for a few days prior to the scan as these can affect FP-CIT uptake. Within each scan, patients are exposed to the same as one year's background radiation exposure.

SPECT tests differ from PET scans in that the tracer remains in the blood stream rather than being absorbed by surrounding tissues (as per PET), thereby limiting the images to

[15] Also known as CBD. This is a rare and progressive NDD involving the cerebral cortex and the basal ganglia. The symptoms typically begin in people from 50–70 years of age, and the average disease duration is six years. It is characterised by marked disorders in movement and cognitive dysfunction, and is classified as one of the Parkinson plus syndromes.
[16] A DatSCAN is the trade name for SPECT imaging of FP-CIT

areas where blood flows. SPECT scans are cheaper (cost for one scan: £1,000 [185]) and more readily available than higher resolution PET scans (cost for one scan: over £5,000). The amount of radiation that the body is exposed to is less than an x-ray or CAT scan, however, SPECT is still an invasive imaging technique.

Research has revealed that SPECT imaging of the striatal dopamine transfer activation can be used in addition to clinical assessment to diagnose PD. Decreased FP-CIT binding in the contralateral striatum to the Parkinsonian limb is associated with PD. SPECT imaging can detect early phases of PD, given that even before the appearance of motor symptoms PD patients have 40-50% reduction in the striatal dopamine transfer activation [186]–[188]. Individuals with early PD symptoms, such as olfactory deficit [189] or RBD [84], [176], also reveal abnormal SPECT scans. Findings indicate that FP-CIT SPECT scans can differentially diagnose PSP, CBD, MSA, and dementia with LB [17] relative to PD. Nevertheless, this research typically does not use participants within the early phases of these diseases [188], [190]–[193], hence, limiting its use as a clinical tool. SPECT imaging can result in false positives for individuals without PD but with other degenerative Parkinsonian conditions (Table 2.5) [194]. NICE and SIGN guidelines suggest using FP-CIT SPECT imaging for confirmation of PD diagnosis for patients who potentially have another disorder causing a non-degenerative Parkinsonism or non-Parkinsonian tremor.

Table 2.5 - Conditions with abnormal FP-CIT SPECT scan; adapted from [194]

Abnormal scan	Normal scan
PD	Essential tremor
PSP	Dystonic tremor
MSA	Drug-induced Parkinsonism
CBD	Toxin-induced Parkinsonism
DLB	Normal pressure hydrocephalus
HD	Psychogenic Parkinsonism

Limitations include poor specificity, patients exposed to radiation, cost, potential subjectivity, and poor scan accessibility of clinicians both within countries and on a

[17] Also known as LBD. This is a type of dementia that shares symptoms with both Alzheimer's disease and Parkinson's disease.

worldwide level as few countries possess scans. Further, evaluation of scans can be subjective and dependent on the radiologist's experience, yet, even though there is no established 'cut off' point for abnormality, there is good inter-rater (kappa 0.82-0.92) and intra-rater (0.92-1.00) agreements [188], [192]. Computational methods are being developed and validated to automate this process and reduce subjectivity.

2.3.13.3.4. Dopamine Active Transporter (DAT)

A DAT brain scan examines the function of dopamine transporters and is typically used to explore nigro-striatal dopaminergic neurodegeneration linked to Parkinsonian Syndrome, revealing good specificity [188], [192], [195]–[198]. 123I Ioflupane (radioactive tracer) is used to bind to the DAT in the striatum and then SPECT visualises the level of transporter available. These transporters typically reuptake dopamine from the synaptic cleft and are integral to preserving the presynaptic neuron. PD patients reveal 50-70% less DAT transporters [199]–[201].

A limitation of this approach is that it involves injecting patients with a radioactive tracer (conveying a small risk to patients, like a CAT scan) 3-5 hours before the DAT scan using a gamma camera. The scan lasts 30-45 min and participants are required to remain still during this period (often problematic for PD patients). Figure 2.7 (a) depicts example DAT brain images, (b) represents a patient being injected with a radioactive tracer, and (c) portrays a participant undergoing a DAT scan.

Figure 2.7 - DAT Scan; taken from [202]

DAT scans cannot discriminate between Parkinsonian Syndrome disorders [203], [204]. Yet, DAT scans are useful to examine patients with ambiguous Parkinsonian Syndrome

symptoms, enabling clinicians to distinguish between disorders of essential tremor or drug-induced Parkinsonian Syndrome, as the latter is not associated with reduction in nigro-striatal dopamine transporter) whereas Parkinsonian Syndrome does reveal DAT degeneration [192], [195], [197], [203], [205], [206]. Hence, DAT scans are a relevant tool in clinical assessment and control of patient medication and treatment [207].

2.3.13.3.5. Electroencephalography (EEG)

EEG involves the measurement of brain electrical activity using electrodes positioned on the scalp. In 1924, the first EEG recordings on humans were published. Between 1935 and 1936, Pauline and Hallowell Davis recorded the first event-related potentials on awake humans (published in 1939). The data collected is the electrical signal from many neurons, as illustrated in Figure 2.8. Via the use of averaging methods, event-related potentials enable the extraction of detailed sensory, cognitive, and motor events. The amplitude of a normal EEG has a range of -100 and +100 microvolts. The EEG recorded at the scalp is the total measure of the activity from numerous neurones and various frequency patterns may be identified. There are four waves within the normal EEG: (1) delta rhythm at 1-4 Hz, (2) theta rhythm at 4-8 Hz, (3) alpha rhythm at 8-12 Hz, and (4) beta rhythm at 13-20 Hz. In order to analyse the EEG output, the amplitude and the frequency of waves is used. When individuals are relaxed, wave amplitude is typically larger and frequency is lower, whereas the reverse occurs when individuals are excited.

Figure 2.8 - EEG Scan; taken from [208]

Approximately 20 electrodes are placed on an individual, using the Jasper '10/20' system [209], in which electrodes are located with regards to the location of the frontal, central, temporal, parietal, and occipital lobes. On the left side of the head there is an odd number of sites, on the right side there is an even number of sites, and midline electrodes are referred to as 'z'. The most frequently used type of event-related potential recording is made between a scalp electrode positioned at an area of interest and a reference electrode positioned at a site that is relatively unaffected by the electrical activity of experimental interest. Recordings are based on the difference in voltage between each exploring electrode and the common reference electrode. Participants need to remain still as any facial movement (e.g., eye and jaw movements) can result in fluctuating electrical fields across the scalp. As such, eye movements are measured with the EEG so that trials with large eye movements can be removed or corrected from the data analysis. Current systems can record from up to 128 sites simultaneously using a *geodesic dense array sensor net*.

By creating event-related potential time-stamped maps, following the stimulus event, the relative times at which specific brain regions are activated can be identified.

EEG is currently used widely in research, given that it is non-invasive, able to detect millisecond-level changes in electrical activity, and has high temporal resolution. Disadvantages of EEG include low spatial resolution, low signal to noise ratio, poor measurement of neural activity within the upper brain sections (the cortex), and that it is time consuming to connect participants to the EEG.

2.3.13.3.6. Magneto encephalography (MEG)

MEG measures the magnetic fields produced by brain's electrical activity, using highly sensitive devices labelled superconducting quantum interference devices. MEG is an imaging technique that is used frequently in research and in clinical settings, as portrayed in Figure 2.9, and has applications in aiding surgeons to locate pathology and in aiding researchers to examine the function of specific brain regions, amongst other applications. MEG has high temporal and spatial resolution, although it needs sensitive apparatus and methods that reduce environmental magnetic interference.

Figure 2.9 - MEG Scan; taken from [210]

2.3.13.3.7. Near infrared spectroscopy (NIRS)

NIRS is an optical technique for measuring the brain's blood oxygenation. NIRS involves shining light in the near infrared part of the spectrum (700-900nm) through the skull and examining the degree of attenuation in the remerging light (see Figure 2.10). The level of light attenuation depends on blood oxygenation and, hence, NIRS can provide an indirect measure of brain activity.

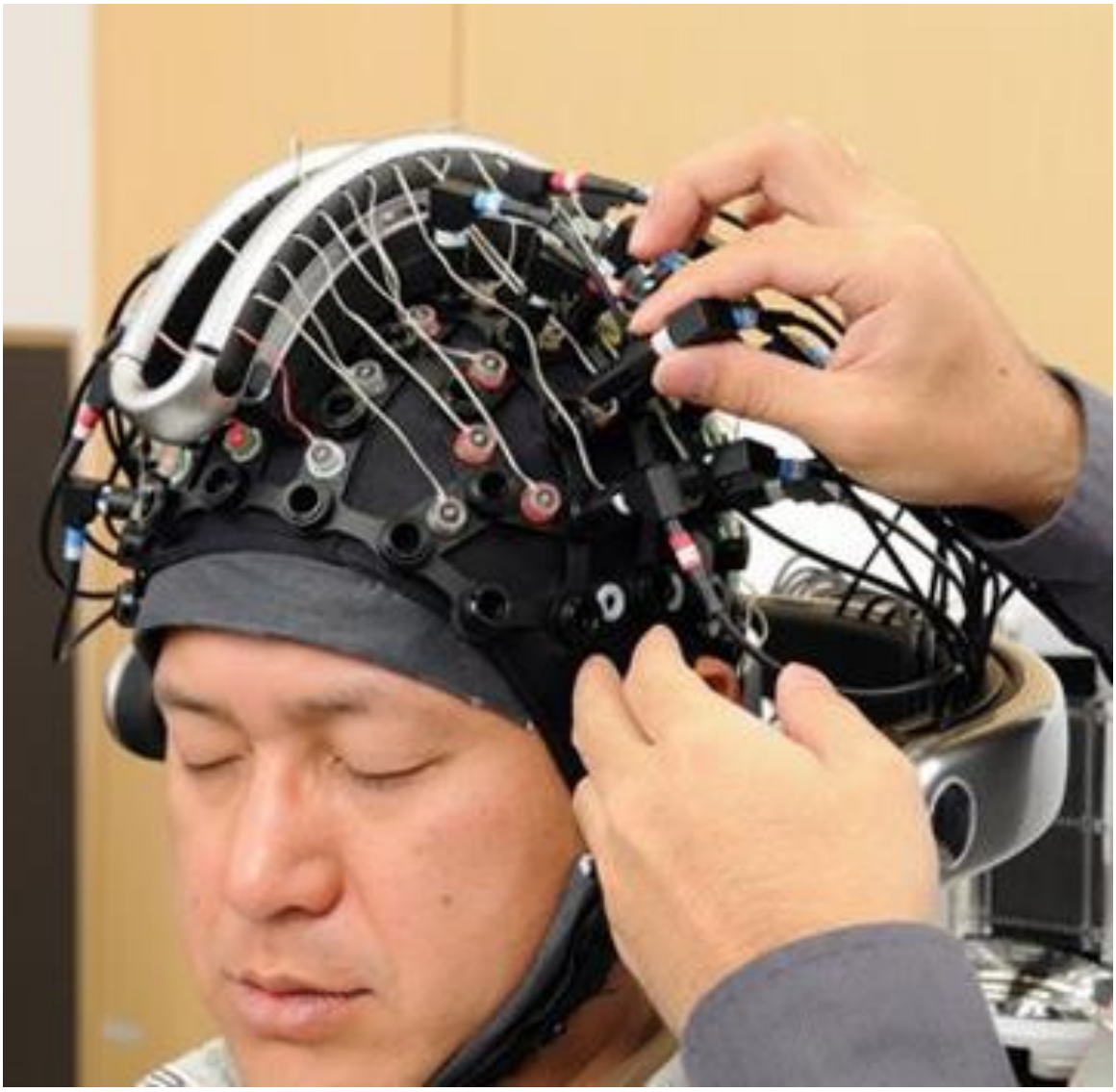

Figure 2.10 - NIRS; taken from [211]

2.3.13.3.8. Magnetic Resonance Imaging (MRI)

MRI has furthered current knowledge significantly over the last three decades, contributing to better disease diagnoses and to the development of treatments. Prior to MRI, methods involving ionising radiation (e.g., x-rays, CAT, and PET) were commonly used although, unlike MRI, these methods are not safe due to the radiation and, also, they are not applicable to the large range of tissue that is currently examined using MRI. Benefits of MRI over CAT include identifying blood flow, cryptic vascular malformations, demyelinating diseases, whilst having no beam-hardening artefacts (as per CAT). For

instance, the posterior fossa can be more easily captured via MRI relative to CAT. As such, in the 1980s, MRI became the standard medical imaging tool. In addition, MRI creates higher resolution images relative to PET. MRI is comparable to an x-ray placed within a large doughnut-shaped magnet. Figure 2.11 is tentatively dated to 1980 and shows Sir Peter Mansfield, in shirt and tie, with his colleagues a couple of years after they completed their first MRI whole-body scans of Sir Peter himself. They had completed the first MRI images of living tissue (the fingers of research student Andrew Maudsley) in 1976, resulting in large advances in the detection of body tissue [212].

Figure 2.11 - The team behind some of the whole-body MRI scans; taken from [213]

In contrast to CAT, structural MRI brain scans have been used to differentially diagnose PD relative to degenerative Parkinsonian conditions (e.g., MSA, PSP, and CBD). Nevertheless, structural changes are visible in the later stages of PD and at this point a PD diagnosis has usually already been confirmed based on clinical assessment. Research has used PET, SPECT, and MRI to classify PD relative to patients with other movement disorders. Neuroimaging methods are costly, which restricts their use in typical clinical

settings [82], [214]–[216], although a MRI scan costs (cost for one scan: £400-600) significantly less than a PET or SPECT scan. Functional MRI (fMRI) is when participant's neural activity is examined whilst the participant completes a given task (e.g., a perception, language or memory task). fMRI will be discussed further in Chapter 3.

2.4. Summary

Difficulties associated with differentially diagnosing PD abound, although these are reduced with greater clinician training/experience and disease progression [107], [217]. Community investigations have revealed that PD diagnoses have 15-26% false positive rates and this is particularly problematic in early stages of PD. PD is typically misdiagnosed with other degenerative Parkinsonian conditions and with late onset tremor [104], [217]. Likewise, research on pathology has indicated that 10-24% of PD diagnoses are not confirmed pathologically and PSP, MSA, and AD are the most frequent revised diagnoses. Further, clinical trials employing rigorous diagnostic assessment criteria have revealed that 4-15% of early stage PD diagnoses do not have abnormal DAT scans. Misdiagnosis hinders patient care and medication. Imaging methods (e.g., FP-CIT SPECT scans and DAT) can provide additional data to confirm a PD diagnosis, although some of these methods have clear drawbacks including low specificity, exposing patients to radiation, cost, and accessibility to clinicians. Hence, fast and non-invasive methods with high accessibility (e.g., able to be administered in clinical environments or even at patient's own home, such as portable MRI scanner trucks) are clear next steps for future research. The research provided in this thesis introduces a novel and highly accurate brain imaging (resting state functional magnetic resonance imaging) biomarker for PD.

Chapter 3. Functional Magnetic Resonance Imaging (fMRI)

fMRI captures brain activity by measuring changes linked to blood flow. Given that cerebral blood flow and neuronal activation are associated with one another, active brain regions also have increased blood flow. The research presented in this thesis is based on rs-fMRI data. Hence, this chapter outlines the physics underlying MRI and fMRI images and provides a review of the steps involved in fMRI data analysis: data preprocessing and processing.

3.1. MRI

MRI is a technique of imaging the body that is both non-invasive and non-ionising; this method is widely used in research and clinical environments. Relative to other imaging techniques, MRI produces high soft tissue contrast (such as diffusion of water with diffusion weighted imaging, blood flow measurement, and metabolite concentration mapping with chemical shift imaging) and can be used flexibly on a range of tissue properties to supplement anatomical readouts [218]. MRI is particularly widespread in both neurology and neurosurgery as it generates high definition images of the brain, spinal cord and vascular anatomy, focusing on the axial plane, sagittal plane and coronal plane, as can be seen in Figure 3.1.

MRI images are derived from the magnetisation characteristics of atomic nuclei. A strong, uniform, external magnetic field is applied, which aligns the protons that usually have a random orientation within the water nuclei of tissue. Subsequently, a burst of radio frequency (RF) energy is applied to this alignment/magnetisation to modify the alignment. The nuclei release RF energy as they regress to their original resting alignment via relaxation processes. Following a specific time interval after applying the initial RF, the released signals are then calculated. The signal is collected and encoded in the frequency domain (k-space). Reconstruction via multi-dimensional Fourier transformations generates the final MRI image. Various types of images can be generated by changing the administered and measured sequence of RF pulses. Repetition Time (TR) corresponds to the time between successive pulse sequences administered to a given slice. Echo time (TE)

corresponds to the time between the delivery of the RF pulse and the receipt of the echo signal.

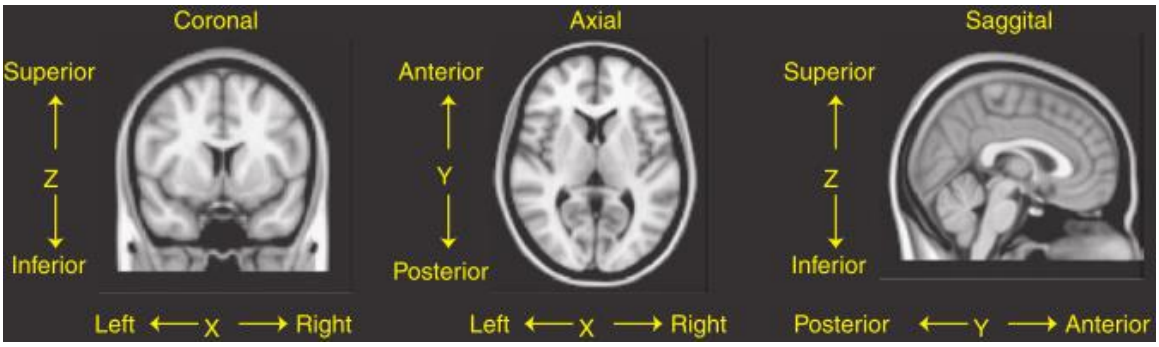

Figure 3.1 - The axial plane, sagittal plane and coronal plane used in the standard coordinate space for MRI; taken from [219]

MRI signal relies on both the local proton (or nuclei) density (with greater proton density relating to greater signal intensity) and the relaxation processes (reduction of transverse/detectable magnetisation). There are two types of relaxation processes that can influence the signal, leading to complementary contrast mechanisms: (1) T1 longitudinal relaxation time: the time constant related to the rate excited spinning protons return to equilibrium and realign with the external magnetic field. T1-weighted images are generated via brief TE and TR times, and the image contrast and brightness are linked to the T1 tissue characteristics. Brain images of T1 contrast contain dark (low signal intensity) ventricles, grey matter is represented as grey (medium intensity) and white matter (WM) is represented as bright/white (high intensity). (2) T2 transverse relaxation time: the time constant that regulates the rate at which excited spinning protons regain equilibrium and lose phase coherence with one another. T2-weighted images are generated by longer TE and TR times, and the image contrast and brightness are related to the T2 tissue characteristics. T2-weighted images reveal bright/white ventricles, with grey matter represented as grey and WM represented by dark regions.

MRI sequences typically involve T1-weighted and T2-weighted scans, as well as Fluid Attenuated Inversion Recovery (termed Flair). These scans can be distinguished by examining the CSF, which is dim in T1-weighted images and light in T2-weighted images. The Flair sequence is generated by very long TE and TR times and it is useful in

identifying pathology as it creates images in which abnormalities are emphasised by being light whilst normal CSF remains dim. Table 3.1 and Table 3.2 respectively depict typical MRI sequences and the shading of key brain tissues. Figure 3.2 presents a T1-weighted, T2-weighted and Flair brain image.

Table 3.1 - Most common MRI sequences and their approximate TR and TE times

	TR (ms)	TE (ms)
T1-weighted (short TR and TE)	500	14
T2-weighted (long TR and TE)	4000	90
Flair (very long TR and TE)	9000	114

Table 3.2 - How tissues are observed in different MRI sequences

Tissue	T1-Weighted	T2-Weighted	Flair
CSF	Dark	Bright	Dark
WM	Light	Dark grey	Dark grey
Cortex	Grey	Light grey	Light grey
Fat (within bone marrow)	Bright	Light	Light
Inflammation (infection, demyelination)	Dark	Bright	Bright

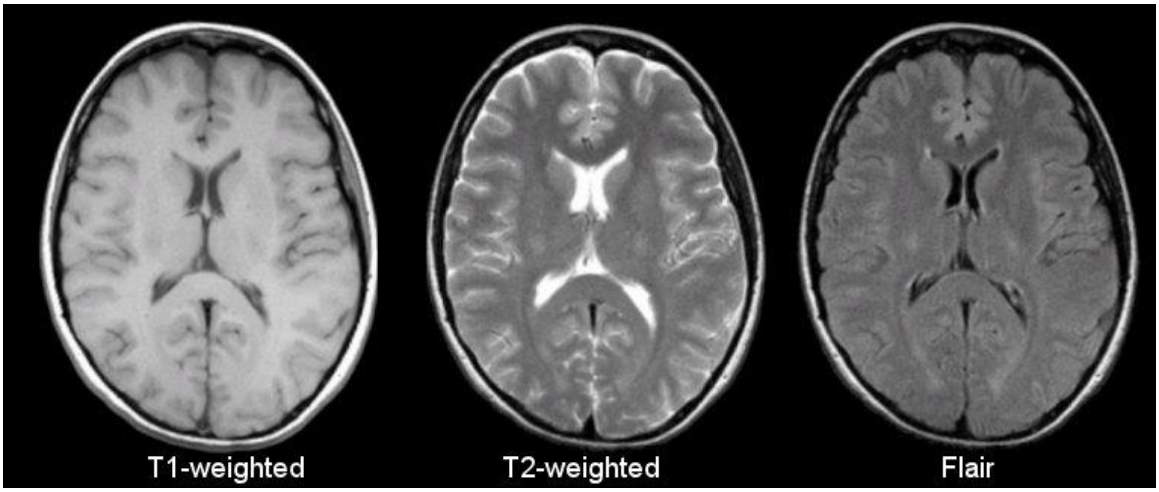

Figure 3.2 - A comparison between T1, T2, and Flair

MRI involves low signal to noise ratio (SNR) because of low energy in the resonance process. The signal intensity is dependent on the number of nuclei generating the signal. Augmenting the signal and/or reducing noise are two strategies to heighten SNR. Higher field magnets produce a larger magnetisation and boost the signal, although these can result in greater image artefacts because of magnetic field inhomogeneities. Decreasing thermal noise generated by the detector electronics (via cryogenic coils) diminishes overall noise and can double and even triple SNR [2]. Hence boosting magnetic field strength and using a cryogenic detector coil enhances SNR, resulting in improved MRI image accuracy.

3.1.1. The BOLD Effect and Hemodynamic Response

Research by Ogawa in 1990 on tubular hypointensities in the rat cortex revealed the contrast technique in MRI of the paramagnetic effect of deoxygenated blood: gradient echo MRI sequences are related to T2* contrast but spin echo sequences are related to T2 [220]. Ogawa posited this effect can be applied to detect hemodynamic response using MRI via the blood oxygenation level dependant contrast (BOLD). This approach is currently widely used as it provides a non-invasive method of detecting the functional response to a stimulus, with a high spatial resolution over the entire brain. BOLD fMRI is superior to other imaging methods (e.g., SPECT, PET) since bolus tracking or radioactive compounds are not required.

The BOLD response to brain activity is an intricate process. Brain activation is related to heightened glucose and oxygen use, although the BOLD signal is characterised by a rise in the fMRI signal within putatively active areas, suggesting greater blood oxyhemoglobin to deoxyhemoglobin ratio. Following the presentation of a stimulus, there is a minor reduction in signal corresponding to a reduced blood oxygenation to deoxyhemoglobin ratio. Subsequently, the signal heightens, then plateaus, and finally falls to baseline values with a minor undershoot. The imbalance between oxygen consumption, cerebral blood volume and flow reaction, and mechanical resistance in the capillaries all contribute to the final shape of the response.

3.2. fMRI

fMRI was designed in the 1990s and has now become one of the standard and most frequently used methods for examining neural activity, which is represented in Figure 3.3. Carp used a search string on pubmed.org for counting the number of human fMRI studies per year [221], substituting YYYY with the year of interest, and the research presented in this thesis has updated Carp's search to reveal the current state of human research using fMRI. As evident in Figure 3.3, there has been an explosion of fMRI research, with over 2300 papers published in 2015.

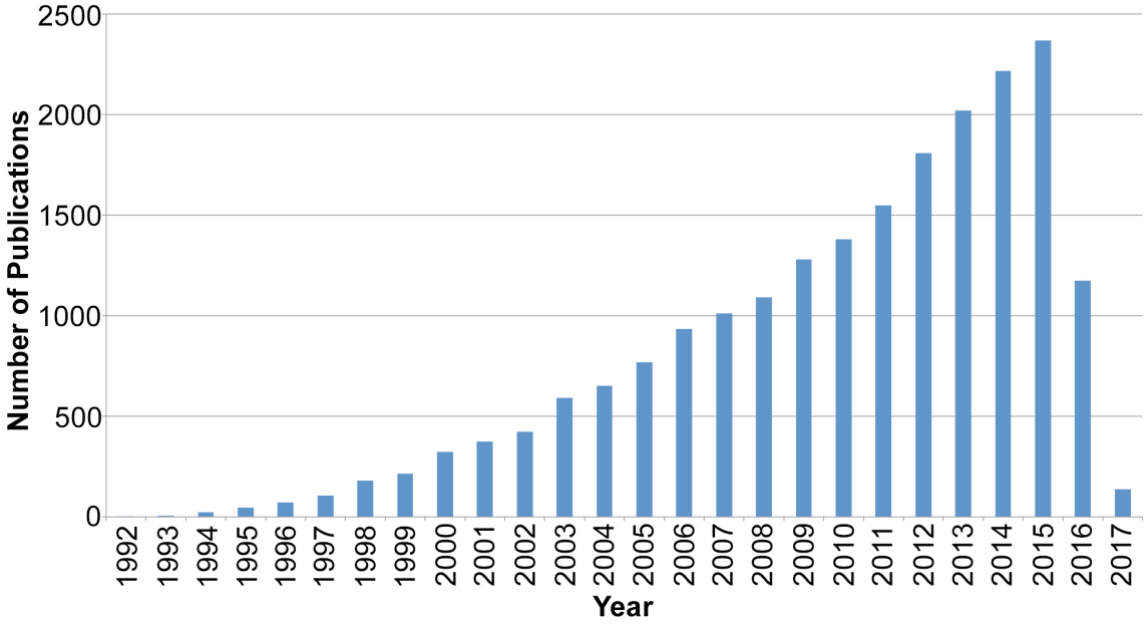

Figure 3.3 - Frequency of Publications on Human fMRI Studies (Until June 2017)

fMRI research has been conducted examining almost all regions of brain science, with a particular emphasis on examining brain regions corresponding to specific functions. For example, findings have revealed that the brain is compartmentalised, containing specialised regions for face perception [222], moral behaviour [223], amongst multiple others. Furthermore, the brain always has some level of neural activity [224], which suggests that at some point it might be possible to interact with vegetative coma patients via monitoring their brain activity [225]. Structure–function mapping advances knowledge regarding how the brain processes information and interactions between specific brain regions when completing tasks (e.g., recall task). As such, the overall aim underlying brain mapping involves identifying associations between neuronal substrates, their connections, and their functional relevance.

fMRI measures changes in blood oxygenation and flow in response to neural activity. fMRI allows researchers to map the activation levels throughout the brain, whilst individuals are engaged in a particular task (e.g., memory task) or at rest. Brain activation is indirectly measured as active regions require more oxygen and have increased blood flow, which modifies the magnetic field. These changes in magnetic field (BOLD) are measured by fMRI. Following neuronal activity, the blood flow changes between various

hundred milliseconds and various seconds. As blood flow changes are directly related to neural activity, the time-course of activity following specific tasks can be mapped. It is important to note that differences in the hemodynamic time-constants across both people and cortical regions limits the temporal resolution to about a second. Regarding fMRI data analysis, signal and image processing methods are required to generate from the raw data a statistical map revealing the brain regions that are activated in response to a given task. Various software packages (including several open source) are available for fMRI data processing and analysis; the most commonly used are FSL (the FMRIB Software Library) [226]–[228] and SPM (Statistical Parametric Mapping) [229]–[231].

Benefits of fMRI include that it is non-invasive, safe, and generates brain images with good spatial resolution and somewhat good temporal resolution, relative to methods such as PET. A further advantage of both fMRI and MRI is that multiple measures can be obtained from these systems. Individuals can be scanned for approximately one hour, producing measures of brain structure and function. The structural measures include measures of anatomy, for example, grey matter or T1 image (see Figure 3.4a), measures of WM (tractography), for example, diffusion tensor imaging or diffusion weighted imaging (see Figure 3.4b, which shows cortex, white and grey matter, third and lateral venticles, putamen, frontal sinus and superior sagittal sinus), measures of vasculature, time of flight imaging, or magnetic resonance angiography (see Figure 3.4c, which depicts carotid and vertebral circulation). Functional measures include measures of task-related activity across the brain or in particular regions (see Figure 3.5), measures of brain connectivity (see Figure 3.6), which are functional relationships across regions, and measures of relationships between the brain and the body, for example, physiological connectivity.

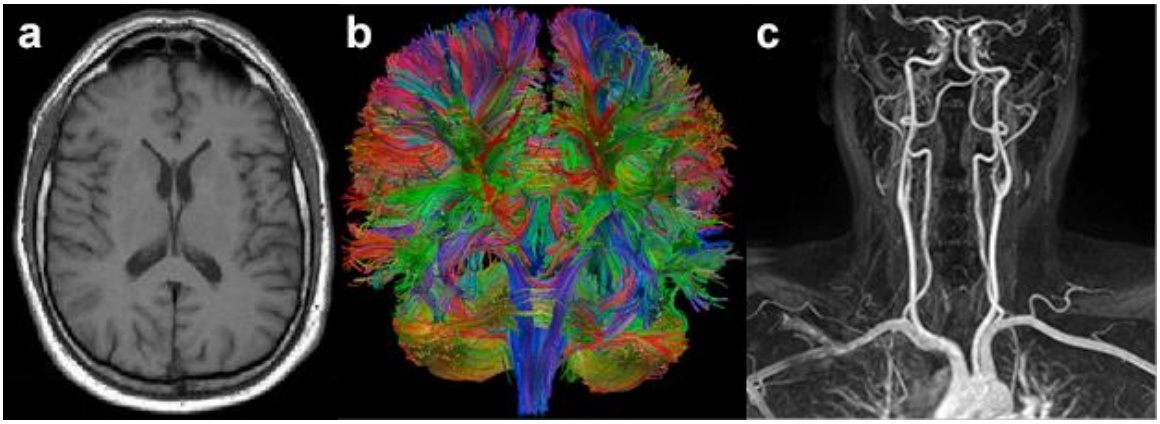

Figure 3.4 - (a) T1-weighted MR image of the brain; taken from [232]. (b) Diffusion Tensor Imaging; taken from [233]. (c) Normal neck magnetic resonance angiography; taken from [234]

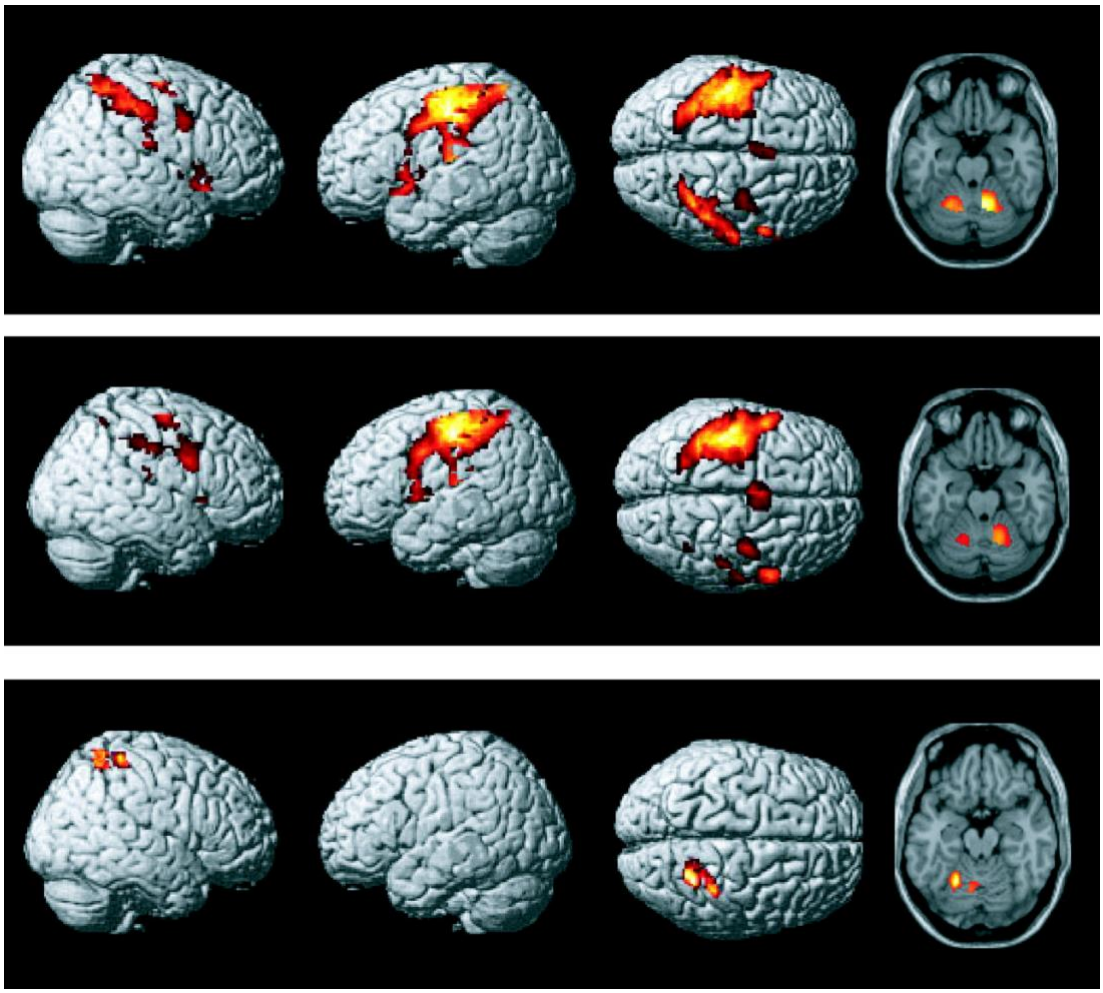

Figure 3.5 - Group maps generated from random effect analysis showing task-related activation during an fMRI study; taken from [235]

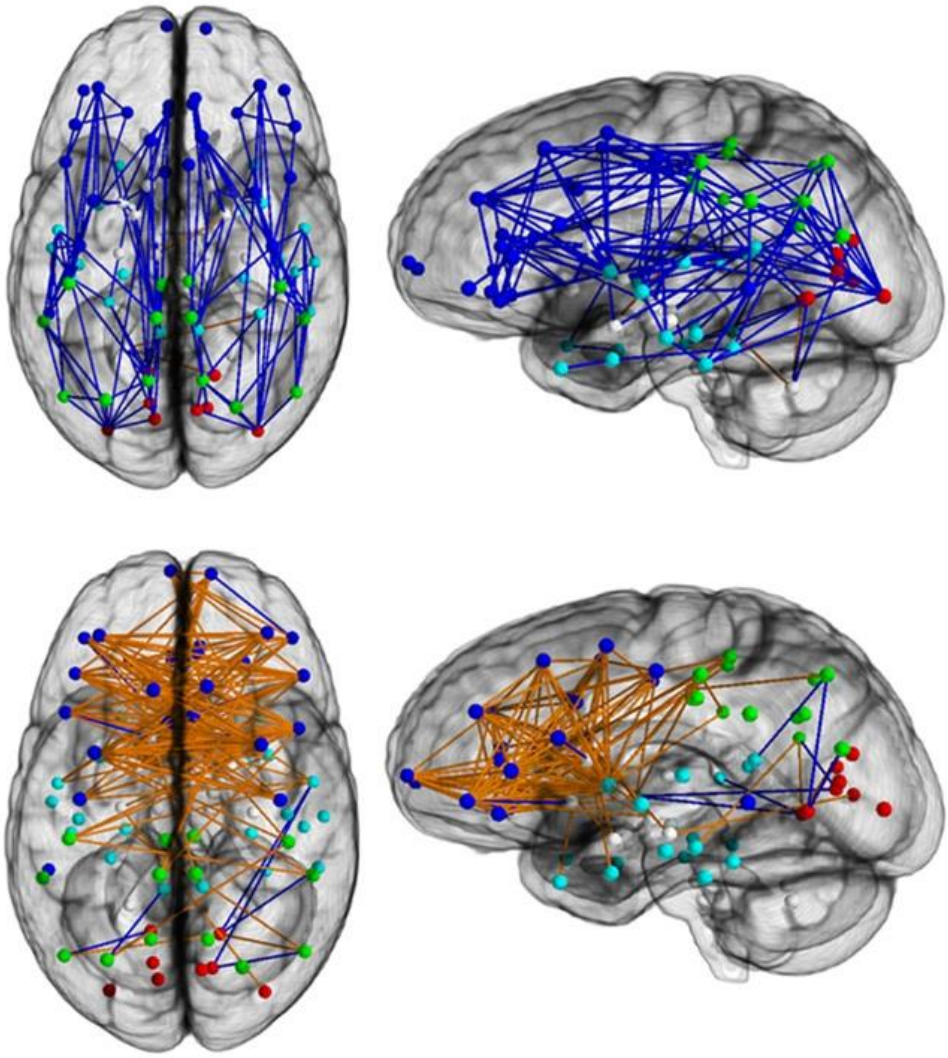

Figure 3.6 - Brain networks show increased connectivity from front to back and within one hemisphere in men (upper) and left to right in women (lower); taken from [236]

Brain imaging can be roughly separated into two major categories: structural brain imaging (which focuses on the study of brain structure, the diagnosis of disease and injury) and functional brain imaging. Modalities of performing structural imaging include CAT, MRI, and PET. Functional brain imaging can be used to study both cognitive and affective processes. Modalities include PET, fMRI, EEG and MEG. Each brain imaging modality provides a different type of measurement of the brain and they also have advantages and disadvantages with respect to spatial resolution, temporal resolution, and invasiveness. fMRI provides a good balance between these properties and has become the dominant functional imaging modality in the past decade.

3.3. Resting-State fMRI (rs-fMRI)

rs-fMRI is a relatively novel and effective approach for examining regional interactions taking place when an individual is not taking part in an explicit task. rs-fMRI is easier for neurologically impaired patients, relative to task-based fMRI as the latter involves cooperation yet many participants (e.g., PD or dementia patients) may find it challenging to follow task instructions within a scanner. The clinical implications of rs-fMRI are in its infancy [237].

rs-fMRI measures spontaneous low frequency fluctuations (0.1 Hz) in the BOLD signal to investigate the functional architecture of the brain. The significance of these fluctuations was first identified by Biswal et al. [238] in a study that involved participants not taking part in any cognitive, language, or motor tasks. Then, using a standard block design fMRI, in which the same participants completed a bilateral finger-tapping task, findings revealed a seed zone in the left somatosensory cortex with a correlation between the BOLD time course of the seed region and of all other brain regions. In addition, findings revealed that the left somatosensory cortex was strongly associated with homologous regions in the contralateral hemisphere. Hence, the two regions shared elements of the BOLD signal (functionally connected) when not engaged in finger-tapping, as represented in Figure 3.7.

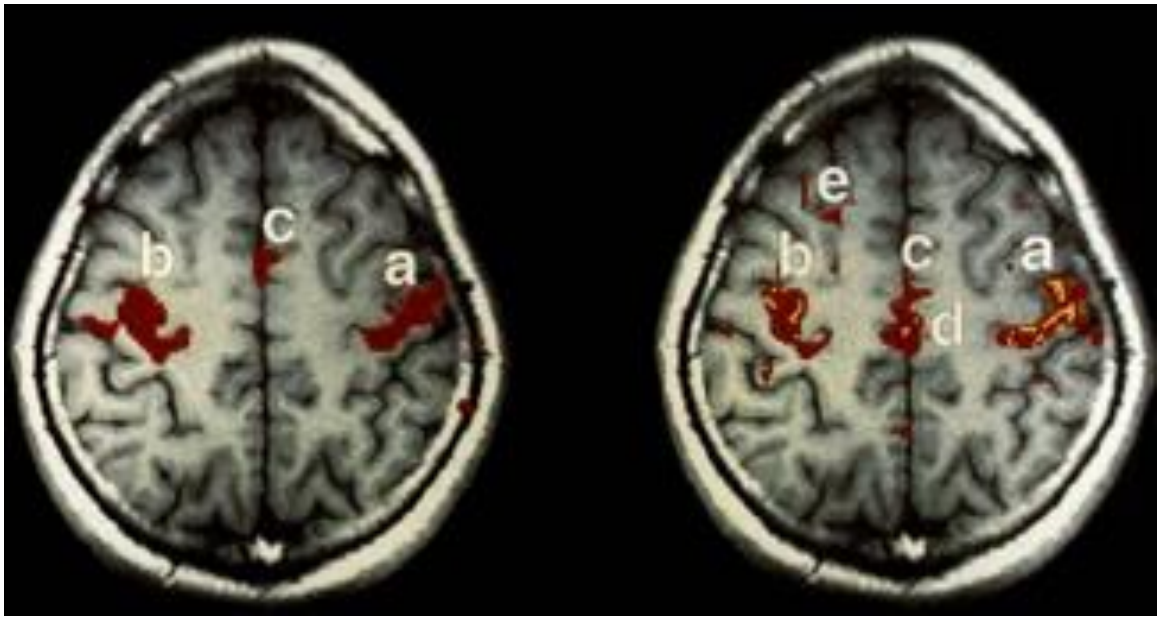

Figure 3.7 - rs-fMRI research, revealing spatial activation map to finger-tapping (left) and cross-correlation of signal fluctuation concerning a seed in the right motor cortex during the baseline (right); taken from [238]

Researchers posit the BOLD signal corresponds to neuronal activity, nevertheless, the basic molecular, cellular, and biophysical mechanisms involved are unknown. Increasing popularity in rs-fMRI research underscores the need to further explore the relationship between electrophysiological and BOLD signal fluctuations. Research has revealed that the networks identified via rs-fMRI have also been identified via EEG and MEG [239], which suggests that electrophysiological factors trigger rs-fMRI signal.

rs-fMRI is related to electrophysiology, yet, research is required to further examine the relationship between global BOLD signal and local electrophysiological recording, and also the source of signal fluctuation. Even though rs-fMRI is based on assumptions that there is a neuronal source underpinning the signal, to date there is no model outlining the mechanisms behind this, which produces challenges regarding controlling spurious noise. Hence, signal cannot be only attributed to neuronal factors. A key disadvantage of rs-fMRI involves the difficulties interpreting signal, which can be exacerbated by the fluctuating modulation of neuronal signal caused by an altered vasculature response in participants suffering from diseases or who have been administered differing anaesthetic medications and dosages.

Research has used rs-fMRI to identify functional networks for neurosurgical procedures and to identify epileptogenic networks in epileptic patients. Furthermore, research has revealed that resting-state network characteristics can differentiate AD patients and patients with minor cognitive impairment from healthy controls [13], [240]. It is possible that patients with disorders of consciousness and psychiatric patients can gain from this technique. The current research examined PD classification using rs-fMRI data.

3.3.1. Default Mode Network (DMN)

rs-fMRI research has identified approximately 12 functional networks (regions correlating with one-another in their temporal signal), capturing sensory and cognitive processes, which remain stable over different participant groups and studies. Functional networks are presented in Figure 3.8 [241]. For instance, somatosensory motor network is responsible for human sensory processing and comprises of primary somatosensory cortex, secondary somatosensory cortex, primary motor cortex and secondary motor cortex, two to three networks in the visual cortex, and one network in the auditory cortex. Sensory networks are bilateral with primary and secondary cortical regions. Higher order cognitive networks comprise of the left and right attention networks and contain the lateral posterior cingulate, middle frontal and orbital, superior parietal, middle temporal cortex, and the DMN, which includes medial prefrontal cortex (mPFC), posterior cingulate cortex (PCC), left inferior parietal cortex (LIPC), and right inferior parietal cortex (RIPC) [242].

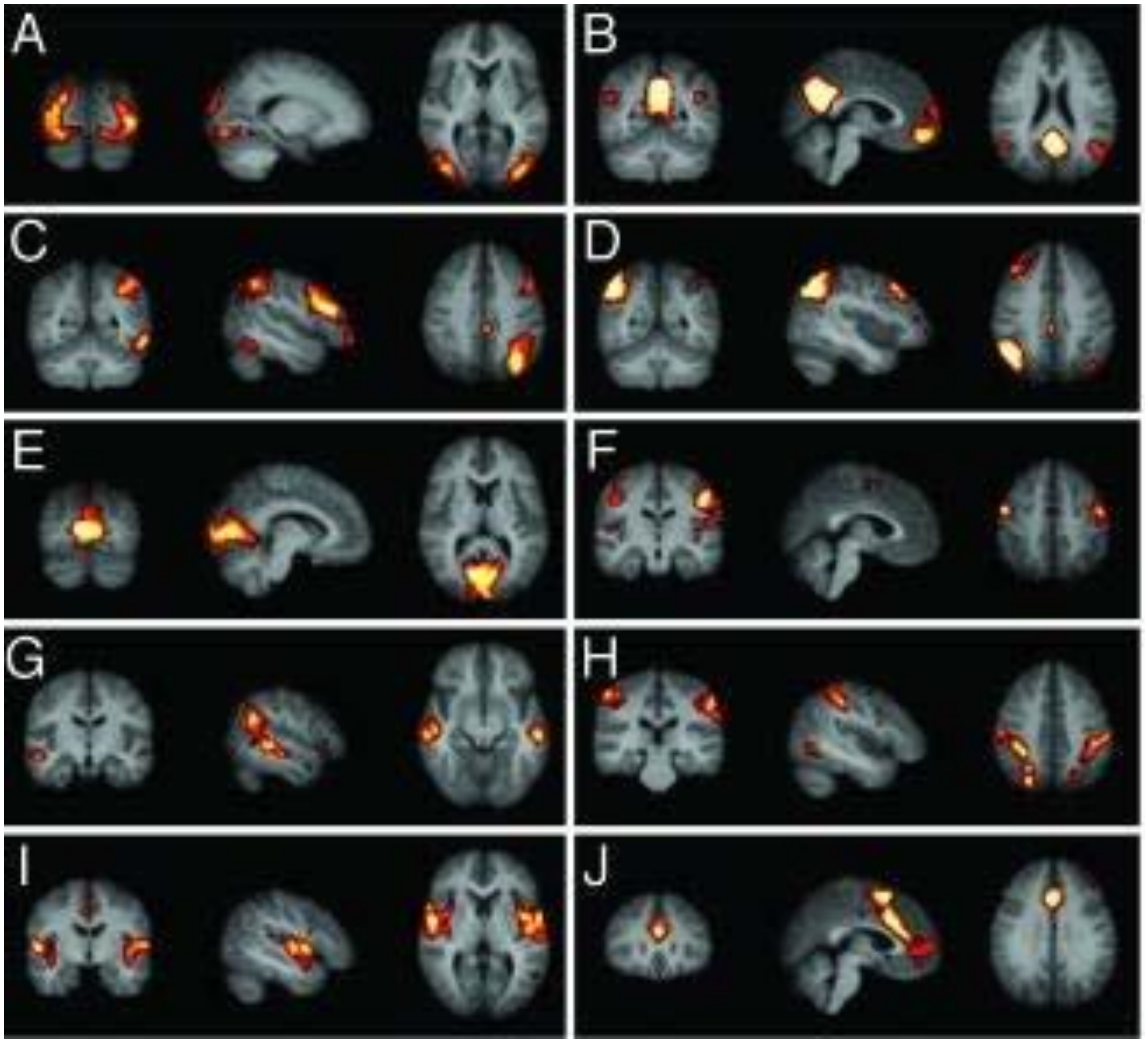

Figure 3.8 - The human resting-state functional network organisation: (B) the default mode network, (C) and (D) right and left attention network; taken from [241]

Raichle et al. [243] identified DMN from PET data [244]. This research involved the analysis of data from healthy participants resting with their eyes shut, which revealed specific brain areas were most active at rest and their activity declined when participants were engaged in cognitive tasks. DMN was identified by Greicius et al. [242] using fMRI and has been verified in numerous studies using a wide variation of analysis methods. Hence, by examining the associations between variations of the BOLD signal, fMRI can identify certain regions that are activated in synchrony each other [237]. It is relevant to note that one assumption underlying this method is the robustness of functional connectivity analysis as a tool that reproduces fundamental aspects of brain organisation via several cognitive states [245].

Much research has examined the DMN as one of the most important resting-state neural networks (RSNs), revealing that it is involved in inner voice, self-reference, and consciousness [246]. The DMN is highly detectable in rs-fMRI given that it is metabolically active when there are no task demands [243] and inversely related to regions involved in task performance (i.e., DMN signal amplitude decreases as signal increases in areas involved in completing a given task) [247].

Beyond the topological organisation that are based on anatomical structures and anti-correlative links between DMN and task driven regions, signal properties include temporal structure. Correlation between brain regions within a network contains signal fluctuations (0.01 Hz to 0.15 Hz frequency), yet, little is known regarding the underpinnings of these fluctuations and their frequency band. Researchers posit the hemodynamic response (captured via BOLD fMRI) adds a low frequency transformation to the high frequency neuronal activity, resulting in this temporal property. Further, rs-fMRI signal has a $1/f$ frequency distribution with less frequencies corresponding to greater power relative to larger frequencies. Random noise (termed Gaussian or white noise) has a flat power distribution over all frequency bands. Hence, despite little research having examined low frequency fluctuations, biologically relevant signal can still be distinguished from random, system-generated, noise.

Functional networks are arranged in terms of biological areas that are separate for sensory systems and higher order cognitive systems. Research reveals that networks portray cognitive properties of patients with diseases and healthy control participants. For example, greater DMN connectivity is related to schizophrenic patients with positive symptoms [248]. Research has revealed links between rs-fMRI networks and comparable networks captured via EEG and MEG [239] with similar patterns of activation following task-based fMRI [224], which suggests that rs-fMRI networks are based on electrophysiological factors. Hence, rs-fMRI research indicates that networks link local neuronal activity and cognition.

3.4. fMRI Data Analysis

fMRI data analysis constitutes an enormous data problem as every brain volume examined contains approximately 100,000 different voxel measurements. Every experiment can capture hundreds of brain volumes per participant, with typically 30 participants in fMRI research to enable population inference. Therefore, the volume of data being analysed is huge, especially considering that fMRI studies typically collect structural (T1-weighted images) data as well as functional (T2*-weighted images) data. Other factors contributing to the challenges involved in fMRI data analysis, beyond the large data issue, include the signal of interest being weak and the presence of temporal and spatial noise structure. fMRI data analysis involves preprocessing followed by processing, which is represented in the pipeline shown in Figure 3.9.

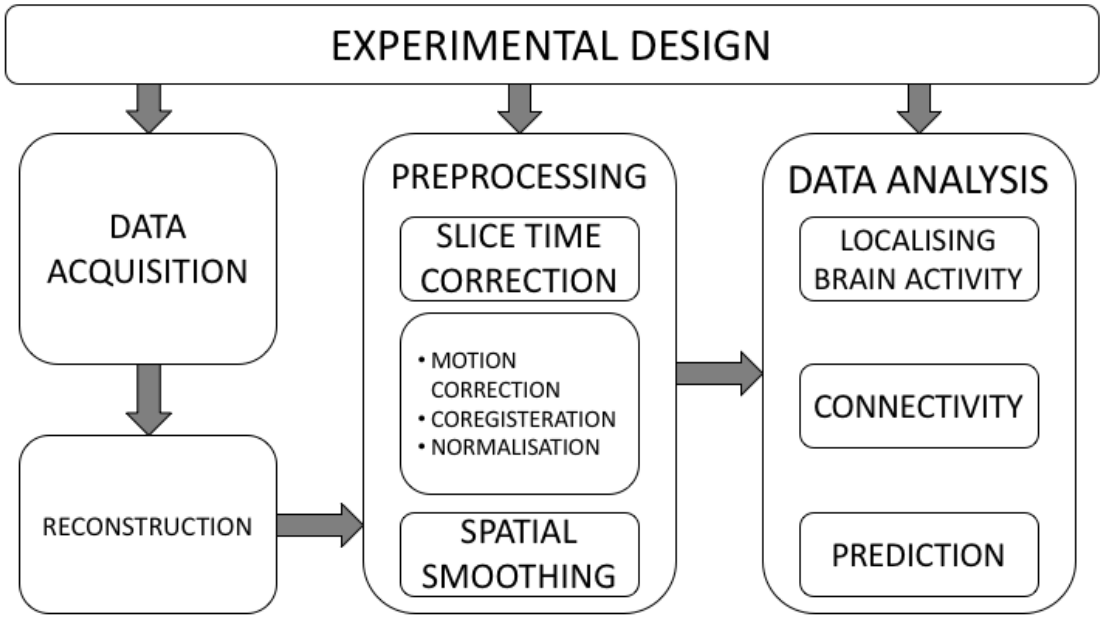

Figure 3.9 - fMRI data processing pipeline

fMRI data analysis includes three main goals:

1. **Localisation** is a brain mapping approach that involves identifying the active brain regions during a given task/psychological event/behaviour.

2. **Connectivity** involves measuring how brain areas are functionally related to each other and is typically represented in terms of correlations between brain regions over times. Three widely used connectivity methods are functional connectivity (correlations over time, for instance, seed-based connectivity examines all the brain areas that are correlated with a specific brain region) [249], effective connectivity (this includes various families of statistical models, such as path analysis and mediation models, Granger causality models, and dynamic causal modelling), and multivariate connectivity (data reduction methods, such as principle components analysis and independent components analysis but also graphical models such as constructing visualisation of the nodes and edges among brain areas and calculating the descriptive properties of these regions to infer behaviour) [250].
3. **Prediction** involves using brain activity to generate hypotheses regarding perceptions, behaviour, health status, amongst other factors, with applications to neurological disorders including AD and PD. These predictions can be validated over participants within a study and across studies.

3.4.1. Preprocessing

There are various fundamental steps involved when preprocessing fMRI data. The research presented in this thesis used CONN (version 17.c) [251], [252] and SPM12 (version 6906 - Wellcome Department of Imaging Neuroscience, London, UK) [230] software packages in MATLAB to complete the fMRI data preprocessing. The preprocessing and processing steps are presented in the same order in which they take place in the SPM12 template scripts.

3.4.1.1. Slice Timing Correction

fMRI data are typically collected via two-dimensional MRI acquisition and each slice is acquired individually. Slices can be acquired in ascending or descending order, or interleaved (each odd slice is acquired sequentially resulting in half of the slices being acquired and subsequently the other half of slices are acquired as depicted in Figure 3.10. In Figure 3.10, slices are acquired in the order 1-3-5-7-2-4-6-8 and the times on the right

represent the relative time the data in the slice is acquired, given a repetition time of 2 seconds.

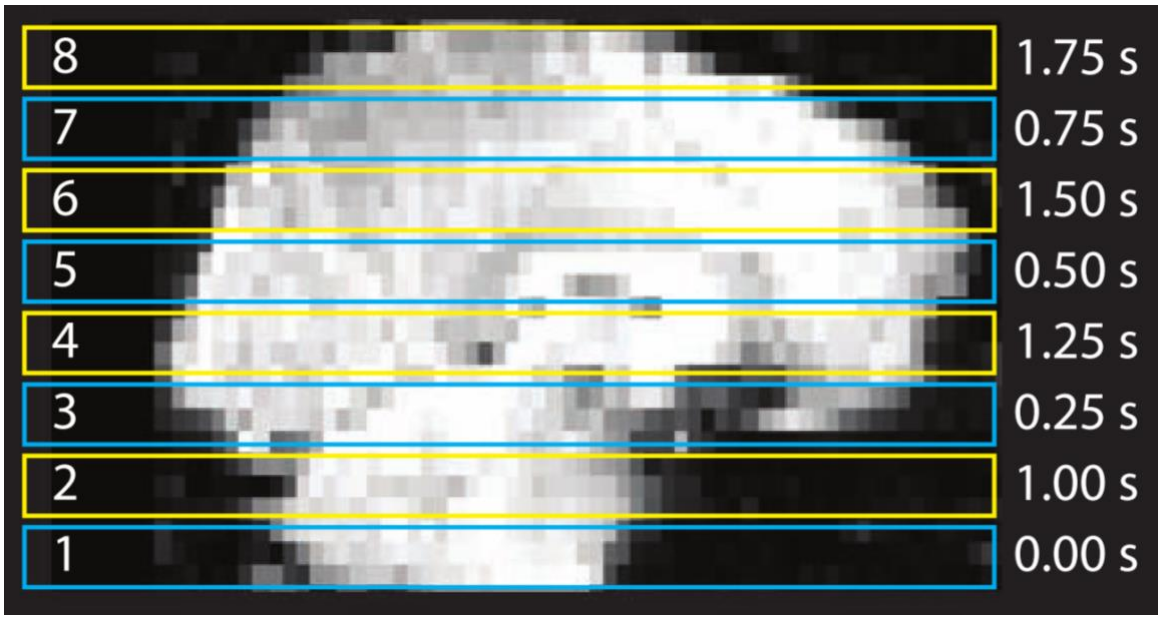

Figure 3.10 - An example of slice timing in an interleaved MRI acquisition; taken from [219]

Hence, different slices are acquired at differing time points and the temporal difference between two slices can be up to various seconds (as a function of the repetition time, or TR of the pulse sequence). Figure 3.11a represents three curves that depict samples from the same hemodynamic response at the times related to the slices represented in Figure 3.10. The slices acquired later in the volume reveal an ostensibly earlier response at every time point since the hemodynamic response has begun prior to the acquisition time. The timing of events/trials are imputed into a statistical model that depicts the expected signal resulting from the task, yet, a core assumption is that data was acquired at the same time. If left uncorrected, this time difference could be problematic as the hemodynamic response has a temporal factor. Hence, temporal differences in slice acquisition can pose challenges for data analysis. Slice timing correction [253] addresses this limitation and involves selecting a reference slice and sinc interpolating the data from all slices to correspond with this reference slice timing, as illustrated in Figure 3.11b. Slice timing correction methods use sinc interpolation, resulting in reduced smoothing in the signal. In Figure 3.11b, the blue line represents the original timeseries from a single voxel in the slice acquired at the

start of each volume acquisition and the red line represents the interpolated time course after correcting this slice to correspond to the centre slice, acquired halfway through the volume, at time TR/2. Slice timing correction can be applied if the precise timing of acquisition is known. This timing changes based on the scanner and pulse sequence used, but the physics support personnel can provide this information.

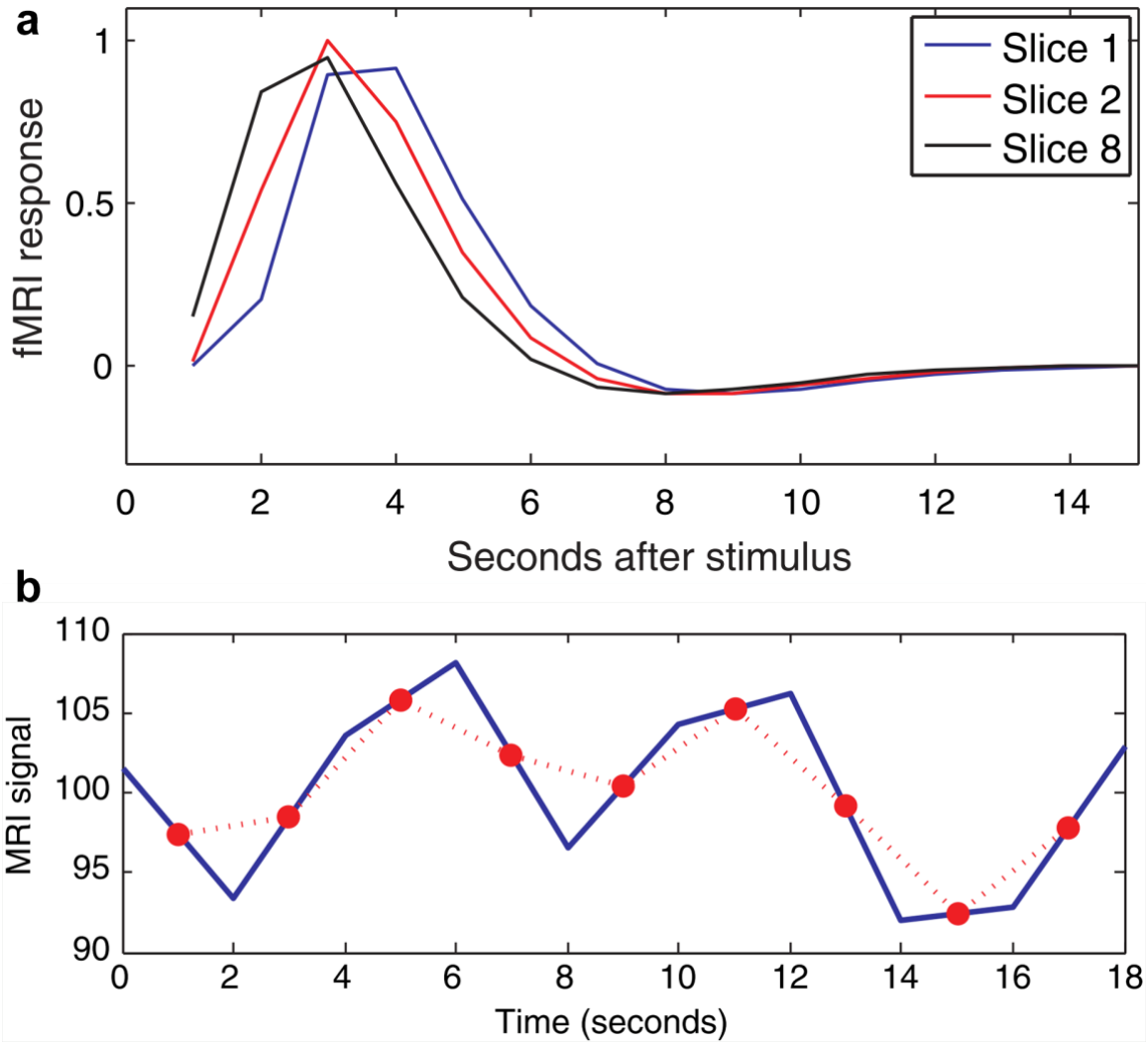

Figure 3.11 - (a) The impact of slice timing on the acquired data. (b) An example of slice timing correction using linear interpolation; taken from [219]

Some limitations of slice timing correction include (1) sinc interpolation, which can lead to artefacts present within one image being spread throughout the timeseries (problematic when considering interactions between slice timing and head motion). (2) When using short repetition times (TR $\leq$ 2 seconds) with interleaved acquisitions, event-related

analysis is robust against slice timing error; this is especially the case when applying spatial smoothing to the data, as data from bordering slices (acquired 1/2 TR from each another) are combined to result in a slice timing error of only TR/2. (3) Statistical models with temporal derivatives control for some timing misspecification and diminish the effect of slice timing discrepancies [219].

3.4.1.2. Motion Correction

All participants move their heads during scanning even minimally (e.g., swallowing), nevertheless, slight movements can significantly impact on the data collected. Figure 3.12 illustrates the impact of movement on brain images; the images reveal activation for a blocked design motor task relative to a resting baseline. The left panel of Figure 3.12 shows a severe example of motion-related artefacts, as this individual had Tourette syndrome, labelled a "flaming brain" artefact. The right panel of Figure 3.12 depicts a common example where movement is represented as activation along one edge of the brain, illustrating motion in that direction, which is related to the task.

Head movement results in two fundamental limitations: (1) Location mismatch of following images in the timeseries (labelled bulk motion). Standard motion correction techniques reduce error by realigning the images in the timeseries to one reference image. Head movement during acquisition is corrected in order for every image voxel to depict the equivalent 3D part of tissue at each time point. SPM applies a 6-parameter affine transformation (rotation and translation about and in the x, y, and z axes). Bulk motion can have a strong impact on activation maps, typically at the image boundaries since significant changes in image intensity arise when a voxel without brain tissue in it at one time point subsequently includes brain tissue resulting from movement. These artefacts differ according to the type of movement, as shown in Figure 3.12. For instance, a ring of positive or negative activation depicts movement along the inferior–superior axis, positive activation on one side of the head and negative activation on the other side depicts movement along the left-right axis, or large areas of positive or negative activation in the orbitofrontal cortex depict rotation along the left-right axis. Artefacts are, also, typically present at the edges of the ventricles.

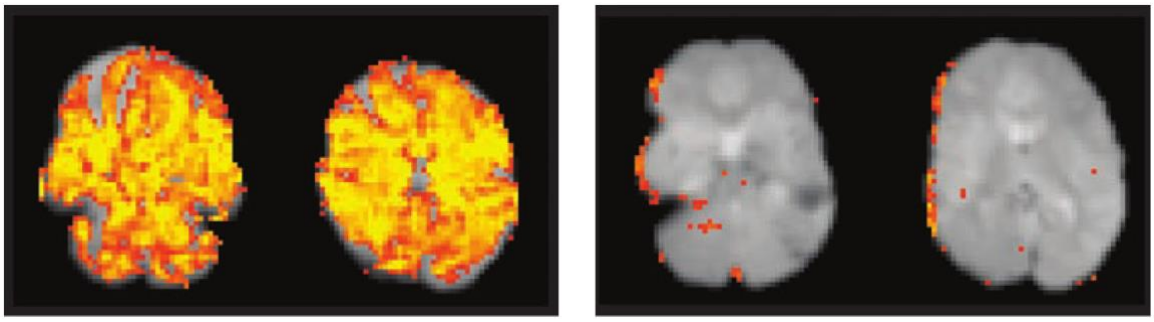

Figure 3.12 - Examples of the impact of head movement on statistical maps; taken from [219]

(2) Head movement can change the MRI signal. Following movement, the protons that move into a voxel from an adjacent slice contain a differing level of excitation from that, which is expected by the scanner, resulting in a reconstructed signal that inaccurately represents the tissue in the voxel (labelled spin history effect) [254]. Spin history effect can generate big variations in the intensity of one slice or a group of slices, resulting in stripes of interchanging bright and dark slices when using interleaved acquisition. Spin history effect cannot be corrected using standard motion correction techniques, although exploratory methods (e.g., independent component analysis) or spin-history corrections can be applied [254], [255].

3.4.1.3. Spatial Normalisation

At times, fMRI data are acquired from one person to gain specific information, e.g., regarding location of a tumour for surgery. Nevertheless, in research, fMRI data is typically acquired from a sample of participants with the aim of generalising findings regarding brain function across participants. Hence, data needs to be collated across participants. A problem is that individual brains vary in size and shape, so they need to be transformed to aligned with each other. The method of spatially transforming data into a common space for data analysis is labelled inter-subject registration or spatial normalisation.

Across individuals, brains reveal strong uniformity regarding structure, yet they vary extensively in terms of size and shape. A normal human brain contains two hemispheres

linked via a corpus callosum whose shape is typically stable across individuals. A set of major sulcal landmarks (such as the central sulcus, sylvian fissure, and cingulate sulcus) and the basal ganglia exist in almost all individuals, yet, variability exists concerning the minute features of brain structure. For instance, individuals possess a transverse gyrus in the superior temporal lobe (labelled Heschl's gyrus) that is linked to the primary auditory cortex, although the number and size of these gyri differ across people [256], [257] with brain structure differences still occurring in identical twins [258]. Spatial normalisation transforms brain images from every participant to minimise inter-individual variability, enabling group analyses.

A widely used template for spatial normalisation was created by the Montreal Neurological Institute, labelled the MNI template. This template consists of the average of 152 normal MRI scans. The MNI template was created to facilitate automated registration rather than landmark-based registration (where normalisation to the template involves first locating anatomical landmarks). Spatial normalisation techniques typically involve preprocessing the anatomical images before performing normalisation. These preprocessing methods involve correction of low-frequency artefacts (labelled bias fields), eliminating non-brain tissues, and segmenting the brain into different tissue (grey matter, WM, and CSF).

3.4.1.4. Segmentation

Brain tissue needs to be segmented into separate tissue classes (grey matter, WM, and CSF). Despite these tissues relating to distinct image intensities in a T1-weighted MRI, it is not possible to select the specific intensity values and threshold the image to locate these classes for three reasons: (1) precise brain segmentation is challenging in data processing as MRI images are noisy and even if the mean intensity of grey matter voxels differs from that of WM, their distributions can intersect. (2) Some voxels encompass various tissue types (labelled partial volume effect) and the intensity of these voxels varies widely in relation to their location. (3) Non-uniformities over the imaging field of view can result in the intensity of grey matter in a given area being more similar to that of WM in a different region, relative to grey matter in this other area. Hence, segmentation based on image intensities is problematic.

The unified segmentation approach [259] is a technique for tissue segmentation that is widely used (see [260] for a review of tissue segmentation methods). This technique involves using both spatial normalisation and bias field correction alongside tissue segmentation, to identify the prior probability of voxels with grey or WM based on a probabilistic atlas of tissue types. The prior probability is applied to image data to identify the tissue class. Hence, two voxels with matching intensities can be labelled as differing tissue classes (e.g., if one is in a region with high probability of grey matter and another with high probability of WM).

3.4.1.5. Generating a Brain Mask

In SPM, WM, grey matter, and CSF masks are generated based on a T1 image. The T1 image is converted to standard MNI-152 space via a non-linear, diffeomorphic warp. A deformation field is applied to the T1, generating a version of the T1 in standard MNI-152 space. Then, the WM, grey matter, and CSF segmentation masks are joined, resulting in a full-brain mask.

3.4.1.6. Registration

A widely used spatial registration technique is volume-based registration to a template image, which involve affine linear registration including other methods of nonlinear registration. Typically used templates include the MNI-305 or MNI-152. These templates are averages of brain scans from many individuals that were registered into a common space. The fMRI image is spatially registered to the T1 via an affine transformation. This step corrects for head movement between acquisition of the T1 and the acquisition of the fMRI data.

Deformation fields are then applied to the fMRI data. At this stage, the fMRI data have been coregistered to the T1, therefore, the deformation field generated for the T1 can be applied to the fMRI data. This step involves applying a pre-existing 3D template image to move the transformed fMRI data onto a grid space with comparable voxel sizes to those of

the original fMRI data (but not the T1 data). T1 voxels sizes are smaller and fMRI data files resampled to a T1 grid-space are quite big.

3.4.1.7. Spatial Smoothing

Between-subject fMRI analyses typically involve spatial smoothing (note that this is relevant for 2D cortical surface as well as 3D volumetric approaches). Smoothing filters out high spatial frequency noise (functioning as a low-pass spatial frequency filter), which enhances the signal-to-noise ratio [261]. This boosts the probability of detecting signals in between-subject analyses, given that the exact cortical location of foci is variable across participants. In addition, following smoothing the data are more similar to a continuous field of random values, which is a core assumption of random field theory (a multiple comparisons correction) [262]–[266].

Spatial smoothing comprises of applying a filter to the image to eliminate high-frequency data, resulting in blurred images with reduced spatial resolution. This is completed to "average out" some of the random noise but to still preserve neural activity-related signal. There are three reasons for applying spatial smoothing: (1) when collating data across participants the spatial location of functional regions is variable and not corrected by spatial normalisation. Spatial smoothing diminishes the incongruity across participants. (2) Eliminating high-frequency information (i.e., minor image changes), boosts the signal-to-noise ratio for signals with greater spatial scales. Given that fMRI activation typically occurs over various voxels, the advantages in terms of boosting signal for bigger features overshadow any limitations involved in removing minor features. In addition, the acquisition of small voxels can minimise dropout in regions prone to artefacts and smoothing aids to counter the heightened noise produced when small voxels are used. (3) Certain analyses (e.g., theory of Gaussian random fields [267]) need some level of spatial smoothness.

Standard spatial smoothing techniques involve the convolution of the 3D image with a 3D Gaussian filter (or kernel). The degree of smoothing generated via a Gaussian kernel is governed by the width of the distribution. In image processing, the width of the

distribution is accounted for by the full width at half-maximum (referred to as FWHM) [268]. This measures the width of the distribution at the point where it is at half of its maximum; it is linked to the standard deviation (σ) via the equation $FWHM = 2\sigma\sqrt{2ln(2)}$, or approximately 2.55σ. Greater FWHM result in larger smoothing, as presented in Figure 3.13. The left panel in Figure 3.13 shows three different kernels used for convolution. The identity kernel has a positive value in its central point and zeros elsewhere, whereas the Gaussian kernels spread progressively from the centre. A random signal was created and convolved with each of these kernels and the results are depicted in the right panel of Figure 3.13. Convolution with the identity kernel reproduces the original data, yet convolution with a Gaussian kernel generates smoothing, with larger smoothing for a wider kernel. The default kernel is FWHM of 6 mm in each direction. In the research presented in this thesis, all the functional images were spatially smoothed using a Gaussian kernel with 8 mm [269] FWHM to account for inter-participant variability while maintaining a relatively high spatial resolution.

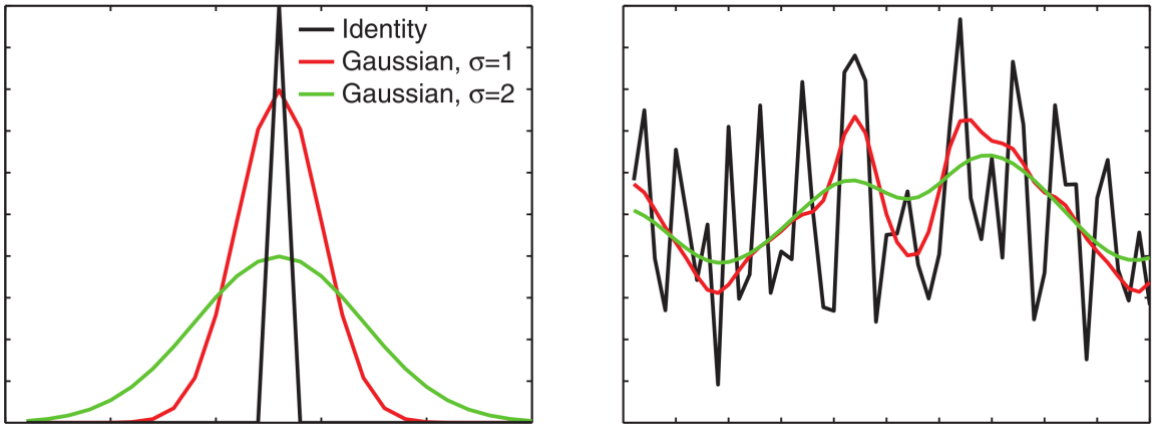

Figure 3.13 - An example of convolution used in spatial smoothing; taken from [219]

Smoothness directly relates to the associations between neighbouring voxels. Hence, the smoothness of an image is not always related to the degree of smoothing applied. An image containing random noise will have little smoothness; yet, MRI images have higher smoothness because of the filtering applied via image reconstruction and because of intrinsic correlations within the image. Following application of smoothness, the smoothness of the image produced is calculated using the following formula:

$$FWHM = \sqrt{FWHM_{intrinsic}^2 + FWHM_{applied}^2}$$

Figure 3.14 illustrates the preprocessing pipeline that was used for preprocessing the rs-fMRI data in the research presented in this thesis. The images are taken from the data analysis steps involved in preprocessing of one of the participants in the study presented in Chapter 6. The parameters of the preprocessing pipeline are also provided in Chapter 5 and Chapter 6.

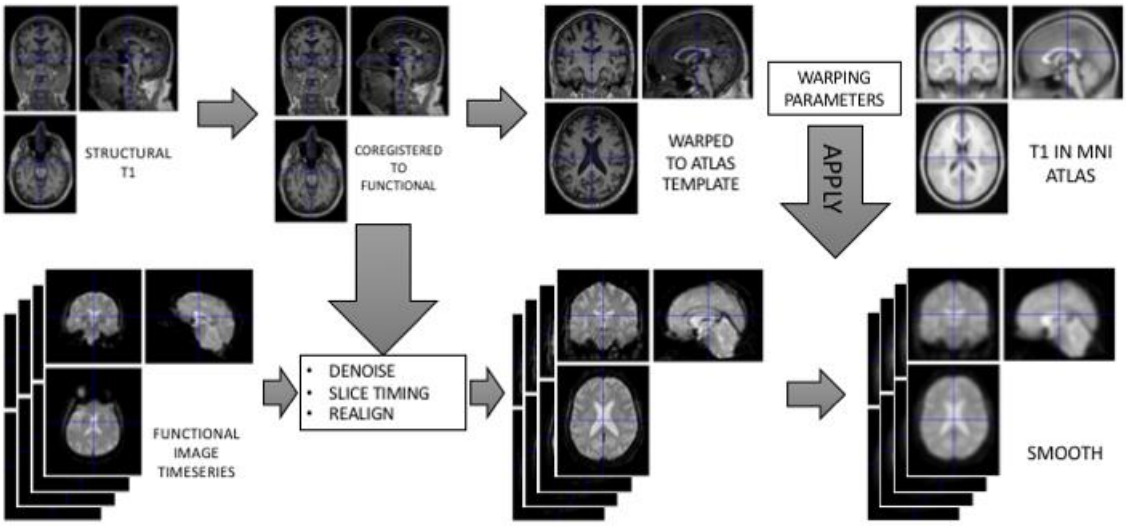

Figure 3.14 - Preprocessing pipeline used in this research

3.5. Processing

3.5.1. Modelling Brain Connectivity

Neuroimaging research currently largely focuses on functional localisation, yet, scientists now are highlighting the relevance of functional integration to account for brain function [270], [271]. This section reviews techniques for analysing brain connectivity using fMRI data. Brain connectivity is essential for identifying interactions between spatially separate brain areas that relate to mental function. There is no strict consensus concerning the analysis of brain connectivity, with analysis techniques frequently being reviewed, modified, and new ones being developed. Nevertheless, this section briefly reviews these brain connectivity analysis techniques as applied to fMRI data.

Studying connectivity in neuroscience involves focusing on three different elements, anatomical connectivity, functional connectivity, and effective connectivity. Anatomical connectivity involves studying the anatomical layout of axons and synaptic connections and focusing on how neural units interact directly with each other. Functional connectivity focuses on the correlation between activity in different brain areas. Moreover, this involves examining the statistical dependencies between measured timeseries (the statistical dependencies of observed neurophysiological responses). Effective connectivity, however, examines the causal influence that one neuronal system exerts over another (at synaptic or neuronal population level). Effective connectivity is investigated in the research presented in this thesis.

3.5.1.1. Effective Connectivity

Effective connectivity models explore the causal relationship and interactions between activity across different brain regions. Modelling methods in brain imaging based on effective connectivity include (1) structural equation modelling (SEM, e.g., multivariate analysis testing for influences among interacting variables), (2) Timeseries analysis (e.g., Granger Causality), (3) methods based on linear regression analysis (e.g. Psychophysical-Interaction analysis and Graphical Causal Models), and (4) methods based on nonlinear dynamic models (e.g., DCM).

SEM is a technique that tests hypotheses regarding causal relationships between variables that are described typically using linear equations (note that nonlinear equations are also feasible; see [272]–[274]). SEM can be sometimes problematic as the standard techniques for testing model fit assume independence between observations, which does not apply to fMRI data. Another primary issue with SEM is that the model is meant to identify causal relations between neuronal signals, but these must be estimated based on indirect and noisy BOLD MRI measurements [275]. This noise in the measurements can result in the identification of spurious relationships between some variables. Furthermore, the transformation from latent neuronal signals to observed hemodynamic signals is quite complex and nonlinear [276]. Although this latter challenge can be addressed using SEM, it is more naturally accommodated within techniques such as DCM, which directly

evaluate the hemodynamic response function to estimate causal relations between neuronal elements.

Granger causality posits that a cause occurs before its effects and, hence, it models causality by exploring the temporal relationship between variables (i.e., data from various variables are modelled over time). Granger causality is seemingly appealing compared to other forms of effective connectivity modelling since it provides a method of testing whole brain effective connectivity without specifying an anatomical network, as is necessary for SEM and DCM. However, Granger causality analysis with fMRI data is challenging due to the temporal characteristics of fMRI data. Granger causality relies upon the relative activity of different regions across time, therefore, it is important to first account for the effects of slice timing. Moreover, Granger causality assumes that the hemodynamic response is similar in its timing characteristics across the brain, which is not the case [277]. These differences overshadow the Granger causality analysis. Moreover, research has shown (using simultaneous electrophysiological recordings and fMRI) that Granger causality on fMRI timeseries does not accurately extract causal influences [278].

Graphical causal models examine the causal relationship (conditional independence relations; i.e., two variables are independent when conditioned on some other variables) between different groups of variables in the graph [274], [279], [280]. From the perspective of regression analyses, the conditioning on a third variable is the same as incorporating this as a covariate within the statistical model. If two variables are correlated with a third variable, then adding that third variable as a regressor eliminates the correlation and results in the variables being independent. One particular challenge to the use of search methods, such as graphical causal modelling, is combining data across multiple participants. It might seem obvious to simply combine the data across participants into a single search, or average the data across participants, however, each of these methods has the potential to obtain incorrect results. This issue has been addressed in the IMaGES [276] method, however, implementing IMaGES is expensive both timely and computationally.

While SEM and graphical causal models are very general methods that are applicable to many different kinds of data, DCM was developed specifically for modelling of causal relations in neuroimaging data (fMRI and EEG/MEG) [281]. DCM is discussed further in the following section, given that it is a key analysis used in the research presented in this thesis.

3.5.2. Dynamic Causal Modelling (DCM)

Research on rs-fMRI is prolific, specifically concerning differences in functional connectivity related to specific brain regions or to specific participant groups. Functional connectivity can explore unusual patterns of distributed activity, yet, causal inferences between neuronal systems cannot be determined [281]. In contrast, effective connectivity (including DCM) is based on a probable model of linked neuronal activity to examine causal interactions between brain regions. DCM is widely used in EEG research [281]–[283]. rs-fMRI data do not include exogenous (activating) inputs, unlike task-based studies; hence, endogenous fluctuations activate in brain regions.

DCM, which is based on Bayesian methods, examines the effective connectivity between brain regions for a specific experimental design. DCM explores effective connectivity (the causal effect of one neuronal system on another) using nonlinear designs to identify a reasonable generative model of measured neural activity (electromagnetic measurements or hemodynamic fMRI measurements).

DCM contains information regarding how neuronal activity results in the measured responses, which allows estimation of the effective connectivity. The research presented in this thesis examines DCM for rs-fMRI, in which deterministic inputs are activating/causing changes in the stimulation of different brain regions. This occurs via a dynamic input–state–output model of several inputs and outputs. The inputs relate to standard stimulus functions linked to the experimental manipulations. An initial challenge involved applying no driving inputs (related to the experimental design) to DCM, as per rs-fMRI [284], though, currently DCM can now be used with rs-fMRI. The state variables include neuronal activities and neurophysiological or biophysical variables that result in

the outputs. The outputs include electromagnetic measurements or hemodynamic fMRI measurements (as per the research presented in this thesis).

DCM is a very powerful tool, yet, it has some limitations. Similar to SEM, the validity of the results in DCM analysis depends upon the anatomical models that are specified and the regions that are used for data extraction. In order to solve this issue, graphical search techniques can be used initially to construct a set of plausible candidate models, and then use DCM to test specific hypotheses about those models. Moreover, DCM is currently only applicable to specific networks (e.g., DMN in the research presented in this thesis) and applying DCM to the whole brain is a very timely and valuable tool in research.

The research presented in this thesis uses spectral DCM [285]. Spectral DCM has advantages over stochastic DCM. Rather than estimating time varying fluctuations in neuronal states (stochastic DCM), in the research presented in this thesis, the parameters of cross spectra were estimated (spectral DCM). Spectral DCM is not only more accurate but also more sensitive to group differences. As such, spectral DCM results in a concise and effective method for estimating the effective connectivity from rs-fMRI timeseries, building on assumptions [269].

The research presented in this thesis explored the effective connectivity between DMN nodes and estimated the corresponding coupling parameters via spectral DCM in combination with different classification methods. The research presented in this thesis applies DCM, substituting high-dimensional fMRI timeseries with a low dimensional vector of parameter estimates. DCM connectivity values are then subjected to classification. This is a novel approach that has not been done previously in the literature. A pipeline of the method used in the research presented in this thesis can be seen in Figure 3.15. Excepting classification for DCM and timeseries, which is a novel element of this research, the processing steps are standard in rs-fMRI data analysis.

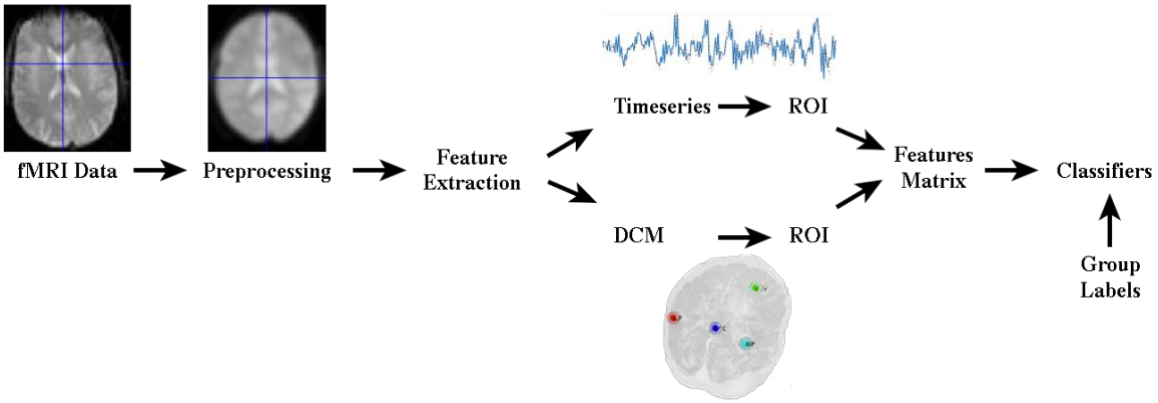

Figure 3.15 - Data analysis pipeline used in this research

3.6. Summary

fMRI is a clinical tool with multiple advantages, generating objective high resolution images. For instance, fMRI is non-invasive, safe to use and does not use radiation such as x-rays, CAT, and PET scans. These qualities make fMRI an effective and efficient tool in aiding clinical diagnosis, providing objective data (relative to less objective data obtained via clinical questionnaires). The research presented in this thesis applies classification to fMRI data to develop an automatic and highly accurate classification method relevant for PD monitoring in Chapter 5 using a controlled clinical experiment, and the methods developed are then applied to the diagnosis of early stage PD in Chapter 6.

Chapter 4. Computational Intelligence

Computational intelligence is the capability of a computer to learn a given task based on data/experimental observation. This chapter outlines key classification models, issues to consider when developing and applying models for real-life/biomedical problems (for instance, classification using class-imbalanced data), using computational intelligent techniques to create classification models, different classification strategies, classification of timeseries data, classification of dynamic data, and CGP classification. These themes are relevant given that the research presented in this thesis applies CGP, SVM, and ANN classification to dynamic class-imbalanced data

4.1. Predictive Modelling

Predictive modelling involves developing mathematical models of anticipated future behaviour, based on previous behaviour. Predictive modelling mainly includes regression and classification modelling. Regression models are used for questions that require predicting a continuous variable, whereas classification models are used for categorical variables. The research presented in this thesis focuses on classification of binary variables (e.g., two dichotomous classes such as PD patients versus health age-matched controls).

4.1.1. Classification

Classifiers are models that predict a nominal or ordinal variable from a limited group of potential options. The prediction is derived from continuous or nominal data (e.g., a fixed-size vector with summary features of the sample). Classifiers generate a mapping between an input feature vector $f = [f_1, f_2, \dots, f_n]$ and an output prediction, p_i. Classification output is reliant on both the model and the type of data used. A common method involves outputting a discrete class prediction as one of the classes in the group of potential classes $\{c_1, c_2, \dots, c_c\}$. This method is particularly useful for basic questions as additional interpretation is not required. Nevertheless, limitations of this method include that it does not provide estimates of certainty and it generates identical output for both a distinctive sample and a sample that is between two classes. Other learning algorithms generate an estimate of certainty and tend to output a vector of continuous values $[o_1, o_2, \dots, o_c]$, per input pattern, with each value relating to a measure of support for a given class. The values

generated can be probabilities, for instance, $\sum_{i=1}^{C} o_i = 1$, or they can be arbitrary units representing a general score. For models that provide an estimate of certainty per category, the overall predicted class is that with the greatest score. Often when classifying dichotomous binary variables, models output only one continuous value (instead of one per class). In this case, a label vote can be generated by thresholding the score regarding an arbitrarily selected level or a pre-determined value. These methods are relevant for clinical contexts in which patients with a specific disease (versus healthy controls) are being identified, as the model score provides a biomarker to facilitate diagnosis.

4.1.1.1. Evaluating Classification Models

There are various methods to evaluate the performance of multiple classifiers. One basic approach involves calculating the classifier's accuracy rate by computing a ratio of accurately predicted data objects, as per:

$$Accuracy = \frac{A_c + B_C}{N}$$

A_c and B_C are the number of correctly predicted instances of classes A and B, respectively, and N represents the total number of patterns. There are two key limitations to using a classifier's *accuracy* as a measure of goodness of fit. Firstly, heavily class-imbalanced data results in many classifiers revealing artificially high accuracy levels and misclassifying trends to the larger class. For instance, in a dataset with 80% cases belonging to class 1 and 20% cases belonging to class 2, a classifier that identifies each sample as class 1 results in 80% accuracy, although it is not useful since this is the *no-information rate*. Whilst classifiers perform with better than no-information rate accuracy, this example underscores the relevance of class distribution given a bias towards the larger class.

Secondly, when evaluating classifiers using accuracy rates, these do not differentiate between Type I and Type II errors, which have distinct disadvantages in real-life scenarios (e.g., misdiagnosis of a healthy individual versus not diagnosing a patient with a disease). Accuracy rates speak directly to a classifier's capacity, despite of all the limitations that they have (see [286] for a review of accuracy rate limitations). However, this depends

highly on the data and the type of classification that is used. Moreover, in some cases, the standard fitness measure in classification is the *overall classification accuracy* (see [287]).

4.1.2. Learning Algorithms

There are numerous learning algorithms that apply classification models to training data, using different standpoints, including linear models, non-linear separators, rule based approaches (such as trees), amongst others. This section reviews the most frequently used methods.

4.1.2.1. Logistic Regression

Logistic regression models fall within the general linear model category by examining research questions that contain a categorical dependent variable (typically a binary variable) [288]. The output includes the log odds of a positive class (labelled an event) and a linear expression between the covariates and the log odds is generated using the maximum likelihood estimation. The coefficients depict the increase in the log odds of an event per one unit increase in the predictor. Linear discriminant analysis involves a somewhat comparable method that generates a linear model by reducing the likelihood of misclassification.

4.1.2.2. Decision Trees

Rule-based models are based on *if... then... else...* conditions. For instance, a model predicting whether Tim will make a cup of coffee could include the following:

IF time spent sleeping is 8 hours or less *AND* a work deadline is fast approaching in 2 hours or less *THEN* Tim will make coffee.

Models can involve a linear sequence of rules, where samples are categorised into a group given that certain conditions/rules are met (e.g., the PART algorithm; [289]). Nevertheless, a frequent method involves decision trees or classification trees (e.g., classification and regression trees [290] and C4.5 [291]). Decision trees involve a hierarchy of rules leading

to a binary tree (although other variations are possible), where *IF* and *ELSE* pathways are depicted by tree branches from corresponding nodes and the class labels are represented by leaves.

Decision trees have three key advantages. Firstly, they provide highly interpretable outcomes that specify the threshold predictor values. Secondly, decision trees enable an inherent method of feature selection, whereby just discriminatory data attributes that are identified by the induction process are present in the final model. Hence, less preprocessing by clinicians (a method that can produce bias) is required. Thirdly, decision trees can be applied to categorical predictors, which involves dividing potential outcomes into two subsets (one per branch). Overfitting is a limitation inherent in decision trees, as these follow certain rules in large trees that can classify small segments of the training data.

4.1.2.3. k-Nearest Neighbours (kNN)

The kNN algorithm predicts classes using data from a training set that includes information on correct class. Using a distance metric (usually Euclidean distance) of the feature vector, the algorithm measures the distance to each data point based on the training data. The *k* training samples containing the lowest distance (*nearest-neighbours*), are assigned a separate class relative to data that is further away. This method is sensitive and flexible given that *k* can, in principle, be any value. Nevertheless, limitations include poor applicability for high dimensionality data and, also, the training set needs to be stored, which can be challenging when working with big data.

4.1.2.4. Support Vector Machines (SVM)

Learning algorithms typically aim to sort each training sample from each class, without focusing on the method they are applying. For instance, linearly separable data contains countless possible discriminatory lines with which to sort the data into the classes. SVM follow a different methodology, by generating the optimal sorting boundary using just samples (labelled *support vectors*) related to each class that are nearest to the perimeter

(i.e. the samples that are most difficult to classify). Predictions are then generated by matching future data to this border. Resemblance between two samples is computed via the dot product, creating a linear classification boundary.

An advantage of SVM is that the similarity function can be substituted by any *kernel function* (via a *kernel trick* that charts two comparison feature vectors to a higher dimension and calculates their resemblance within this space) to generate non-linear boundaries. Linear, radial basis function, and polynomial are three frequently used kernel functions, which give SVM high flexibility. SVM are widely used due to their flexibility and accuracy [292]. There are two key limitations to SVM. Firstly, SVM include high trial-and-error given the possible candidate kernel functions. Secondly, interpretation can be difficult as much weight is given to the support vectors.

4.1.2.5. Ensemble Classifiers

A novel technique involves generating an *ensemble classifier* by joining the outputs from a group of models with low bias error and high variance, which enables misclassifications by one or two models to be cancelled by the majority of the ensemble. Without an ensemble, these minor misclassifications would likely remain undetected. Ensemble classifiers are reviewed in the subsequent sections.

4.1.2.5.1. Diversity Methods

There are various stages within ensemble development in which diversity can be promoted or maintained to enhance ensemble efficacy; these are outlined as follows.

Base Classifier Models

Using various learning algorithms is just one method to generate a diverse range of base classifiers. Indeed, the *no free lunch* theorem [293] specifies that a single best classification model suited to all datasets does not exist. Hence, using multiple learning algorithms enhances the possibility of identifying a strong model. Moreover, each

algorithm uses a novel route for mapping features to an output in a different way, leading to a range of classification rules.

Learning Algorithm Hyper-Parameters

Learning algorithms typically contain at least one tuneable hyper-parameter (e.g., k in kNN models) and a minimum number of observations per split in decision trees. Ensemble diversity is heightened by adopting different versions of each model based on differing hyper-parameter values.

Training Set Manipulation

A widely used diversity preservation method entails training base classifiers on separate subsets of the training set, which facilitates the recognition of clear patterns/trends that are relevant for future predictive modelling. For instance, k-fold cross-validation (CV) is a simple technique in which the training set is split into k equally-sized folds. The model is fitted on $k - 1$ of these subsets, removing one subset at each iteration for validation purposes. Training is performed k times and results in k classifiers, trained on somewhat distinct datasets. k-fold CV is a method applied to the data presented in this thesis. Bootstrapping samples involves a comparable method, in which training subsets are created from the original training set via sampling with replacement. Data samples that may not have been sampled are labelled as out of bag samples and are used for validation.

Data Feature Manipulation

An alternative method involves creating base models on varying subsets of the feature space, which results in classifiers recognising distinct patterns/trends in the training data. One basic approach is to fit each model with one attribute, although distinct or overlapping subsets are applicable too. This method is particularly effective for datasets that do not contain numerous superfluous attributes (unlike datasets in which each feature contains useful information) and it is commonly used for ensemble building.

Modify Output Classes

Research involving multi-class categorisation can use approaches that split datasets in relation to the output classes. For instance, Dietterich and Bakiri's error-correcting output coding technique randomly divides the original classes into two groups and relabels data accordingly [294]. The learning algorithm then uses this new training data to perform standard binary classification. The process of randomly dividing the data and training is replicated multiple times and results in a new classifier at each iteration that can differentiate between these binary groups and, finally, an ensemble is created as a result of these classifiers. Other techniques involve resampling the dataset into two groups (one class and other examples) with the purpose of using these new datasets to train classifiers. This process can then be repeated for each class to result in models that can recognise a given class.

Selection Process

The member selection phase is the last stage of ensemble building, in which diversity can be preserved. There are various methods for ensemble model selection. For instance, the base classifiers can be selected as a function of their individual accuracy or an explicit measure of diversity. Ensemble size is related to its performance, with smaller sizes having less varied opinions.

4.1.2.5.2. Ensemble Building Approaches

This section outlines three frequently used ensemble building methods: bagging, boosting, and random forests.

Bagging

Bagging, a term derived from the two stage method of **B**ootstrap **AGG**regat**ING,** involves iteratively applying one learning algorithm to distinct subsets of the original training set to result in a diverse set of classifiers [295]. These bootstrap replicates are selected at random

with replacement, creating different subsets of data that can include repeated instances of the same of data samples. After training a given number of models, the loop stops and an ensemble is created based on the classifiers generated. The outputs of all members are joined, allowing the prediction of new data patterns. Bagging is effective on unstable predictors with large variance error characteristics (e.g., decision trees).

Boosting

Boosting, similar to bagging, involves manipulating the training set so that models examine distinct samples, resulting in a diverse classifier [296], [297]. A model is trained on a subset selected randomly without replacement from the dataset. Samples are then weighted in relation to classification accuracy and the subsequent fitted model is highly endorsed if it accurately classifies samples that were misclassified by the first model. Hence, an ensemble of classifiers controls against errors made by individual models on given samples. Following each iteration, the weights are recalculated to represent the ensemble's accuracy. The ensemble is finalised when a certain number of iterations is concluded and it can then predict new trends by using the collective results from all the models. Boosting typically provides greater classification accuracy relative to bagging [298], yet accuracy diminishes to a larger extent when artificial noise is incorporated into a dataset for boosting, compared to bagging [299].

Random Forests

Random Forests is an extension of bagging that is used for decision trees [300]. The method involves (1) training many thousand decision trees to bag and (2) applying a feature subsetting method so that base classifiers use multiple features to generate predictions. These diversity preservation methods allow the ensuing ensemble to be robust, accurate, and able to generalise to other datasets, whilst minimising overfitting (which is a common problem with decision trees).

4.2. Computational Intelligence (CI)

CI is a branch of Artificial Intelligence (AI), that includes optimisation algorithms and representations, such as Genetic Programming (GP) [301]–[303], ANN [304]–[308], Particle Swarm Optimisation (PSO) [309], [310], Ant Colony Optimisation (ACO) [311], [312], Fuzzy Logic [313], [314], and Artificial Immune System (AIS) [315]–[317]. These methods are typically motivated by natural phenomena, including Darwinian evolution, the neural networks within mammalian brains, the behaviour of ant colonies, amongst others.

CI methods typically decipher complicated problems without needing expertise from a specific research area and, furthermore, they commonly examine optimising computational constructs. The process typically starts by randomly generating a function to resolve a given problem (the assigned task), which results in low performance accuracy, although over the iterative teaching process performance increases. This method does not contain bias regarding the search process, which would be present if a human were to attempt to solve a problem based on preconceptions/priori experience.

CI has been applied to research on pattern recognition, control engineering, circuit design, and symbolic regression. CI methods are particularly useful when standard methods cannot be applied or when a problem is too demanding to be resolved using analytical/mathematical techniques. The research presented in this thesis explores the application of CI to classification predictive modelling, using GP, ANN, and SVM.

4.2.1. Artificial Neural Networks (ANN)

ANN involve computational models of tightly connected organisations of neurons within the human brain. ANN can be used both as computational tools and to examine the brain's functional connectivity. A key element is their huge dispersed network of individually feeble computational units (neurons) that in conjunction can create high computational power [318]. Indeed, ANN were created to generate a computational resource comparable to the processing power of the human brain, in an attempt to tackle problems that computers typically find challenging but which are easily completed by humans (e.g.,

pattern recognition). Via learning algorithms, these networks are able to find the solution of diverse tasks without needing expertise domain knowledge.

4.2.2. Genetic Programming (GP)

GP is a branch of CI that is based on Darwinian evolution with the objective of automatically producing functioning computer programs. For instance, EAs are global optimisers that iteratively train a population of candidate programs encoded by primitive data strings called genotypes, with behaviour being managed by phenotypes (depicted as operators acting on inputs). Training occurs via an iterative process of identifying good performing individuals and joining these to create novel candidate solutions, hence, elongating the search with every stage. Regarding predictive modelling, the model involves a complex non-linear expression depicted as a data structure (e.g., syntax tree, graph, or list). *Fitness* criteria controls the evolutionary search, instead of fitting a model to the data based on maximum-likelihood estimation.

4.2.2.1. History

Turing, in 1948, posited randomly generated networks can complete tasks using a *genetical search* [319]. Although this theory was proposed in relation to binary networks imitating the human brain, the behaviour of unorganised machines and EAs do overlap. EAs were developed by Barricelli et al. in 1954 as an optimisation algorithm [320], yet they became widely used only when Holland applied these to evolve a list of numbers, depicting chromosomes, in his research on Genetic Algorithms (GA) in 1973 and 1975 [321], [322]. EAs and GA then became widely used as optimising parameters for complicated problems. Cramer, in 1985, explored the evolution of a chromosome via GA that encoded a structure similar to a tree – this research developed GP [323]. Koza in early 1990s, further examined GP, outlining the effective application of GP to many tasks, in addition to outlining new methods [324]–[326]. Modern day advances in computational power (e.g., Graphics Processing Units and supercomputer clusters) has enabled GP to be applied to increasingly difficult tasks including population based searches via EAs.

4.2.2.2. Implementation

This section provides an overview of the practical considerations when applying GP to a classification problem.

4.2.2.2.1. Program Representation

GP includes the iterative population based training methods common to EAs with programs focusing on certain inputs via a group of primitive functions. Evolved individuals contain two factors: (1) the genotype (a string of primitive data types overseeing the respective individual's functional representation); and (2) the phenotype (which commands an individual's behaviour via representing the program using typical computational data structure). The phenotype of an evolved expression is often depicted using a syntax tree, in which branch nodes relate to a specific functions and leaf nodes represent inputs to the program or numerical constants. GP involves domain specific functions, for instance, an expression evolved with GP to address a symbolic regression task can use mathematical operators, whereas a program directing a robot can incorporate other functions including "turn left".

Modifying the function set used by the tree results in functions evolving to become specific to a certain task. Nevertheless, ANN (as another CI method) uses a different strategy in which each node typically has the same function independent of the task. Genotypes enable basic changes to a genetic code, although an interpreter (to deconstruct a given function) is needed before running the program. Figure 4.1 depicts an example arithmetic syntax tree, where a genotype can be expressed as the fundamental functions in pre-order notation: +(−(8, 3), /(9, 5)). To translate this into a string of primitive values for use as a chromosome, it is necessary for the functions to be depicted by basic data types via a lookup table.

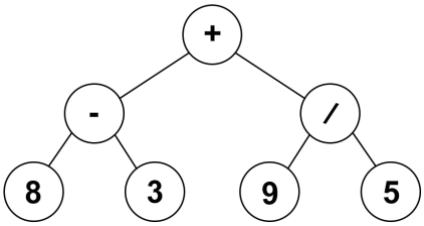

Figure 4.1 - A syntax tree representing the expression $(8 - 3) + (9 \div 5)$

4.2.2.2.2. Evolutionary Computing

Evolutionary computing is a form of computer science that is based on Darwin's theory of evolution [327], involving the examination of non-deterministic search algorithms. Evolutionary computing involves a particular methodology to answer search problems using stochastic[18] heuristic[19] population based[20] optimisation. The population is used to search for a solution via interacting between themselves and/or the environment (Swarm Intelligence; a review of this is beyond the scope of this document) or via a method derived from Darwinian evolution (EAs). Natural evolution posits that individuals' fitness is directly related to their success in competing for limited resources, allowing them to survive and reproduce [328]. Furthermore, when considering stochastic trial-and-error style problem solving processes (also known as generate-and-test), there is a grouping of candidate solutions. The solutions that provide better answers to the problem are more likely to be maintained and to be used to generate additional candidate solutions. As the research presented in this thesis makes use of EAs, this area is discussed in the following section.

Evolutionary Algorithms (EAs)

EAs are a branch of GP that search the solution space of possible programs to identify the optimal configuration for a specific task, which is grounded in evolutionary genetics. A

[18] Involves a random element in the search process.
[19] Uses experience and/or learning to guide the search.
[20] Makes use of many agents working together or independently.

key element of EAs involves *survival of the fittest* (i.e., survival of genes with good traits). Darwin, in 1859, posited that all species evolve through a continuous process involving the acquisition of beneficial traits from their predecessors [327]. Individuals with these inherited beneficial traits have better survival prospects and increased likelihood of procreation. Advances in human genome research has resulted in a gene centric approach to evolution (see [329]). This approach outlines that genes which compete for survival and alleles conveying survival benefits have a greater probability of being inherited by subsequent generations.

EAs contain various heuristics that solve optimisation problems using contexts similar to those of natural evolution. EAs can evolve using varying degrees of abstraction; nevertheless, they are always focused on a group of candidate solutions (*population*) that is improved via an iterative process over *generations*. EAs are an approved set of heuristics, which are flexible to use, relying on minimal *a priori* expert knowledge for the optimisation task. Given a quality function to be maximised, a set of candidate answers can be randomly generated, for instance, sections of the function's domain, and a quality function can be applied to determine abstract *fitness* (larger values/scores are better). Given the fitness values, certain candidates are selected for use in the subsequent generation via recombination (also known as sexual reproduction or crossover) and/or mutation. Recombination is an operator that is applied to two or more selected candidates (the so-called parents or genotype or *chromosomes*), by mixing their genetic material (*genes*), to create one or more new candidates (the children or new chromosomes or the *offspring*). Mutation is applied to one candidate (asexual reproduction) or two candidates (sexual reproduction) and results in one new candidate. Regarding classification problems, fitness relates to the accuracy; note though that as with the selection of function set, the fitness function is application-reliant. For instance, for a population evolved to control a robot to finish specific jobs in the shortest time possible, fitness represents duration of time for each candidate solution to finish these jobs with lower fitness values representing better solutions.

Executing the operations of recombination and mutation on the parents leads to the generation of a novel group of candidates (the offspring), whose fitness is measured. The

offspring subsequently compete (given their fitness and age) against the previous generation in terms of being selected for the next generation. This process can be iterated until a candidate with sufficient quality (a solution) is identified or a previously set computational limit is attained. The user can define the termination conditions, although these conditions normally specify a maximum number of allowed generations and a target fitness value (when an acceptable solution is found).

Each individual's fitness is assessed following every generation and a prearranged number of candidate solutions are chosen from the population (depending on their fitness) as *parents* for the subsequent generation (*selection* stage). Child solutions are generated from the parents through different *breeding* strategies that fine-tune the parent's genotype to create novel genetic material. In nature, evolution enables children to inherit beneficial genes from their parents. Similarly, in EAs this process enables optimisation search to progress to novel areas of the solution space through two operators: *crossover* (joining two parent genotypes for the child) and *mutation* (randomly changing alleles). Advantages of crossover include that they aid the search avoid local optima and expand the degree of exploration, whereas mutation facilitates a local search. Both crossover and mutation can be modified via hyper-parameters (e.g., *mutation rate*, the ratio of crossover to mutation). Figure 4.2 depicts the individual from Figure 4.1 following mutation. Figure 4.3 represents an example of crossover. Child solutions are automatically used for the subsequent generation or, alternatively, they compete against their parents for survival, as a function of the selected *replacement* operator. This method of evaluation, selection, breeding, and replacement, is replicated for a given number of generations and the fittest individual in the last generation is chosen as the winner/output.

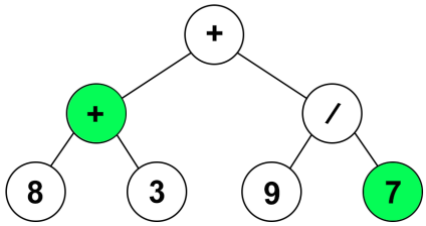

Figure 4.2 - The tree from Figure 4.1 following mutation, now depicting the expression $(8 + 3) + (9 \div 7)$

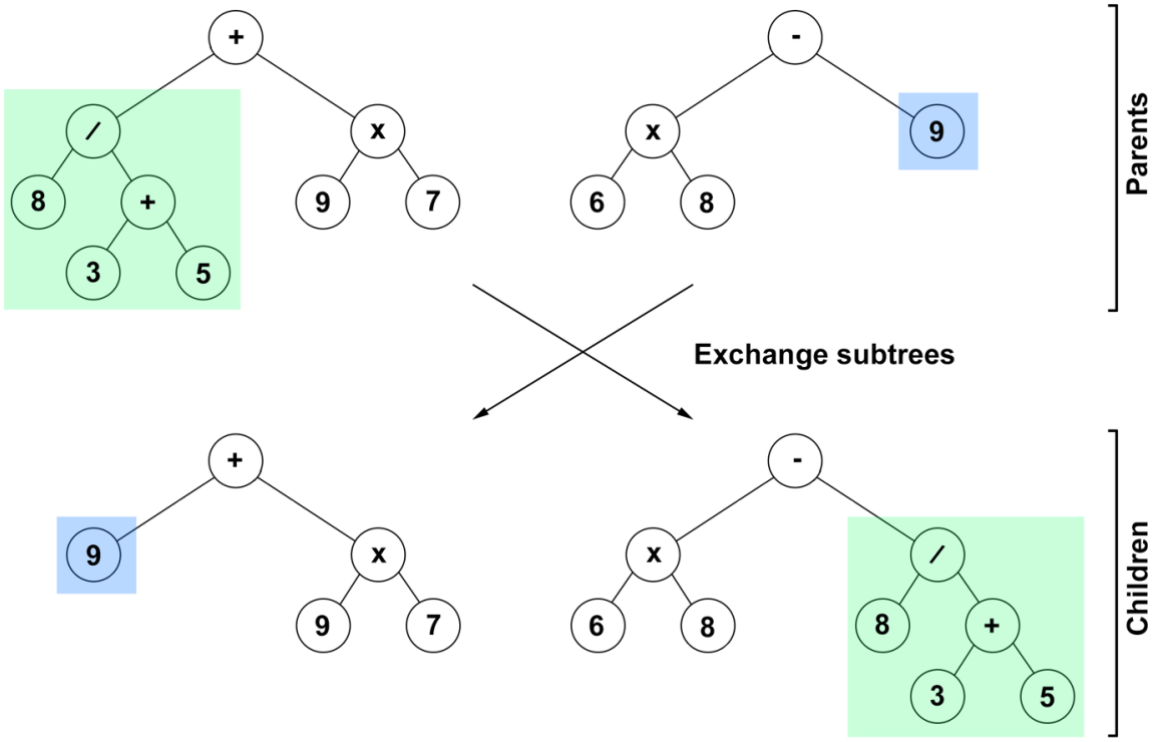

Figure 4.3 - Children share genetic material from both parents when created through crossover

Many elements within both EAs and natural selection are stochastic. For example, even individuals with low fitness have a possibility of reproducing and passing on their genetic material. Concerning recombination, elements from the parents are selected randomly to be recombined. Likewise, with regards to mutation, both the elements to be modified and the elements to replace these are selected randomly. The general scheme of an EA is presented in pseudo code in Figure 4.4 and Figure 4.5.

```
BEGIN
    INITIALISE population with random candidate solutions;
    EVALUATE each candidate;
    REPEAT UNTIL (TERMINATION CONDITION is satisfied) DO
        1   SELECT parents;
        2   RECOMBINE pairs of parents;
        3   MUTATE the resulting offspring;
        4   EVALUATE new candidates;
        5   SELECT individuals for the next generation;
    OD
END
```

Figure 4.4 - The general scheme of an EA in pseudo code; taken from [328]

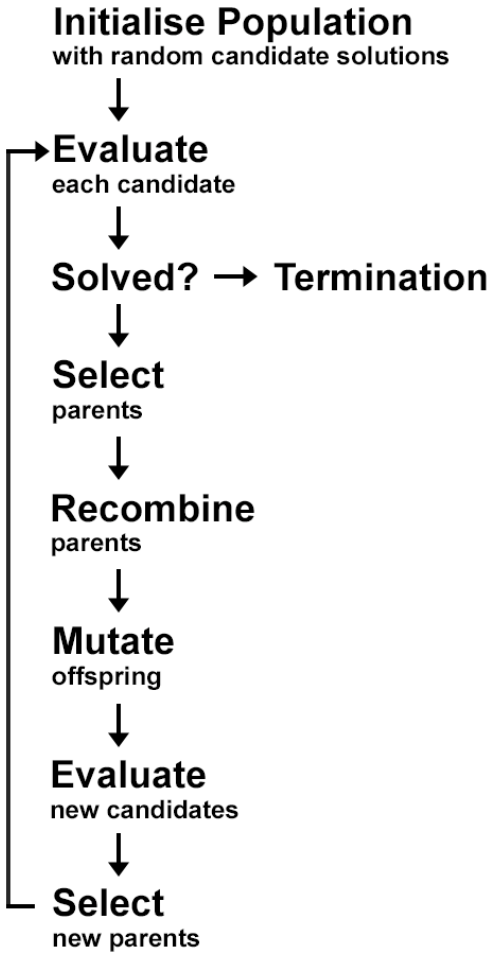

Figure 4.5 - General diagram of how EAs work

Advantages of EAs include their high flexibility (as fitness function and the operators relating to every candidate are domain specific) and their capability to join local and global searches to generate an exhaustive solution space. Hence, EAs, and specifically GP, are widely used in many research areas (e.g., temporal forecasting, symbolic regression, signal processing, classification, and control tasks). EAs enable manipulation of a trade-off between *exploration* and *exploitation*, which can create qualitatively distinct searches (applications typically use a balanced approach of these two). Explorative searches are similar to random search techniques and focus on a big section of the solution space, rather than examining minor sections within this space. Exploitative searches, however, are like hill climbing as they commonly involve configurations enabling just minor changes to the current generation, generating exhaustive searches in minor areas of the search space to find maxima.

4.2.3. Cartesian Genetic Programming (CGP)

CGP [330], [331] is a type of EAs and it is a strand of GP [325], [332], [333], that encodes computational structures as generic cyclic/acyclic graphs/nodes indexed by their Cartesian coordinates. It is called 'Cartesian' because it represents a program using a two-dimensional grid of nodes. A limitation of numerous GP methods is program bloat [334], yet, CGP is resistant to this [335], [336].

The size of CGP chromosomes remain stable even while the number of active nodes may change during evolution, resulting in inconstant length phenotypes (solutions). Users indicate the greatest number of nodes and a percentage of these are then activated. Overestimating the number of nodes can enhance evolution [337] and increase the neutral genetic drift.

As described in Figure 4.6, CGP chromosomes contain three elements (see [338]): (1) function genes (F_i), which characterise indexes in a function look-up-table and define the purpose of all nodes; (2) connection genes (C_i), defining the positions where all nodes obtain its inputs. In traditional acyclic CGP, connection genes can link a specific node to prior nodes or to program inputs. (3) Output genes (O_i) focus on program inputs or internal nodes and specify the program outputs.

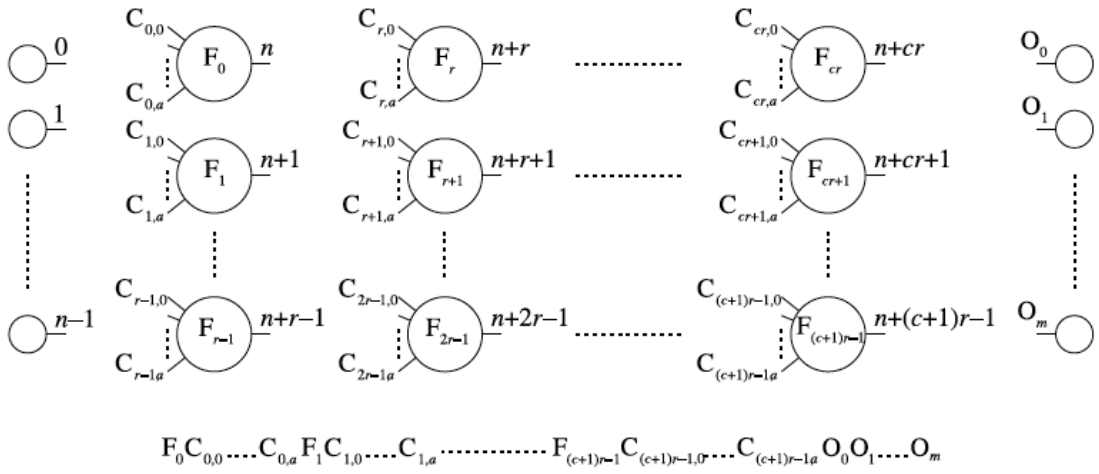

Figure 4.6 - General form of CGP; taken from [330]

All nodes are linked to prior nodes or program inputs. Program inputs do not all necessarily need to be implemented, which allows evolution to identify the relevant inputs. Unlike tree-based GP, an improvement that CGP offers is the reutilisation of node outputs. CGP is designed for multiple-input multiple-output problems and, even though only some nodes are used in the final program output, the unused nodes are inactive and allow neutral genetic drift, which promotes flexible phenotype length.

For the research presented in this thesis, a new open source cross platform CGP library (version 2.4) [339] was used since it is able to evolve symbolic expressions, Boolean logic circuits, and ANN, and it can be extended to different areas. The CGP library enables the control of evolutionary parameters and the application of custom evolutionary stages.

In CGP, programs are represented in the form of directed acyclic graphs. These graphs are represented as a two-dimensional grid of computational nodes. The genes that make up the genotype in CGP are integers that represent where a node derives its data, what actions the node executes on the data, and where the output data needed by the user is located. During the decoding of the genotype, several nodes might be ignored when these node outputs are not required in the calculation of output data. When nodes are ignored, these nodes and their genes are labelled *non-coding*. Following the decoding of a genotype, the program is labelled a phenotype. The genotype in CGP has a determined length. The phenotype length (concerning the number of computational nodes) can range from zero nodes (occurring

when all the program outputs were directly connected to program inputs) to the maximum number of nodes present in the genotype (occurring when all nodes in the graph are necessary).

The genotype–phenotype mapping used in CGP is one of its fundamental and distinguishing elements. The types of computational node functions used in CGP are determined by the user and are presented in a function look-up table. In CGP, every node in the directed graph represents a set function and is encoded by certain genes. One gene is the address of the computational node function in the function look-up table, which is labelled a function gene. The remaining node genes determine from where the node derives its data. These genes represent addresses in a data structure (normally an array), labelled connection genes. Nodes take their inputs in a feed-forward manner from either the output of nodes in a previous column or from a program input (labelled a terminal). The number of connection genes a node has is selected to be the maximum number of inputs (labelled the arity) that any function in the function look-up table has. The program data inputs are given the absolute data addresses 0 to $n_i - 1$, where n_i is the number of program inputs. The data outputs of nodes in the genotype are given addresses sequentially, column by column, starting from n_i to $n_i + L_n - 1$, where L_n is the user-determined upper bound of the number of nodes, as shown in Figure 4.6. Figure 4.7 represents an example of a CGP program.

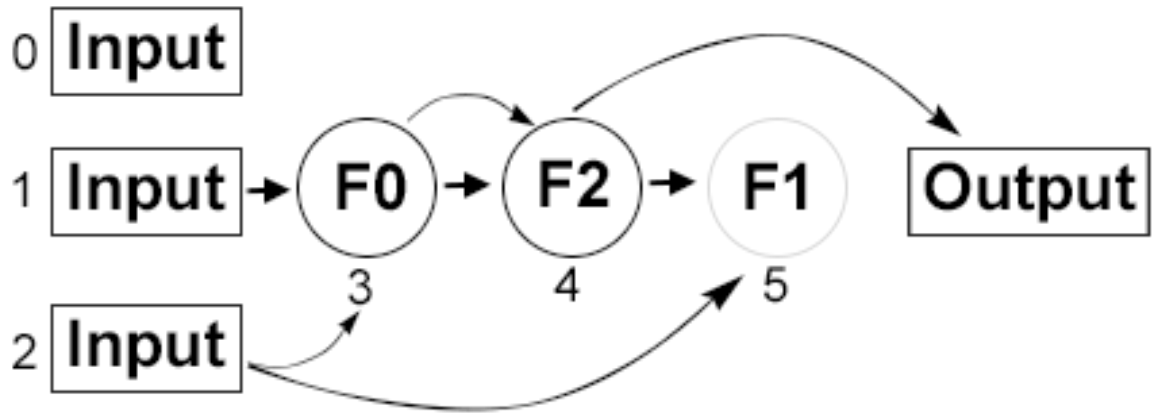

Figure 4.7 - An example CGP program; taken from [340]

4.2.3.1. Recurrent Cartesian Genetic Programming (RCGP)

CGP enables connection genes to link to prior nodes within a graph (including inputs). RCGP [338], [340] extends CGP by enabling cyclic connections, such as feedback, to any node (even itself) and to program inputs. Figure 4.8 illustrates an example of a RCGP program. A potentially unlimited number of connection genes leads to mutations generating the same number of recurrent and feed-forward connections [340]. Yet, many times, it is not necessary for 50% of connections to be recurrent, hence, the inclusion of recurrent connection probability (a parameter to regulate the likelihood that a mutation to a connection gene produces a recurrent connection). For example, 10% recurrent connection probability leads to 10% of connection gene mutations generating cyclic connections and a 0% recurrent connection probability results in traditional CGP with just acyclic connections. RCGP chromosomes perform in the same way as traditional CGP chromosomes, though differences lie concerning the program output(s). Specifically, in RCGP these can be identified by the inputs and by the internal nodes, even before the outputs have been generated [340]. As such, similar to the settings of recursive equations, the nodes begin with zero output until generating the new output. Further exploration of the use of other non-zero initial node values in terms of transfer functions [338] is required, yet, this is beyond the scope of this thesis.

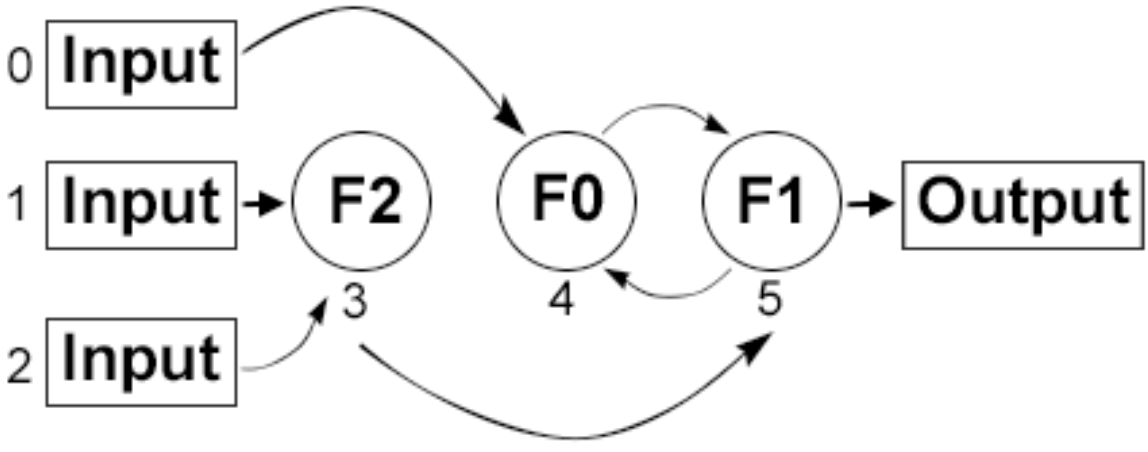

Figure 4.8 - An example RCGP program; taken from [340]

All studies presented in this thesis are executed and run using the open source cross platform CGP-Library (version 2.4) [339]. To evaluate chromosome fitness, a custom

fitness function is used, which assigns a score equal to the total number of values from the training sequence (100), subtracted by the number of accurate predicted values in sequence, minus 0.01 for additional accurate values following a predicted inaccurate value. Higher fitness values indicate a worse solution.

4.3. Computational Intelligence Applied to Predictive Modelling

GP and ANN can be used as learning algorithms for training classification models. CI algorithms generate classifiers via randomly creating initial models and then enhancing these through an iterative process. The last models tend to be non-linear and provide a different method relative to standard statistical modelling techniques, although the CI search techniques are highly computationally costly compared to standard statistical methods.

A key benefit of CI is the greater flexibility in the model representation, training techniques, and assumptions regarding data, which makes CI superior to standard predictive modelling methods for varied complex tasks. For instance, some statistical models make assumptions regarding the data (e.g., logistic regression does not function appropriately with collinear data attributes) and typically require big sample sizes to avoid violating the statistical assumptions. In contrast, CI algorithms require a measure of model goodness, with no statistical assumptions, and typically involve highly flexible model representations that enable them to be applied to data presented in different formats. These advantages have resulted in CI methods being applied to wide ranging research areas (e.g., complex multidimensional data from image recognition).

In general, there are multiple key differences between machine learning/CI classifiers and statistical-based classifiers, five of which are outlined as follows [341]. Firstly, no single prediction method yields high performance under every context, as per the *no free lunch theorem* [342]. For instance, statistical-based classifiers focus largely on predictive modelling. In contrast, machine learning classifiers typically involve explanatory modelling (also referred to as descriptive modelling) techniques, being data-driven instead of examining an experimental hypothesis. Explanatory modelling is largely applied to

testing causal theory, whereas descriptive modelling focuses on "*representing the data structure in a compact manner*" [343].

Secondly, machine learning techniques can be validated and there are various validation approaches that provide an unbiased estimate of the accuracy of the selected prediction rule [344], for instance, CV that is applied to the research presented in this thesis. Regarding statistical approaches, there is no independent dataset for validation, hence, this method relies on the CV error estimates of the various candidate techniques.

Thirdly, statistical modelling requires various assumptions to be fulfilled before running the analysis and the distribution of dependent or independent variable needs to be identified. This is not the case for machine learning algorithms, which have far less assumptions and enable contexts in which there may not be any continuity of boundary.

Fourthly, machine learning techniques include a large variation of tools. For example, Online Learning tools predict and learn concurrently from big data, and Random Forest and Gradient Boosting work quickly with big data. Indeed, machine learning is highly efficient for data containing both many attributes and many observations. Statistical modelling approaches are typically used for smaller datasets with fewer attributes to prevent overfitting.

Finally, fewer assumptions in predictive modelling results in greater predictive power. Machine learning is not based on human expertise or assumptions, rather, it uses iterations to identify patterns within the data, which typically lead to high predictive power. Machine learning involves no *a priori* assumptions regarding the relationship between the variables, neither does it assume that there is one best model to account for the data. This approach is completely data-driven: all the data is inputted and the machine then empirically outlines any relationships within the data (main effects and/or lower order interaction effects).

Nevertheless, the flexibility inherent in machine learning algorithms enable many sets of inputs that may not be logically related to one another to be applied, which can increase the risk of overfitting (identifying spurious correlations) relative to traditional statistical

models. Hence, CV and multiple techniques of penalised models have been developed over the last 25-30 years to counter any limitations linked to overfitting. In contrast, statistical models are derived from coefficient estimation and the modeller must have expert knowledge concerning the relationship and nature of the variables before running the analysis. Indeed, statistical learning models assume that there is a fundamental *data generating model*, and inputs must be logically linked to the independent variable. The absence of any *a priori* assumption/logical link is often negatively deemed *fishing for an answer* [345].

4.4. Generalisation

A fundamental objective to predictive modelling involves their performance to new datasets, rather than just their performance on a training dataset. *Overfitting* occurs when classifiers reveal good accuracy for a training dataset but performance is significantly lower when applied to novel datasets. Overfitting can be accounted for by dividing the dataset into a training (typically 50% of the original dataset) and test set (also typically 50% of the original dataset). The model uses the training set to learn and, afterwards, the unbiased performance is examined on the test set. Good test set accuracy indicates that the model has high generalisation properties. In some cases, the testing set is further divided into validation and test so that the generalisation ability can be assessed during the search for a solution [346]. A validation set is further useful for early stopping (before overfitting occurs in training) [347]. This method has been used for classification in the research presented in this thesis (see [348], [349] for other examples). Splitting a dataset into test and training sets randomly can create bias given that the test set might not fully represent the population data. Hence, rather than randomly dividing the data, *k-fold cross-validation* is used instead.

CV is an easy and well-established method for estimating prediction error. For large enough datasets, some of this data can be used as a validation set to evaluate the performance of a prediction model. *k*-fold CV takes some of the data to fit the model and another separate set of the data to test the model. The data is divided into *k* approximately equal sets. Figure 4.9 illustrates an example of *k*-fold CV for $k = 10$, where the red subset

represents the test set. Regarding the *k*th part (test set), the model is fit to the other *k*−1 parts and the prediction error of the fitted model is estimated when predicting the *k*th part. Likewise, the same process is applied to $k = 1, 2, ..., k$ and combine the *k* estimates of prediction error. $k = 5$ or 10 are the most accepted and widely used values [350]–[353].

Figure 4.9 - An example of 10-fold cross-validation

10-fold CV involves splitting the data into *10* equally sized *folds* and iterating through each fold in which each is used as a validation set whilst the *9* folds are used as training sets. The average validation fold accuracy score is calculated, providing a rigorous evaluation of a model that has been fitted on distinct splits of the dataset.

k-fold CV has been widely used mainly for two reasons: (1) as a testing method, which gives a nearly unbiased estimate of the generalisation power of the model by avoiding *overfitting*; (2) as a model selection technique. The research presented in this thesis used *k*-fold CV mainly for the first reason. As outlined previously, the process involves training *k* models for each fold, and then training one final model over the entire training set. This process is repeated 10 times (runs) for each fold and the average accuracy over the *k* folds is reported in the end.

The *bias-variance decomposition of error* [354] accounts for overfitting by positing that model error is the total of three values: (1) the *irreducible error*, which represents unavoidable noise in the dataset; (2) *bias* from the overall fit of the model to the data; and (3) error resulting from *variance*, which is the capability of model adaptability to change in relation to the predictor values. Figure 4.10 represents two distinct models classifying binary data with has two continuous predictors. A distinct non-linear relationship between the explanatory variables and the class is visible. Although the decision boundary as presented in Figure 4.10a includes bias error, it is suitable for classifying new data. The classification function depicted in Figure 4.10b detects minor inconsistencies in the

training set, resulting in a model with low bias error but, also, poor generalisation to new datasets.

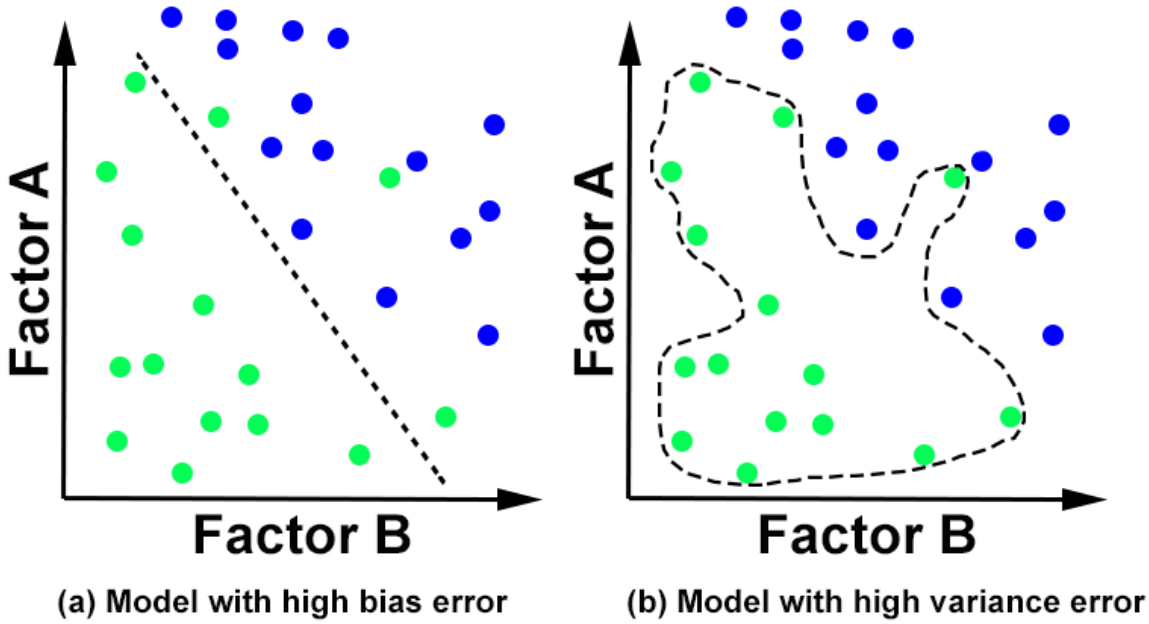

Figure 4.10 - Comparison of two models' decision functions with divergent error characteristics

Models include noise from variance error and bias error, although learning algorithms can display differing amounts of each with the aim of reducing overall error. For instance, decision trees often display good accuracy when sorting training set samples (low bias error), although accuracy for sorting new data is lower (high variance error). Both kNN learning algorithms and SVM directly manipulate both variance error and bias error. For example, in kNN learning algorithms, low values of k correspond to models with high variance and low bias error and increasing k results in larger bias error. Likewise, in SVM, parameter c manipulates the margin of the boundary between these errors, with smaller c resulting in less cost linked to misclassifying training points and a greater margin (lower variance error but higher bias error). A model that generalises well to new datasets needs to have low variance error. Typically, given that minor improvements in bias error strongly enhance model generalisation to new datasets, models with low bias error and high variance error can be enhanced more easily, relative to the reverse situation.

CI classifiers (like traditional statistical learning algorithms) can result in overfitting to the training data and low generalisation. Overfitting is aggravated by very specific non-linear classification expressions generated by protracted training methods. Over-training can be avoided using similar methods as those outlined for predictive modelling, in addition to techniques selected for a given learning algorithm. For instance, classifiers trained via iterative learning algorithms (e.g., EAs) can apply the *early stopping* method to decrease overfitting, which halts the training run before the predetermined number of iterations has been completed. When to stop is regulated by accuracy of the fittest model on another validation set (a reduction in accuracy signals overfitting and reduced generalisation). Alternatively, a smaller generation limit can be allowed for problems with reduced sample size, in which splitting the data into a validation set is not practical.

A limitation implicit in CI learning algorithms is their rigorous search that, when joined with dynamically sized classifiers (e.g., expression trees), can generate very precise discrimination rules including more terms than a standard linear model and large interaction depth. Just as decision trees result in overfitting to the training set, likewise expression trees can overfit when trained using EAs. Magnitude of GP evolved expressions does not influence overfitting [355], yet augmenting the number of hidden neurons in ANN can result in overfitting [356].

4.5. Classification of Timeseries Data

The research presented in this thesis involves the classification of timeseries data from both a clinical experiment (Chapter 5) and from PD patients (Chapter 6). Predictive modelling involves charting a set of data features (consisting of continuous or nominal data) to an output class prediction. Given the increase of different data types, research has now examined the classification of sequenced data. Xing et al. established a three-factor taxonomy of methods of sequence classification: (1) model-based methods generate predictions based on mathematical models of phenomena; (2) distance-based methods compare the resemblance between two sequences, for instance, via SVM with a specific kernel function for the input data; and (3) feature-based methods rely on domain data to

detect discriminatory elements of the sequence [357]. Research reveals that feature-based methods typically produce greater classification accuracy [358], [359].

Classification of timeseries data mostly requires *a priori* information to diminish the number of possible summary features and to select only those features with probable discriminatory power. Researchers have automated the feature selection process. For instance, Deng et al. designed the Time Series Forest algorithm, to recognise distinct elements within data trends [360]. A prediction is generated by dividing the input timeseries data into intervals and contrasting the findings of two intervals being input into one of three functions (mean, standard deviation, and slope of a regression line) at each node. The training process aims to divide the data with the objective of being highly discriminatory. A limitation of this method is that interpretability of the resulting models is poor, with still many potentially redundant features that may need to be removed before additional modelling. Fulcher et al. outlined a method to visually represent discrepancies in timeseries data from applications, containing over 9000 summary features (e.g., statistical summaries, signal processing algorithms, and information theory approaches) [361]. This method enables effective comparison of sequences acquired from multiple sources, yet, the data still requires dimensionality reduction before performing classification.

Feature design classification techniques diminish the dimensionality of raw sequence data into wieldy-sized feature vectors that can be applied to standard modelling. Research has begun to automatise this process, yet, initial findings reveal that automatisation generates numbers of features with poor interpretability. An objective feature design algorithm needs to analyse data without relying on expert knowledge. CI methods are applicable when there is both a big search space and no *a priori* knowledge regarding potential solutions.

GP is relevant in these situations given its flexibility in representation and behaviour. Sharman et al. applied GP to evolve adaptive signal processing algorithms via arithmetic operators, time-delay functions, and a stack for retaining useful output values, although they did not use classification [362]. Similarly, Parallel Architecture Discovery and Orchestration [363], [364], involves a classification framework for input signals using diverse forms (e.g., timeseries data, images, and video files), which is based on domain

specific function sets and a shared stack. Zeus [365] is applicable to timeseries classification, in which every chromosome depicts a candidate feature vector and every gene represents a function chosen from a group of arithmetic operators and statistical means of aggregation. Nevertheless, Eads et al. revealed that modelling timeseries data via SVM without feature designs, relative to Zeus, results in classifiers with higher accuracy.

4.6. Classification with Dynamical Systems

Methods for modelling timeseries data typically include summarising the raw data into distinct groups of feature sets via aggregate functions (e.g., mean, standard deviation, or signal processing tools), which can be automised with GP. Indeed, some research on the automatisation of timeseries data includes the use of stacks and time delay functions, yet GP can present challenges regarding implementing and interpreting. Nevertheless, there are some CI algorithms, labelled computational dynamical systems (coined by Stepney in [366]; e.g., Recurrent Neural Network, Gene Regulatory Network, [367], [368], and Artificial Biochemical Network [369]) that have been effectively applied to timeseries data. These methods involve modelling with networks of the body's biological mechanisms, enabling recurrent connections related to short term memory as well as to output feedback.

A traditional predictive model based on GP involves mapping inputs (a feature vector) to an output (a label prediction or a group of scores/probabilities) via a mathematical function. The output is constant for a specific group of inputs. In contrast, dynamical systems generate stateful programs that can be executed online and which can modify themselves in relation to varying input data. Hence, dynamical systems are often used in control problems and temporal forecasting [370]–[377].

Dynamical systems can also be used for classification problems [378]–[380]. Yet most research modelling timeseries data focuses on forecasting and not differentiating samples into separate classes. For dynamical systems, classification of sequence data needs to take into account the method of giving the inputs to the system as recurrent networks are less

flexible than GP and typically allow only one sample per timestep (similar to forecasting). Hence, fitting a training set containing various timeseries can be onerous.

Dynamical systems have been used to model PD using finger tapping movement data, based on a feature design technique with a windowed GP method [381]. Both methods were inputted into an ensemble aimed at predicting PD (versus healthy controls). The research outlined in this thesis extends this study by examining the classification of rs-fMRI data for PD patients and prodromal PD patients from healthy age-matched controls. This is done using CGP and RCGP (as a dynamic classifier) as the main classification techniques and using the rs-fMRI data (particularly large and extremely dynamic data) as inputs to the classifiers.

4.7. Imbalanced Data

Learning algorithms typically assume approximately equal class distributions. Hence, these algorithms may function with low accuracy when faced with significantly imbalanced datasets, labelled between-class imbalance. For instance, the research presented in this thesis used heavily between-class imbalanced data from 102 PD patients versus eight controls. Misclassification in the medical domain can have grave implications (e.g., non-diagnosis of cancer [382]). As such, highly accurate classifiers that function with between-class imbalanced data to identify a minority class (e.g., AD patients) are necessary. In these cases, the overall accuracy or error rate may not be sufficient and other metrics (e.g., receiver operating characteristics curves, precision-recall curves, and cost curves) may better represent the performance of algorithms with imbalanced data.

Small sample sizes with significantly imbalanced class distributions are particularly common in clinical research due to challenges in recruiting patient samples and constrict learning due to two limitations. Firstly, reduced sample sizes result in problems linked to absolute rarity and within-class imbalances. Secondly, algorithms frequently do not generalise inductive rules across the sample. Algorithm performance is restricted by the limitations implicit in generating conjunctions across many features with reduced samples, which can result in overfitting (which occurs when the rules produced are overly precise).

Indeed, the class-imbalance problem is a relative problem that depends on (1) the degree of class-imbalance; (2) the complexity of the concept represented by the data; (3) the overall size of the training set; and (4) the classifier involved. More specifically, the higher the degree of class-imbalance, the higher the complexity of the concept. The smaller the total size of the training set, the greater the effect of class imbalances in classifiers sensitive to the problem [383].

Research has revealed that the performance of certain classifiers trained on specific imbalanced data can be similar to the performance of these same classifiers trained on the same data that has been modified to have approximately equal class distributions [383], [384]. Yet, for most imbalanced data, solutions for learning predominantly focus on modifying the data sample to achieve a sample that has balanced class distributions, which enhances the overall classification accuracy relative to the original imbalanced sample [385]–[387]. There are four key solutions for imbalanced learning based on sampling methods, which are summarised briefly below: random oversampling and undersampling, informed undersampling, synthetic sampling with data generation, and adaptive synthetic sampling.

4.7.1. Random Oversampling and Undersampling

The numbers within each class distribution can be modified to achieve balanced data. Oversampling involves adding additional data to the original dataset whereas undersampling involves taking away data. Random oversampling is achieved by incorporating a group sampled from the lesser class. For a group of randomly selected minority examples, the number of samples in the original set should be increased by replicating the selected examples and adding them to the whole set. Therefore, the number of total examples is augmented and the class distribution balance of the set is adjusted accordingly. For undersampling, a set of majority class examples are randomly selected and removed from the set.

Both oversampling and undersampling have limitations that constrain the performance of learning algorithms [388]–[390]. Given that oversampling includes additional simulated

data, this can result in overfitting [389] in which classifiers create precise rules to account for the various replications of the data. Overfitting is apparent when training accuracy is high but testing and validation accuracies are substantially lower [388]. A limitation inherent in undersampling is that eliminating data can lead the classifier to neglect key patterns within the majority class.

4.7.2. Informed Undersampling

Informed undersampling minimises the impact of removing data that occurs in standard random undersampling. The EasyEnsemble, BalanceCascade, and kNN algorithms are examples of informed undersampling [391]. For instance, regarding eliminating data from the majority class, EasyEnsemble uses unsupervised independent random sampling with replacement, whilst BalanceCascade involves supervised and systematic identification of data to eliminate. kNN involves four techniques to undersample [392]: (1) NearMiss-1 removes data whose mean are most similar to the three *nearest* minority class data points; (2) NearMiss-2 removes data whose mean are most similar to the three *furthest* minority class data points; (3) NearMiss-3 removes a certain number of data whose mean are most similar to *different* minority class data points; and (4) Most Distant removes data whose mean are *furthest* from the three *nearest* minority class data points. Findings reveal that NearMiss-2 provides more accurate results relative to the remaining kNN methods.

4.7.3. Synthetic Sampling with Data Generation (SMOTE)

Simulated data does not have the limitations implicit in oversampling, yet it increases the original dataset to significantly enhance algorithm learning. For example, SMOTE uses the feature space resemblances between minority data to generate simulated data. Findings reveal high accuracy for SMOTE in diverse applications [393]. Still, limitations include over-generalisation and low variance [394]. Specifically, SMOTE can result in heightened inter-class overlap [394] as it creates the same number of synthetic data samples per original minority example but does not take into account proximity of other data examples.

4.7.4. Adaptive Synthetic Sampling (ADASYN)

Adaptive sampling methods do not have the limitations associated with SMOTE and examples include Border-line-SMOTE [395] and ADASYN [396] algorithms. For instance, ADASYN is an extension of SMOTE that systematically and adaptively generates varying numbers of synthetic data based on minority and majority class distributions to create class balanced data [396], focusing on generating synthetic data close to the inter-class boundary. The research presented in this thesis applies ADASYN to generate class balanced data for classification.

4.8. CGP Classification

Classification has been widely investigated in machine learning and in data mining [397], [398], and involves predicting the value of a categorical attribute (the class) from values of different attributes (predicting attributes). A search algorithm is used to induce a classifier from a group of correctly classified data instances (training set). A different group of correctly classified data instances (test set) is used to examine the quality of the classifier obtained. Various types of models (e.g., decision trees or rules) can depict the classifiers. Genetic programming is an evolutionary learning method with good prospects for classification.

The training set selects the optimum parameters for a specific model and provides an unbiased estimate of cost function when evaluating specific parameters. Selecting the parameters optimises the estimate of the cost function, given the training set, which biases this estimate. The selected parameters perform best on the training set and, therefore, the performance of these parameters in the training set is quite optimistic. For the research presented in this thesis, a different set has been used for validation (referred to as validation set) by dividing the test set into validation and test sets. Validation sets are used in order to help with avoiding overfitting, as discussed in the generalisation section 4.4, and to select the best model. Model evaluation provides a representative estimate of the cost function as it selects the best performing model using the validation data, which biases this estimate, resulting in an optimistic estimate of model performance. Following training

of models (using training data) and model selection (using validation data), the final step involves identifying performance of the selected model (using test data), providing an unbiased estimate of performance.

Various types of models (e.g., decision trees or rules) can depict classifiers. CGP is a graph based EAs system with good prospects for classification. For the research presented in this thesis, multiple outputs were used in the CGP library, one for each class. Advantages associated with this method include the following:

- When using a threshold, its position is often arbitrarily chosen, however, its position is likely to affect the results. This issue is removed when using multiple outputs.
- When multiple outputs are used, these can involve using different factors in classification rather than the same factors, as per the other class.
- The value of each output gives a confidence related to its corresponding class, rather than the produced class.
- When using many classes (more than two), the order in which they are assigned to the regions of the single output is important. For instance, whether "Class A" flow more naturally into "Class B" or "Class C", however, when using multiple outputs this is not a concern.
- When using many classes (more than two), the region of values that represent each class become smaller as the number of the types of classes increases. This is likely to make the task more challenging as the number of classes grows.

CGP has multiple advantages relative to other classification methods. These include the reuse of sub-functions, having no bloat [336], and being able to report the classification model in a mathematical formula/expression as well as a graphical diagram. Figure 4.11 and Figure 4.12 present examples of CGP classification trees; dotted lines in Figure 4.12 represent implicit reuse of sub-expressions within the network and numbers indicate offsets in the matching window. These advantages are conducive to understanding the underlying elements of the classification model in more detail, which is not an option in most other classification methods.

Classification approaches apply machine learning classifiers to differentiate disease subtypes, often applying this to clinical data (e.g., patients versus healthy controls [399], [400]). The research presented in this thesis examines the use of classification as a diagnostic tool to improve current methods of PD diagnosis.

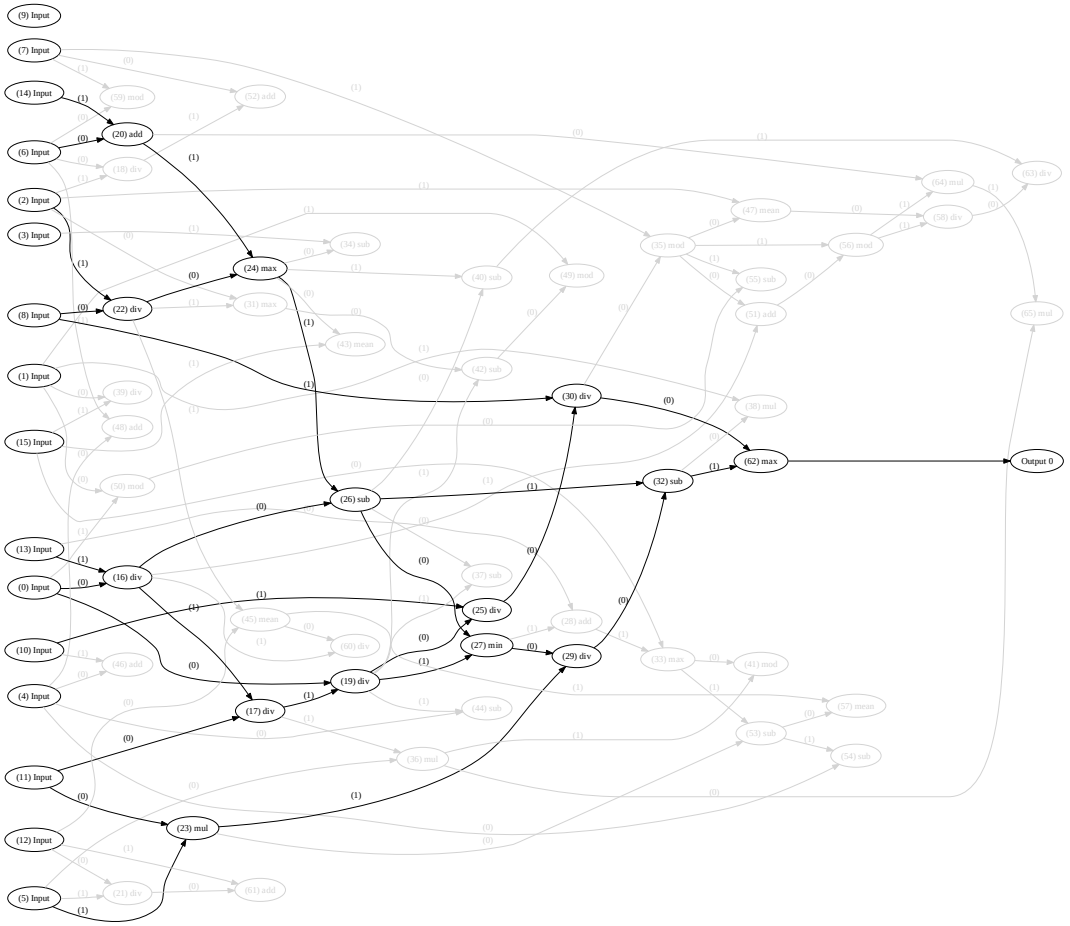

Figure 4.11 - An example of CGP classification tree (graph)

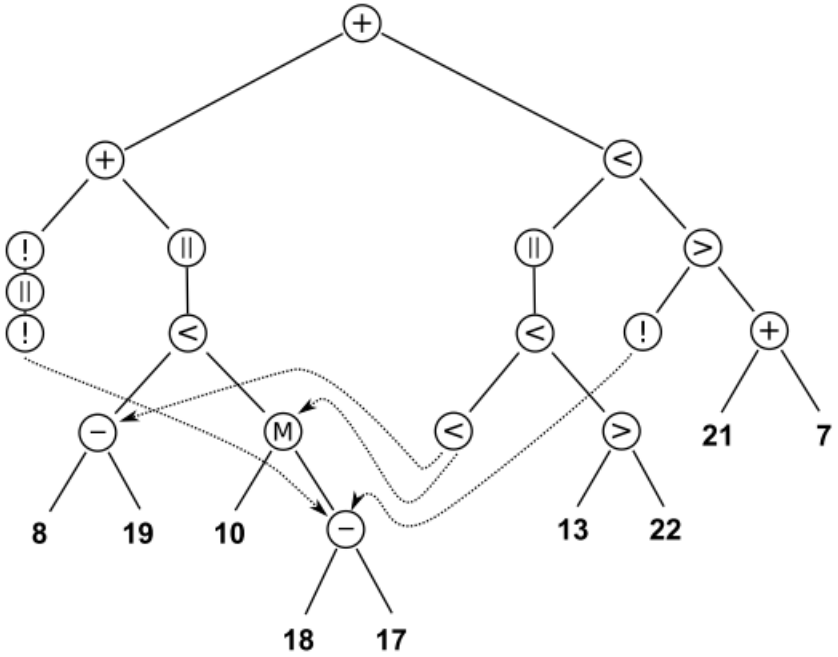

Figure 4.12 - An example of best evolved classifier; taken from [401]

4.9. Research Applying EAs to Classification of Medical Data

Research has explored the classification of PD patients using learning algorithms, yet, these have not used EAs. EAs offer a novel approach to disease classification, which is examined in this thesis. EAs are optimising algorithms based on Darwinian evolutionary theory. CGP is a subtype of EAs that as a norm evolves directed acyclic computational structures of nodes. RCGP is an extension of CGP, which enables cyclic or feedback connections. CGP and RCGP are used in the research presented in this thesis as the main methods of classification. CGP and RCGP have never been applied to neuroimaging data, including rs-fMRI, which is explored in this thesis.

A key benefit of using EAs alongside an expressive dynamical representation is the ability to explore a wide area of classifiers. In addition, since these classifiers do not rely on expert knowledge, they can identify trends that might not be detected by experts and contribute to furthering expert knowledge. For instance, evolved classifiers and their distributions have provided the following scientific contributions: the differential effect of dominance on diagnostic accuracy, the over-representation of certain trends of acceleration in the movements of PD patients, and amplitude and frequency blends with diagnostic power. Regardless of how efficient these classifiers are, a limitation of this method is the lack of knowledge underpinning how these algorithms function, rendering them often unfathomable to experts. Hence, these classifiers are a valuable tool in guiding and/or supporting a medical diagnosis, yet, an automated diagnosis cannot be approved unless the clinician is confident regarding the biological underpinnings of the diagnosis.

4.10. Computational Intelligence Approaches for Diagnosing Parkinson's Disease

4.10.1. Kinematic Research

Technological advances have resulted in more medical research using recorded data to broaden current understanding of the mechanisms underpinning diseases, for instance, positional data has been instrumental in furthering knowledge on Parkinsonian movement disorders. Prior to the development of wearable sensors, medical research used invasive technology. For example, Benecke et al. (1987) measured elbow and finger position during

motion and revealed that PD patients (relative to healthy controls) had difficulties switching between different sequential motor movements. Specifically, Benecke et al. used a potentiometer and a strain gauge to record kinematic data while capturing electromyography signals from a body part [402]. Likewise, Agostino et al. (1992) asked patients to trace patterns while grasping a device with two potentiometers to examine sequential arm movements in individuals with PD, HD, and Dystonia, revealing abnormal sequential movements for patients with these disorders [403]. These tests were not conducted within patients' home setting, where movement symptoms would occur, but in a clinical environment using awkward and intrusive equipment.

Currently, sensors are relatively small and provide multiple advantages for research, such as diminished awareness of the sensors and reduced discomfort, resulting in a more tranquil test setting. Much research uses positional and rotational sensors to measure the affected body part's spatial location. For example, a component of assessing motor decline involves a finger-tapping task in which patients tap their thumb and forefinger numerous times whilst a medic evaluates performance. Agostino et al. used optoelectronic sensors to generate 3D finger-tapping motion data of a single finger and of four fingers tapping on a thumb, taking the number of finger taps, size, pauses and time as relevant variables. Findings revealed worse motor disorders for single finger taps relative to whole hand taps for PD patients [404], [405]. More recent research has used finger-tapping [406], reach and grasp [407], and drawing tasks [408], which focus on the application of EAs for different methods of diagnosing PD or confirming the diagnosis of PD.

A standardised movement task (such as finger-tapping) can be advantageous as it is easily replicable for subsequent research and provides a familiar setting for participants. Nevertheless, finger-tapping tasks typically reveal low inter- and intra-rater reliability, which is typically attributed to the limited capacity of medics to evaluate performance solely via observation [93]. Specialised equipment may be better equipped to detect early stages of cognitive and motor decline, relative to medics. A new approach involves capturing patient movements via computational techniques and using EAs to assess performance. A recent study revealed that using EAs to assess performance data from a finger-tapping task alone differentiated PD patients from age-matched controls with an

accuracy of ~95% [409]. As such, data from traditional motor exams combined with new computational techniques are a valid and relatively novel approach in diagnosing PD.

Table 4.1 presents research examining movement features, which underscores the relevance that neurodegenerative researchers have placed on investigating motor disorders via the use of transducer technology to capture precise movement data. This research typically uses a standard feature extraction and statistical analysis method to examine the data, which relies on both expert knowledge and an exploratory hypothesis. In contrast, research using data mining often comprises of unbiased data analysis to detect significant relationships and trends.

Table 4.1 - Kinematic research

Authors	Date	Research
Manson et al. [410]	2000	Ambulatory dyskinesia monitors
Berardelli et al. [411]	2001	Effect of sensory cues on bradykinesia
Hoff et al. [412]	2001	Objectively assessing LIDs
Bonato et al. [413]	2004	Measuring longitudinal motor fluctuations
Iansek et al. [414]	2006	Improving freezing of gait with external cues
Chee et al. [415]	2009	Reducing step length of freezing of gait
Rodrigues et al. [416]	2009	The slowing of successive finger tap motions
Yokoe et al. [417]	2009	Different features of finger tapping
Patel et al. [418]	2009	The use of quantifiable data to estimate severity of symptoms
Espay et al. [419]	2011	How bradykinesia responds to dopaminergic medication
Heldman et al. [93]	2011	Examining the Modified Bradykinesia Rating Scale reliability with quantifiable measures
Ling et al. [420]	2012	Reducing hypokinesia in PD patients and PSP

4.10.2. Non-movement PD studies

Non-movement data is also relevant in PD research. For instance, many imaging modalities have been used to examined the neural activity of PD patients. Indeed, the research presented in this thesis examines rs-fMRI brain images of PD patients, prodromal PD patients, and healthy control participants. Alternative tasks have been developed to explore the impact of motor symptoms on many functions, including investigating tremor in a shape tracing task [421], examining the influence of PD treatment on the sequence

effect and ‘true’ bradykinesia via a pegboard test [422], and exploring the speech patterns of PD patients (relative to healthy controls) by focusing on vocal disorders when pronouncing vowel sounds [423].

4.10.3. Computational Modelling

Advances in computer software technology have enabled learning algorithms to be trained on big data with sensible time frames, producing intricate computational models of NDDs. For example, Keijsers et al. asked patients suffering from LID to wear a sensor whilst completing their normal daily tasks, capturing 92 movement features that were modelled with ANN. The features involved summaries of the processed separation signal (in time and frequency domains), based on prior information of the frequency bands in which dyskinesia are often present [424]. Patel et al. examined dyskinesia and bradykinesia using accelerometer data from which multiple features were identified. Signal processing techniques were used for data analysis, such as analysing the frequency spectrum and running cross-correlation, and then using SVM to model these features. Patel et al. used prior expertise for the data extraction process to identify areas of the frequency spectrum sensitive to variations in Parkinsonian symptoms [418], [425].

Lones et al. developed two classification methods (sliding window genetic programming expressions and artificial biochemical networks) that identified PD patients relative to healthy controls with ~95% accuracy using movement data [409]. Furthermore, Ericsson et al. used SPECT images of the basal ganglia of PD patients, to model 17 summary data features via SVM, based on expert knowledge [426]. Lones et al. developed classifiers based on EAs to differentiate LID movements from participants’ daily activities [427]. Lones et al.’s research has been translated into LID-monitor, which is a device that has been developed in the spinout company formed by the researchers, ClearSky Medical Diagnostics. LID-monitor is already being used in multiple NHS centres (including Harrogate District Hospital, Leeds Teaching Hospital, amongst others) and in hospitals in other countries (including China).

These described studies generated accurate models and are relevant in clinical settings to aid PD diagnosis and monitoring. Most of the research presented in this section used prior expert understanding of the disease to extract summary features from the raw data, which were then input into predictive model learning algorithms. Nevertheless, an automated and interpretable feature extraction method for data mining would further advance knowledge by objectively modelling kinematic symptoms and providing relevant information to clinicians. The research presented in this thesis explores an automated and interpretable technique for modelling brain imaging data of PD patients and prodromal PD patients.

4.11. Summary

Statistical predictive modelling provides multiple methods for predicting outcomes using known data, which are applicable for the development of an automated diagnosis of PD, as outlined in the research presented in this thesis. Specifically, CGP classifiers reveal good performance, even with small sample sizes, and can generate models that obtain high accuracy rates. This method is applicable to data with numerous features, such as timeseries data. A highly effective method of classifying timeseries data involves feature design techniques to limit the dimensionality into discrete, practicable feature vectors. Nevertheless, many feature design methods have poor interpretability, hence, *a priori* knowledge is valuable to aid interpretation. For the analysis of rs-fMRI data, this is already accounted for in the preprocessing steps. Therefore, there is no need to use any feature design/selection techniques before subjecting the timeseries data to classification. Moreover, the research presented in this thesis uses DCM analysis of rs-fMRI data as a novel tool that has not been previously explored in classification.

Chapter 5. Towards Monitoring Parkinson's Disease Following Drug Treatment: CGP Classification of rs-fMRI Data

This chapter develops a technique for accurate monitoring of PD using EAs on rs-fMRI data for participants prescribed Modafinil (typically prescribed for PD patients to relieve physical fatigue). Modafinil is a "smart drug" and a fatigue-reducing medication that boosts attention and memory. Approximately half of PD patients report fatigue-related symptoms [124], hence, Modafinil is typically prescribed to alleviate these symptoms [123]–[126]. A crucial overarching research question is: Can methods for differentiating clinical groups using EAs on rs-fMRI data be developed, based on a controlled clinical experiment and focusing on a medication (Modafinil) typically administered to PD patients? This experiment examines rs-fMRI data in which healthy young adults were administered a single dose of 100 mg of Modafinil (versus a placebo). This experiment used healthy participants rather than PD patients in order to initially develop accurate monitoring of individuals administered a common PD medication, with implications for PD patient monitoring. The activity of RSN and the intrinsic connectivity are explored and these data are subjected to a classifier to examine the physiological impact of Modafinil.

This chapter addresses two further research questions: (1) Can CGP be applied to the monitoring of participants administered Modafinil (versus a control group) using rs-fMRI data? And (2) are timeseries analyses and DCM analyses relevant for classification? This research is exploratory, being the first to examine Modafinil classification to differentiate participant groups. Developing a method to differentiate participants is clinically relevant with implications for patient diagnosis (classification of patients relative to healthy controls, as presented in Chapter 6) and drug treatment monitoring.

Research on rs-fMRI [238], [428] is prolific [429], specifically with regards to differentiating participant groups or brain states based on functional connectivity. For instance, Esposito et al. examined the effects of Modafinil on functional connectivity using rs-fMRI [430]. rs-fMRI is a valuable element of fMRI data acquisition, given that it is well-suited to investigating changes in functional connectivity [431] and can explore certain research questions that could not be addressed by relying solely on task-related fMRI data. Numerous cortical circuits can be examined simultaneously using rs-fMRI.

Moreover, confounding variables, for instance, inter-individual variability in task performance, is limited when examining participants that are at rest [432].

BOLD fMRI was primarily used in early rs-fMRI studies, examining the activity of certain brain regions at rest [433]. Findings revealed that rest activity is related to at least ten [241], [434]–[436] functional RSN [241], these include the DMN [243], the Salience Network, the Fronto Parietal Control (FPC) network (lateralised in both hemispheres), the primary Sensory Motor Network, the Exstrastriate Visual System, and the Dorsal Attention Network [241]. DMN is studied as a norm within rs-fMRI research and, hence, is examined in the current study.

Functional connectivity allows researchers to examine abnormal patterns of activity, yet, it does not enable researchers to identify the original source of effective connectivity (i.e., the impact of one neuronal system on another) [281]. DCM is based on a neuronal model of coupled neuronal states. Friston et al. developed DCM to fully examine the effective connectivity underpinning functional connectivity, focusing on rs-fMRI [285]. The current research applies classification to the DCM analysis of rs-fMRI data.

5.1. Research Overview and Aims

A core research objective is to develop methods for differentiating clinical groups following drug treatment, based on a controlled clinical experiment. Specifically, this experiment explored the question: Can accurate monitoring of PD be achieved using EAs on rs-fMRI data for patients prescribed Modafinil? This research involved an analysis of rs-fMRI data taken from OpenfMRI database (accession number: ds000133; [437]), in which healthy young adults were administered a single dose of 100 mg of Modafinil (versus a placebo). In the current research, the activity of the RSN and the functional connectivity were examined, and this data was subjected to a classifier to explore the physiological impact of Modafinil.

This research applied EAs, specifically CGP and RCGP, for the classification of rs-fMRI using DCM and timeseries analyses. The timeseries values and DCM values from the rs-

fMRI data are subjected to supervised classification. CGP and RCGP have not previously been used for the classification of brain imaging data. Moreover, this study explores an additional novel question: is DCM analysis useful for classification? Previous research has not explored the applicability of DCM values in classification.

A common limitation in medical research involves recruiting equal sample sizes of patients and healthy controls, often leading to class-imbalanced data (e.g., unequal groups of controls versus patients). The OpenfMRI dataset used in this experiment is heavily class-imbalanced, with many more control participants relative to participants administered Modafinil. Hence, this research explored the applicability of classification methods to class-imbalanced data, with implications for the transferability of medical research based on limited and imbalanced sample sizes.

5.2. Method

5.2.1. Participants

Twenty-six male participants were recruited with an age range of 25–35 years. Participants were right-handed (tested using the Edinburgh Handedness inventory) [438], with similar educational level, and no history of psychiatric, neurological or medical (hypertension, cardiac disorders, and epilepsy) conditions as identified by the Millon test and by clinical examination. The participants provided written consent and the study was approved by the ethics committee of University of Chieti (PROT 2008/09 COET on 14/10/2009) and conducted in accordance with the Helsinki Declaration [430], [437].

5.2.2. Procedure

Participants were told to consume their normal amount of nicotine and caffeine and to refrain from consuming alcohol 12 hours prior to the study. Participants were administered Modafinil (100 mg) or placebo. The study was double blind and both the Modafinil and placebo pills looked identical. Following consumption of the drug, participants were given an fMRI scan.

Data from all participants were processed and separated into their corresponding treatment groups (Modafinil and placebo). Data from one control participant was excluded, as this data was too poor quality to be analysed. Pre-session data was also categorised as placebo/control as this research only examines the effect of Modafinil on brain functionality. The Modafinil group contains 39 participants: 13 participants tested in one session with three runs. The control group contains 111 participants: 12 participants tested in two sessions with three runs plus 13 participants tested in one session with three runs.

5.2.3. rs-fMRI Acquisition

rs-fMRI BOLD data was separated in three runs (duration: 4 min each) and then high resolution T1 anatomical images were acquired. Participants were instructed to focus on the centre of a grey screen that was projected on a LCD screen, viewed via a mirror located over the participant's head. The participant's head was placed in an eight-channel coil with foam padding to reduce involuntary head movements. BOLD functional imaging was performed with a Philips Achieva 3T Scanner (Philips Medical Systems, Best, The Netherlands), using T2*-weighted echo planar imaging (EPI) free induction decay (FID) sequences and applying the following parameters: Echo Time (TE) 35 ms, matrix size 66×66, Field of View (FoV) 256 mm, in-plane voxel size 464 mm, flip angle 75°, slice thickness 4 mm and no gaps. 140 functional volumes consisting of 30 transaxial slices were acquired per run with a volume Repetition Time (TR) of 1671 ms. High resolution structural images were acquired at the end of the three rs-fMRI runs through a 3D MPRAGE sequence employing the following parameters: sagittal, matrix 256×256, FoV 256 mm, slice thickness 1 mm, no gaps, in-plane voxel size $1\ mm \times 1\ mm$, flip angle 12°, TR = 9.7 ms and TE = 4 ms [430], [437].

5.2.4. Imaging Data Analysis

5.2.4.1. Preprocessing

The imaging data analyses were done using the CONN (version 17.c) [251], [252] and SPM12 (version 6906 - Wellcome Department of Imaging Neuroscience, London, UK) [230] software packages based on MATLAB. Preprocessing included 4D NIFTI (an

Analyze-style data format, proposed by the NIFTI Data Format Working Group as a "short-term measure to facilitate inter-operation of fMRI") import, 4D to 3D NIFTI conversion, and reduction of the spatial distortion using the Field Map toolbox in SPM12 [230]. Anatomical data was segmented and both anatomical and functional data were normalised. All the functional images were motion corrected and coregistered to participants' own high-resolution anatomical image. The participants' anatomical images were normalised to the standard T1 template in the MNI space, as provided by SPM12. Then the normalisation parameters of each participant were applied to the functional images to normalise all the functional images into the MNI space. The EPI data was unwarped (using field-map images) to compensate for the magnetic field inhomogeneities, realigned to correct for motion, and slice-time corrected to the middle slice. The normalisation parameters from the T1 stream were then applied to warp the functional images into MNI space. All the functional images were spatially smoothed using a Gaussian kernel with 8 mm [269] FWHM to account for inter-participant variability while maintaining a relatively high spatial resolution. Linear and quadratic detrending of the fMRI signal was applied, which involved covarying out WM and CSF signal. WM and CSF signals were predicted for each volume from the mean value of WM and CSF masks, derived by thresholding SPM's tissue probability maps at 0.8. The data was bandpass filtered (0.008–0.1 Hz).

The preprocessing pipeline can be seen in Figure 3.14 and the MATLAB batch script for preprocessing is presented below:

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% SPATIAL PREPROCESSING
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
data_path = [data_folder_name, '/sub-', ...
    subjectno, '/ses-', sessionprepost, '/'];

f = spm_select('FPList', fullfile(data_path,['func/Run0', ...
    runno]), '^sub.*\.nii$');
a = spm_select('FPList', fullfile(data_path,'anat'), ...
    '^sub.*\.nii$');

cd([data_path, 'func/Run0', runno]);

clear matlabbatch
```

```
% Realign
%-----------------------------------------------------------------
matlabbatch{1}.spm.spatial.realign.estwrite.data = ...
    {cellstr(f)};
matlabbatch{1}.spm.spatial.realign.estwrite.roptions.which ...
    = [0 1];

% Coregister
%-----------------------------------------------------------------
matlabbatch{2}.spm.spatial.coreg.estimate.ref    = ...
    cellstr(spm_file(f(1,:),'prefix','mean'));
matlabbatch{2}.spm.spatial.coreg.estimate.source = cellstr(a);

% Segment
%-----------------------------------------------------------------
matlabbatch{3}.spm.spatial.preproc.channel.vols  = cellstr(a);
matlabbatch{3}.spm.spatial.preproc.channel.write = [0 1];
matlabbatch{3}.spm.spatial.preproc.warp.write    = [0 1];

% Normalise: Write
%-----------------------------------------------------------------
matlabbatch{4}.spm.spatial.normalise.write.subj.def      = ...
    cellstr(spm_file(a,'prefix','y_','ext','nii'));
matlabbatch{4}.spm.spatial.normalise.write.subj.resample = ...
    cellstr(f);
matlabbatch{4}.spm.spatial.normalise.write.woptions.vox  = ...
    [3 3 3];

matlabbatch{5}.spm.spatial.normalise.write.subj.def      = ...
    cellstr(spm_file(a,'prefix','y_','ext','nii'));
matlabbatch{5}.spm.spatial.normalise.write.subj.resample = ...
    cellstr(spm_file(a,'prefix','m','ext','nii'));
matlabbatch{5}.spm.spatial.normalise.write.woptions.vox  = ...
    [1 1 1];

% Smooth
%-----------------------------------------------------------------
matlabbatch{6}.spm.spatial.smooth.data = cellstr(spm_file ...
    (f,'prefix','w'));
matlabbatch{6}.spm.spatial.smooth.fwhm = [8 8 8];

spm_jobman('run',matlabbatch);
clear matlabbatch;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% PROCESSING
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%



%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Defining the GLM
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
mkdir('GLM');
cd([data_path, 'func/Run0', runno, '/GLM']);

matlabbatch{1}.spm.stats.fmri_spec.dir = {[data_path, ...
```

```
    'func/Run0', runno, '/GLM']};
matlabbatch{1}.spm.stats.fmri_spec.timing.units = 'scans';
matlabbatch{1}.spm.stats.fmri_spec.timing.RT = TR;
matlabbatch{1}.spm.stats.fmri_spec.timing.fmri_t = 16;
matlabbatch{1}.spm.stats.fmri_spec.timing.fmri_t0 = 8;
matlabbatch{1}.spm.stats.fmri_spec.sess.scans = cellstr...
    (spm_select('FPList', fullfile(data_path,['func/Run0', ...
    runno]), '^sw.*\.nii$'));
matlabbatch{1}.spm.stats.fmri_spec.sess.cond = struct( ...
    'name', {}, 'onset', {}, 'duration', {}, 'tmod', {}, ...
    'pmod', {}, 'orth', {});
matlabbatch{1}.spm.stats.fmri_spec.sess.multi = {''};
matlabbatch{1}.spm.stats.fmri_spec.sess.regress = struct( ...
    'name', {}, 'val', {});
matlabbatch{1}.spm.stats.fmri_spec.sess.multi_reg = {''};
matlabbatch{1}.spm.stats.fmri_spec.sess.hpf = 128;
matlabbatch{1}.spm.stats.fmri_spec.fact = struct('name', ...
    {}, 'levels', {});
matlabbatch{1}.spm.stats.fmri_spec.bases.hrf.derivs = [0 0];
matlabbatch{1}.spm.stats.fmri_spec.volt = 1;
matlabbatch{1}.spm.stats.fmri_spec.global = 'None';
matlabbatch{1}.spm.stats.fmri_spec.mthresh = 0.8;
matlabbatch{1}.spm.stats.fmri_spec.mask = {''};
matlabbatch{1}.spm.stats.fmri_spec.cvi = 'AR(1)';
matlabbatch{2}.spm.stats.fmri_est.spmmat(1) = ...
    cfg_dep('fMRI model specification: SPM.mat File', ...
    substruct('.','val', '{}',{1}, '.','val', '{}',{1}, ...
    '.','val', '{}',{1}), substruct('.','spmmat'));
matlabbatch{2}.spm.stats.fmri_est.write_residuals = 0;
matlabbatch{2}.spm.stats.fmri_est.method.Classical = 1;

spm_jobman('run',matlabbatch);
clear matlabbatch;

% Now we will also need to extract signal from CSF and white
% matter (WM) to be used as confound. First, we are going to
% extract the WM (Pons) time series, which we will use as one
% of the nuisance variables.

matlabbatch{1}.spm.util.voi.spmmat = {[data_path, ...
    'func/Run0',  runno, '/GLM/SPM.mat']};
matlabbatch{1}.spm.util.voi.adjust = NaN;
matlabbatch{1}.spm.util.voi.session = nosess;
matlabbatch{1}.spm.util.voi.name = 'WM';
matlabbatch{1}.spm.util.voi.roi{1}.sphere.centre = [0 -24 -33];
matlabbatch{1}.spm.util.voi.roi{1}.sphere.radius = 4;
matlabbatch{1}.spm.util.voi.roi{1}.sphere.move.fixed = 1;
matlabbatch{1}.spm.util.voi.roi{2}.mask.image = {[data_path,...
    'func/Run0',  runno,  '/GLM/mask.nii']};
matlabbatch{1}.spm.util.voi.roi{2}.mask.threshold = 0.8;
matlabbatch{1}.spm.util.voi.expression = 'i1 & i2';

spm_jobman('run',matlabbatch);
clear matlabbatch;

% We do the same thing to extract CSF (from one of the
% ventricles) signal with a sphere centred on [1 -39 3].
```

```
% This will create the file VOI_CSF_1.mat.

matlabbatch{1}.spm.util.voi.spmmat = {[data_path, ...
    'func/Run0',  runno, '/GLM/SPM.mat']};
matlabbatch{1}.spm.util.voi.adjust = NaN;
matlabbatch{1}.spm.util.voi.session = nosess;
matlabbatch{1}.spm.util.voi.name = 'CSF';
matlabbatch{1}.spm.util.voi.roi{1}.sphere.centre = [1 -39 3];
matlabbatch{1}.spm.util.voi.roi{1}.sphere.radius = 4;
matlabbatch{1}.spm.util.voi.roi{1}.sphere.move.fixed = 1;
matlabbatch{1}.spm.util.voi.roi{2}.mask.image = {[data_path,...
    'func/Run0',  runno,  '/GLM/mask.nii']};
matlabbatch{1}.spm.util.voi.roi{2}.mask.threshold = 0.8;
matlabbatch{1}.spm.util.voi.expression = 'i1 & i2';

spm_jobman('run',matlabbatch);
clear matlabbatch;

% Next we need to adjust the SPM with the covariates.

matlabbatch{1}.spm.stats.fmri_spec.dir = {[data_path, ...
    'func/Run0', runno, '/GLM']};
matlabbatch{1}.spm.stats.fmri_spec.timing.units = 'scans';
matlabbatch{1}.spm.stats.fmri_spec.timing.RT = TR;
matlabbatch{1}.spm.stats.fmri_spec.timing.fmri_t = 16;
matlabbatch{1}.spm.stats.fmri_spec.timing.fmri_t0 = 8;
matlabbatch{1}.spm.stats.fmri_spec.sess.scans = cellstr ...
    (spm_select('FPList', fullfile(data_path,['func/Run0', ...
    runno]), '^sw.*\.nii$'));
matlabbatch{1}.spm.stats.fmri_spec.sess.cond = struct( ...
    'name', {}, 'onset', {}, 'duration', {}, 'tmod', {}, ...
    'pmod', {}, 'orth', {});
matlabbatch{1}.spm.stats.fmri_spec.sess.multi = {''};
matlabbatch{1}.spm.stats.fmri_spec.sess.regress = struct( ...
    'name', {}, 'val', {});
matlabbatch{1}.spm.stats.fmri_spec.sess.multi_reg = ...
    cellstr(spm_select('FPList', fullfile(data_path,[ ...
    'func/Run0',runno], '/GLM/'), '^VOI_CSF_.*\.mat$'));
matlabbatch{1}.spm.stats.fmri_spec.sess.multi_reg = cellstr ...
    (spm_select('FPList', fullfile(data_path,['func/Run0', ...
    runno], '/GLM/'), '^VOI_WM_.*\.mat$'));
matlabbatch{1}.spm.stats.fmri_spec.sess.multi_reg = cellstr ...
    (spm_select('FPList', fullfile(data_path,['func/Run0', ...
    runno], '/'), '^rp_sub-.*\.txt$'));
matlabbatch{1}.spm.stats.fmri_spec.sess.hpf = 128;
matlabbatch{1}.spm.stats.fmri_spec.fact = struct('name', {},...
    'levels', {});
matlabbatch{1}.spm.stats.fmri_spec.bases.hrf.derivs = [0 0];
matlabbatch{1}.spm.stats.fmri_spec.volt = 1;
matlabbatch{1}.spm.stats.fmri_spec.global = 'None';
matlabbatch{1}.spm.stats.fmri_spec.mthresh = 0.8;
matlabbatch{1}.spm.stats.fmri_spec.mask = {''};
matlabbatch{1}.spm.stats.fmri_spec.cvi = 'AR(1)';
matlabbatch{2}.spm.stats.fmri_est.spmmat(1) = cfg_dep...
    ('fMRI model specification: SPM.mat File', substruct(...
    '.','val', '{}',{1}, '.','val', '{}',{1}, '.','val', ...
    '{}',{1}), substruct('.','spmmat'));
```

```
matlabbatch{2}.spm.stats.fmri_est.write_residuals = 0;
matlabbatch{2}.spm.stats.fmri_est.method.Classical = 1;

spm_jobman('run',matlabbatch);
clear matlabbatch;
```

5.2.4.2. Processing

5.2.4.2.1. Timeseries

Functional connectivity in DMN is well studied. Hence, this research took as regions of interest (nodes) the most commonly reported four major parts of DMN, as shown in Figure 5.1: medial prefrontal cortex (mPFC, centred at 3, 54, −2), posterior cingulate cortex (PCC, centred at 0, −52, 26), left inferior parietal cortex (LIPC, centred at − 50, − 63, 32), and right inferior parietal cortex (RIPC, centred at 48, − 69, 35). For each participant, the volumes of interest were defined as spheres centred at those coordinates mentioned above with an 8 mm radius and with a mask threshold of 0.5. The first eigenvectors were extracted after removing the effect of head motion and low frequency drift. This vector is stored for each region as timeseries. The MATLAB script for extracting the PCC timeseries is presented below; in order to extract the timeseries for other brain regions (LIPC, RIPC, and mPFC), the word "PCC" are substituted accordingly.

```
matlabbatch{1}.spm.util.voi.spmmat = {[data_path, ...
    'func/Run0',  runno, '/GLM/SPM.mat']};
matlabbatch{1}.spm.util.voi.adjust = NaN;
matlabbatch{1}.spm.util.voi.session = nosess;
matlabbatch{1}.spm.util.voi.name = 'PCC';
matlabbatch{1}.spm.util.voi.roi{1}.sphere.centre = [0 -52 26];
matlabbatch{1}.spm.util.voi.roi{1}.sphere.radius = 8;
matlabbatch{1}.spm.util.voi.roi{1}.sphere.move.fixed = 1;
matlabbatch{1}.spm.util.voi.roi{2}.mask.image = {[data_path,...
    'func/Run0',  runno,  '/GLM/mask.nii']};
matlabbatch{1}.spm.util.voi.roi{2}.mask.threshold = 0.5;
matlabbatch{1}.spm.util.voi.expression = 'i1 & i2';

spm_jobman('run',matlabbatch);
clear matlabbatch;
```

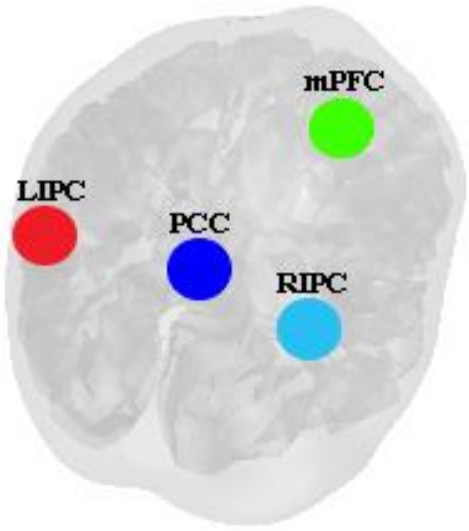

Figure 5.1 - The four DMN regions of interest used in this research

5.2.4.2.2. Dynamic Causal Modelling (DCM)

The spectral DCM analyses were conducted using the DCM12 implemented in SPM12. The regions of interest of DCM analyses were defined according to the peak of the DMN independent component maps, as presented in Figure 5.1. The main purpose of the current DCM analysis was to investigate the endogenous/intrinsic effective connectivity and to examine the causal interactions across these regions. The modelled low frequency fluctuations were set as driving inputs to all four nodes, and different models were defined by considering a full connection for all nodes. Expected posterior model probabilities and exceedance probabilities were computed. The intrinsic connectivity parameters (16 values that were stored in DCM.Ep.A matrix, referring to all parameters of intrinsic/effective connectivity [281]; Table 5.1 presents an example of an intrinsic connectivity matrix) from each participant were subjected to classification using CGP.

Table 5.1 - An example of DCM.Ep.A, intrinsic connectivity matrix

-0.19	0.17	0.07	0.12
0.15	-0.18	0.04	0.13
0.11	0.08	-0.38	-0.11
0.10	0.12	-0.04	-0.17

5.2.5. Cartesian Genetic Programming

As discussed in Chapter 4, CGP [331] is a strand of GP [325], [332], that encodes computational structures as generic cyclic/acyclic graphs. For this research, a new cross platform open source CGP library [339] was used since it is able to evolve symbolic expressions, Boolean logic circuits, and ANN, and it can be extended to diverse research areas. The CGP library enables the control of evolutionary parameters and the application of custom evolutionary stages.

5.2.5.1. Classification

To have equal class representation, data from each class was divided randomly into subsets of 70% (training), 15% (validation), and 15% (test). The geometry of the programs in the population (chromosomes) has fifty nodes with a function set of four mathematical operations (+, -, ×, ÷), multiple inputs (according to the datasets), and one output (class 1 for Modafinil participant group, class 0 for the control participant group). At each generation, the fittest chromosome is selected and the next generation is formed with its mutated versions (mutation rate = 0.1). Evolution stops when 15000 iterations are reached. To obtain statistical significance, the classification was done in 10 runs for each combination of inputs and the accuracy was averaged over the runs. The results (the winning chromosome, the networks, and the accuracy values) were stored for each run individually.

5.2.5.1.1. Classification of Timeseries

rs-fMRI is a widespread tool for exploring the functionality of the brain, using volume timeseries data. These scans contain abundant data; hence, obtaining relevant and useful data from raw scans (i.e., high dimensional datasets) can be difficult. Machine learning algorithms provide various tools that create datasets with less dimensions and more useful data, although challenges persist regarding how to select relevant data and how to maintain the interpretability of this data. This can result in losing important properties of the raw

data, although dealing with such a large number of features can be computationally expensive and very time consuming.

In the current experiment, the RCGP algorithm is used to classify the features that appeared across time in participant scans. The number of timeseries values is 145, i.e., for each of the four regions in the DMN, there is a vector of 145 values. Analysing/classifying the timeseries values was conducted in three different ways in terms of inputs to the classifier:

1. The timeseries values for each region separately were used as inputs to the classifier, to classify with 145 features per region (relating to DMN regions: LIPC, RIPC, PCC, and mPFC) and per participant.
2. The timeseries values together in four columns (one per region) were used as inputs to the classifier to classify the data with four columns (corresponding to the four DMN regions) of 145 features per participant.
3. The timeseries values inserted together in one column were used as inputs to the classifier to classify the data with 580 features in one vector for each participant. The order of inputting the timeseries values for each DMN region to form the final vector was consistent between participants.

Classification was completed in 10 runs for each combination of inputs and the accuracy was averaged over the runs. The same inputs were used to classify the data using ANN in MATLAB (Pattern Recognition and Classification Application – nprtool with 10 hidden layers) and also SVM in MATLAB for comparison/validation. RCGP was used with 10% probability for the recurrent connections. A complete pipeline of the preprocessing and processing of the data is presented in Figure 3.15.

The MATLAB script for the SVM classification method is included below:

```
%% Import data
clc; close all; clear

load DCM_Ep_A_total.mat
load Target_DCM.mat
```

```
totaldata = table();

for i = 1:size(DCM_Ep_A_total,2)
    totaldata{:,i} = DCM_Ep_A_total(:,i);
end
totaldata.Target = Target_DCM;

runs = 10;      % Number of runs to repeat the classification
AccuTrain = zeros(runs,1);
AccuTest = zeros(runs,1);
lossKnn = zeros(runs,1);
lossDa = zeros(runs,1);
results = zeros(runs,2);
for runno = 1:runs
    %% Split the data into training and test sets
    pt = cvpartition(totaldata.Target,'HoldOut',0.3);

    dataTrain = totaldata(training(pt),:);
    dataTest = totaldata(test(pt),:);

    %% Create a model
    mdl = fitcsvm(dataTrain,'Target','KernelFunction','gaussian');

    %% Calculate the loss errTrain and errTest
    [label,score] = predict(mdl,dataTest);
    % label represents the classification of each row in inputs; Each row
    % corresponds to a row in X, which is a new observation. The first
column
    % contains the scores for the observations being classified in the
    % negative class, and the second column contains the scores
observations
    % being classified in the positive class.
    errTrain = resubLoss(mdl);
    errTest = loss(mdl,dataTest);

    AccuTrain(runno,1) = 100*(1-errTrain);
    AccuTest(runno,1) = 100*(1-errTest);

    %% Create models and calculate loss
    % Create partition
    k = 10;     % Number of Folds
    part = cvpartition(totaldata.Target,'KFold',k);

    % k-NN
    mdlKnn = fitcknn(totaldata,'Target','NumNeighbors',5,...
        'CVPartition',part);
    lossKnn(runno,1) = kfoldLoss(mdlKnn);

    % Discriminant analysis
    mdlDa = fitcdiscr(totaldata,'Target','CVPartition',part);
    lossDa(runno,1) = kfoldLoss(mdlDa);

    %% Create a table to hold the results
```

```
    results(runno,:) = [(1-lossKnn(runno,1))*100 (1-
lossDa(runno,1))*100];

end

%% Take the mean of the results
meanTrain = mean(AccuTrain);
meanTest = mean(AccuTest);
sdTrain = std(AccuTrain);
sdTest = std(AccuTest);
meanresults = mean(results);
sdresults = std(results);
```

5.2.5.1.2. Classification of DCM

The classification was run in the CGP Library with 16 inputs (all the DCM values sorted by region and presented as only one vector per participant). To facilitate comparison, the classification with data from the same participants was run using ANN and SVM, both run in MATLAB.

5.2.6. Adaptive Synthetic Sampling

As previously mentioned, the Modafinil group contained 39 participants and the control group contained 111 participants, resulting in highly imbalanced data. Hence, ADASYN was used to make the data balanced. After the process, the minor group for each combination (in the training set) had a higher number of participants, which made the data balanced for CV. The validation and test sets were kept the same.

5.2.7. *k*-Fold Cross-Validation

10-fold CV was used to evaluate the classification accuracy using an unbiased estimate of the generalisation accuracy [351]. CV is beneficial as it allows the generation of independent test sets with enhanced reliability. With *10*-fold CV, typically one (of *10*) subset is the test set and remaining nine subsets are training sets. These sets are then rotated so that each set is used to test the data once. One repetition of the *10*-fold CV does

not produce sufficient classification accuracies for comparison, therefore, *10*-fold CV is repeated 10 independent times and the mean accuracy over all *10* trials is calculated.

Since the main classification methodology in this research involved dividing the data into three different subsets (training, validation, and test), the data was divided into *10* subsets and, each time, one of the *10* subsets were used as the test set, another one for validation, and the remaining eight as a training set. The data was divided using stratified random sampling enabling the sample proportion in each data subset to be the same as that in the original data (i.e., equal class distribution in the subsets as per the original data). Hence, the data was split for each class initially and then the classes were mixed to form the completed set. This was done for each combination of inputs to both CGP (for DCM values) and RCGP (for timeseries values). In this study, the data for each combination of inputs was divided into three parts of 80% (training), 10% (validation), and 10% (test).

5.3. Results

This study examined the classification of 39 participants administered Modafinil versus 111 control participants. The analysis (classification) focused on organising features to be used as inputs to the classifier in CGP and also in RCGP, implemented using the CGP Library. To validate the results, the analysis/classification was additionally completed using ANN and SVM, both in MATLAB.

5.3.1. Classification of Timeseries

Initially, the timeseries values for each region were used as inputs to the classifier individually. Therefore, the data was classified with 145 features from each region per participant. The same procedure was completed separately for each DMN region (PCC, mPFC, RIPC, and LIPC). Then, the timeseries values were used as inputs to the classifier to classify the data in four columns (relating to the four DMN regions) of 145 features per participant. Finally, the timeseries values together in one column were used as inputs to the classifier to classify the data with 580 features in one vector for each participant. The results after 10 runs for each combination were averaged and are summarised in Table 5.2.

The classification was also done using ANN and SVM. For SVM, only the training and test sets were considered for classification.

Table 5.2 - Classification results for the timeseries values

		Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
Classification results for each DMN region				
PCC	RCGP	73.75 (2.98)	75.52 (1.59)	74.28 (1.22)
	ANN	83.74 (7.78)	73.48 (12.88)	72.18 (8.01)
	SVM	100 (0)	NA	73.81 (0.50)
mPFC	RCGP	71.88 (5.00)	74.80 (1.57)	73.52 (1.38)
	ANN	78.07 (9.02)	71.75 (9.90)	64.78 (9.93)
	SVM	100 (0)	NA	74.00 (0.46)
RIPC	RCGP	74.40 (2.96)	76.51 (2.57)	71.26 (5.94)
	ANN	81.92 (7.63)	72.18 (7.46)	70.00 (8.06)
	SVM	100 (0)	NA	73.71 (0.49)
LIPC	RCGP	72.79 (3.52)	74.79 (1.55)	72.53 (3.22)
	ANN	85.88 (10.20)	70.88 (11.96)	70.45 (8.65)
	SVM	100 (0)	NA	73.90 (0.49)
Classification results for all the DMN regions (4 inputs)				
	RCGP	73.63 (2.11)	75.92 (2.16)	73.99 (2.94)
	ANN	73.83 (0.17)	74.49 (0.67)	74.24 (0.49)
	SVM	74.25 (0.05)	NA	73.95 (0.05)
Classification results for all the DMN regions (1 input)				
	RCGP	74.68 (1.60)	76.55 (2.82)	74.57 (1.82)
	ANN	73.97 (0.13)	74.11 (0.37)	74.01 (0.45)
	SVM	74.00 (0.00)	NA	74.00 (0.00)

Findings revealed that the Modafinil group were successfully classified from the control group with a maximum accuracy of 74.57% using RCGP (minimum accuracy: 71.26%). The results from the other two classification techniques (ANN and SVM) validated this finding as they were very similar: 64.78-74.24% for ANN and SVM in all the different combinations of inputs.

Unlike for the DCM analyses, mixed ANOVAs were not conducted to evaluate the correspondence between participant group and timeseries features, given that there were 580 features per participant, which would not be interpretable.

5.3.2. Classification of Dynamic Causal Modelling (DCM)

Classification using the CGP Library was executed with 16 inputs (all the DCM values sorted by region and presented as one vector per participant) and 1 output (class 1 for the Modafinil group and class 0 for the control group). The results were then averaged over 10 runs and are presented in Table 5.3. The classification was done using ANN and SVM. For SVM, only the training and test sets were considered for classification.

Table 5.3 - Classification results for DCM values

	Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
CGP	75.63 (7.13)	80.35 (6.18)	73.89 (7.70)
ANN	79.99 (5.57)	73.93 (7.94)	67.39 (6.57)
SVM	73.81 (0.50)	NA	73.81 (0.50)

Findings revealed that the Modafinil group were successfully classified from the control group with 73.89% accuracy using CGP. The results from the other two classification techniques (ANN and SVM) validated this finding as they were very similar: 67.39% for ANN and 73.81% for SVM.

Two examples of the CGP classification network trees/graphs can be seen in Figure 5.2 and Figure 5.3. This reflects one of the core advantages of CGP in terms of providing a *white box solution*, giving more information on the inputs used and enhanced knowledge concerning the final solution obtained in classification, which is not easily possible (if at all) with ANN and SVM classification techniques. These networks and their respective mathematical expressions are very complex and often difficult to interpret. Nevertheless, these networks can provide highly useful information. For example, in Figure 5.2, only half of the inputs have been used to arrive at the final classification. Similarly, Figure 5.3 reveals that only 11 of the inputs are used and the network evolved, in this case, is arguably somewhat simpler than that depicted in Figure 5.2. Future work can explore constraining the geometry of the classifier (at the cost of classification accuracy) in order to simplify the networks and enhance interpretability.

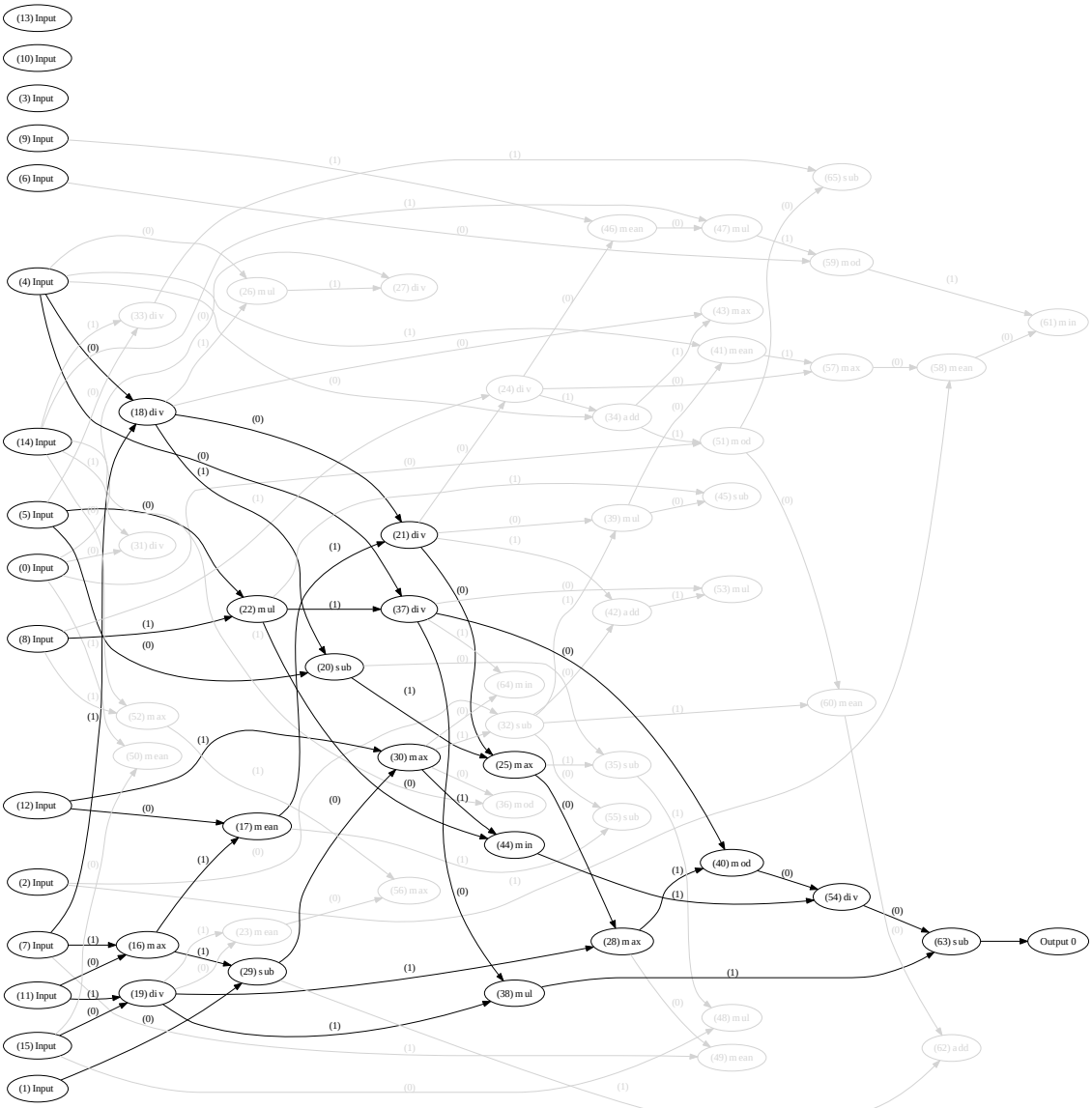

Figure 5.2 - CGP classification tree for the classification of Modafinil vs. control; example 1

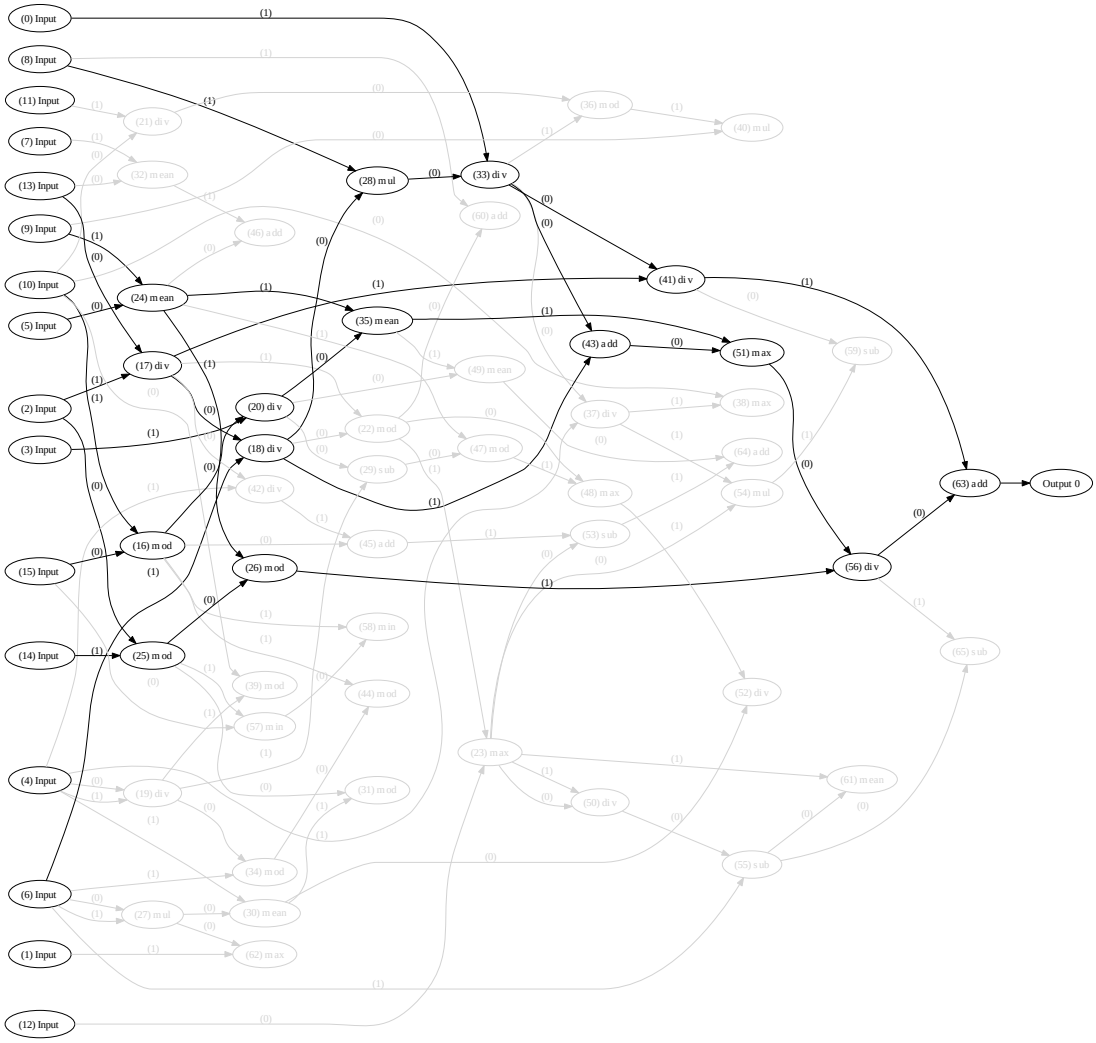

Figure 5.3 - CGP classification tree for the classification of Modafinil vs. control; example 2

To evaluate the correspondence between participant group and DCM features, a mixed 2×16 ANOVA between the participant group (Modafinil and control) and DCM features (16 inputs per participant) was conducted. A Greenhouse-Geisser correction was used as the model violated sphericity. The ANOVA revealed a significant main effect of DCM features $\left(F(4.39, 210.64) = 228.90, MSE = 2.71, p < .001, \eta_p^2 = .83\right)$. There was no significant interaction effect between participant group and DCM features $\left(F(4.39, 210.64) = 0.83, MSE = 0.01, p = .516, \eta_p^2 = .02\right)$ and no significant main effect of group $\left(F(1, 48) = 1.12, MSE = 0.00, p = .294, \eta_p^2 = .02\right)$. This main effect of features represents the key finding as it indicates that the features in general are essential, whereas information on participant group *per se* is not.

5.3.3. *k*-Fold Cross-Validation

In order to evaluate the performance of the classifier, *k*-fold CV was conducted on all the different combinations of inputs for both DCM and timeseries values.

5.3.3.1. Cross-Validation for RCGP for Timeseries

The inputs were divided into *10* folds with 80% of the data used for training, 10% for validation, and 10% for test. After the artificial data samples were synthesised for the minor class in the training set, CV was repeated for 10 runs and the results were averaged, as shown in Table 5.4. Findings indicated that the Modafinil group were successfully classified from the control group with a maximum accuracy of 63.13% using RCGP in CV (minimum accuracy: 51.67%).

Table 5.4 - Cross-validation results for the timeseries values

		Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
Classification results for each DMN region				
PCC	RCGP	58.74 (5.37)	70.36 (0.83)	63.13 (2.86)
mPFC	RCGP	57.01 (5.13)	67.26 (3.52)	58.69 (3.99)
RIPC	RCGP	55.14 (4.91)	68.51 (4.04)	58.09 (5.80)
LIPC	RCGP	56.97 (4.41)	70.05 (2.41)	58.09 (4.31)
Classification results for all the DMN regions (4 inputs)				
	RCGP	44.63 (5.99)	71.74 (10.16)	51.67 (8.30)
Classification results for all the DMN regions (1 input)				
	RCGP	60.21 (3.59)	69.86 (2.53)	60.73 (3.96)

5.3.3.2. Cross-Validation for CGP for DCM

The inputs were divided into *10* folds with 80% of the data used for training, 10% for validation, and 10% for test. After the artificial data samples were synthesised for the minor class in the training set, CV was repeated for 10 runs and the results were averaged, as shown in Table 5.5. Findings revealed that the Modafinil group were successfully classified from the control group with 59.55% accuracy using CGP in CV.

Table 5.5 - Cross-validation results for DCM values

	Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
CGP	62.55 (2.87)	76.61 (2.90)	59.55 (6.53)

5.4. Discussion

This research develops methods for differentiating clinical groups using EAs on rs-fMRI data, based on a controlled clinical experiment. Specifically, this experiment examined the monitoring of participants administered Modafinil (versus a control group) using rs-fMRI data. These findings build on previous research exploring the application of EAs for monitoring PD patients following treatment with Levodopa (LID-monitor) [427].

A distinctive element of this research involved the use of DCM analysis for classification. CGP classification method was used for timeseries analyses and DCM analyses and these findings were validated with two other widely used classification methods (ANN and SVM). Findings were further validated using *k*-fold CV technique. Given that the data was highly imbalanced, ADASYN was used to balance the data before performing CV, providing a more equal class distribution within the training set (i.e., balanced numbers of Modafinil and control participants). Findings revealed a maximum classification accuracy of 75%. An important finding was that there was almost no difference in the classification accuracies between timeseries and DCM data. Moreover, EAs, specifically CGP, have not previously been used for classification of brain imaging data; hence, this research provided a novel application of CGP (on rs-fMRI data). CGP provided equivalent performance accuracy when compared with ANN and SVM classification methods. A relevant question for future research is whether CGP is relevant for the classification of PD patients (rather than healthy participants, as per the current research) administered Modafinil relative to non-medicated PD patients who are experiencing fatigue.

A key aspect of this research involved the application of a dynamical method of classification (RCGP). CGP has been previously used in the classification of biomedical data but not brain imaging data. This research applied CGP to rs-fMRI data, a timely approach given that fMRI is a non-invasive method that generates images with high spatial resolution and good temporal resolution and it is widely used in medical facilities [439]–[442] (e.g., for diagnosis).

fMRI data is dynamic as biomedical data from the nervous system are complex, nonlinear and nonstationary. Nevertheless, research on the classification of biomedical timeseries data typically uses static classifiers (e.g., SVM [443] and feed-forward neural networks [444]; see also [445], [446]). Other research has compared dynamical to static classifiers for biomedical data, revealing better discrimination and increased diversity in dynamical classifiers [409]. The current research used two static classifiers (SVM and ANN) to validate the findings from a dynamical classifier (RCGP) for the classification of timeseries data, which revealed comparable findings between all three methods. Whilst

dynamical classifiers might theoretically better represent the nature of biomedical imaging data, future research is needed to further examine this question.

This research used a novel approach, applying classification to DCM data. Findings revealed that DCM data for classification provided comparable accuracy across all three classifiers, relative to timeseries data, which is surprising as other research has revealed that classification often provides better accuracies with raw unprocessed data. Timeseries data is raw fMRI data and includes over 100 features. In contract, DCM data is processed, representing the effective connectivity (the causal effect) of one neuronal region on another and, in this case, contained only 16 features. These findings underscore the relevance of DCM data for classification, even though this data is processed and contains much less features relative to timeseries data.

Following *k*-fold CV, classification accuracy for timeseries values decreased to 52-63% and for DCM values accuracy was reduced to 60%; these findings held across all participant groups. This is due to the fact that this research used heavily class-imbalanced data containing 39 Modafinil participants and 111 control participants. Yet, standard classification methods typically assume balanced class distributions. Imbalanced data significantly reduces classification accuracy [447] given that the classifier cannot be trained efficiently to distinguish the differences between features in the two classes. A widely used solution involves modifying the data to obtain a sample with balanced class distributions, which often increases the overall classification accuracy compared to the original imbalanced sample [385]–[387]. Nevertheless, this solution is not perfect given that the balanced data is only used in the training set, which compromises the accuracy in the validation and test sets as the classifier that was trained for the CV is a different classifier with a different accuracy level than that used in the classification part. The current research used ADASYN to balance the training set. The fact that findings revealed accuracies of approximately up to 75% in the classification of Modafinil timeseries and DCM data speaks to the robustness of CGP as a classification method, even for highly imbalanced data as in this research.

5.5. Conclusion

To conclude, this research explored the classification of participants administered Modafinil (versus a control group) using novel data (DCM values) and a novel classification method (CGP). This demonstrates the power of the technique to monitor PD patients in response to the medication they are receiving as their medication is adjusted. These results add to previous literature examining EAs for monitoring PD patients following Levodopa treatment [427]. Classification accuracy for DCM analyses was compared to that of timeseries analyses. Moreover, two other classification methods (ANN and SVM) were used to validate the findings, including employing *k*-fold CV. Findings revealed a maximum accuracy of 75% for CGP. Classification accuracy was equivalent for DCM analyses and timeseries analyses and, further, accuracy did not differ by method (CGP, ANN, or SVM). These findings underscore the relevance of CGP as a novel classification method for dynamic brain imaging data, for both processed data (as per DCM values) and raw data (i.e., timeseries). Hence, the methods developed are potentially relevant for drug treatment monitoring and for differentiating between clinical groups, for instance, the diagnosis of patients relative to healthy controls. Applying these methods developed to a clinical sample of PD patients and examining the transferability of these tools is the focus of the research presented in the subsequent chapter.

Chapter 6. Early Stage Diagnosis of Parkinson's Disease: CGP Classification of rs-fMRI Data

As previously discussed in Chapter 2, currently, there is poor differential diagnosis of NDDs with high rates of misdiagnosis and low test-retest reliability [77], [104]; indeed, PD has rates of misdiagnosis of 15-26%. Hence, this research examines a key question: Can early stage PD (before motor symptoms are apparent, i.e., prodromal PD) be diagnosed using EAs on rs-fMRI data? Early diagnosis of PD is fundamental in providing patients with palliative care during the early phases before motor symptoms are present, enabling effective disease management and maintaining patient quality of life. Moreover, once a neuroprotective drug to treat PD is developed, the early diagnosis of PD would have even greater clinical implications [448].

This chapter explores an automatic and non-invasive method of confirming the diagnosis of PD, specifically, examining the classification of participants diagnosed with PD, prodromal PD participants, and healthy age-matched controls using rs-fMRI data. This research involves an analysis of rs-fMRI data taken from the Parkinson's Progression Markers Initiative database (PPMI; http://www.ppmi-info.org/data). The PPMI [158] is a landmark, large-scale, comprehensive, observational, international, and multi-centre study that recruits *de novo* (early-untreated) PD patients, prodromal PD patients, and age-matched healthy participants (among other participant groups) to identify PD progression biomarkers. This study applies EAs, specifically CGP and RCGP, for the classification of rs-fMRI in PD using DCM and timeseries analyses. The timeseries values and DCM values from the rs-fMRI data are subjected to supervised classification and the findings are validated with two other commonly used classification methods (ANN and SVM) as well as employing *k*-fold CV.

A key aim of this chapter is to identify the applicability of CGP and RCGP classification for both timeseries and DCM analyses regarding the analysis of PD data. CGP and RCGP have not previously been used in the classification of brain imaging data. This study examines an additional novel question: is DCM analysis useful for classification for PD data? Previous research has not explored the applicability of DCM values in classification and, to date, little research has applied DCM to PD data [15]–[20]. Hence, by doing both, this research develops automatic procedures for identifying PD brain imaging preclinical biomarkers, which can be used for aiding/confirming early PD diagnosis.

Furthermore, a typical difficulty when conducting medical research involves recruiting equal sample sizes of patients and healthy controls, often resulting in class-imbalanced data (e.g., unequal groups of controls versus patients). The PPMI database is heavily class-imbalanced, with many more PD patients relative to prodromal PD and control participants. Hence, this chapter explored the applicability of classification methods to class-imbalanced data, with key implications for the transferability of medical research based on limited and imbalanced sample sizes.

6.1. Method

6.1.1. Participants

PPMI is a longitudinal study where participants underwent a comprehensive longitudinal follow-up schedule of clinical, imaging and biospecimen assessments. There were eight healthy controls (all male, mean age = 68, *SD* = 3.16), 18 prodromal PD patients (13 male, mean age = 68, *SD* = 4.03), and 102 early PD patients (71 male, mean age = 63, *SD* = 7.83). The overall age range was 50-75 years. 3 Tesla rs-fMRI, dopamine transporter (DAT) imaging, and MRI scans were acquired for all participants.

PD participants were recruited at disease threshold (diagnosis within two years and untreated for PD) and were required to have an asymmetric resting tremor or asymmetric bradykinesia or two of bradykinesia (resting tremor and rigidity). DAT deficit was acquired for PD participants. Healthy participants had no significant neurologic dysfunction, no first degree family member with PD, and they obtained a MoCA > 26. The study was approved by the institutional review board of all participating sites. Written informed consent was obtained from all participants.

6.1.2. rs-fMRI Acquisition

A standardised MRI protocol included acquisition of whole-brain structural and functional scans on 3 Tesla Siemens Trio Tim MR system (for more information see http://www.ppmi-info.org/). 3D T1 structural images were acquired in a sagittal orientation using a MPRAGE GRAPPA protocol with Repetition Time (TR) 2300 ms, Echo Time

(TE) 2.98 ms, Field of View (FoV) 256 mm, Flip Angle (FA) 9° and 1 mm^3 isotropic voxel. For each participant, 212 BOLD echo-planar rs-fMRI images (40 slices each, ascending direction) were acquired during an 8 min, 29 s scanning session (acquisition parameters: TR = 2400 ms, TE = 25 ms, FoV = 222 mm, FA = 80° and 3.3 mm^3 isotropic voxels). Participants were instructed to rest quietly, keeping their eyes open, and were asked not to fall asleep.

6.1.3. Imaging Data Analysis

6.1.3.1. Preprocessing

The preprocessing steps were the same as described in Chapter 5, excepting that preprocessing included DICOM to 3D NIFTI conversion.

6.1.3.2. Processing

Timeseries and DCM analyses were obtained as described in Chapter 5.

6.1.4. Cartesian Genetic Programming (CGP)

6.1.4.1. Classification

This research applied the same classification method as described in Chapter 5, excepting that there was one output for the classifier (class 1 for one group, class 0 for the other one).

6.1.4.1.1. Classification of Timeseries and DCM

This research applied same classification method using timeseries and DCM values as data inputs as described in Chapter 5, excepting that for timeseries values the number of data inputs is 210 (corresponding to the number of image slices). When inputting the timeseries values for the four DMN regions of interest to the classifier in one vector, the number of features is 840 per participant. Classification was completed for each of the three participant groups (PD, prodromal PD, and control) against the remaining two.

6.1.5. Adaptive Synthetic Sampling (ADASYN)

As previously mentioned, the control group contained eight participants, the prodromal PD group contained 18 patients, and the PD group contained 102 patients. Therefore, the data was highly imbalanced. Following the same steps outlined in Chapter 5, ADASYN was applied to make the data balanced for training the classifier.

6.1.6. *k*-Fold Cross-Validation

This research applied the same *k*-fold CV method as described in Chapter 5, excepting that for both PD participants and prodromal PD participants relative to control participants, CV was completed using *9* folds (due to the small number of samples in the control group).

6.2. Results

This study examined classification of 102 PD participants, 18 prodromal PD participants, and eight healthy age-matched controls. The analysis (classification) focused on organising features to be used as inputs to the classifier in CGP and also in RCGP and was implemented using the CGP Library. To validate these findings, the analysis/classification was additionally completed using ANN in MATLAB (Pattern Recognition and Classification Application – nprtool with 10 hidden layers) and SVM in MATLAB.

6.2.1. Classification of Timeseries

Initially, the timeseries values for each region were used as inputs to the classifier individually. Therefore, the data was classified with 210 features from each region per participant. The same procedure was repeated separately for each DMN region (PCC, mPFC, RIPC, and LIPC). Then, the timeseries values were used as inputs to the classifier to classify the data with four columns (relating to the four DMN regions) of 210 features per participant. Finally, the timeseries values together in one column, were used as inputs to the classifier to classify the data with 840 features in one vector for each participant. The results after 10 runs for each combination were averaged and are presented for each category in Table 6.1, Table 6.2, and Table 6.3. The classification was also completed

using ANN and SVM. For SVM, only the training and test sets were considered for classification.

Table 6.1 - Classification results for the timeseries values (PD vs. controls)

		Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
Classification results for each DMN region				
PCC	RCGP	92.06 (2.67)	91.60 (2.91)	91.95 (2.66)
	ANN	92.73 (0.07)	92.72 (0.32)	92.74 (0.31)
	SVM	92.73 (0.00)	NA	92.73 (0.00)
mPFC	RCGP	92.03 (2.68)	91.73 (3.02)	92.01 (2.88)
	ANN	92.69 (0.06)	92.85 (0.41)	92.76 (0.40)
	SVM	92.73 (0.00)	NA	92.73 (0.00)
RIPC	RCGP	91.13 (3.97)	92.41 (3.64)	91.46 (5.15)
	ANN	92.68 (0.11)	92.85 (0.50)	92.90 (0.39)
	SVM	92.73 (0.00)	NA	92.73 (0.00)
LIPC	RCGP	91.66 (2.55)	91.75 (2.76)	91.68 (3.46)
	ANN	92.69 (0.10)	92.89 (0.39)	92.61 (0.38)
	SVM	92.73 (0.00)	NA	92.73 (0.00)
Classification results for all the DMN regions (4 inputs)				
	RCGP	91.89 (2.77)	92.03 (3.34)	91.79 (3.45)
	ANN	92.66 (0.13)	92.96 (0.45)	92.74 (0.46)
	SVM	92.78 (0.01)	NA	92.73 (0.00)
Classification results for all the DMN regions (1 input)				
	RCGP	89.95 (8.53)	91.84 (2.77)	92.09 (2.68)
	ANN	92.75 (0.08)	92.74 (0.25)	92.59 (0.07)
	SVM	92.73 (0.00)	NA	92.73 (0.00)

Table 6.2 - Classification results for the timeseries values (PD vs. prodromal PD)

		Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
Classification results for each DMN region				
PCC	RCGP	85.64 (2.03)	85.94 (2.23)	85.87 (2.24)
	ANN	85.02 (0.18)	84.80 (0.82)	85.02 (0.48)
	SVM	85.00 (0.00)	NA	85.00 (0.00)
mPFC	RCGP	85.10 (1.67)	85.97 (2.19)	85.49 (2.35)
	ANN	85.00 (0.13)	84.66 (0.33)	85.25 (0.69)
	SVM	85.00 (0.00)	NA	85.00 (0.00)
RIPC	RCGP	84.93 (1.22)	87.54 (1.12)	80.64 (11.25)
	ANN	84.97 (0.07)	85.10 (0.28)	84.93 (0.21)
	SVM	85.00 (0.00)	NA	85.00 (0.00)
LIPC	RCGP	75.00 (22.97)	85.85 (2.36)	85.56 (3.15)
	ANN	84.99 (0.15)	84.91 (0.62)	85.05 (0.47)
	SVM	85.00 (0.00)	NA	85.00 (0.00)
Classification results for all the DMN regions (4 inputs)				
	RCGP	84.69 (2.87)	87.54 (4.01)	85.02 (3.51)
	ANN	85.06 (0.15)	84.74 (0.64)	84.95 (0.46)
	SVM	85.06 (0.02)	NA	85.00 (0.01)
Classification results for all the DMN regions (1 input)				
	RCGP	85.47 (1.86)	86.27 (2.78)	85.71 (2.43)
	ANN	84.99 (0.03)	85.12 (0.34)	84.90 (0.26)
	SVM	85.00 (0.00)	NA	85.00 (0.00)

Table 6.3 - Classification results for the timeseries values (prodromal PD vs. controls)

		Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
Classification results for each DMN region				
PCC	RCGP	69.46 (14.76)	78.74 (7.12)	62.41 (18.16)
	ANN	69.05 (0.42)	69.16 (2.09)	68.53 (1.44)
	SVM	69.23 (0.01)	NA	69.22 (0.02)
mPFC	RCGP	66.82 (15.66)	76.32 (8.82)	66.49 (17.38)
	ANN	69.72 (0.79)	69.53 (1.78)	69.32 (1.75)
	SVM	69.87 (0.13)	NA	69.73 (0.26)
RIPC	RCGP	55.93 (25.99)	83.26 (10.26)	55.15 (21.75)
	ANN	69.42 (0.32)	69.53 (1.62)	70.11 (2.02)
	SVM	69.70 (0.14)	NA	69.23 (0.25)
LIPC	RCGP	67.02 (14.68)	78.66 (16.73)	70.62 (13.40)
	ANN	69.18 (0.50)	69.17 (1.70)	68.92 (1.56)
	SVM	69.23 (0.00)	NA	69.23 (0.00)
Classification results for all the DMN regions (4 inputs)				
	RCGP	62.45 (12.09)	72.86 (10.24)	62.46 (22.83)
	ANN	69.35 (0.74)	68.96 (1.83)	69.75 (1.16)
	SVM	74.32 (0.27)	NA	70.76 (0.52)
Classification results for all the DMN regions (1 input)				
	RCGP	64.15 (15.10)	70.04 (12.18)	54.64 (30.34)
	ANN	69.27 (0.15)	68.77 (0.74)	69.35 (0.66)
	SVM	69.25 (0.01)	NA	69.21 (0.06)

As depicted in Table 6.1, findings revealed that PD patients were successfully classified from healthy controls with a maximum of 92.09% accuracy using RCGP (minimum accuracy: 91.46%). The results from the other two classification techniques (ANN and SVM) validated this finding as they were very similar: between 92.59% and 92.90% for ANN and SVM in all the different combinations of inputs.

Table 6.2 illustrates that PD patients were successfully classified from prodromal PD patients with a maximum accuracy of 85.87% using RCGP (minimum accuracy: 80.64%). The results from the other two classification techniques (ANN and SVM) validated this finding as they were very similar: between 84.90% and 85.25% for ANN and SVM in all the different combinations of inputs.

The results revealed, as represented in Table 6.3, that prodromal PD patients were successfully classified from healthy controls with a maximum accuracy of 70.62% using RCGP (minimum accuracy: 54.64%). The results from the other two classification techniques (ANN and SVM) validated this finding as they were very similar: between 68.53% and 70.76% for ANN and SVM in all the different combinations of inputs.

Unlike for the DCM analyses, mixed ANOVAs were not conducted to evaluate the correspondence between participant group and timeseries features given that there were 840 features per participant, which would not be interpretable.

6.2.2. Classification of Dynamic Causal Modelling (DCM)

Classification using CGP implemented in CGP Library was executed with 16 inputs (all the DCM values sorted by region and presented as only one vector per participant) and 1 output (class 1 for one group, class 0 for another one). The results were then averaged over 10 runs and are presented for each category in Table 6.4. The classification was also done using ANN and SVM. For SVM, only the training and test sets were considered for classification.

Table 6.4 - Classification results for DCM values

	Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
PD vs. controls			
CGP	91.00 (5.56)	93.23 (3.83)	90.87 (4.41)
ANN	93.42 (1.97)	93.52 (7.57)	91.18 (8.86)
SVM	92.86 (0.68)	NA	92.86 (0.68)
PD vs. prodromal PD			
CGP	80.01 (7.79)	90.82 (3.85)	79.12 (12.36)
ANN	86.06 (2.47)	83.32 (8.29)	83.32 (7.39)
SVM	85.36 (0.58)	NA	85.36 (0.58)
Prodromal PD vs. controls			
CGP	74.11 (27.42)	90.11 (12.31)	75.21 (23.01)
ANN	78.34 (16.44)	70.00 (22.97)	65.00 (26.87)
SVM	68.42 (0.00)	NA	68.42 (0.00)

As illustrated in Table 6.4, findings revealed that PD patients were successfully classified from healthy controls with 90.87% accuracy using CGP. The results from the other two classification techniques (ANN and SVM) validated this finding as they were very similar: 91.18% for ANN and 92.86% for SVM. The results also revealed that PD patients were classified from prodromal PD patients with 79.12% accuracy using CGP. The results from ANN and SVM again validate these findings with 83.32% and 85.36% accuracy rates, respectively. Finally, the results indicated that prodromal PD patients were classified from healthy controls with 75.21% accuracy using CGP. The results from ANN and SVM also validate these findings with 65.00% and 68.42% accuracy rates, respectively.

Examples of the CGP classification trees/graphs can be seen in Figure 6.1 and Figure 6.2 for PD versus controls, Figure 6.3 and Figure 6.4 for PD versus prodromal PD, and Figure 6.5 and Figure 6.6 for prodromal PD versus controls. This represents a fundamental benefit of CGP, generating a *white box solution* that enhances interpretability of the classification network (not always possible with ANN and SVM classification methods). Despite the complexity inherent in these networks, they can provide crucial relevant information. For instance, in Figure 6.1 and Figure 6.2, approximately half of the inputs have been used to arrive at the final classification. These two networks are rather simpler than those represented in Figure 6.3, Figure 6.4, Figure 6.5, and Figure 6.6.

Figure 6.1 - CGP classification tree for the classification of PD vs. controls; example 1

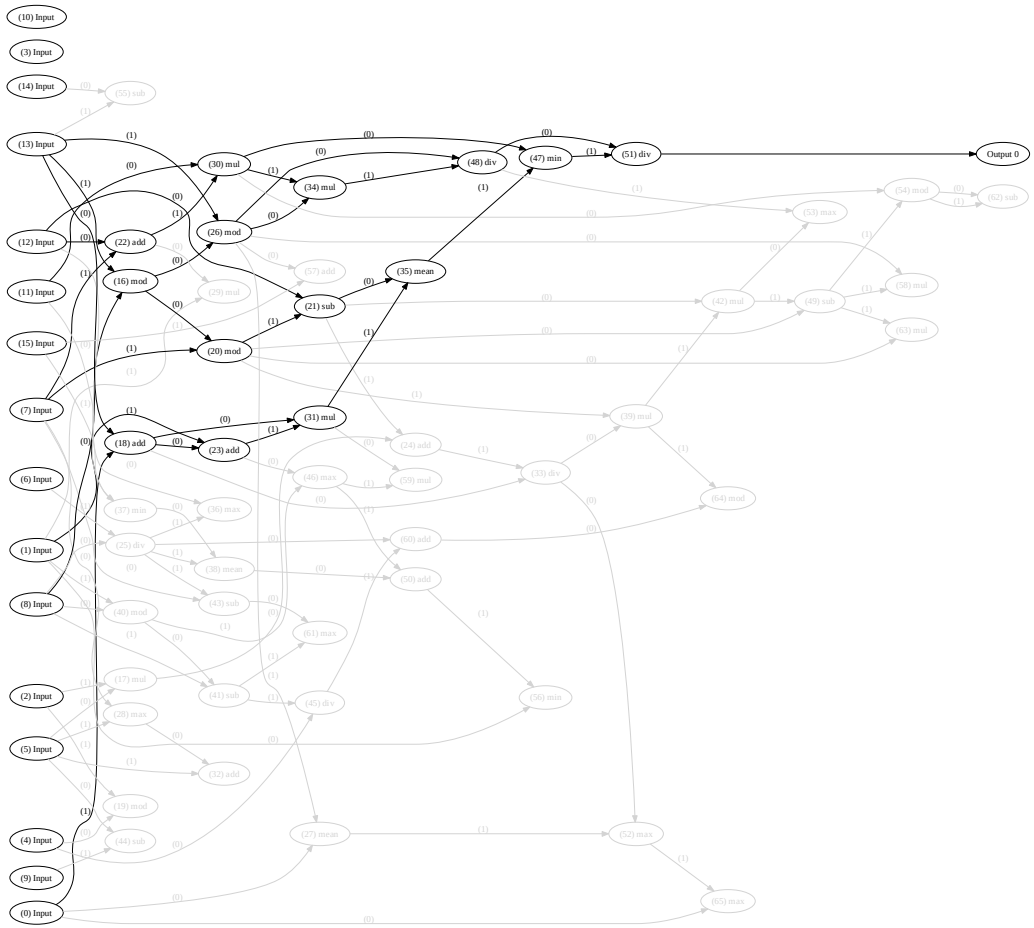

Figure 6.2 - CGP classification tree for the classification of PD vs. controls; example 2

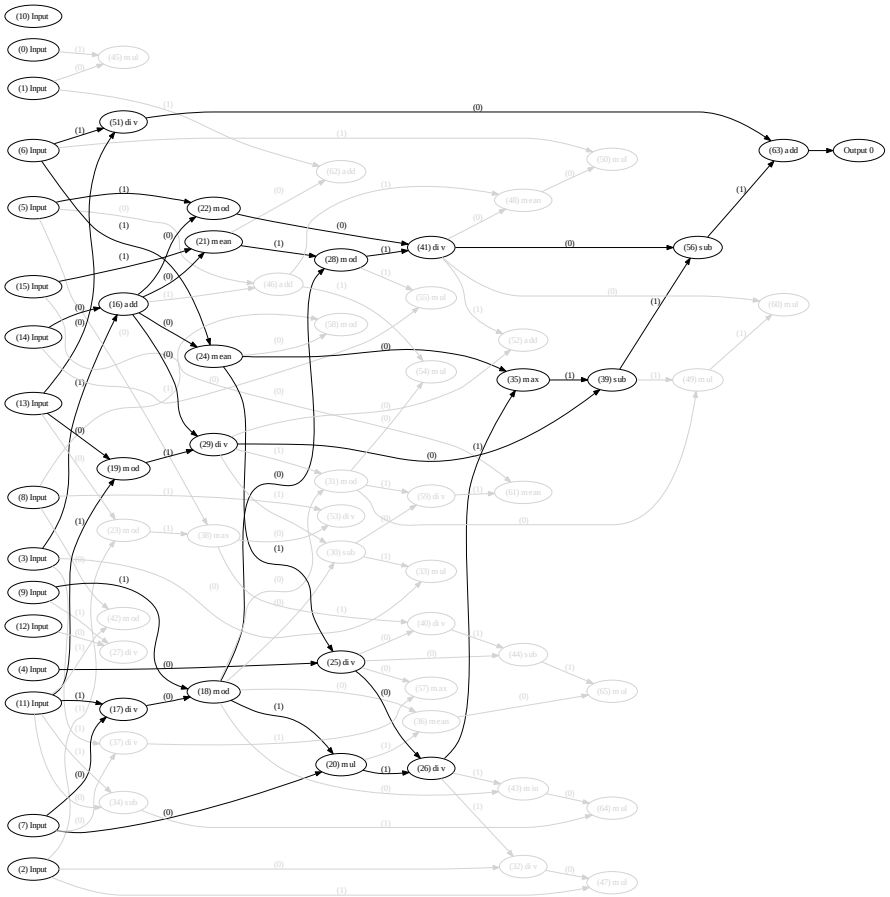

Figure 6.3 - CGP classification tree for the classification of PD vs. prodromal PD; example 1

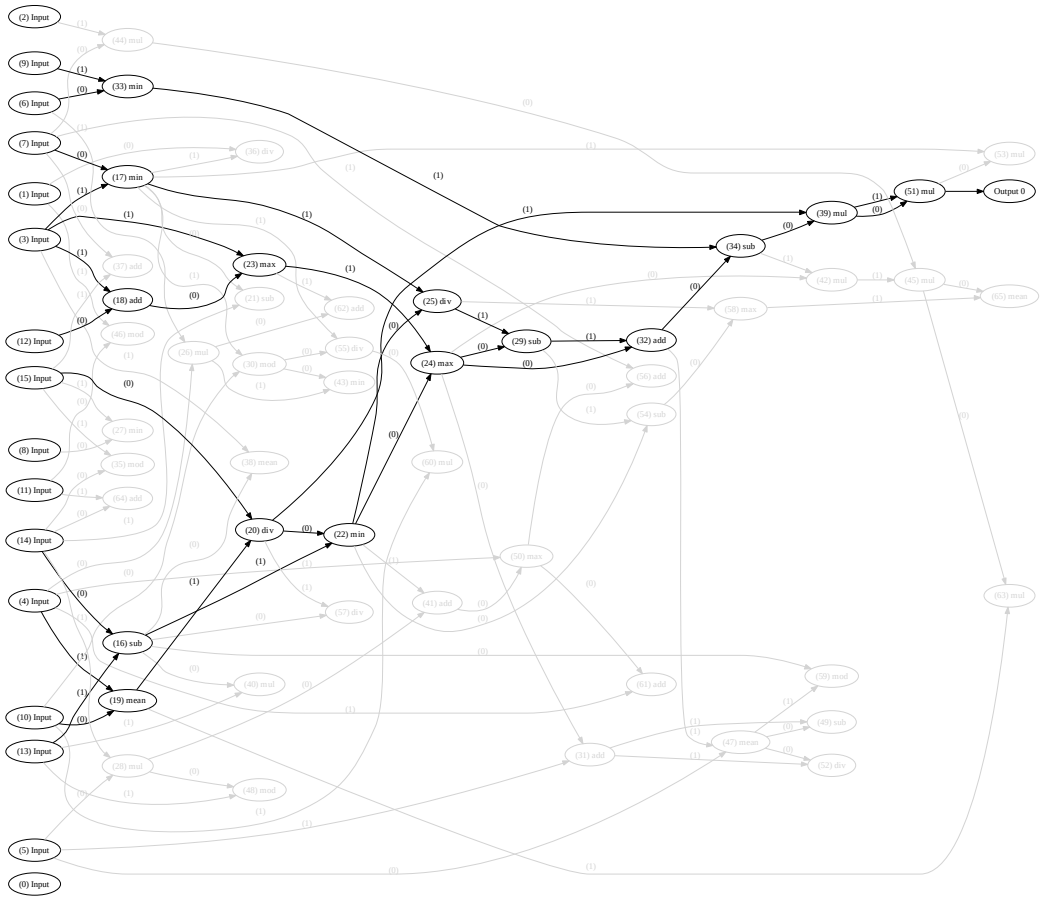

Figure 6.4 - CGP classification tree for the classification of PD vs. prodromal PD; example 2

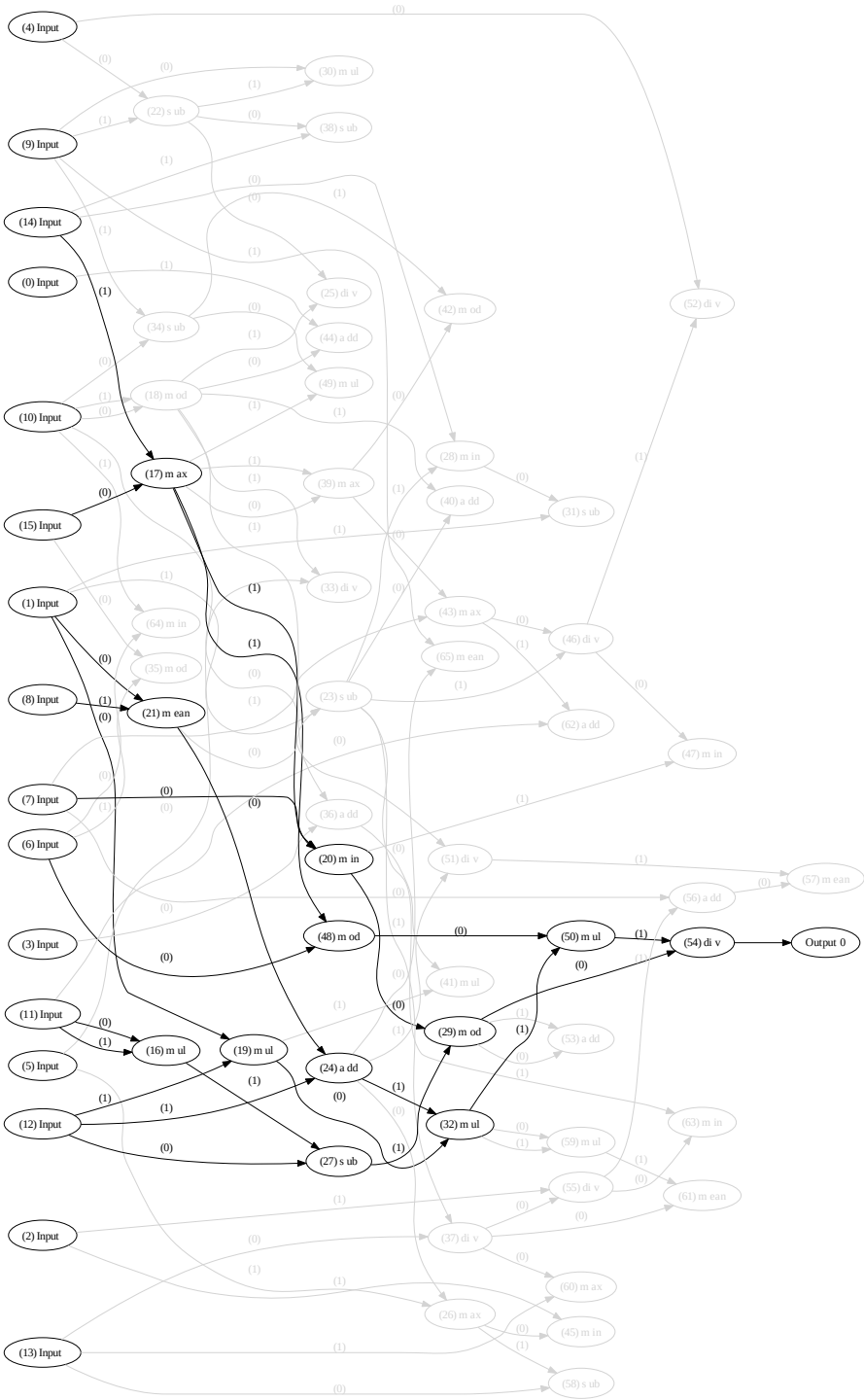

Figure 6.5 - CGP classification tree for the classification of prodromal PD vs. controls; example 1

Figure 6.6 - CGP classification tree for the classification of prodromal PD vs. controls; example 2

To evaluate the correspondence between participant group and DCM features, three mixed ANOVAs were conducted. In each ANOVA only two groups (e.g., PD patients and control) were considered rather than all three groups (PD patients, prodromal PD patients, and healthy control) at once, to mimic the classification method used in which, again, only two groups were considered at any point. For all three ANOVAs, a Greenhouse-Geisser correction was used as the models violated sphericity.

Firstly, to examine the correspondence between PD and control participants and DCM features, these were subjected to a mixed 2×16 ANOVA between the participant group (PD patients and healthy control) and DCM features (16 inputs per participant). The ANOVA revealed a significant main effect of DCM features $\left(F(5.47, 590.68) = 64.25, MSE = 2.17, p < .001, \eta_p^2 = .37\right)$. There was no significant interaction effect between participant group and DCM features $\left(F(5.47, 590.68) = 2.08, MSE = 0.07, p = .060, \eta_p^2 = .02\right)$ and no significant main effect of group $\left(F(1, 108) = 0.02, MSE = 0.00, p = .879, \eta_p^2 = .00\right)$.

Secondly, to explore the correspondence between PD patients and prodromal PD participants and DCM features, these were subjected to a mixed 2×16 ANOVA between the participant group (PD patients and prodromal PD) and DCM features (16 inputs per participant). The ANOVA revealed a significant main effect of DCM features $\left(F(5.66, 667.89) = 112.36, MSE = 3.68, p < .001, \eta_p^2 = .49\right)$. There was no significant interaction effect between participant group and DCM features $\left(F(5.66, 667.89) = 1.22, MSE = 0.04, p = .296, \eta_p^2 = .01\right)$ and no significant main effect of group $\left(F(1, 118) = 0.17, MSE = 0.00, p = .680, \eta_p^2 = .00\right)$.

Finally, to investigate the correspondence between prodromal PD and control participants and DCM features, these were subjected to a mixed 2×16 ANOVA between the participant group (prodromal PD and control) and DCM features (16 inputs per participant). The ANOVA revealed a significant main effect of DCM features $\left(F(5.15, 123.58) = 44.51, MSE = 1.71, p < .001, \eta_p^2 = .65\right)$. There was no significant

interaction effect between participant group and DCM features $\left(F(5.15, 123.58) = 1.16, MSE = 0.05, p = .331, \eta_p^2 = .05\right)$ and no significant main effect of group $\left(F(1, 24) = 0.01, MSE = 0.00, p = .913, \eta_p^2 = .00\right)$.

Over the three ANOVAs, findings consistently revealed a main effect of DCM features with no significant main effect of participant group and no significant interaction effect. This represents the key finding as it indicates that the features in general are essential, whereas information on participant group *per se* is not.

6.2.3. *k*-Fold Cross-Validation

To evaluate the performance of the classifier, *k*-fold CV was conducted on all the different combinations of inputs for both DCM and timeseries values.

6.2.3.1. Cross-Validation for RCGP for Timeseries

The inputs were divided into folds with 80% of the data used for training, 10% for validation, and 10% for test. After the artificial data samples were synthesised for the minor class in the training set (using ADASYN), CV was repeated for 10 runs and the results were averaged, as depicted in Table 6.5, Table 6.6, and Table 6.7. Findings revealed that PD patients were successfully classified from healthy controls with a maximum accuracy of 91.22% using RCGP in CV (minimum accuracy: 87.55%, see Table 6.5). PD patients were successfully classified from prodromal PD patients with a maximum accuracy of 82.99% using RCGP in CV (minimum accuracy: 79.54%, see Table 6.6). Prodromal PD patients were successfully classified from healthy controls with a maximum accuracy of 68.28% using RCGP in CV (minimum accuracy: 62.58%, see Table 6.7).

Table 6.5 - Cross-validation results for the timeseries values (PD vs. controls)

		Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
Classification results for each DMN region				
PCC	RCGP	88.95 (8.14)	92.09 (0.45)	88.42 (8.03)
mPFC	RCGP	85.76 (8.49)	92.60 (1.13)	87.55 (6.40)
RIPC	RCGP	91.45 (1.40)	92.76 (1.48)	90.99 (1.50)
LIPC	RCGP	90.92 (1.70)	92.02 (0.48)	91.22 (0.51)
Classification results for all the DMN regions (4 inputs)				
	RCGP	91.54 (0.86)	92.49 (1.16)	90.48 (2.16)
Classification results for all the DMN regions (1 input)				
	RCGP	90.58 (2.61)	92.37 (0.99)	90.45 (1.62)

Table 6.6 - Cross-validation results for the timeseries values (PD vs. prodromal PD)

		Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
Classification results for each DMN region				
PCC	RCGP	82.91 (2.46)	85.20 (1.06)	81.74 (3.18)
mPFC	RCGP	82.44 (2.50)	84.53 (0.28)	82.56 (2.19)
RIPC	RCGP	81.96 (4.15)	84.97 (1.23)	79.54 (7.56)
LIPC	RCGP	82.81 (2.17)	85.19 (0.89)	82.59 (1.65)
Classification results for all the DMN regions (4 inputs)				
	RCGP	81.82 (6.60)	84.94 (0.85)	80.96 (8.07)
Classification results for all the DMN regions (1 input)				
	RCGP	82.68 (2.55)	85.18 (1.11)	82.99 (1.88)

Table 6.7 - Cross-validation results for the timeseries values (prodromal PD vs. controls)

		Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
Classification results for each DMN region				
PCC	RCGP	66.90 (4.33)	75.93 (5.80)	64.20 (7.01)
mPFC	RCGP	70.91 (7.12)	81.62 (6.70)	67.15 (10.77)
RIPC	RCGP	68.53 (7.74)	77.25 (4.18)	62.58 (11.13)
LIPC	RCGP	65.87 (10.02)	79.52 (5.36)	68.28 (10.15)
Classification results for all the DMN regions (4 inputs)				
	RCGP	66.71 (5.64)	79.78 (5.75)	63.53 (9.79)
Classification results for all the DMN regions (1 input)				
	RCGP	63.49 (4.82)	79.94 (6.26)	64.64 (12.77)

6.2.3.2. Cross-Validation for CGP for DCM

The inputs were divided into folds with 80% of the data used for training, 10% for validation, and 10% for test. After the artificial data samples were synthesised for the minor class in the training set (using ADASYN), CV was repeated for 10 runs and the results were averaged, as shown in Table 6.8.

Table 6.8 - Cross-validation results for DCM values

	Training % (*SD*)	Validation % (*SD*)	Test % (*SD*)
PD vs. controls			
CGP	77.37 (10.28)	85.63 (3.97)	76.45 (8.43)
PD vs. prodromal PD			
CGP	64.17 (9.56)	75.38 (6.36)	63.28 (6.01)
Prodromal PD vs. controls			
CGP	59.91 (6.42)	88.45 (9.58)	53.81 (13.85)

The results revealed that PD patients were successfully classified from the healthy controls with an accuracy of 76.45% using CGP in CV. PD patients were classified from the prodromal PD participants with 63.28% accuracy using CGP in CV. Finally, the findings indicated that prodromal participants were classified from healthy controls with 53.81% accuracy using CGP in CV.

6.3. Discussion

This research examined the classification of participants diagnosed with PD, prodromal PD participants, and healthy age-matched controls using rs-fMRI data. A key research question was: Can early stage PD (prodromal PD) be diagnosed using EAs on rs-fMRI data? Another distinctive element of this research involved (1) the application of EAs (CGP) as a classification method and (2) DCM analysis for classification. CGP classification was used for DCM analyses as well as timeseries analyses and the findings were validated with two other commonly used classification methods (ANN and SVM). A crucial benefit of EAs, specifically CGP, is that they provide a *white box solution* giving more information on the inputs used and enhanced knowledge concerning the final solution obtained in classification, compared to ANN and SVM. Findings were additionally validated using *k*-fold CV technique on the data.

The data was highly imbalanced, hence, ADASYN was used to balance the data before performing CV, resulting in an equal class distribution within the training set (i.e., balanced numbers of PD, prodromal PD, and control participants). Across timeseries and DCM analyses, findings revealed that PD versus control participants were classified with a maximum accuracy of 92%, PD versus prodromal PD participants with a maximum accuracy of 86%, and prodromal PD versus control participants with a maximum accuracy of 75%. These findings are notable as early diagnosis of PD (before motor symptoms) is in its infancy with high rates of misdiagnosis, impacting on patient treatment and quality of life. This finding embodies the most important research output from this thesis.

Findings further revealed almost no difference in the classification accuracies between timeseries and DCM data. In addition, findings revealed that CGP almost always provided equivalent performance accuracy when compared with ANN and SVM classification methods. Hence, these findings underscore the relevance of DCM analyses for classification and CGP as a novel classification tool for brain imaging data.

A novel question addressed by this research was: is DCM analysis useful for classification? Although DCM has recently become a widespread tool to model effective

connectivity in neuroimaging data [19], [449]–[452], no research has examined classification using DCM analysis for any type of fMRI data, including rs-fMRI. Moreover, few studies have conducted DCM analysis on PD data [15]–[20]. Hence, this research aimed to further existing knowledge on DCM as applied to PD. An interesting finding was little difference in the classification accuracies between timeseries and DCM analyses (1-6% difference in the maximum accuracies). The findings, firstly, speak to the relevance of DCM data for classification and, secondly, broaden the small literature on DCM as applied to PD data. DCM analysis was conducted on the DMN, yet, new research [453] has examined DCM across multiple brain networks and currently whole-brain DCM of fMRI data is an ongoing project by the SPM developers lab. As such, exploring DCM across specific regions influenced by PD (e.g., basal ganglia) is an exciting avenue for future research.

Classification of rs-fMRI data is a key theme examined by this research, specifically, this research explored the applicability of CGP and RCGP to classify rs-fMRI data. Previous research on the classification of fMRI data has largely focused on statistical modelling techniques (e.g., independent components analysis, multivariate pattern analysis; [342], [343], [454]). These latter approaches are mostly predictive (hypothesis driven), yet, machine learning methods (as per this research) are explanatory modelling techniques (mostly data driven). Machine learning approaches are advantageous given that (1) they can be validated, providing an unbiased estimate of accuracy [344]; (2) they are typically based on fewer assumptions (relative to statistical-based techniques); and (3) they can predict and learn concurrently from large datasets, whereas statistical modelling are typically used for small datasets to avoid overfitting [343], [454].

The approach used in this research was distinctive as EAs, specifically CGP, have not been used for classification of brain imaging data. CGP provided approximately equivalent performance accuracy when compared with ANN and SVM classification methods across all participant groups (PD, prodromal PD, and control). Further research can examine CGP as applied to task-based fMRI, which would enable researchers to explore a number of sophisticated questions including finger-tapping tasks (unlike rs-fMRI), which is used in behavioural measures and is currently a leading method for confirming PD diagnosis.

Following *k*-fold CV, classification accuracy for timeseries values decreased by only a few percentage. Yet, for DCM values, accuracy was reduced to 76% for the PD versus control participant group, although these findings held across all participant groups. The reduced accuracy is due to having used class-imbalanced data as, whilst there was a large sample of PD patients, sample sizes of the prodromal PD and control groups were comparably smaller. Moreover, the number of features included in the classification was significantly smaller in DCM values in comparison with the timeseries values.

Imbalanced data limits classification accuracy [383], [447], [455] as most classification methods assume balanced class distributions, creating two problems. Firstly, high class-imbalanced data results in limited training set sample size as classifiers often treat imbalanced data as though they have small sample sizes (since some classes may have small numbers). Secondly, learning focuses on classes with larger sample sizes, rather than focusing on discriminating between the sizes of the classes in the data and the characteristics of the actual data (rather than the synthesised data). Solutions to imbalanced data typically include changing/generating data to obtain equal class-distributions, which tends to improve classification accuracy [385]–[387]. Nonetheless, given that the synthetic balanced data is applied only to the training sets, classification accuracy is limited for the validation and test sets as before the classifier that was trained for the CV was a different classifier with a different accuracy level than that used in the classification part. This is an inevitable problem associated with imbalanced data, regardless of the number of folds involved in CV.

In this research, ADASYN was used, which is a method of generating synthetic data to create balanced class-distributions for the training set (i.e., generating balanced numbers of PD, prodromal PD, and control participants). ADASYN was applied to the three classification groups: PD versus control, PD versus prodromal PD, and prodromal PD versus control. Findings revealed maximum classification accuracies of 71-92% across all groups when using timeseries and DCM data. Despite the reduced accuracy following CV, results revealed that classification was reliable for both timeseries and DCM values. Hence, classification on DCM and timeseries values can be used as brain-imaging

biomarkers for PD and the current findings underscore the relevance of CGP as a classification method, even for highly imbalanced data.

This study involved an analysis of a large open dataset (PPMI) [158]. Contributing to open science (e.g., making data publically available) shares resources from one project to other research, such as the re-analysis of biomedical raw data to examine new predictions [456], as per the current research. Advantages of open science include data sharing, saving resources (e.g., money and time), exchanging expert viewpoints, reducing fraud and *p*-hacking, training purposes, amongst others. Here, in addition to the reproducibility and resource-sharing advantages provided by open data, this research was able to access a sample of 128 participants (excluding participants who did not fulfil the eligibility requirements). Most previous research using fMRI data acquire limited samples, typically 20 participants ([12], [407], [457], [458]; although 77 participants were tested in [13]). Having such a large sample size lends confidence to the generalisability of these findings.

6.4. Conclusion

To conclude, this chapter explored the classification of rs-fMRI data in participants diagnosed with PD, prodromal PD participants, and healthy age-matched controls using novel data (DCM analyses) and a novel classification method (CGP). Classification accuracy for DCM analyses was compared to that of timeseries analyses, and two other classification methods (ANN and SVM) were used to validate the findings, as well as employing *k*-fold CV. Across DCM and timeseries analyses, findings revealed maximum accuracies of 86% for classification of prodromal PD patients versus PD patients, and 92% for PD patients versus healthy controls using CGP. Early diagnosis of PD (before motor symptoms appear) is rife with challenges and current methods have high rates of misdiagnosis. This research further revealed a maximum accuracy of 75% in differentiating prodromal PD patients from healthy controls, with medical implications for disease management, patient treatment, and patient quality of life. Hence, this finding is the most important research output from this thesis. Furthermore, classification accuracy was approximately equivalent for (1) DCM analyses and timeseries analyses and (2) different classification methods (CGP, ANN, or SVM). CGP has distinct advantages

regarding the information linked to the solutions they generate. Therefore, these findings speak to the relevance of CGP as a novel classification tool for two types of brain imaging data (DCM and timeseries rs-fMRI analyses). This research developed automatic procedures for identifying PD brain imaging preclinical biomarkers, which are fundamental in improving accuracy of PD diagnosis methods. Furthermore, these findings highlight the applicability of DCM analyses for classification.

Chapter 7. Discussion

Within the last chapter of this thesis, a summary of the empirical findings is presented and the main themes arising from these are discussed. Subsequently, potential limitations and future directions are outlined.

7.1. Research Aims and Summary of Findings

The research presented in this thesis examined two principal research questions, which are both novel and exploratory with medical implications in relation to monitoring and diagnosis for PD:

Research question 1:

> *Can accurate monitoring of PD be achieved using EAs on rs-fMRI data for patients prescribed Modafinil?*

This research question is investigated in Chapter 5, using a controlled experiment to explore the classification of brain imaging data from participants administered a fatigue-reducing medication (Modafinil) typically prescribed to PD patients, versus a control group. Participants were classified with a maximum accuracy of 75% for CGP (classification of timeseries values). The methods developed are relevant for medication monitoring and were then applied to the classification of clinical data (PD diagnosis) in Chapter 6.

Research question 2:

> *Can early stage PD be diagnosed using EAs on rs-fMRI data?*

This research question is explored in Chapter 6, which married brain imaging and classification approaches to identify a biomarker for PD. Findings revealed maximum accuracy rates of 75% for prodromal PD patients versus healthy controls (classification of DCM values), 86% for PD patients versus prodromal PD patients (classification of timeseries values), and 92% for PD patients versus healthy controls (classification of timeseries values). Hence, this research develops automatic methods for classifying PD

brain imaging preclinical biomarkers. The fact that prodromal PD patients were differentiated from healthy controls with such a high accuracy rate is notable and represents the most valuable finding from this thesis, given that current methods of diagnosing PD have low reliability and low accuracy, particularly for early stages of PD in which motor symptoms are not present, as per prodromal PD.

In addition, Chapter 5 and Chapter 6 explored a novel research objective: to examine the applicability of CGP and RCGP classification for both timeseries and DCM analyses in rs-fMRI since CGP and RCGP have not previously been used in the classification of brain imaging data, nor has research examined the applicability of DCM in classification. Chapter 5 and Chapter 6 both examined CGP classification of rs-fMRI data, applying classification to DCM values and timeseries values. Findings were validated with ANN and SVM classification methods as well as employing *k*-fold CV. Across both chapters, findings revealed almost no difference in the classification accuracies between timeseries and DCM data. Furthermore, CGP provided equivalent performance accuracy relative to ANN and SVM. CGP has distinct advantages as it provides a *white box solution*, enabling researchers to have more information regarding the elements underpinning the classification process. Therefore, this research underscores the applicability of DCM analyses for classification and CGP as a novel classification tool for brain imaging data.

A typical limitation in medical research involves recruiting low numbers of patients, often resulting in class-imbalanced data (e.g., unequal numbers of controls versus patients). The research presented in this thesis investigated the applicability of classification methods to two datasets with heavily class-imbalanced data, which imitates the conditions prolific in medical research, with implications for the transferability of medical research using small and unequal sample sizes. The current research used ADASYN to balance the training set for the datasets used in both Chapter 5 and Chapter 6. Across both studies, findings revealed high classification accuracy rates, underscoring the robustness of CGP for the classification of medical data, even for heavily imbalanced data.

7.2. Key Novel Aspects

This thesis developed novel automatic methods for identifying PD brain imaging preclinical biomarkers using a data-driven exploratory approach, that can facilitate clinical diagnosis, monitoring and investigation of PD. A crucial novelty of this work involves the classification of prodromal PD. Prodromal PD patients were classified with maximum accuracy rates of 75% relative to healthy age-matched control participants. Prodromal PD diagnosis typically has poor accuracy rates and high levels of misdiagnosis with other Parkinsonian diseases. This finding is fundamental given that the early diagnosis of prodromal PD, occurring 5-20 years prior to motor symptoms, is essential for efficient disease management.

Further novelties examined in this thesis include the exploration of CGP and RCGP as applied to both timeseries and DCM analyses. Note that CGP, RCGP, and DCM have not previously been applied to the classification of fMRI data. Findings revealed that CGP yielded equivalent performance accuracy compared to ANN and SVM, yet, it provides core benefits including a *white box solution* that gives researchers more detailed information regarding the derived solutions. In addition, timeseries and DCM analyses provided equivalent classification accuracies. DCM is advantageous as it specifies the effective, causal, connectivity between brain regions, hence, yielding relevant information for clinicians examining PD deficits. As such, these novel findings indicate that DCM can be applied to classification and, furthermore, CGP is a new classification technique for brain imaging data. Finally, the exploration of these novel research points with class-imbalanced datasets (following the application of ADASYN to reduce the impact of class-imbalanced data on classification accuracy) facilitates the generalisation of the research findings to wider medical contexts, given that clinical research typically recruits unequal sized samples of patients and control participants.

7.3. Themes

In this section, three themes are discussed: (1) biomedical imaging data for classification with applications for diagnosis; (2) EAs for classification; and (3) DCM values (relative to timeseries values) as applied to classification.

7.3.1. Biomedical Imaging Data for Classification

Classification is a relevant tool for diagnosis. Current methods of PD diagnosis focus largely on movement data, and have misdiagnosis rates of up to 15-26% by general neurologists [7], [8]. Yet, early diagnosis is key to providing adequate treatment given that treatment is less efficient as PD progresses. Hence, this thesis explored the use of rs-fMRI data for classification, focusing on PD patients (versus prodromal PD participants, and healthy age-matched controls; Chapter 6) and participants administered Modafinil (versus a control group; Chapter 5). Findings revealed high classification accuracies: 92% for PD versus control participants, 86% for PD versus prodromal PD participants, 75% for prodromal PD versus control participants, and 75% for participants administered Modafinil versus a control group.

fMRI methods have multiple benefits, including being non-invasive (relative to other brain imaging techniques, e.g. PET or CAT) and time-efficient (collecting data for Chapter 5 took only 4 minutes per participant), though it is costly (relative to questionnaire methods, yet, substantially cheaper than many other brain imaging techniques). Timeseries values provide over 100 features automatically taken from the raw data, whereas movement data mostly requires processing before extracting features (typically 20-30 features). DCM values involve far less features (only 16 values) relative to timeseries values and is based on Bayesian techniques.

7.3.2. Evolutionary Algorithms (EAs)

Research examining classification of fMRI data have typically used statistical modelling techniques, for instance, independent components analysis and multivariate pattern

analysis [342], [343], [454]. These statistical approaches are often hypothesis driven, whereas EAs primarily use data-driven explanatory modelling techniques. The latter have numerous benefits including validation (providing an unbiased estimate of accuracy [344]), relying on less assumptions compared to statistical-based methods, and predicting and learning simultaneously from big datasets (to prevent overfitting statistical methods often use small datasets [343], [454]).

EAs have been successfully applied to big datasets, such as biomedical data. EAs engage in a solution search space, modifying candidate solutions within the data via a number of steps (labelled mutation and crossover). This process is iterative given that EAs start with a set of random solutions and, following each iteration, a new set is generated via the application of mutation and crossover to a subset (selected based on their fitness level) of the earlier set of random solutions.

GP is an EA that maximises executable expressions [459]. Research has revealed that CGP (as used in Chapter 5 and Chapter 6) is competitive relative to other GP methods [330] and has multiple advantages in comparison to other classification methods. These include the reuse of sub-functions, having no bloat [336], and being able to report the classification model in a mathematical formula/expression as well as a graphical diagram (see Figure 4.11 and Figure 4.12 for examples), which provides a deeper understanding regarding the underlying elements of the classification model, in comparison to other widely used classification methods including ANN and SVM. A further advantage of CGP is that the geometry of CGP can be constrained in order to form simpler networks, even at the cost of less accurate classifiers. Hence, EAs, including CGP, enable researchers to control the balance between developing a good classifier versus better understanding of the data inputs in the classification.

EAs have been applied to PD movement data, obtaining classification accuracy of ~95% [409]. Yet, CGP has not previously been applied to the classification of brain imaging data. The research presented examined the classification of brain imaging data using EAs. Specifically, Chapter 5 and Chapter 6 examined the applicability of CGP and RCGP to classify rs-fMRI data. Findings revealed successful classification of PD patients (Chapter

6, relative to prodromal PD and healthy controls) and of participants administered Modafinil (Chapter 5, relative to a control group) with high accuracy rates. Hence, these results suggest that CGP classification can be used as a tool for identification of brain imaging biomarkers associated with PD and with the ingestion of Modafinil for PD drug treatment monitoring.

EAs are commonly used techniques for generating classifiers [287], [460]–[462]. A novel element of the research presented in this thesis involves the use of dynamical classifiers (e.g., artificial biochemical networks). Biomedical signals (e.g., signals derived from the human nervous system) are dynamic as these are complex, nonlinear and nonstationary. A benefit of using RCGP is its dynamic nature. Previous work on the classification of biomedical timeseries data typically has used static classifiers (e.g., SVM [443] and feed-forward neural networks [444]; see also [445], [446] for examples of research that has examined the use of static classifiers to dynamical features). Nevertheless, fMRI data is dynamic (being complex, nonlinear and nonstationary). Hence, it is possible that a dynamic classifier is better suited to biomedical imaging data. Further research can examine the performance of dynamic classifiers relative to static classifiers.

7.3.3. Dynamic Causal Modelling (DCM)

DCM analysis is advantageous as it provides information on the causal links between different brain regions. Recently DCM has gained in popularity in the analysis of neuroimaging data [19], [449]–[452], yet, previous research has not investigated whether DCM values are useful in classification. Further, only limited research has examined DCM as applied to PD data [15]–[20]. The research presented in this thesis applied DCM values to classification for PD brain imaging data (Chapter 6) and for participants administered Modafinil (Chapter 5). Indeed, Chapter 5 and Chapter 6 adopted a novel approach by using DCM to examine the endogenous/intrinsic effective connectivity and the causal interactions across the DMN regions and, subsequently, the DCM values were subjected to classification. Findings across both Chapter 5 and Chapter 6 revealed that DCM values for classification provided comparable accuracy across all three classifiers, compared to timeseries values.

Other research indicates that classification accuracies are typically higher when using raw unprocessed data (as per timeseries values). The timeseries values used in this research included over 100 features per participant, whereas the DCM values included only 16 features per participant and consisted of processed data (comprising of the causal effect of one neuronal region on another). Findings from Chapter 5 and Chapter 6 speak to the applicability of DCM data for classification, regardless of this being processed with few features (compared to timeseries data). Moreover, Chapter 6 findings extend the limited research on DCM in relation to PD data.

7.4. Potential Limitations

In this section, two potential limitations are outlined: (1) the use of heavily imbalanced data with implications on classification accuracy; and (2) the classification of datasets including only brain imaging data, rather than including various sources of behavioural data, such as movement data.

7.4.1. CGP Classification Networks

A core advantage of CGP is that it can provide a *white box solution*, giving researchers detailed information on how the inputs have been used to train the classifiers in order to derive outputs. Hence, this research aimed to achieve more specific brain imaging biomarkers, considering that there are 16 features in a DCM connectivity matrix. This was supposed to be accomplished via detecting a clear relationship between the classification networks concerning input nodes (DCM connectivity values) and their use in training the resulting classifiers. However, despite exploring a large number of the networks generated, there were no apparent similarities regarding how the algorithms employed the inputs for classification of the brain imaging data. Moreover, the research examined modifying the classification parameters for CGP and RCGP, revealing no change in the findings despite differing classification accuracies. Future research can indeed explore the possibility of detecting classification networks that have greater similarity (i.e., greater similarity in *white box solutions*) in terms of using DCM connectivity values as inputs with the objective of obtaining high classification accuracy rates.

7.4.2. Class-Imbalanced Data

Imbalanced datasets are relatively common with regards to the classification of biomedical data [447], [463]. The minority class typically includes data from the key group (rather than the control group) due to recruitment and cost limitations [463] and this data may include extreme cases [464], further impeding accurate classification. The research presented included heavily class-imbalanced data: Chapter 5 used 39 participants administered Modafinil and 111 participants in the control participant group; Chapter 6 used a sample of eight healthy controls, 18 prodromal PD patients, and 102 early PD patients. Nevertheless, the strong results achieved in both Chapter 5 and Chapter 6 are notable given the imbalanced nature of the data and can be better generalised to many clinical datasets, which are typically class-imbalanced (given limitations in recruiting clinical samples).

Imbalanced data reduces classification accuracy [383], [447], [455] given that classification methods typically assume balanced class distributions [464]. Four implications of class-imbalanced data are as follows: (1) training set sample size is reduced since classifiers typically consider that imbalanced data has limited sample sizes; (2) learning is biased towards classes with bigger sample sizes; (3) classification rules predicting the majority class (e.g., for Chapter 5, the control participant group) tend to be specialised, hence, with low coverage, and so general rules (predicting the minority class) are prioritised; and (4) the classifier may incorrectly label small groups of minority class data as noise and reject these. Further, classification of the minority class can be worsened by the presence of genuine noise data, given the small sample size available, as there is less accurate data available to train the classifier [465].

Solutions commonly involve changing/generating data to obtain more balanced class-distributions, resulting in heightened classification accuracy [385]–[387]. In Chapter 5 and Chapter 6, ADASYN was used to generate synthetic data leading to balanced class-distributions for the training set (e.g., for Chapter 5 this involved generating balanced numbers of participants in the Modafinil and control participant groups in the training set). A limitation of generating synthetic balanced data is that it is applied only to the training

sets. The classifier trained for CV is a different classifier with a different accuracy level than that used in the classification part. Therefore, classification accuracy for the validation and test sets is typically reduced. In Chapter 5 and Chapter 6, classification accuracy was reduced following CV, yet, findings still underscored the relevance of CGP as a robust classification method, including for class-imbalanced data.

Imbalanced data remains an important limitation that is very common in medical datasets, as used in the research presented. Chapter 6 used open data acquired from the PPMI dataset. This dataset is a growing source of data; hence, the class-imbalanced limitation can be addressed in time, considering that every year a large amount of data is collected globally from approved centres.

7.4.3. fMRI Data

Chapter 5 and Chapter 6 examined CGP classification of rs-fMRI. The data was acquired from open sources, hence, this limited the scope of the data available to only brain imaging data. A key future direction would be to integrate both brain imaging and behavioural data. This is particularly relevant for classification of NDDs in which movement is often implicated.

Indeed, movement characteristics have been widely studied in research regarding how EAs can be used to induce classifiers able to discriminate PD patients from healthy controls or from other NDDs. For instance, Lones et al. developed two classification techniques: sliding window genetic programming expressions and artificial biochemical networks [409]. These classifiers identified PD patients (relative to healthy controls) with ~95% accuracy based on movement data. In contrast, Long et al. examined only brain imaging data (multimodal imaging and multi-level measurements) and successfully classified early PD patients (relative to healthy controls) with an accuracy of 87% [458]. As such, these promising classification powers suggest that using a combination of rs-fMRI and behavioural tasks (e.g., finger-tapping [406], reach and grasp [407], and drawing tasks [408]) may provide a non-invasive and more accurate approach, improving the clinical diagnosis of early PD and potentially other NDDs.

7.5. Future Directions

In this section, two areas for future research are outlined: (1) Classification of other neurological/neurodegenerative diseases; and (2) whole-brain DCM analysis.

7.5.1. Classification of Neurological/Neurodegenerative Diseases

The research presented in this thesis involved the classification of PD patients (Chapter 6) and participants administered Modafinil for accurate PD drug monitoring (Chapter 5) and identified a brain imaging biomarker for both PD and Modafinil by focusing on changes to the DMN whilst participants were at rest. Research reveals that changes to the DMN may signal several neurological disorders, for instance, AD [466], [467], attention deficit hyperactivity disorder [468], [469] and depression [470]–[472] amongst others. Hence, an important direction for future research involves identifying whether CGP classification is applicable to other neurological/neurodegenerative diseases, such as AD.

Currently, PD has misdiagnosis rates of 15-26% [7], [8]. Given that the research presented in this thesis obtained high classification accuracy rates, findings can be applied to medical settings with the objective of confirming disease diagnosis and accurate drug treatment monitoring. Indeed, this research can be applied to other neurological/neurodegenerative diseases in order to examine whether the intrinsic connectivity as a result of DCM analysis can be used as a biomarker for diagnostic purposes.

The use of resting-state scans (as per the research reported in this thesis) is well suited to individuals with neurological/neurodegenerative diseases for two reasons. Firstly, task-based studies include more variables, requiring greater control. With rs-fMRI, group differences are more clearly identified since differences between scans that do not arise from participants' mental state are reduced. Secondly, participants with certain NDDs (e.g., dementia) may not be able to correctly execute the tasks involved in a task-based study, yet resting-state studies can still include these participants.

Chapter 6 examined the classification of PD patients, prodromal PD patients, and healthy controls using binary classification (e.g., classification of PD patients relative to prodromal PD patients). A timely next step involves the classification of multiple PD subgroups simultaneously. Indeed, rather than having binary classification, future research could examine developing classifiers with multiple outputs (one output per PD subtype).

7.5.2. Whole-Brain Dynamic Causal Modelling

DCM analysis focuses only on specific brain regions, selected in relation to a given research question. This research examined DCM analysis on the DMN. The DMN has been typically examined in rs-fMRI research with much research linking changes in the DMN to biomarkers for neurological conditions [466]–[472]. Nevertheless, expanding the brain regions included in DCM analysis to explore large – even whole-brain – neural networks is a relevant and timely approach that can broaden our understanding of the mechanisms triggering cognitive processes (e.g., consciousness) for both individuals with neurological/neurodegenerative disorders (e.g., identifying biomarkers for neurological disorders) and healthy individuals.

7.6. Overall Conclusions

This thesis investigates two overarching themes: (1) the development of methods for accurate monitoring of PD medication using EAs on rs-fMRI data for participants administered Modafinil; and (2) the diagnosis of early stage PD (prodromal PD) using EAs on rs-fMRI data. Classification was conducted on timeseries values as well as DCM values. Additional novel features of this research include the application of CGP and RCGP classification for brain imaging data and, also, the use of DCM values in classification. Findings revealed maximum accuracy rates of 75% for prodromal PD patients versus healthy controls, 86% for PD patients versus prodromal PD patients, and 92% for PD patients versus healthy controls. These findings were validated with two classification methods (ANN and SVM) and revealed that CGP provided equivalent performance accuracy relative to ANN and SVM. Further, there was almost no difference in the classification accuracies between timeseries and DCM data. These findings attest the

significance of DCM analyses for classification and CGP as a new classification method for brain imaging data.

As such, this research identified a brain imaging biomarker for PD patients, with particularly high accuracy for diagnosing the most challenging of PD subtypes – prodromal PD – early PD patients without motor symptoms. This finding represents the most valuable research output from this thesis. Standard methods for the diagnosis of prodromal PD reveal low reliability with high rates of misdiagnosis, underscoring the clinical relevance of the current findings in developing an automatic method for classifying PD brain imaging preclinical biomarkers. Indeed, implications abound regarding confirmation of PD diagnosis – a relevant and timely approach given current overall PD misdiagnosis levels of 15-26% [7], [8]. Finally, the research presented in this thesis underscores the importance developing current classification methods for improving the differential diagnosis of neurological/neurodegenerative diseases, including subtypes within these (e.g., PD subtypes). In addition, methods examining the concurrent classification of multiple different disease subgroups (rather than binary classification) are a much-needed step for future research.

List of Abbreviations

ACO	Ant Colony Optimisation
AD	Alzheimer's Disease
ADASYN	Adaptive Synthetic Sampling
AI	Artificial Intelligence
AIS	Artificial Immune System
ANN	Artificial Neural Networks
Bagging	Bootstrap Aggregating
BOLD	Blood-Oxygen Level Dependent
BRAIN	Bradykinesia Akinesia Incoordination Test
CAT	Computed Axial Tomography
CBD	Corticobasal Degeneration
CDR	Clinical Dementia Rating
CGP	Cartesian Genetic Programming
CI	Computational Intelligence
CONN	Functional Connectivity Toolbox
CSF	Cerebrospinal Fluid
CT	See CAT
CV	Cross-validation
DAT	Dopamine Transporter Imaging
DCM	Dynamic Causal Modelling
DMN	Default Mode Network
EAs	Evolutionary Algorithms
EEG	Electroencephalography
EPI	Echo Planar Imaging
EU	European Union
FA	Flip Angle
FFT	Fast Fourier Transform
FID	Free-Induction Decay
Flair	Fluid Attenuated Inversion Recovery
fMRI	Functional Magnetic Resonance Imaging
FoV	Field of View
FPC	Fronto Parietal Control
FSL	FMRIB Software Library
FWHM	Full Width at Half-Maximum
GA	Genetic Algorithms
GP	Genetic Programming
GWAS	Genome Wide Association Surveys
HD	Huntington's Disease

kNN	k-Nearest Neighbours
LB	Lewy Body
LID	Levodopa-Induced Dyskinesia
LIPC	Left Inferior Parietal Cortex
MBRS	Modified Bradykinesia Rating Scale
MDS-UPDRS	Movement Disorder Society sponsored revision of Unified Parkinson's Disease Rating Scale
MEG	Magneto Encephalography
MoCA	Montreal Cognitive Assessment
mPFC	Medial Prefrontal Cortex
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
MSA	Multiple System Atrophy
NDDs	Neurodegenerative Diseases
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NINDS	National Institute of Neurological Disorders and Stroke
NIRS	Near Infrared Spectroscopy
PCC	Posterior Cingulate Cortex
PD	Parkinson's Disease
PET	Positron Emission Tomography
PPMI	Parkinson's Progression Markers Initiative
PSO	Particle Swarm Optimisation
PSP	Progressive Supranuclear Palsy
RBD	Rapid Eye Movement Sleep Behaviour Disorder
RCGP	Recurrent CGP
RF	Radio Frequency
RIPC	Right Inferior Parietal Cortex
rs-fMRI	Resting State fMRI
RSN	Resting State Neural Network
SD	Standard Deviation
SEM	Structural Equation Modelling
SIGN	Scottish Intercollegiate Guidelines Network
SPECT	Single Photon Emission Computed Tomography
SMOTE	Synthetic Sampling with Data Generation
SNR	Signal to Noise Ratio
SPM	Statistical Parametric Mapping
SVM	Support Vector Machines
TE	Echo Time
TR	Repetition Time
UHDRS	Unified Huntington's Disease Rating Scale

UK	United Kingdom
UPDRS	Unified Parkinson's Disease Rating Scale
UPSIT	University of Pennsylvania Smell Identification Test
USA	United States of America
WM	White Matter

References

[1] World Health Organization, “World Health Statistics – Large Gains in Life Expectancy,” 2014. [Online]. Available: http://www.who.int/mediacentre/news/releases/2014/world-health-statistics-2014/en/. [Accessed: 17-Sep-2015].

[2] The University of California Berkley (USA) and Max Planck Institute for Demographic Research (Germany), “The Human Mortality Database,” 2014. [Online]. Available: http://www.mortality.org/. [Accessed: 17-Sep-2015].

[3] United Nations, Department of Economics and Social Affairs, and Population Division, “World Population Prospects 2015 - Key Findings & Advance Tables,” New York, 2015.

[4] G. Levy, “The relationship of Parkinson disease with aging.,” *Arch. Neurol.*, vol. 64, no. 9, pp. 1242–1246, 2007.

[5] J. a. Driver, G. Logroscino, J. M. Gaziano, and T. Kurth, “Incidence and remaining lifetime risk of Parkinson disease in advanced age,” *Neurology*, vol. 72, no. 5, pp. 432–438, 2009.

[6] World Health Organization, “Draft twelfth general programme of work Report of the Programme , Budget and Administration Committee of the Executive Board to the Sixty-sixth World Health Assembly,” 2013.

[7] A. J. Hughes, S. E. Daniel, L. Kilford, and A. J. Lees, “Accuracy of clinical diagnosis of idiopathic Parkinson’s disease: a clinico-pathological study of 100 cases.,” *J. Neurol. Neurosurg. Psychiatry*, vol. 55, no. 3, pp. 181–184, 1992.

[8] A. H. Rajput, B. Rozdilsky, and A. Rajput, “Accuracy of clinical diagnosis in parkinsonism - a prospective study.,” *Can. J. Neurol. Sci.*, vol. 18, no. 3, pp. 275–8,

1991.

[9] J. B. Rowe, “Connectivity analysis is essential to understand neurological disorders,” *Front. Syst. Neurosci.*, vol. 4, no. September, pp. 1–13, 2010.

[10] Y. Zhou, J. H. J. Dougherty, K. F. Hubner, B. Bai, R. L. Cannon, and R. K. Hutson, “Abnormal connectivity in the posterior cingulate and hippocampus in early Alzheimer’s disease and mild cognitive impairment.,” *Alzheimers. Dement.*, vol. 4, no. 4, pp. 265–270, 2008.

[11] G. Wu, D. Zhang, D. Shen, P. Yan, K. Suzuki, and F. Wang, *Machine learning in medical imaging*. Springer, 2013.

[12] A. Tusche, J. Smallwood, B. C. Bernhardt, and T. Singer, “Classifying the wandering mind: Revealing the affective content of thoughts during task-free rest periods,” *Neuroimage*, vol. 97, pp. 107–116, 2014.

[13] E. Challis, P. Hurley, L. Serra, M. Bozzali, S. Oliver, and M. Cercignani, “Gaussian process classification of Alzheimer’s disease and mild cognitive impairment from resting-state fMRI,” *Neuroimage*, vol. 112, pp. 232–243, 2015.

[14] A. Tessitore *et al.*, “Resting-state brain connectivity in patients with Parkinson’s disease and freezing of gait,” *Parkinsonism Relat. Disord.*, vol. 18, no. 6, pp. 781–787, Jul. 2012.

[15] J. Kahan *et al.*, “Therapeutic Subthalamic Nucleus Deep Brain Stimulation Reverses Cortico-Thalamic Coupling during Voluntary Movements in Parkinson’s Disease,” *PLoS One*, vol. 7, no. 12, 2012.

[16] D. M. Herz, H. R. Siebner, O. J. Hulme, E. Florin, M. S. Christensen, and L. Timmermann, “Levodopa reinstates connectivity from prefrontal to premotor cortex during externally paced movement in Parkinson’s disease,” *Neuroimage*, vol. 90, pp. 15–23, 2014.

[17] C. L. Rae *et al.*, "Atomoxetine restores the response inhibition network in Parkinson's disease," *Brain*, vol. 139, no. 8, pp. 2235–2248, 2016.

[18] D. M. Herz *et al.*, "Abnormal dopaminergic modulation of striato-cortical networks underlies levodopa-induced dyskinesias in humans," *Brain*, vol. 138, no. 6, pp. 1658–1666, 2015.

[19] J. B. Rowe, L. E. Hughes, R. A. Barker, and A. M. Owen, "Dynamic causal modelling of effective connectivity from fMRI: Are results reproducible and sensitive to Parkinson's disease and its treatment?," *Neuroimage*, vol. 52, no. 3, pp. 1015–1026, 2010.

[20] R. J. Moran, "Deep brain stimulation for neurodegenerative disease: A computational blueprint using dynamic causal modeling," in *Progress in Brain Research*, vol. 222, 2015, pp. 125–146.

[21] A. Bradford, M. E. Kunik, P. Schulz, S. P. Williams, and H. Singh, "Missed and Delayed Diagnosis of Dementia in Primary Care: Prevalence and Contributing Factors," *Alzheimer Dis. Assoc. Disord.*, vol. 23, no. 4, pp. 306–314, Oct. 2009.

[22] V. Kotagal *et al.*, "Factors associated with cognitive evaluations in the United States," *Neurology*, vol. 84, no. 1, pp. 64–71, 2015.

[23] L. Boise, M. B. Neal, and J. Kaye, "Dementia assessment in primary care: Results from a study in three managed care systems," *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.*, vol. 59, no. 6, pp. 621–626, 2004.

[24] L. Boise, R. Camicioli, D. L. Morgan, J. H. Rose, and L. Congleton, "Diagnosing Dementia: Perspectives of Primary Care Physicians," *Gerontologist*, vol. 39, no. 4, pp. 457–464, Aug. 1999.

[25] S. Kansara, A. Trivedi, S. Chen, J. Jankovic, and W. Le, "Early diagnosis and therapy of Parkinson's disease: can disease progression be curbed?," *J. Neural*

Transm., vol. 120, no. 1, pp. 197–210, Jan. 2013.

[26] D. J. Selkoe and J. Peter J Lansbury, "Alzheimer's Disease Is the Most Common Neurodegenerative Disorder," *Basic Neurochem. Mol. Cell. Med. Asp.*, 1999.

[27] I. Mcdowell *et al.*, "The Canadian Study of Health and Aging: risk factors for Alzheimer's disease in Canada.," *Neurology*, vol. 44, no. 11, pp. 2073–80, Nov. 1994.

[28] R. Brookmeyer, S. Gray, and C. Kawas, "Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset," *Am. J. Public Health*, vol. 88, no. 9, pp. 1337–1342, 1998.

[29] P. A. Osterrieth, "Le test de copie d'une figure complexe: contribution à l'étude de la perception et de la mémoire.," *Arch. Psychol. (Geneve).*, vol. 28, pp. 286–356, 1944.

[30] A. Rey, "L'examen psychologique dans les cas d'encéphalopathie traumatique," *Arch. Psychol. (Geneve).*, vol. 28, pp. 215–285, 1941.

[31] A. Hazell and S. L. Smith, "Towards an objective assessment of Alzheimer's disease: the application of a novel evolutionary algorithm in the analysis of figure copying tasks," *GECCO-2008 Work. MedGEC Med. Appl. Genet. Evol. Comput.*, pp. 2073–2080, 2008.

[32] T. Pringsheim, K. Wiltshire, L. Day, J. Dykeman, T. Steeves, and N. Jette, "The incidence and prevalence of Huntington's disease: A systematic review and meta-analysis," *Mov. Disord.*, vol. 27, no. 9, pp. 1083–1091, Aug. 2012.

[33] M. Papoutsi, I. Labuschagne, S. J. Tabrizi, and J. C. Stout, "The cognitive burden in Huntington's disease: Pathology, phenotype, and mechanisms of compensation," *Mov. Disord.*, vol. 29, no. 5, pp. 673–683, Apr. 2014.

[34] Huntington Study Group, “Unified Huntington’s disease rating scale: Reliability and consistency,” *Mov. Disord.*, vol. 11, no. 2, pp. 136–142, Mar. 1996.

[35] J. B. Carroll, G. P. Bates, J. Steffan, C. Saft, and S. J. Tabrizi, “Treating the whole body in Huntington’s disease,” *Lancet Neurol.*, vol. 14, no. 11, pp. 1135–1142, Nov. 2015.

[36] A. W. Willis, “Parkinson disease in the elderly adult,” *Mo. Med.*, vol. 110, no. 5, pp. 406–410, 2013.

[37] L. Bertram and R. E. Tanzi, “The genetic epidemiology of neurodegenerative disease,” *J. Clin. Invest.*, vol. 115, no. 6, pp. 1449–1457, 2005.

[38] T. Pringsheim, N. Jette, A. Frolkis, and T. D. L. Steeves, “The prevalence of Parkinson’s disease: A systematic review and meta-analysis,” *Mov. Disord.*, vol. 29, no. 13, pp. 1583–1590, 2014.

[39] H. L. Martin and P. Teismann, “Glutathione - a review on its role and significance in Parkinson’s disease,” *FASEB J.*, vol. 23, no. 10, pp. 3263–3272, Oct. 2009.

[40] R. J. Wurtman, “Personalized medicine strategies for managing patients with Parkinsonism and cognitive deficits,” *Metabolism*, vol. 62, no. SUPPL.1, pp. S27–S29, Jan. 2013.

[41] E. Gelpi *et al.*, “Multiple organ involvement by alpha-synuclein pathology in Lewy body disorders,” *Mov. Disord.*, vol. 29, no. 8, pp. 1010–1018, Jul. 2014.

[42] D. B. Miller and J. P. O’Callaghan, “Biomarkers of Parkinson’s disease: Present and future,” *Metabolism.*, vol. 64, no. 3, pp. S40–S46, 2015.

[43] National Collaborating Centre for Chronic Conditions Great Britain, *Parkinson’s Disease*. 2006.

[44] J. Parkinson, *An Essay on the Shaking Palsy*. Whittingham and Rowland, 1817.

[45] J.-M. Charcot, “De la paralysie agitante,” in *Oeuvres Complètes (t 1) Leçons sur les maladies du système nerveux*, Paris, 1872, pp. 155–188.

[46] P. A. Kempster, B. Hurwitz, and A. J. Lees, “A new look at James Parkinson’s Essay on the Shaking Palsy.,” *Neurology*, vol. 69, no. 5, pp. 482–5, Jul. 2007.

[47] A. Björklund and S. B. Dunnett, “Dopamine neuron systems in the brain: an update,” *Trends Neurosci.*, vol. 30, no. 5, pp. 194–202, May 2007.

[48] O. Hornykiewicz, “The discovery of dopamine deficiency in the parkinsonian brain.,” *J. Neural Transm. Suppl.*, no. 70, pp. 9–15, 2006.

[49] W. Birkmayer and O. Hornykiewicz, “The L-3,4-dioxyphenylalanine (DOPA)-effect in Parkinson-akinesia,” *Wien. Klin. Wochenschr.*, vol. 73, pp. 787–8, Nov. 1961.

[50] W. Birkmayer and O. Hornykiewicz, “The effect of l-3,4-dihydroxyphenylalanine (=DOPA) on akinesia in parkinsonism,” *Parkinsonism Relat. Disord.*, vol. 4, no. 2, pp. 59–60, Aug. 1998.

[51] G. C. Cotzias, P. S. Papavasiliou, and R. Gellene, “Modification of Parkinsonism — Chronic Treatment with L-Dopa,” *N. Engl. J. Med.*, vol. 280, no. 7, pp. 337–345, Feb. 1969.

[52] K. S. P. McNaught, P. Jenner, and C. Olanow, “Protein mishandling: Role of the ubiquitin proteasome system in the pathogenesis of Parkinson’s disease.,” in *Parkinson’s disease and movement disorders.*, 5th ed., J. Jankovic and E. Tolosa, Eds. Philadelphia: Lippincott Williams and Wilkins, 2007, pp. 33–49.

[53] T. Pan, S. Kondo, W. Le, and J. Jankovic, “The role of autophagy-lysosome pathway in neurodegeneration associated with Parkinson’s disease,” *Brain*, vol. 131,

no. 8, pp. 1969–1978, 2008.

[54] D. Salat and E. Tolosa, “Levodopa in the treatment of Parkinson’s disease: current status and new developments.,” *J. Parkinsons. Dis.*, vol. 3, no. 3, pp. 255–69, 2013.

[55] R. L. Nussbaum and C. E. Ellis, “Alzheimer’s Disease and Parkinson’s Disease,” *N. Engl. J. Med.*, vol. 348, no. 14, pp. 1356–1364, Apr. 2003.

[56] L. M. L. de Lau and M. M. B. Breteler, “Epidemiology of Parkinson’s disease.,” *Lancet. Neurol.*, vol. 5, no. June, pp. 525–535, 2006.

[57] T. J. Collier, N. M. Kanaan, and J. H. Kordower, “Ageing as a primary risk factor for Parkinson’s disease: evidence from studies of non-human primates,” *Nat. Rev. Neurosci.*, vol. 12, no. 6, pp. 359–366, 2011.

[58] M. C. de Rijk *et al.*, “Prevalence of parkinsonism and Parkinson’s disease in Europe: the EUROPARKINSON Collaborative Study.,” *J. Neurol. Neurosurg. Psychiatry*, vol. 62, no. 1, pp. 10–15, Jan. 1997.

[59] H. Braak, K. Del Tredici, U. Rüb, R. A. . de Vos, E. N. . Jansen Steur, and E. Braak, “Staging of brain pathology related to sporadic Parkinson’s disease,” *Neurobiol. Aging*, vol. 24, no. 2, pp. 197–211, Mar. 2003.

[60] T. Gasser, “Genetics of Parkinson’s disease,” *Clin. Genet.*, vol. 54, no. 4, pp. 259–265, Oct. 1998.

[61] J. W. Langston, “Current theories on the cause of Parkinson’s disease,” *J. Neurol. Neurosurg. Psychiatry*, vol. 52(Suppl), no. February, pp. 13–17, 1989.

[62] R. Betarbet, T. B. Sherer, G. MacKenzie, M. Garcia-Osuna, A. V. Panov, and J. T. Greenamyre, “Chronic systemic pesticide exposure reproduces features of Parkinson’s disease,” *Nat. Neurosci.*, vol. 3, no. 12, pp. 1301–1306, 2000.

[63] P. Jenner and C. W. Olanow, “Understanding cell death in Parkinson’s disease.,” *Ann. Neurol.*, vol. 44, no. 3 Suppl 1, pp. S72–S84, 1998.

[64] J. M. Fearnley and A. J. Lees, “Ageing and Parkinson’S Disease: Substantia Nigra Regional Selectivity,” *Brain*, vol. 114, no. 5, pp. 2283–2301, 1991.

[65] L. Findley *et al.*, “Direct economic impact of Parkinson’s disease: A research survey in the United Kingdom,” *Mov. Disord.*, vol. 18, no. 10, pp. 1139–1145, 2003.

[66] K. Whetten-Goldstein, F. Sloan, E. Kulas, T. Cutson, and M. Schenkman, “The burden of Parkinson’s disease on society, family, and the individual.,” *Journal of the American Geriatrics Society*, vol. 45, no. 7. pp. 844–9, 1997.

[67] C. LePen, S. Wait, F. Moutard-Martin, M. Dujardin, and M. Ziégler, “Cost of illness and disease severity in a cohort of French patients with Parkinson’s disease,” *Pharmacoeconomics*, vol. 16, no. 1, pp. 59–69, 1999.

[68] R. C. Dodel *et al.*, “Cost of illness in Parkinson disease. A retrospective 3-month analysis of direct costs,” *Nervenarzt*, vol. 68, no. 12, pp. 978–984, Dec. 1997.

[69] R. C. Dodel, K. Berger, and W. H. Oertel, “Health-Related Quality of Life and Healthcare Utilisation in Patients with Parkinson’s Disease,” *Pharmacoeconomics*, vol. 19, no. 10, pp. 1013–1038, 2001.

[70] C. Vossius, J. P. Larsen, C. Janvin, and D. Aarsland, “The economic impact of cognitive impairment in Parkinson’s disease,” *Mov. Disord.*, vol. 26, no. 8, pp. 1541–1544, Jul. 2011.

[71] P. A. Kempster, S. S. O’Sullivan, J. L. Holton, T. Revesz, and A. J. Lees, “Relationships between age and late progression of Parkinson’s disease: A clinico-pathological study,” *Brain*, vol. 133, no. 6, pp. 1755–1762, 2010.

[72] “MDS-UPDRS Training Program & Exercise,” 2017. [Online]. Available:

https://mds.movementdisorders.org/updrs/. [Accessed: 26-Nov-2017].

[73] "Parkinson's disease - Symptoms," 2016. [Online]. Available: https://www.nhs.uk/conditions/parkinsons-disease/symptoms/. [Accessed: 26-Nov-2017].

[74] J. Jankovic, "Parkinson's disease: clinical features and diagnosis," *J. Neurol. Neurosurg. Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008.

[75] J. Pagonabarraga and J. Kulisevsky, "Cognitive impairment and dementia in Parkinson's disease.," *Neurobiol. Dis.*, vol. 46, no. 3, pp. 590–6, Jun. 2012.

[76] W. R. G. Gibb and A. J. Lees, "The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease.," *J. Neurol. Neurosurg. Psychiatry*, vol. 51, no. 6, pp. 745–752, 1988.

[77] N. P. S. Bajaj, V. Gontu, J. Birchall, J. Patterson, D. G. Grosset, and A. J. Lees, "Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study.," *J. Neurol. Neurosurg. Psychiatry*, vol. 81, no. 11, pp. 1223–1228, Nov. 2010.

[78] M. M. Hoehn and M. D. Yahr, "Parkinsonism : onset, progression, and mortality Parkinsonism: onset, progression, and mortality," *Neurology*, vol. 17, no. 5, pp. 427–442, 1967.

[79] C. M. Hawkes, "Diagnosis and treatment of Parkinson's disease. Anosmia is a common finding.," *BMJ Br. Med. J.*, vol. 310, no. 6995, p. 1995, 1995.

[80] S. Fahn, "Description of Parkinson's Disease as a Clinical Syndrome," *Ann. N. Y. Acad. Sci.*, vol. 991, no. 1, pp. 1–14, Jan. 2006.

[81] J. L. Cummings, C. Henchcliffe, S. Schaier, T. Simuni, A. Waxman, and P. Kemp, "The role of dopaminergic imaging in patients with symptoms of dopaminergic

system neurodegeneration," *Brain*, vol. 134, no. 11, pp. 3146–3166, 2011.

[82] S. Sharma *et al.*, "Biomarkers in Parkinson's disease (recent update)," *Neurochem. Int.*, vol. 63, no. 3, pp. 201–229, 2013.

[83] L. V. Kalia, S. K. Kalia, and A. E. Lang, "Disease-modifying strategies for Parkinson's disease," *Mov. Disord.*, vol. 30, no. 11, pp. 1442–1450, 2015.

[84] K. Stiasny-Kolster *et al.*, "Combination of 'idiopathic' REM sleep behaviour disorder and olfactory dysfunction as possible indicator for α-synucleinopathy demonstrated by dopamine transporter FP-CIT-SPECT," *Brain*, vol. 128, no. 1, pp. 126–137, 2005.

[85] S. Fahn and R. Elton, "Members of the UPDRS Development Committee Unified Parkinson's Disease Rating Scale," *Recent Dev. Park. Dis.*, vol. 2, pp. 153–163, 1987.

[86] C. G. Goetz *et al.*, "Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Mov. Disord.*, vol. 23, no. 15, pp. 2129–2170, 2008.

[87] C. G. Goetz *et al.*, "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan.," *Mov. Disord.*, vol. 22, no. 1, pp. 41–7, Jan. 2007.

[88] L. Henderson *et al.*, "Scales for rating motor impairment in Parkinson's disease: studies of reliability and convergent validity.," *J. Neurol. Neurosurg. Psychiatry*, vol. 54, no. 1, pp. 18–24, 1991.

[89] D. A. Bennett, K. M. Shannon, L. A. Beckett, C. G. Goetz, and R. S. Wilson, "Metric properties of nurses' ratings of parkinsonian signs with a modified Unified Parkinson's Disease Rating Scale," *Neurology*, vol. 49, no. 6, pp. 1580–1587, 1997.

[90] R. Camicioli, S. J. Grossmann, P. S. Spencer, K. Hudnell, and W. Kent Anger, "Discriminating mild parkinsonism: Methods for epidemiological research," *Mov. Disord.*, vol. 16, no. 1, pp. 33–40, 2001.

[91] P. Martínez-Martín, A. Gil-Nagel, L. M. Gracia, J. B. Gómez, J. Martínez-Sarriés, and F. Bermejo, "Unified Parkinson's disease rating scale characteristics and structure," *Mov. Disord.*, vol. 9, no. 1, pp. 76–83, Jan. 1994.

[92] Royal College of Physicians and Association of British Neurologists, *Local adult neurology services for the next decade*, no. June. London: Royal College of Physicians, 2011.

[93] D. A. Heldman *et al.*, "The modified bradykinesia rating scale for Parkinson's disease: Reliability and comparison with kinematic measures," *Mov. Disord.*, vol. 26, no. 10, pp. 1859–1863, Aug. 2011.

[94] A. Kishore *et al.*, "Unilateral versus bilateral tasks in early asymmetric Parkinson's disease: Differential effects on bradykinesia," *Mov. Disord.*, vol. 22, no. 3, pp. 328–333, 2007.

[95] A. J. Espay, D. E. Beaton, F. Morgante, C. A. Gunraj, A. E. Lang, and R. Chen, "Impairments of speed and amplitude of movement in Parkinson's disease: A pilot study," *Mov. Disord.*, vol. 24, no. 7, pp. 1001–1008, 2009.

[96] Z. S. Nasreddine *et al.*, "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment," *J Am Geriatr Soc*, vol. 53, no. 4, pp. 695–699, 2005.

[97] J. C. Dalrymple-Alford *et al.*, "The MoCA: Well-suited screen for cognitive impairment in Parkinson disease," *Neurology*, vol. 75, no. 19, pp. 1717–1725, 2010.

[98] S. Hoops *et al.*, "Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease," *Neurology*, vol. 73, no. November 24, pp. 1738–45,

2009.

[99] D. J. Gill, A. Freshman, J. A. Blender, and B. Ravina, “The Montreal Cognitive Assessment as a screening tool for cognitive impairment in Parkinson’s disease,” *Mov. Disord.*, vol. 23, no. 7, pp. 1043–1046, 2008.

[100] All Party Parliament Group for Parkinson’s Disease, “Please mind the gap: Parkinson’s disease services today,” 2009.

[101] National Institute for Health and Care Excellence, “Parkinson’s disease in adults (NICE guideline [NG71]),” 2017. [Online]. Available: https://www.nice.org.uk/guidance/ng71. [Accessed: 26-Nov-2017].

[102] Scottish Intercollegiate Guidelines Network, “Diagnosis and pharmacological management of Parkinson’s disease: A national clinical guideline,” 2010.

[103] C. A. Davie, “A review of Parkinson’s disease.,” *Br. Med. Bull.*, vol. 86, pp. 109–27, 2008.

[104] A. Schrag, Y. Ben-Shlomo, and N. Quinn, “How valid is the clinical diagnosis of Parkinson’s disease in the community?,” *J. Neurol. Neurosurg. Psychiatry*, vol. 73, no. 5, pp. 529–534, Nov. 2002.

[105] C. B. Levine, K. R. Fahrbach, A. D. Siderowf, R. P. Estok, V. M. Ludensky, and S. D. Ross, “Diagnosis and Treatment of Parkinson’s Disease: A Systematic Review of the Literature,” *AHRQ*, vol. 03, no. 57, p. 306, 2003.

[106] A. J. Hughes, S. E. Daniel, and A. J. Lees, “Improved accuracy of clinical diagnosis of Lewy body Parkinson’s disease,” *Neurology*, vol. 57, no. 8, pp. 1497–1499, Oct. 2001.

[107] A. J. Hughes, S. E. Daniel, Y. Ben-Shlomo, and A. J. Lees, “The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service.,”

Brain a J. Neurol., vol. 125, no. Pt 4, pp. 861–870, 2002.

[108] J. Jankovic, A. H. Rajput, M. P. McDermott, and D. P. Perl, "The evolution of diagnosis in early Parkinson disease," *Arch. Neurol.*, vol. 57, no. 3, pp. 369–372, 2000.

[109] "The National Service Framework for Long-term Conditions," London, 2005.

[110] National Institute for Health and Care Excellence, "Parkinson's disease in adults (NICE guideline [NG71]) - What medicines might help?," 2017. [Online]. Available: https://www.nice.org.uk/guidance/ng71/ifp/chapter/What-medicines-might-help. [Accessed: 26-Nov-2017].

[111] D. J. Brooks, "Optimizing levodopa therapy for Parkinson's disease with levodopa/carbidopa/entacapone: Implications from a clinical and patient perspective," *Neuropsychiatr. Dis. Treat.*, vol. 4, no. 1 A, pp. 39–47, 2008.

[112] W. Poewe, A. Antonini, J. C. Zijlmans, P. R. Burkhard, and F. Vingerhoets, "Levodopa in the treatment of Parkinson's disease: an old drug still going strong.," *Clin. Interv. Aging*, vol. 5, pp. 229–238, 2010.

[113] P. A. LeWitt, "Levodopa for the Treatment of Parkinson's Disease," *N. Engl. J. Med.*, vol. 359, no. 23, pp. 2468–76, 2008.

[114] B. Thanvi, N. Lo, and T. Robinson, "Levodopa-induced dyskinesia in Parkinson's disease: clinical features, pathogenesis, prevention and treatment," *Postgrad. Med. J.*, vol. 83, no. 980, pp. 384–388, 2007.

[115] O. Rascol, D. J. Brooks, A. D. Korczyn, P. P. De Deyn, C. E. Clarke, and A. E. Lang, "A five-year study of the incidence of dyskinesia in patients with early Parkinson's disease who were treated with ropinirole or levodopa," *N. Engl. J. Med.*, vol. 342, no. 20, pp. 1484–1491, 2000.

[116] A. H. Rajput *et al.*, “Clinical-pathological study of levodopa complications,” *Mov. Disord.*, vol. 17, no. 2, pp. 289–296, 2002.

[117] J. M. Rabey, T. Prokhorov, A. Miniovitz, E. Dobronevsky, and C. Klein, “Effect of quetiapine in psychotic Parkinson’s disease patients: A double-blind labeled study of 3 months’ duration,” *Mov. Disord.*, vol. 22, no. 3, pp. 313–318, Feb. 2007.

[118] J. L. Juncos *et al.*, “Quetiapine improves psychotic symptoms and cognition in Parkinson’s disease,” *Mov. Disord.*, vol. 19, no. 1, pp. 29–35, Jan. 2004.

[119] H. H. Fernandez, J. H. Friedman, C. Jacques, and M. Rosenfeld, “Quetiapine for the treatment of drug-induced psychosis in Parkinson’s disease,” *Mov. Disord.*, vol. 14, no. 3, pp. 484–487, May 1999.

[120] I. F. Hussain, C. M. Brady, M. J. Swinn, C. J. Mathias, and C. J. Fowler, “Treatment of erectile dysfunction with sildenafil citrate (Viagra) in parkinsonism due to Parkinson’s disease or multiple system atrophy with observations on orthostatic hypotension,” *J. Neurol. Neurosurg. Psychiatry*, vol. 71, no. 3, pp. 371–374, Sep. 2001.

[121] J.-M. Senard, C. Brefel-Courbon, O. Rascol, and J.-L. Montastruc, “Orthostatic Hypotension in Patients with Parkinson’s Disease,” *Drugs Aging*, vol. 18, no. 7, pp. 495–505, 2001.

[122] R. A. Wright *et al.*, “A double-blind, dose-response study of midodrine in neurogenic orthostatic hypotension,” *Neurology*, vol. 51, no. 1, p. 120 LP-124, Jul. 1998.

[123] R. G. Elbers, H. W. Berendse, and G. Kwakkel, “Treatment of Fatigue in Parkinson Disease,” *JAMA*, vol. 315, no. 21, pp. 2340–2341, 2016.

[124] R. G. Elbers, J. Verhoef, E. E. van Wegen, H. W. Berendse, and G. Kwakkel, “Interventions for fatigue in Parkinson’s disease,” *Cochrane Database Syst. Rev.*,

no. 10, pp. 1–80, Oct. 2015.

[125] J.-S. Lou *et al.*, "Using Modafinil to Treat Fatigue in Parkinson Disease," *Clin. Neuropharmacol.*, vol. 32, no. 6, pp. 305–310, Nov. 2009.

[126] H. L. Tyne, J. Taylor, G. A. Baker, and M. J. Steiger, "Modafinil for Parkinson's disease fatigue," *J. Neurol.*, vol. 257, no. 3, pp. 452–456, Mar. 2010.

[127] M. J. Minzenberg, A. J. Watrous, J. H. Yoon, S. Ursu, and C. S. Carter, "Modafinil shifts human locus coeruleus to low-tonic, high-phasic activity during functional MRI," *Science (80-.).*, vol. 322, no. 5908, pp. 1700–1702, 2008.

[128] A. Raggi, G. Plazzi, G. Pennisi, D. Tasca, and R. Ferri, "Cognitive evoked potentials in narcolepsy: A review of the literature," *Neurosci. Biobehav. Rev.*, vol. 35, no. 5, pp. 1144–1153, 2011.

[129] N. Dawson, R. J. Thompson, A. McVie, D. M. Thomson, B. J. Morris, and J. A. Pratt, "Modafinil reverses phencyclidine-induced deficits in cognitive flexibility, cerebral metabolism, and functional brain connectivity," *Schizophr. Bull.*, vol. 38, no. 3, pp. 457–474, 2012.

[130] M. Kahbazi, A. Ghoreishi, F. Rahiminejad, M.-R. Mohammadi, A. Kamalipour, and S. Akhondzadeh, "A randomized, double-blind and placebo-controlled trial of modafinil in children and adolescents with attention deficit and hyperactivity disorder," *Psychiatry Res.*, vol. 168, no. 3, pp. 234–237, 2009.

[131] N. D. Volkow *et al.*, "Evaluating dopamine reward pathway in ADHD: clinical implications.," *Jama*, vol. 302, no. 10, pp. 1084–91, 2009.

[132] A. D. Kalechstein, J. J. Mahoney, J. H. Yoon, R. Bennett, and R. De La Garza, "Modafinil, but not escitalopram, improves working memory and sustained attention in long-term, high-dose cocaine users," *Neuropharmacology*, vol. 64, pp. 472–478, 2013.

[133] A. C. Dean, R. J. Sevak, J. R. Monterosso, G. Hellemann, C. A. Sugar, and E. D. London, “Acute modafinil effects on attention and inhibitory control in methamphetamine-dependent humans.,” *J. Stud. Alcohol Drugs*, vol. 72, no. 6, pp. 943–53, 2011.

[134] D. C. Turner, T. W. Robbins, L. Clark, A. R. Aron, J. Dowson, and B. J. Sahakian, “Cognitive enhancing effects of modafinil in healthy volunteers,” *Psychopharmacol.*, vol. 165, no. 3, pp. 260–269, 2003.

[135] W. G. Ondo, R. Fayle, F. Atassi, and J. Jankovic, “Modafinil for daytime somnolence in Parkinson’s disease: double blind, placebo controlled parallel trial,” *J. Neurol. Neurosurg. Psychiatry*, vol. 76, no. 12, pp. 1636–1639, Dec. 2005.

[136] Sarah Knapton, “‘Smart drug’ taken by one in four students really does boost performance,” *The Telegraph*, 2015. [Online]. Available: http://www.telegraph.co.uk/news/science/science-news/11812682/Smart-drug-taken-by-one-in-four-students-really-does-boost-performance.html. [Accessed: 25-Mar-2017].

[137] Graeme Paton, “Universities warned of ‘explosion’ in use of smart drugs,” *The Telegraph*, 2013. [Online]. Available: http://www.telegraph.co.uk/education/universityeducation/10340244/Universities-warned-of-explosion-in-use-of-smart-drugs.html. [Accessed: 25-Mar-2017].

[138] H. Greely *et al.*, “Towards responsible use of cognitive-enhancing drugs by the healthy,” *Nature*, vol. 456, no. 7223, pp. 702–705, 2008.

[139] H. Moldofsky, “A randomized trial of the long-term, continued efficacy and safety of modafinil in narcolepsy,” *Sleep Med.*, vol. 1, no. 2, pp. 109–116, 2000.

[140] S. H. Kollins, E. K. MacDonald, and C. R. Rush, “Assessing the abuse potential of methylphenidate in nonhuman and human subjects: A review,” *Pharmacol. Biochem. Behav.*, vol. 68, no. 3, pp. 611–627, 2001.

[141] N. D. Volkow *et al.*, “Effects of modafinil on dopamine and dopamine transporters in the male human brain: Clinical implications,” *Jama*, vol. 301, no. 11, pp. 1148–1154, 2009.

[142] C. W. Olanow and J. A. Obeso, “The significance of defining preclinical or prodromal Parkinson’s disease,” *Mov. Disord.*, vol. 27, no. 5, pp. 666–669, 2012.

[143] D. D. Truong, R. Bhidayasiri, and E. Wolters, “Management of non-motor symptoms in advanced Parkinson disease,” *J. Neurol. Sci.*, vol. 266, no. 1–2, pp. 216–228, 2008.

[144] I. Ferrer, I. López-Gonzalez, M. Carmona, E. Dalfó, A. Pujol, and A. Martínez, “Neurochemistry and the non-motor aspects of PD,” *Neurobiol. Dis.*, vol. 46, no. 3, pp. 508–526, 2012.

[145] M. M. Mielke and W. Maetzler, “A ‘bird’s eye’ view on the current status and potential benefits of blood biomarkers for Parkinson’s disease.,” *Biomark Med*, vol. 8, no. 2, pp. 225–227, Feb. 2014.

[146] J. H. McKinnon, B. M. Demaerschalk, J. N. Caviness, K. E. Wellik, C. H. Adler, and D. M. Wingerchuk, “Sniffing out Parkinson disease: can olfactory testing differentiate parkinsonian disorders?,” *Neurologist*, vol. 13, no. 6, pp. 382–385, 2007.

[147] A. Y. Deutch, “Parkinson’s disease redefined,” *Lancet Neurol.*, vol. 12, no. 5, pp. 422–423, 2013.

[148] Biomarkers Definitions Working Group, “Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework,” *Clin. Pharmacol. Ther.*, vol. 69, no. 3, pp. 89–95, 2001.

[149] L. Parnetti *et al.*, “Cerebrospinal fluid biomarkers in Parkinson disease,” *Nat. Rev. Neurol.*, vol. 9, no. 3, pp. 131–140, 2013.

[150] E. Saracchi, S. Fermi, and L. Brighina, “Emerging candidate biomarkers for Parkinson’s disease: a review.,” *Aging Dis.*, vol. 5, no. 1, pp. 27–34, 2014.

[151] M. G. Schlossmacher and B. Mollenhauer, “Biomarker research in Parkinson’s disease: objective measures needed for patient stratification in future cause-directed trials.,” *Biomark. Med.*, vol. 4, no. 5, pp. 647–50, 2010.

[152] J. R. Streffer *et al.*, “Prerequisites to launch neuroprotective trials in Parkinson’s disease: An industry perspective,” *Mov. Disord.*, vol. 27, no. 5, pp. 651–655, 2012.

[153] S. T. DeKosky and K. Marek, “Looking Backward to Move Forward: Early Detection of Neurodegenerative Disorders,” *Science (80-.).*, vol. 302, no. 5646, pp. 830–834, 2003.

[154] C. Scherzer, “Interview - Searching for biomarkers in Parkinson’s disease,” *Biomark. Med.*, vol. 3, no. 2, pp. 113–114, Apr. 2009.

[155] H. Chen *et al.*, “Research on the premotor symptoms of Parkinson’s disease: Clinical and etiological implications,” *Environ. Health Perspect.*, vol. 121, no. 11–12, pp. 1245–1252, 2013.

[156] G. Poste, “Bring on the biomarkers,” *Nature*, vol. 469, no. 7329, pp. 156–157, 2011.

[157] D. Mohammadi, “The Harvard Biomarker Study’s big plan,” *Lancet Neurol.*, vol. 12, no. 8, pp. 739–740, 2013.

[158] K. Marek *et al.*, “The Parkinson Progression Marker Initiative (PPMI),” *Prog. Neurobiol.*, vol. 95, no. 4, pp. 629–635, 2011.

[159] S. W. Scholz, T. Mhyre, H. Ressom, S. Shah, and H. J. Federoff, “Genomics and bioinformatics of Parkinson’s disease,” *Cold Spring Harb. Perspect. Med.*, vol. 2, no. 7, pp. 1–16, 2012.

[160] C. Klein and A. Westenberger, “Genetics of Parkinson’s Disease,” *Cold Spring Harb. Perspect. Med.*, vol. 2, no. 1, p. a008888, 2012.

[161] H. Alonso-Navarro, F. Jimenez-Jimenez, E. Garcia-Martin, and J. Agundez, “Genomic and Pharmacogenomic Biomarkers of Parkinson’s Disease,” *Curr. Drug Metab.*, vol. 15, no. 2, pp. 129–181, Mar. 2014.

[162] T. Gasser, “Genomic and proteomic biomarkers for Parkinson disease,” *Neurology*, vol. 72, no. 7 Supplement 2, p. S27 LP-S31, Feb. 2009.

[163] M. J. Edwards, M. Stamelou, N. Quinn, and K. P. Bhatia, *Parkinson’s Disease and other Movement Disorders*. Oxford University Press, 2016.

[164] E. Garbayo, E. Ansorena, and M. J. Blanco-Prieto, “Drug development in Parkinson’s disease: From emerging molecules to innovative drug delivery systems,” *Maturitas*, vol. 76, no. 3, pp. 272–278, 2013.

[165] E. Valera and E. Masliah, “Immunotherapy for neurodegenerative diseases: Focus on α-synucleinopathies,” *Pharmacol. Ther.*, vol. 138, no. 3, pp. 311–322, 2013.

[166] A. Schneeberger, M. Mandler, F. Mattner, and W. Schmidt, “Vaccination for Parkinson’s disease,” *Parkinsonism Relat. Disord.*, vol. 18, pp. S11–S13, 2012.

[167] S. Yadav *et al.*, “Rodent models and contemporary molecular techniques: Notable feats yet incomplete explanations of Parkinson’s disease pathogenesis,” *Mol. Neurobiol.*, vol. 46, no. 2, pp. 495–512, 2012.

[168] K. D. van Dijk *et al.*, “Diagnostic cerebrospinal fluid biomarkers for Parkinson’s disease: A pathogenetically based approach,” *Neurobiol. Dis.*, vol. 39, no. 3, pp. 229–241, 2010.

[169] S. M. Rappaport, D. K. Barupal, D. Wishart, P. Vineis, and A. Scalbert, “The blood exposome and its role in discovering causes of disease,” *Environ. Health Perspect.*,

vol. 122, no. 8, pp. 769–774, 2014.

[170] M. Bogdanov *et al.*, "Metabolomic profiling to develop blood biomarkers for Parkinson's disease," *Brain*, vol. 131, no. 2, pp. 389–396, 2008.

[171] S. M. Rappaport, "Biomarkers intersect with the exposome," *Biomarkers*, vol. 17, no. 6, pp. 483–489, 2012.

[172] J. Nyhlén, R. Constantinescu, and H. Zetterberg, "Problems associated with fluid biomarkers for Parkinson's disease," *Biomark. Med.*, vol. 4, no. 5, pp. 671–681, Oct. 2010.

[173] N. P. Visanji, C. Marras, L. N. Hazrati, L. W. C. Liu, and A. E. Lang, "Alimentary, my dear Watson? The challenges of enteric α-synuclein as a Parkinson's disease biomarker," *Movement Disorders*, vol. 29, no. 4. pp. 444–450, 2014.

[174] I. Devic *et al.*, "Salivary α-synuclein and DJ-1: Potential biomarkers for Parkinson's disease," *Brain*, vol. 134, no. 7, 2011.

[175] D. Berg *et al.*, "Changing the research criteria for the diagnosis of Parkinson's disease: obstacles and opportunities.," *Lancet. Neurol.*, vol. 12, no. 5, pp. 514–24, May 2013.

[176] R. B. Postuma, J. F. Gagnon, M. Vendette, M. L. Fantini, J. Massicotte-Marquez, and J. Montplaisir, "Quantifying the risk of neurodegenerative disease in idiopathic REM sleep behavior disorder," *Neurology*, vol. 72, no. 15, pp. 1296–1300, 2009.

[177] A. J. Noyce, J. P. Bestwick, L. Silveira-Moriyama, C. H. Hawkes, and J. Hardy, "PREDICT-PD: Identifying risk of Parkinson's disease in the community: Methods and baseline results," *J Neurol Neurosurg Psychiatry*, pp. 1–7, 2013.

[178] I. Liepelt-Scarfone *et al.*, "Evaluation of progression markers in the premotor phase of Parkinson's Disease: The progression markers in the premotor phase study,"

Neuroepidemiology, vol. 41, no. 3–4, pp. 174–182, 2013.

[179] S. Lerche *et al.*, "Mild Parkinsonian signs in the elderly - Is there an association with PD? Crossectional findings in 992 individuals," *PLoS One*, vol. 9, no. 3, pp. 1–8, 2014.

[180] D. Berg, K. Marek, G. W. Ross, and W. Poewe, "Defining at-risk populations for Parkinson's disease: Lessons from ongoing studies," *Mov. Disord.*, vol. 27, no. 5, pp. 656–665, 2012.

[181] Imaging Specialists, "CT Scan - CAT Scan | Imaging Specialists of Charleston - MRI, Womens Imaging, Mammogram, Ultrasound, Digital x-ray, Radiology," 2015. [Online]. Available: http://www.imagingsc.com/ct-with-chest-shield/. [Accessed: 19-Nov-2015].

[182] R. Nehru, "Application of Radiation in Medicine and Research (Handout)." 2013.

[183] S. Hellwig *et al.*, "[18 F] FDG-PET is superior to [123 I IBZM-SPECT for the differential diagnosis of parkinsonism," *Neurology.*, vol. 79, no. 13, p. 1314–22., 2012.

[184] P. Piccini, D. J. Burn, R. Ceravolo, D. Maraganore, and D. J. Brooks, "The role of inheritance in sporadic Parkinson's disease: Evidence from a longitudinal study of dopaminergic function in twins," *Ann. Neurol.*, vol. 45, no. 5, pp. 577–582, 1999.

[185] N. Bajaj, R. A. Hauser, and I. D. Grachev, "Clinical utility of dopamine transporter single photon emission CT (DaT-SPECT) with (123 I) ioflupane in diagnosis of parkinsonian syndromes," *J. Neurol. Neurosurg. Psychiatry*, vol. 84, no. 11, pp. 1288–1295, 2013.

[186] H. Bernheimer, W. Birkmayer, O. Hornykiewicz, K. Jellinger, and F. Seitelberger, "Brain dopamine and the syndromes of Parkinson and Huntington Clinical, morphological and neurochemical correlations," *J. Neurol. Sci.*, vol. 20, no. 4, pp.

415–455, Dec. 1973.

[187] M. J. Kaufman and B. K. Madras, “Severe depletion of cocaine recognition sites associated with the dopamine transporter in Parkinson’s-diseased striatum,” *Synapse*, vol. 9, no. 1, pp. 43–49, 1991.

[188] H. T. S. Benamer *et al.*, “Prospective study of presynaptic dopaminergic imaging in patients with mild parkinsonism and tremor disorders: Part 1. Baseline and 3-month observations,” *Mov. Disord.*, vol. 18, no. 9, pp. 977–984, Sep. 2003.

[189] M. M. Ponsen, D. Stoffers, J. Booij, B. L. F. Van Eck-Smit, E. C. Wolters, and H. W. Berendse, “Idiopathic hyposmia as a preclinical sign of Parkinson’s disease,” *Ann. Neurol.*, vol. 56, no. 2, pp. 173–181, 2004.

[190] L. Filippi *et al.*, “123I-FP-CIT in progressive supranuclear palsy and in Parkinson’s disease: a SPECT semiquantitative study,” *Nucl. Med. Commun.*, vol. 27, no. 4, pp. 381–386, 2006.

[191] H. Matsui *et al.*, “Brain perfusion differences between Parkinson’s disease and multiple system atrophy with predominant parkinsonian features,” *Park. Relat. Disord.*, vol. 11, no. 4, pp. 227–232, 2005.

[192] V. L. Marshall *et al.*, “Parkinson’s disease is overdiagnosed clinically at baseline in diagnostically uncertain cases: A 3-year European multicenter study with repeat [123 I]FP-CIT SPECT,” *Mov. Disord.*, vol. 24, no. 4, pp. 500–508, Mar. 2009.

[193] M. Yoshita, “Differentiation of idiopathic Parkinson’s disease from striatonigral degeneration and progressive supranuclear palsy using iodine-123 meta-iodobenzylguanidine myocardial scintigraphy.,” *J. Neurol. Sci.*, vol. 155, no. 1, pp. 60–7, 1998.

[194] D. J. Brooks, “Parkinson’s disease: Diagnosis,” *Parkinsonism Relat. Disord.*, vol. 18, pp. S31–S33, 2012.

[195] H. T. S. Benamer *et al.*, "Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: The [123I]-FP-CIT study group," *Mov. Disord.*, vol. 15, no. 3, pp. 503–510, May 2000.

[196] L. Jaukovic, T. Ilic, M. Dopudja, and B. Ajdinovic, "123I-FP-CIT brain SPECT (DaTSCAN) imaging in the diagnosis of patients with movement disorders: First results," *Vojnosanit. Pregl.*, vol. 69, no. 2, pp. 157–162, 2012.

[197] F. Sixel-Döring, K. Liepe, B. Mollenhauer, E. Trautmann, and C. Trenkwalder, "The role of 123I-FP-CIT-SPECT in the differential diagnosis of Parkinson and tremor syndromes: a critical assessment of 125 cases," *J. Neurol.*, vol. 258, no. 12, pp. 2147–2154, Dec. 2011.

[198] E. Tolosa, T. Vander Borght, and E. Moreno, "Accuracy of DaTSCAN (123I-ioflupane) SPECT in diagnosis of patients with clinically uncertain parkinsonism: 2-Year follow-up of an open-label study," *Mov. Disord.*, vol. 22, no. 16, pp. 2346–2351, Oct. 2007.

[199] R. R. Gainetdinov, S. R. Jones, F. Fumagalli, R. M. Wightman, and M. G. Caron, "Re-evaluation of the role of the dopamine transporter in dopamine system homeostasis," *Brain Res. Rev.*, vol. 26, no. 2–3, pp. 148–153, May 1998.

[200] Y. Ma, V. Dhawan, M. Mentis, T. Chaly, P. G. Spetsieris, and D. Eidelberg, "Parametric mapping of [18F]FPCIT binding in early stage Parkinson's disease: A PET study," *Synapse*, vol. 45, no. 2, pp. 125–133, Aug. 2002.

[201] K. L. Marek *et al.*, "[sup 123 I] beta-CIT/SPECT imaging demonstrates bilateral loss of dopamine transporters in hemi-Parkinson's disease," *Neurology*, vol. 46, no. 1, pp. 231–237, Jan. 1996.

[202] British Nuclear Medicine Society, "BNMS - DAT Brain Scan," *British Nuclear Medicine Society*, 2016. [Online]. Available: https://www.bnms.org.uk/patients-information-sheets/dat-brain-scan.html. [Accessed: 08-Jan-2018].

[203] W. Poewe and C. Scherfler, “Role of dopamine transporter imaging in investigation of parkinsonian syndromes in routine clinical practice,” *Mov. Disord.*, vol. 18, no. S7, pp. S16–S21, Oct. 2003.

[204] L. D. Perju-Dumbrava *et al.*, “Dopamine transporter imaging in autopsy-confirmed Parkinson’s disease and multiple system atrophy,” *Mov. Disord.*, vol. 27, no. 1, pp. 65–71, Jan. 2012.

[205] R. A. Hauser and D. G. Grosset, “[123I]FP-CIT (DaTscan) SPECT Brain Imaging in Patients with Suspected Parkinsonian Syndromes,” *J. Neuroimaging*, vol. 22, no. 3, pp. 225–230, Jul. 2012.

[206] P. Manoharan, D. R. S. Jamieson, and R. F. Bury, “Initial clinical experience with [123I]ioflupane scintigraphy in movement disorders,” *Clin. Radiol.*, vol. 62, no. 5, pp. 463–471, May 2007.

[207] K. D. Seifert and J. I. Wiener, “The impact of DaTscan on the diagnosis and management of movement disorders: A retrospective study,” *Am. J. Neurodegener. Dis.*, vol. 2, no. 1, pp. 29–34, 2013.

[208] Vrije Universiteit Brussel, “SOPS - Onderzoek in Sociale Neuro / Research on Social Neuroscience,” 2015. [Online]. Available: http://www.vub.ac.be/EXTO/index.php?option=com_content&task=view&id=95. [Accessed: 19-Nov-2015].

[209] H. H. Jasper, “Report of the committee on methods of clinical examination in electroencephalography,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 10, no. 2, pp. 370–375, May 1958.

[210] NIMH, “Neuroimaging and Mental Illness: A Window Into the Brain,” pp. 1–8, 2010.

[211] A. van den Heuvel, “Think Hello | Next Nature Network,” 2009. .

[212] P. Mansfield and A. A. Maudsley, “Medical imaging by NMR.,” *Br. J. Radiol.*, vol. 50, pp. 188–194, 1977.

[213] K. Nightingale, “Behind the picture: The humble beginnings of MRI - insight | insight,” 2013. [Online]. Available: http://www.insight.mrc.ac.uk/2013/01/10/behind-the-picture-the-humble-beginnings-of-mri/. [Accessed: 19-Nov-2015].

[214] D. Berg, J. Godau, and U. Walter, “Transcranial sonography in movement disorders,” *Lancet Neurol.*, vol. 7, no. 11, pp. 1044–1055, 2008.

[215] P. J. Tuite, S. Mangia, and S. Michaeli, “Magnetic Resonance Imaging (MRI) in Parkinson’s Disease,” *J. Alzheimer’s Dis. Park.*, vol. 03, no. 03, pp. 1–5, 2013.

[216] C. Gaig and E. Tolosa, “When does Parkinson’s disease begin?,” *Mov. Disord.*, vol. 24, no. S2, 2009.

[217] J. Meara, B. K. Bhowmick, and P. Hobson, “Accuracy of diagnosis in patients with presumed Parkinson’s disease,” *Age Ageing*, vol. 28, no. 2, pp. 99–102, 1999.

[218] M. Rudin, “Molecular Imaging: Basic Principles and Applications in Biomedical Research,” in *Molecular Imaging*, IMPERIAL COLLEGE PRESS, 2013.

[219] R. A. Poldrack, J. A. Mumford, and T. E. Nichols, *Handbook of functional MRI data analysis*, vol. 4. Cambridge University Press, 2011.

[220] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, “Brain magnetic resonance imaging with contrast dependent on blood oxygenation.,” *Proc. Natl. Acad. Sci.*, vol. 87, no. 24, pp. 9868–9872, Dec. 1990.

[221] J. Carp, “The secret lives of experiments: methods reporting in the fMRI literature.,” *Neuroimage*, vol. 63, no. 1, pp. 289–300, Oct. 2012.

References

[222] V. Bruce and A. Young, *Face Perception*. Psychology Press, 2013.

[223] J. D. Greene, R. B. Sommerville, L. E. Nystrom, J. M. Darley, and J. D. Cohen, "An fMRI investigation of emotional engagement in moral judgment.," *Science*, vol. 293, no. 5537, pp. 2105–8, Sep. 2001.

[224] S. M. Smith *et al.*, "Correspondence of the brain's functional architecture during activation and rest.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 31, pp. 13040–5, 2009.

[225] M. M. Monti *et al.*, "Willful Modulation of Brain Activity in Disorders of Consciousness," *N. Engl. J. Med.*, vol. 362, no. 7, pp. 579–589, 2010.

[226] M. W. Woolrich *et al.*, "Bayesian analysis of neuroimaging data in FSL," *Neuroimage*, vol. 45, no. 1, pp. S173–S186, Mar. 2009.

[227] M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, and S. M. Smith, "FSL," *Neuroimage*, vol. 62, no. 2, pp. 782–790, Aug. 2012.

[228] S. M. Smith *et al.*, "Advances in functional and structural MR image analysis and implementation as FSL," *Neuroimage*, vol. 23, pp. S208–S219, Jan. 2004.

[229] K. J. Friston, A. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith, and R. S. J. Frackowiak, "Statistical parametric maps in functional imaging: A general linear approach," *Hum. Brain Mapp.*, vol. 2, no. 4, pp. 189–210, 1994.

[230] K. J. Friston, W. D. Penny, J. Ashburner, S. J. Kiebel, and T. Nichols, *Statistical Parametric Mapping: The Analysis of Funtional Brain Images*, 1st ed. Elsevier/Academic Press, 2006.

[231] J. Ashburner *et al.*, "Statistical Parametric Mapping toolbox." Wellcome Trust Centre for Neuroimaging, London, 2014.

[232] MR-TIP, “MRI Images - Brain MRI Images T1 - MR-TIP.com,” 2015. [Online]. Available: http://www.mr-tip.com/serv1.php?type=img&img=Brain MRI Images T1. [Accessed: 19-Nov-2015].

[233] A. Siddique, “Personalized Brain Mapping: Diffusion Tensor Imaging Limits Tumor Surgery Risk,” 2013. [Online]. Available: http://www.medicaldaily.com/personalized-brain-mapping-diffusion-tensor-imaging-limits-tumor-surgery-risk-244859. [Accessed: 19-Nov-2015].

[234] P. Adamczyk and D. S. Liebeskind, “Neuroimaging: Vascular Imaging: Computed Tomographic Angiography, Magnetic Resonance Angiography, and Ultrasound | Clinical Gate,” 2015. [Online]. Available: http://clinicalgate.com/neuroimaging-vascular-imaging-computed-tomographic-angiography-magnetic-resonance-angiography-and-ultrasound/. [Accessed: 19-Nov-2015].

[235] P. Pantano *et al.*, “A longitudinal fMRI study on motor activity in patients with multiple sclerosis,” *Brain*, vol. 128, no. 9, pp. 2146–2153, Aug. 2005.

[236] M. Ingalhalikar *et al.*, “Sex differences in the structural connectome of the human brain.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 2, pp. 823–8, 2014.

[237] M. H. Lee, C. D. Smyser, and J. S. Shimony, “Resting-state fMRI: a review of methods and clinical applications,” *AJNR Am J Neuroradiol*, vol. 34, no. 10, pp. 1866–1872, 2013.

[238] B. B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, “Functional connectivity in the motor cortex of resting human brain using echo-planar MRI.,” *Magn. Reson. Med.*, vol. 34, no. 4, pp. 537–541, 1995.

[239] M. J. Brookes *et al.*, “Investigating the electrophysiological basis of resting state networks using magnetoencephalography,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 40, pp. 16783–16788, Oct. 2011.

[240] A. Khazaee, A. Ebrahimzadeh, and A. Babajani-Feremi, “Identifying patients with Alzheimer’s disease using resting-state fMRI and graph theory,” *Clin. Neurophysiol.*, vol. 126, no. 11, pp. 2132–2141, 2015.

[241] J. S. Damoiseaux *et al.*, “Consistent resting-state networks across healthy subjects.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 37, pp. 13848–53, 2006.

[242] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, “Functional connectivity in the resting brain: a network analysis of the default mode hypothesis,” *Proc.Natl.Acad.Sci.U.S A*, vol. 100, no. 0027–8424 (Print) LA–eng PT–Journal Article PT–Research Support, Non–U.S. Gov’t PT–Research Support, U.S. Gov’t, P.H.S SB–IM, pp. 253–258, 2003.

[243] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, “A default mode of brain function.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 2, pp. 676–682, 2001.

[244] D. a Gusnard and M. E. Raichle, “Searching for a baseline: Functional imaging and the resting human brain,” *Nat. Rev. Neurosci.*, vol. 2, no. 10, pp. 685–694, Oct. 2001.

[245] M. de la Iglesia-Vaya, J. Molina-Mateo, M. J. Escarti-Fabra, A. S. Kanaan, and L. Martí-Bonmati, “Brain Connections – Resting State \nfMRI Functional Connectivity,” *Nov. Front. Adv. Neuroimaging*, 2013.

[246] R. L. Buckner and D. C. Carroll, “Self-projection and the brain,” *Trends Cogn. Sci.*, vol. 11, no. 2, pp. 49–57, Feb. 2007.

[247] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, “From The Cover: The human brain is intrinsically organized into dynamic, anticorrelated functional networks,” *Proc. Natl. Acad. Sci.*, vol. 102, no. 27, pp. 9673–9678, Jul. 2005.

[248] A. G. Garrity, K. McKiernan, D. Lloyd, K. A. Kiehl, and V. D. Calhoun, “Aberrant ‘Default Mode’ Functional Connectivity in Schizophrenia,” *Am J Psychiatry*, vol. 1643, no. March, pp. 450–457, 2007.

[249] M. Roy, D. Shohamy, N. Daw, M. Jepma, G. E. Wimmer, and T. D. Wager, “Representation of aversive prediction errors in the human periaqueductal gray,” *Nat. Neurosci.*, vol. 17, no. 11, pp. 1607–1612, 2014.

[250] I. Cribben, R. Haraldsdottir, L. Y. Atlas, T. D. Wager, and M. A. Lindquist, “Dynamic connectivity regression: Determining state-related changes in brain connectivity,” *Neuroimage*, vol. 61, no. 4, pp. 907–920, 2012.

[251] S. Whitfield-Gabrieli and A. Nieto-Castanon, “Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks,” *Brain Connect.*, vol. 2, no. 3, pp. 125–141, 2012.

[252] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: Uses and interpretations,” *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010.

[253] R. N. A. Henson, C. Büchel, O. Josephs, and K. J. Friston, “The slice-timing problem in event-related fMRI,” *Neuroimage*, vol. 9, p. 125, 1999.

[254] K. J. Friston, S. C. Williams, R. Howard, R. S. J. Frackowiak, and R. Turner, “Movement-Related effects in fMRI time-series,” *Magn. Reson. Med.*, vol. 35, no. 3, pp. 346–355, Mar. 1996.

[255] L. Muresan, R. Renken, J. B. T. M. Roerdink, and H. Duifhuis, “Automated Correction of Spin-History Related Motion Artefacts in fMRI: Simulated and Phantom Data,” *IEEE Trans. Biomed. Eng.*, vol. 52, no. 8, pp. 1450–1460, Aug. 2005.

[256] J. Rademacher *et al.*, “Probabilistic Mapping and Volume Measurement of Human Primary Auditory Cortex,” *Neuroimage*, vol. 13, no. 4, pp. 669–683, Apr. 2001.

[257] V. B. Penhune, R. J. Zatorre, J. D. MacDonald, and A. C. Evans, “Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans.,” *Cereb. Cortex*, vol. 6, no. 5, pp. 661–72, 1996.

[258] A. W. Toga and P. M. Thompson, “Genetics of Brain Structure and Intelligence,” *Annu. Rev. Neurosci.*, vol. 28, no. 1, pp. 1–23, Jul. 2005.

[259] J. Ashburner and K. J. Friston, “Unified segmentation,” *Neuroimage*, vol. 26, no. 3, pp. 839–851, 2005.

[260] L. P. Clarke *et al.*, “MRI segmentation: Methods and applications,” *Magn. Reson. Imaging*, vol. 13, no. 3, pp. 343–368, Jan. 1995.

[261] K. M. Petersson, T. E. Nichols, J.-B. Poline, and A. P. Holmes, “Statistical limitations in functional neuroimaging II. Signal detection and statistical inference,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 354, no. 1387, pp. 1261–1281, Jul. 1999.

[262] A. Andrade *et al.*, “Detection of fMRI activation using Cortical Surface Mapping,” *Hum. Brain Mapp.*, vol. 12, no. 2, pp. 79–93, Feb. 2001.

[263] K. M. Petersson, T. E. Nichols, J.-B. Poline, and A. P. Holmes, “Statistical limitations in functional neuroimaging. I. Non-inferential methods and statistical models,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 354, no. 1387, pp. 1239–1260, Jul. 1999.

[264] K. J. Worsley, M. Andermann, T. Koulis, D. MacDonald, and A. C. Evans, “Detecting changes in nonisotropic images,” *Hum. Brain Mapp.*, vol. 8, no. 2–3, pp. 98–101, 1999.

[265] K. J. Worsley, S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans, “A unified statistical approach for determining significant signals in images of cerebral activation,” *Hum. Brain Mapp.*, vol. 4, no. 1, pp. 58–73, 1996.

[266] K. J. Worsley, A. C. Evans, S. Marrett, and P. Neelin, “A Three-Dimensional Statistical Analysis for CBF Activation Studies in Human Brain,” *J. Cereb. Blood Flow Metab.*, vol. 12, no. 6, pp. 900–918, Nov. 1992.

[267] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications*, vol. 104. CRC press, 2005.

[268] K. J. Friston *et al.*, “Analysis of fMRI Time-Series Revisited,” *Neuroimage*, vol. 2, no. 1, pp. 45–53, Mar. 1995.

[269] A. Razi, J. Kahan, G. Rees, and K. J. Friston, “Construct validation of a DCM for resting state fMRI,” *Neuroimage*, vol. 106, pp. 1–14, Feb. 2015.

[270] A. R. McIntosh, “Towards a network theory of cognition,” *Neural Networks*, vol. 13, no. 8–9, pp. 861–870, Nov. 2000.

[271] K. J. Friston, “Models of Brain Function in Neuroimaging,” *Annu. Rev. Psychol.*, vol. 56, no. 1, pp. 57–87, Feb. 2005.

[272] K. A. Bollen, *Structural equations with latent variables*, vol. 10. New York: Wiley, 1989.

[273] P. M. Bentler and J. A. Stein, “Structural equation models in medical research,” *Stat. Methods Med. Res.*, vol. 1, no. 2, pp. 159–181, Aug. 1992.

[274] B. Shipley, *Cause and Correlation in Biology: A User’s Guide to Path Analysis, Structural Equations, and Causal Inference*, vol. 1, no. 2001. Cambridge: Cambridge University Press, 2000.

[275] W. D. Penny, K. E. Stephan, A. Mechelli, and K. J. Friston, “Modelling functional integration: a comparison of structural equation and dynamic causal models,” *Neuroimage*, vol. 23, no. SUPPL. 1, pp. S264–S274, Jan. 2004.

[276] J. D. Ramsey, S. J. Hanson, C. Hanson, Y. O. Halchenko, R. A. Poldrack, and C. Glymour, “Six problems for causal inference from fMRI,” *Neuroimage*, vol. 49, no. 2, pp. 1545–1558, Jan. 2010.

[277] F. M. Miezin, L. Maccotta, J. M. Ollinger, S. E. Petersen, and R. L. Buckner, “Characterizing the Hemodynamic Response: Effects of Presentation Rate, Sampling Procedure, and the Possibility of Ordering Brain Activity Based on Relative Timing,” *Neuroimage*, vol. 11, no. 6, pp. 735–759, Jun. 2000.

[278] O. David *et al.*, “Identifying Neural Drivers with Functional MRI: An Electrophysiological Validation,” *PLoS Biol.*, vol. 6, no. 12, pp. 2683–2697, Dec. 2008.

[279] P. Spirtes, C. N. Glymour, and R. Scheines, *Causation, prediction, and search*, 2nd ed., vol. 45, no. 3. Cambridge, Massachusetts: MIT press, 2000.

[280] J. Pearl, *Causality: Models, Reasoning, and Inference*, vol. 21, no. 4. Cambridge: Cambridge University Press, 2000.

[281] K. J. Friston, L. Harrison, and W. Penny, “Dynamic causal modelling,” *Neuroimage*, vol. 19, no. 4, pp. 1273–1302, Aug. 2003.

[282] K. J. Friston, A. Bastos, V. Litvak, K. E. Stephan, P. Fries, and R. J. Moran, “DCM for complex-valued data: Cross-spectra, coherence and phase-delays,” *Neuroimage*, vol. 59, no. 1, pp. 439–455, 2012.

[283] R. J. Moran, M. Symmonds, K. E. Stephan, K. J. Friston, and R. J. Dolan, “An in vivo assay of synaptic function mediating human cognition,” *Curr. Biol.*, vol. 21, no. 15, pp. 1320–1325, 2011.

[284] K. E. Stephan and K. J. Friston, “Analyzing effective connectivity with functional magnetic resonance imaging,” *Wiley Interdiscip. Rev. Cogn. Sci.*, vol. 1, no. 3, pp. 446–459, 2010.

[285] K. J. Friston, J. Kahan, B. Biswal, and A. Razi, “A DCM for resting state fMRI,” *Neuroimage*, vol. 94, pp. 396–407, Jul. 2014.

[286] F. J. Provost, T. Fawcett, and R. Kohavi, “The case against accuracy estimation for comparing induction algorithms.,” in *ICML*, 1998, vol. 98, pp. 445–453.

[287] U. Bhowan, M. Johnston, M. Zhang, and X. Yao, “Evolving diverse ensembles using genetic programming for classification with unbalanced data,” *IEEE Trans. Evol. Comput.*, vol. 17, no. 3, pp. 368–386, Jun. 2013.

[288] D. R. Cox, “The Regression Analysis of Binary Sequences,” *J. R. Stat. Soc.*, vol. 20, no. 2, pp. 215–242, 1958.

[289] E. Frank and I. H. Witten, “Generating accurate rule sets without global optimization,” *Proc. Fifteenth Int. Conf. Mach. Learn.*, pp. 144–151, 1998.

[290] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*. CRC press, 1984.

[291] J. R. Quinlan, *C4.5: programs for machine learning*. Elsevier, 1993.

[292] C. Cortes and V. Vapnik, “Support Vector Networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[293] D. H. Wolpert and W. G. Macready, “No free lunch theorems for optimization,” *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.

[294] T. G. Dietterich and G. Bakiri, “Solving Multiclass Learning Problems via Error-Correcting Output Codes,” *Jouranal Artifical Intell. Res.*, vol. 2, pp. 263–286, 1995.

[295] L. Breiman, “Bagging predictors,” *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.

[296] L. G. Valiant, “A theory of the learnable,” *Commun. ACM*, vol. 27, no. 11, pp. 1134–1142, Nov. 1984.

[297] M. Kearns and L. Valiant, “Cryptographic limitations on learning Boolean formulae and finite automata,” *J. ACM*, vol. 41, no. 1, pp. 67–95, Jan. 1994.

[298] Y. Freund and R. R. E. Schapire, “Experiments with a New Boosting Algorithm,” *Int. Conf. Mach. Learn.*, pp. 148–156, 1996.

[299] T. G. Dietterich, “Ensemble Methods in Machine Learning,” in *Multiple classifier systems*, 2000, pp. 1–15.

[300] L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[301] W. Banzhaf, P. Nordin, R. E. Keller, and F. D. Francone, *Genetic Programming: An Introduction*, vol. 1. Morgan Kaufmann San Francisco, 1998.

[302] J. R. Koza, *Genetic Programming II, Automatic Discovery of Reusable Subprograms*. MIT Press, Cambridge, MA, 1992.

[303] J. R. Koza, “Genetic programming,” *Encycl. Comput. Sci. Technol.*, no. 1989, pp. 1–26, 1997.

[304] V. Vemuri, “Artificial neural networks,” 1988.

[305] D. Niebur, “Artificial Neural Networks,” 1995.

[306] R. J. Schalkoff, *Artificial neural networks*, vol. 1. McGraw-Hill New York, 1997.

[307] J. J. Hopfield, “Artificial neural networks,” *IEEE Circuits Devices Mag.*, vol. 4, no. 5, pp. 3–10, 1988.

[308] S. Suzuki, “Artificial neural networks,” *Math. Neural Networks Model. Algorithms*

Appl., vol. 8, p. 349, 1997.

[309] M. Clerc, "Standard Particle Swarm Optimisation," *hal.archives-ouvertes.fr*, 2012.

[310] T. Zeugmann *et al.*, "Particle Swarm Optimization," in *Encyclopedia of Machine Learning*, Boston, MA: Springer US, 2011, pp. 760–766.

[311] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: optimization by a colony of cooperating agents," *IEEE Trans. Syst. Man Cybern. Part B*, vol. 26, no. 1, pp. 29–41, 1996.

[312] M. Dorigo, V. Maniezzo, A. Colorni, and M. Dorigo, "Positive Feedback as a Search Strategy," *Tech. Rep. 91-016*, no. June, pp. 1–20, 1991.

[313] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic*, vol. 4. Prentice hall New Jersey, 1995.

[314] A. Lotfi Zadeh, "FUZZY SETS," in *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*, vol. Volume 6, WORLD SCIENTIFIC, 1996, pp. 394–432.

[315] D. Dasgupta, "Using immunological principles in anomaly detection," *Proc. Artif. Neural Networks Eng. (ANNIE'96), St. Louis, USA*, 1996.

[316] J. O. Kephart, "A biologically inspired immune system for computers," in *Artificial Life IV: proceedings of the fourth international workshop on the synthesis and simulation of living systems*, 1994, pp. 130–139.

[317] S. Forrest, S. A. Hofmeyr, A. Somayaji, and T. A. Longstaff, "A sense of self for Unix processes," in *Proceedings 1996 IEEE Symposium on Security and Privacy*, 1996, pp. 120–128.

[318] S. Haykin, *Neural networks: A comprehensive foundation*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1994.

References

[319] A. M. Turing, “Intelligent machinery. report for national physical laboratory. reprinted in ince, dc (editor). 1992. mechanical intelligence: Collected works of am turing,” 1948.

[320] N. A. Barricelli, “Esempi numerici di processi di evoluzione,” *Methodos*, vol. 1, no. 21–22, pp. 45–68, 1954.

[321] J. H. Holland, “Genetic algorithms and the optimal allocation of trials,” *SIAM J. Comput.*, vol. 2, no. 2, pp. 88–105, 1973.

[322] J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1975.

[323] N. L. Cramer, “A representation for the adaptive generation of simple sequential programs,” in *Proceedings of the first international conference on genetic algorithms*, 1985, pp. 183–187.

[324] J. R. Koza, *Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems*. Stanford University, Department of Computer Science Stanford, CA, 1990.

[325] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT press, 1992.

[326] J. R. Koza, “Genetic programming II: Automatic discovery of reusable subprograms,” *Cambridge, MA, USA*, 1994.

[327] C. Darwin, *On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life*. J. Murray, 1859.

[328] A. E. Eiben and J. E. Smith, *Introduction to Evolutionary Computing*. 2015.

[329] R. Dawkins, *The selfish gene*, 40th anniv. Oxford university press, 2016.

[330] J. F. Miller, *Cartesian Genetic Programming*, 1st ed. Berlin, Heidelberg: Springer, 2011.

[331] J. F. Miller and P. Thomson, "Cartesian Genetic Programming," in *Proceedings of the Third European Conference on Genetic Programming (EuroGP)*, 2000, vol. 1820, pp. 121–132.

[332] W. Banzhaf, P. Nordin, R. Keller, and F. Francone, *Genetic programming: an introduction*. San Francisco: dpunkt.verlog; Morgan Kaufmann Publishers, Inc., 1998.

[333] R. Poli, W. B. Langdon, and N. F. Mcphee, *A Field Guide to Genetic Programing*. Published via http://lulu.com and freely available at http://www.gp-field-guide.org.uk, 2008.

[334] S. Silva and E. Costa, "Dynamic limits for bloat control in genetic programming and a review of past and current bloat theories," *Genet. Program. Evolvable Mach.*, vol. 10, no. 2, pp. 141–179, Jun. 2009.

[335] J. F. Miller, "What bloat? cartesian genetic programming on boolean problems," *Genet. Evol. Comput. Conf. Late Break. Pap.*, no. July, pp. 295–302, 2001.

[336] A. J. Turner and J. F. Miller, "Cartesian Genetic Programming: Why No Bloat?," in *Genetic Programming: 17th European Conference, EuroGP 2014*, M. Nicolau, K. Krawiec, M. I. Heywood, M. Castelli, P. García-Sánchez, J. J. Merelo, V. M. Rivas Santos, and K. Sim, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 222–233.

[337] V. Vassilev and J. F. Miller, "The advantages of landscape neutrality in digital circuit evolution," *Evolvable Syst. from Biol. to Hardw.*, vol. 6, 2000.

[338] A. J. Turner and J. F. Miller, "Recurrent Cartesian Genetic Programming Applied to Famous Mathematical Sequences," in *Proceedings of the Seventh York Doctoral*

Symposium on Computer Science & Electronics, 2014, pp. 37–46.

[339] A. J. Turner and J. F. Miller, "Introducing a cross platform open source Cartesian Genetic Programming library," *Genet. Program. Evolvable Mach.*, vol. 16, no. 1, pp. 83–91, Mar. 2015.

[340] A. J. Turner and J. F. Miller, "Recurrent Cartesian Genetic Programming," *Parallel Probl. Solving from Nat. - PPSN XIII*, vol. 8672, pp. 476–486, 2014.

[341] A.-L. Boulesteix and M. Schmid, "Machine learning versus statistical modeling," *Biometrical J.*, vol. 56, no. 4, pp. 588–593, Jul. 2014.

[342] D. H. Wolpert, "The Supervised Learning No-Free Lunch Theorems," *Proc. 6th Online World Conf. Soft Comput. Ind. Appl.*, vol. 50 Suppl, pp. 25–42, 2001.

[343] G. Shmueli, "To Explain or to Predict?," *Stat. Sci.*, vol. 25, no. 3, pp. 289–310, Aug. 2010.

[344] J. Kruppa *et al.*, "Probability estimation with machine learning methods for dichotomous and multicategory outcome: Theory," *Biometrical J.*, vol. 56, no. 4, pp. 534–563, Jul. 2014.

[345] J. P. Simmons, L. D. Nelson, and U. Simonsohn, "False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant," *Psychol. Sci.*, vol. 22, no. 11, pp. 1359–1366, Nov. 2011.

[346] A. J. Turner and J. F. Miller, "Cartesian genetic programming encoded artificial neural networks," in *Proceeding of the fifteenth annual conference on Genetic and evolutionary computation conference - GECCO '13*, 2013, p. 1005.

[347] M. A. Lones, J. E. Alty, J. Cosgrove, D. R. S. Jamieson, and S. L. Smith, "Going through directional changes: evolving human movement classifiers using an event based encoding," in *Proceedings of the Genetic and Evolutionary Computation*

Conference Companion on - GECCO '17, 2017, pp. 1365–1371.

[348] Xin Yao and Yong Liu, "Making use of population information in evolutionary artificial neural networks," *IEEE Trans. Syst. Man Cybern. Part B*, vol. 28, no. 3, pp. 417–425, Jun. 1998.

[349] M. Brameier and W. Banzhaf, "Evolving teams of predictors with linear genetic programming," *Genet. Program. Evolvable Mach.*, vol. 2, no. 4, pp. 381–407, 2001.

[350] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. New York, NY: Springer series in statistics Springer, Berlin, 2001.

[351] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *The International Joint Conference on Artificial Intelligence (IJCAI)*, 1995, vol. 14, no. 2, pp. 1137–1145.

[352] P. Burman, "A comparative study of ordinary cross-validation, v-fold cross-validati on and the repeated learning testing methods," *Biometrika*, vol. 76, no. 3, pp. 503–514, 1989.

[353] P. Burman, "Estimation of Optimal Transformations Using v-Fold Cross Validation and Repeated Learning-Testing Methods," *Indian J. Stat. Ser. A*, vol. 52, no. 3, pp. 314–345, 1990.

[354] S. Geman, E. Bienenstock, and R. Doursat, "Neural Networks and the Bias/Variance Dilemma," *Neural Comput.*, vol. 4, no. 1, pp. 1–58, Jan. 1992.

[355] W. B. Langdon and B. F. Buxton, "Genetic Programming for Improved Receiver Operating Characteristics," *Mult. Classif. Syst.*, no. July, pp. 68–77, 2001.

[356] H. A. Abbass, "A Memetic Pareto Evolutionary Approach to Artificial Neural Networks," *Aust. Jt. Conf. Artif. Intell.*, pp. 1–12, 2001.

[357] Z. Xing, J. Pei, and E. Keogh, “A brief survey on sequence classification,” *ACM SIGKDD Explor. Newsl.*, vol. 12, no. 1, p. 40, Nov. 2010.

[358] D. Y. Harvey and M. D. Todd, “Automated Feature Design for Numeric Sequence Classification by Genetic Programming,” *IEEE Trans. Evol. Comput.*, vol. 19, no. 4, pp. 474–489, Aug. 2015.

[359] A. Bagnall, L. Davis, J. Hills, and J. Lines, “Transformation Based Ensembles for Time Series Classification,” in *Proceedings of the 2012 SIAM International Conference on Data Mining*, vol. 12, Philadelphia, PA: Society for Industrial and Applied Mathematics, 2012, pp. 307–318.

[360] H. Deng, G. Runger, E. Tuv, and M. Vladimir, “A time series forest for classification and feature extraction,” *Inf. Sci. (Ny).*, vol. 239, pp. 142–153, Aug. 2013.

[361] B. D. Fulcher, M. A. Little, and N. S. Jones, “Highly comparative time-series analysis: the empirical structure of time series and their methods,” *J. R. Soc. Interface*, vol. 10, no. 83, pp. 20130048–20130048, Apr. 2013.

[362] K. C. Sharman, “Evolving signal processing algorithms by genetic programming,” in *1st International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications (GALESIA)*, 1995, vol. 1995, pp. 473–480.

[363] A. Teller and M. Veloso, “Program Evolution for Data Mining,” *Int. J. Expert Syst. 8-3*, p. 213–236, 1995.

[364] A. Teller and M. Veloso, “PADO: A New Learning Architecture for Object Recognition,” pp. 1–34, 1995.

[365] D. R. Eads *et al.*, “Genetic Algorithms and Support Vector Machines for Time Series Classification,” 2002, p. 74.

[366] S. Stepney, “Nonclassical Computation — A Dynamical Systems Perspective,” in *Handbook of Natural Computing*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 1979–2025.

[367] N. Jakobi, “Harnessing morphogenesis,” in *On Growth, Form and Computers*, Elsevier, 2003, pp. 392–404.

[368] W. Banzhaf, “Artificial Regulatory Networks and Genetic Programming,” in *Genetic Programming Theory and Practice*, R. Riolo and B. Worzel, Eds. Boston, MA: Springer US, 2003, pp. 43–61.

[369] M. A. Lones *et al.*, “Artificial biochemical networks: Evolving dynamical systems to control dynamical systems,” *IEEE Trans. Evol. Comput.*, vol. 18, no. 2, pp. 145–166, 2014.

[370] T. Quick, C. L. Nehaniv, K. Dautenhahn, and G. Roberts, “Network-Driven Control Systems Introduction: Genetic Regulatory Networks and,” in *Advances in Artificial Life*, 7th Europe., W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim, and J. Ziegler, Eds. New York: Springer-Verlag Berlin Heidelberg, 2003, pp. 266–277.

[371] M. A. Lones, A. M. Tyrrell, S. Stepney, and L. S. D. Caves, “Controlling Complex Dynamics with Artificial Biochemical Networks,” in *ECAL*, New York: Springer-Verlag Berlin Heidelberg, 2011, pp. 159–170.

[372] M. A. Lones *et al.*, “Discriminating normal and cancerous thyroid cell lines using implicit context representation Cartesian genetic programming,” *2010 IEEE World Congr. Comput. Intell. WCCI 2010 - 2010 IEEE Congr. Evol. Comput. CEC 2010*, 2010.

[373] R. J. Williams and D. Zipser, “A Learning Algorithm for Continually Running Fully Recurrent Neural Networks,” *Neural Comput.*, vol. 1, no. 2, pp. 270–280, Jun. 1989.

[374] K. J. Hunt, D. Sbarbaro, R. Żbikowski, and P. J. Gawthrop, “Neural networks for

control systems—A survey," *Automatica*, vol. 28, no. 6, pp. 1083–1112, Nov. 1992.

[375] M. Nicolau, M. Schoenauer, and W. Banzhaf, "Evolving Genes to Balance a Pole," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6021 LNCS, pp. 196–207, May 2010.

[376] C.-M. Kuan and T. Liu, "Forecasting exchange rates using feedforward and recurrent neural networks," *J. Appl. Econom.*, vol. 10, no. 4, pp. 347–364, Oct. 1995.

[377] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction," *IEEE Trans. Neural Networks*, vol. 5, no. 2, pp. 240–254, Mar. 1994.

[378] R. B. Allen and C. A. Kamm, "A Recurrent Neural Network for Word Identification from Continuous Phoneme Strings," *Adv. Neural Inf. Process. Syst. 3*, pp. 206–212, 1991.

[379] T. L. Burrows and M. Niranjan, "The use of recurrent neural networks for classification," in *Proceedings of IEEE Workshop on Neural Networks for Signal Processing*, 1994, pp. 117–125.

[380] M. Hüsken and P. Stagge, "Recurrent neural networks for time series classification," *Neurocomputing*, vol. 50, pp. 223–235, Jan. 2003.

[381] M. A. Lones, S. L. Smith, A. M. Tyrrell, J. E. Alty, and D. R. S. Jamieson, "Characterising neurological time series data using biologically motivated networks of coupled discrete maps," *BioSystems*, vol. 112, no. 2, pp. 94–101, 2013.

[382] R. B. Rao, S. Krishnan, and R. S. Niculescu, "Data mining for improved cardiac care," *ACM SIGKDD Explor. Newsl.*, vol. 8, no. 1, pp. 3–10, 2006.

[383] N. Japkowicz and S. Stephen, "The class imbalance problem: A systematic study,"

Intell. Data Anal., vol. 6, no. 5, pp. 429–449, 2002.

[384] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data," *ACM SIGKDD Explor. Newsl. - Spec. issue Learn. from imbalanced datasets*, vol. 6, no. 1, pp. 20–29, 2004.

[385] G. Weiss and F. Provost, "The effect of class distribution on classifier learning: an empirical study," *Rutgers Univ*, 2001.

[386] J. Laurikkala, "Improving identification of difficult small classes by balancing class distribution," *Proc. 8th Conf. AI Med. Eur. Artif. Intell. Med.*, pp. 63–66, 2001.

[387] A. Estabrooks, T. Jo, and N. Japkowicz, "A Multiple Resampling Method for Learning from Imbalanced Data Sets," *Comput. Intell.*, vol. 20, no. 1, pp. 18–36, 2004.

[388] R. C. Holte, L. Acker, and B. Porter, "Concept Learning and the Problem of Small Disjuncts.," *Ijcai*, pp. 813–818, 1989.

[389] D. Mease, A. J. Wyner, and A. Buja, "Boosted classification trees and class probability/quantile estimation," *J. Mach. Learn. Res.*, vol. 8, pp. 409–439, 2007.

[390] C. Drummond and R. C. Holte, "C4.5, class imbalance, and cost sensitivity: why under-sampling beats over-sampling," *Work. Learn. from Imbalanced Datasets II*, pp. 1–8, 2003.

[391] X.-Y. Liu, J. Wu, and Z.-H. Zhou, "Exploratory Undersampling for Class Imbalance Learning," *IEEE Trans. Syst. Man Cybern.*, vol. 39, no. 2, pp. 539–550, 2009.

[392] J. Zhang and I. Mani, "kNN Approach to Unbalanced Data Distributions: A Case Study involving Information Extraction," *Work. Learn. from Imbalanced Datasets II ICML Washingt. DC 2003*, pp. 42–48, 2003.

References

[393] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

[394] B. Wang and N. Japkowicz, "Imbalanced Data Set Learning with Synthetic Samples," *Proc. IRIS Mach. Learn. Work.*, pp. 1–25, 2004.

[395] H. Han, W. Wang, and B. Mao, "Borderline-SMOTE : A New Over-Sampling Method in," pp. 878–887, 2005.

[396] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning," in *Proceeding of the International Joint Conference on Neural Networks*, 2008, pp. 1322–1328.

[397] J. Han and M. Kamber, *Data mining: Concepts and Techniques*, Second Edi. San Francisco: Elsevier Inc., 2006.

[398] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*, First Edit. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2005.

[399] J. Dukart, K. Mueller, H. Barthel, A. Villringer, O. Sabri, and M. L. Schroeter, "Meta-analysis based SVM classification enables accurate detection of Alzheimer's disease across different clinical centers using FDG-PET and MRI," *Psychiatry Res. - Neuroimaging*, vol. 212, no. 3, pp. 230–236, 2013.

[400] D. Zhang, Y. Wang, L. Zhou, H. Yuan, D. Shen, and Alzheimer's Disease Neuroimaging Initiative, "Multimodal Classification of Alzheimer's Disease and Mild Cognitive Impairment," *Neuroimage*, vol. 55, no. 3, pp. 856–867, 2011.

[401] S. L. Smith and M. A. Lones, "Implicit Context Representation Cartesian Genetic Programming for the assessment of visuo-spatial ability," in *2009 IEEE Congress on Evolutionary Computation*, 2009, pp. 1072–1078.

[402] R. Benecke, J. C. Rothwell, J. P. R. Dick, B. L. Day, and C. D. Marsden, "Disturbance of sequential movements in patients with Parkinson's disease," *Brain*, vol. 110, no. 2, pp. 361–379, 1987.

[403] R. Agostino, A. Berardelli, A. Formica, N. Accornero, and M. Manfredi, "Sequential Arm Movements in Patients With Parkinson's Disease, Huntington's Disease and Dystonia," *Brain*, vol. 115, no. 5, pp. 1481–1495, 1992.

[404] R. Agostino, A. Currà, M. Giovannelli, N. Modugno, M. Manfredi, and A. Berardelli, "Impairment of individual finger movements in Parkinson's disease," *Mov. Disord.*, vol. 18, no. 5, pp. 560–565, May 2003.

[405] R. Agostino, A. Berardelli, A. Currà, N. Accornero, and M. Manfredi, "Clinical impairment of sequential finger movements in Parkinson's disease," *Mov. Disord.*, vol. 13, no. 3, pp. 418–421, May 1998.

[406] M. A. Lones *et al.*, "Evolving classifiers to inform clinical assessment of Parkinson's disease," *Proc. 2013 IEEE Symp. Comput. Intell. Healthc. e-Health, CICARE 2013 - 2013 IEEE Symp. Ser. Comput. Intell. SSCI 2013*, pp. 76–82, 2013.

[407] C. Picardi, J. Cosgrove, S. L. Smith, D. R. S. Jamieson, and J. E. Alty, "Objective Assessment of Cognitive Impairment in Parkinson's Disease Using Evolutionary Algorithm," in *Applications of Evolutionary Computation: 20th European Conference, EvoApplications 2017, Amsterdam, The Netherlands, April 19-21, 2017, Proceedings, Part I*, G. Squillero and K. Sim, Eds. Cham: Springer International Publishing, 2017, pp. 109–124.

[408] M. A. Lones and S. L. Smith, "Medical Applications of Evolvable Hardware," pp. 253–271, 2015.

[409] M. A. Lones *et al.*, "Evolving classifiers to recognize the movement characteristics of parkinson's disease patients," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 559–576, 2014.

[410] A. J. Manson, P. Brown, J. D. O'Sullivan, P. Asselman, D. Buckwell, and A. J. Lees, "An ambulatory dyskinesia monitor," *J. Neurol. Neurosurg. Psychiatry*, vol. 68, no. 2, pp. 196–201, Feb. 2000.

[411] A. Berardelli, J. C. Rothwell, P. D. Thompson, and M. Hallett, "Pathophysiology of bradykinesia in Parkinson's disease," *Brain*, vol. 124, no. 11, pp. 2131–2146, Nov. 2001.

[412] J. I. Hoff, A. A. v/d Plas, E. A. H. Wagemans, and J. J. van Hilten, "Accelerometric assessment of levodopa-induced dyskinesias in Parkinson's disease," *Mov. Disord.*, vol. 16, no. 1, pp. 58–61, Jan. 2001.

[413] P. Bonato, D. M. Sherrill, D. G. Standaert, S. S. Salles, and M. Akay, "Data mining techniques to detect motor fluctuations in Parkinson's disease," in *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2004, vol. 4, pp. 4766–4769.

[414] R. Iansek, F. Huxham, and J. McGinley, "The sequence effect and gait festination in Parkinson disease: Contributors to freezing of gait?," *Mov. Disord.*, vol. 21, no. 9, pp. 1419–1424, Sep. 2006.

[415] R. Chee, A. Murphy, M. Danoudis, N. Georgiou-Karistianis, and R. Iansek, "Gait freezing in Parkinson's disease and the stride length sequence effect interaction," *Brain*, vol. 132, no. 8, pp. 2151–2160, Aug. 2009.

[416] J. P. Rodrigues, F. L. Mastaglia, and G. W. Thickbroom, "Rapid slowing of maximal finger movement rate: fatigue of central motor control?," *Exp. Brain Res.*, vol. 196, no. 4, pp. 557–563, Jul. 2009.

[417] M. Yokoe, R. Okuno, T. Hamasaki, Y. Kurachi, K. Akazawa, and S. Sakoda, "Opening velocity, a novel parameter, for finger tapping test in patients with Parkinson's disease," *Parkinsonism Relat. Disord.*, vol. 15, no. 6, pp. 440–444, Jul. 2009.

[418] S. Patel *et al.*, “Monitoring Motor Fluctuations in Patients With Parkinson’s Disease Using Wearable Sensors,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 864–873, Nov. 2009.

[419] A. J. Espay *et al.*, “Differential response of speed, amplitude, and rhythm to dopaminergic medications in Parkinson’s disease,” *Mov. Disord.*, vol. 26, no. 14, pp. 2504–2508, Dec. 2011.

[420] H. Ling, L. A. Massey, A. J. Lees, P. Brown, and B. L. Day, “Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson’s disease,” *Brain*, vol. 135, no. 4, pp. 1141–1153, Apr. 2012.

[421] N. M. Aly, J. R. Playfer, S. L. Smith, and D. M. Halliday, “A novel computer-based technique for the assessment of tremor in Parkinson’s disease,” *Age Ageing*, vol. 36, no. June, pp. 395–399, 2007.

[422] S. Y. Kang *et al.*, “Characteristics of the sequence effect in Parkinson’s disease,” *Mov. Disord.*, vol. 25, no. 13, pp. 2148–2155, Oct. 2010.

[423] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, “Novel Speech Signal Processing Algorithms for High-Accuracy Classification of Parkinson’s Disease,” *IEEE Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1264–1271, May 2012.

[424] N. L. W. Keijsers, M. W. I. M. Horstink, and S. C. A. M. Gielen, “Automatic assessment of levodopa-induced dyskinesias in daily life by neural networks.,” *Mov. Disord.*, vol. 18, no. 1, pp. 70–80, Jan. 2003.

[425] S. Patel *et al.*, “Home monitoring of patients with Parkinson’s disease via wearable technology and a web-based application,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010, vol. 2010, pp. 4411–4414.

[426] A. Ericsson, M. N. Lonsdale, K. Astrom, L. Edenbrandt, and L. Friberg, “Decision

Support System for the Diagnosis of Parkinson's Disease," in *Scandinavian Conference on Image Analysis*, vol. 3540, Springer, 2005, pp. 740–749.

[427] M. A. Lones, J. E. Alty, P. Duggan-Carter, A. J. Turner, D. R. S. Jamieson, and S. L. Smith, "Classification and characterisation of movement patterns during levodopa therapy for Parkinson's disease," in *Proceedings of the 2014 conference companion on Genetic and evolutionary computation companion - GECCO Comp '14*, 2014, pp. 1321–1328.

[428] B. B. Biswal, J. Van Kylen, and J. S. Hyde, "Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps.," *NMR Biomed.*, vol. 10, no. 4–5, pp. 165–170, 1997.

[429] J. S. Damoiseaux and M. D. Greicius, "Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity," *Brain Struct. Funct.*, vol. 213, no. 6, pp. 525–533, 2009.

[430] R. Esposito *et al.*, "Acute Effects of Modafinil on Brain Resting State Networks in Young Healthy Subjects," *PLoS One*, vol. 8, no. 7, pp. 1–10, 2013.

[431] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, "The brain's default network: Anatomy, function, and relevance to disease," *Ann. N. Y. Acad. Sci.*, vol. 1124, pp. 1–38, 2008.

[432] L. K. Ferreira and G. F. Busatto, "Resting-state functional connectivity in normal brain aging," *Neurosci. Biobehav. Rev.*, vol. 37, no. 3, pp. 384–400, 2013.

[433] A. S. Fleisher, A. Sherzai, C. Taylor, J. B. S. Langbaum, and R. B. Buxton, "Resting-state BOLD networks versus task-associated functional MRI for distinguishing Alzheimer's disease risk groups," *Neuroimage*, vol. 47, no. 4, pp. 1678–1690, 2009.

[434] G. Deco, V. K. Jirsa, and A. R. McIntosh, "Emerging concepts for the dynamical

organization of resting-state activity in the brain.," *Nat. Rev. Neurosci.*, vol. 12, no. 1, pp. 43–56, 2011.

[435] D. Mantini, M. G. Perrucci, G. C. Del, G. L. Romani, and M. Corbetta, "Electrophysiological signatures of resting state networks in the human brain," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 32, pp. 13170–13175, 2007.

[436] M. P. van den Heuvel and H. E. Hulshoff Pol, "Exploring the brain network: A review on resting-state fMRI functional connectivity," *Eur. Neuropsychopharmacol.*, vol. 20, no. 8, pp. 519–534, 2010.

[437] N. Cera, A. Tartaro, and S. L. Sensi, "Modafinil alters intrinsic functional connectivity of the right posterior insula: A pharmacological resting state fMRI study," *PLoS One*, vol. 9, no. 9, pp. 1–12, 2014.

[438] R. C. Oldfield, "The assessment and analysis of handedness: The Edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, 1971.

[439] B. Sen *et al.*, "Classification of Obsessive-Compulsive Disorder from Resting-State fMRI," pp. 3606–3609, 2016.

[440] Y.-W. Sun *et al.*, "Abnormal functional connectivity in patients with vascular cognitive impairment, no dementia: A resting-state functional magnetic resonance imaging study," *Behav. Brain Res.*, vol. 223, no. 2, pp. 388–394, 2011.

[441] M. R. Arbabshirani, E. Castro, and V. D. Calhoun, "Accurate classification of schizophrenia patients based on novel resting-state fMRI features," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, vol. 2014, pp. 6691–6694.

[442] J. Kahan *et al.*, "Resting state functional MRI in Parkinson's disease: The impact of deep brain stimulation on 'effective' connectivity," *Brain*, vol. 137, no. 4, pp. 1130–1144, 2014.

[443] W. A. Chaovalitwongse, R. S. Pottenger, S. Wang, Y. J. Fan, and L. D. Iasemidis, “Pattern- and network-based classification techniques for multichannel medical data signals to improve brain diagnosis,” *IEEE Trans. Syst. Man, Cybern. Part ASystems Humans*, vol. 41, no. 5, pp. 977–988, 2011.

[444] L. Guo, D. Rivero, and A. Pazos, “Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks,” *J. Neurosci. Methods*, vol. 193, no. 1, pp. 156–163, 2010.

[445] J. V. Marcos, R. Hornero, D. Álvarez, F. del Campo, C. Zamarrón, and M. López, “Utility of multilayer perceptron neural network classifiers in the diagnosis of the obstructive sleep apnoea syndrome from nocturnal oximetry,” *Comput. Methods Programs Biomed.*, vol. 92, no. 1, pp. 79–89, 2008.

[446] J. Frank, S. Mannor, and D. Precup, “Activity and Gait Recognition with Time-Delay Embeddings Time-Delay Embeddings,” *AAAI Conf. Artif. Intell.*, pp. 407–408, 2010.

[447] H. He and E. A. Garcia, “Learning from imbalanced data,” *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, 2009.

[448] F. Stocchi and C. W. Olanow, “Obstacles to the development of a neuroprotective therapy for Parkinson’s disease.,” *Mov. Disord.*, vol. 28, no. 1, pp. 3–7, Jan. 2013.

[449] D. Dima, K. E. Stephan, J. P. Roiser, K. J. Friston, and S. Frangou, “Effective Connectivity during Processing of Facial Affect: Evidence for Multiple Parallel Pathways,” *J. Neurosci.*, vol. 31, no. 40, pp. 14378–14385, 2011.

[450] P. Allen *et al.*, “Cingulate activity and fronto-temporal connectivity in people with prodromal signs of psychosis,” *Neuroimage*, vol. 49, no. 1, pp. 947–955, 2010.

[451] R. J. Moran *et al.*, “Alterations in brain connectivity underlying beta oscillations in parkinsonism,” *PLoS Comput. Biol.*, vol. 7, no. 8, 2011.

[452] A. K. Rehme, S. B. Eickhoff, L. E. Wang, G. R. Fink, and C. Grefkes, "Dynamic causal modeling of cortical activity from the acute to the chronic stage after stroke," *Neuroimage*, vol. 55, no. 3, pp. 1147–1158, 2011.

[453] A. Razi *et al.*, "Large-scale DCMs for resting state fMRI," *Netw. Neurosci.*, pp. 1–41, 2017.

[454] L. Breiman, "Statistical Modeling: The Two Cultures," *Stat. Sci.*, vol. 16, no. 3, pp. 199–231, 2001.

[455] R. Caruana, "Learning from imbalanced data: Rank metrics and extra tasks," *Proc. Am. Assoc. Artif. Intell. Conf.*, no. January 2000, pp. 51–57, 2000.

[456] H. A. Piwowar and T. J. Vision, "Data reuse and the open data citation advantage," *PeerJ*, vol. 1, p. e175, 2013.

[457] M. F. Dirkx *et al.*, "The Cerebral Network of Parkinson's Tremor: An Effective Connectivity fMRI Study," *J. Neurosci.*, vol. 36, no. 19, pp. 5362–5372, 2016.

[458] D. Long *et al.*, "Automatic Classification of Early Parkinson's Disease with Multi-Modal MR Imaging," *PLoS One*, vol. 7, no. 11, pp. 1–9, 2012.

[459] N. F. Mcphee, R. Poli, and W. B. Langdon, *Field Guide to Genetic Programming*. Published via http://lulu.com and freely available at http://www.gp-field-guide.org.uk, 2008.

[460] S. M. Winkler, M. Affenzeller, and S. Wagner, "Using enhanced genetic programming techniques for evolving classifiers in the context of medical diagnosis," *Genet. Program. Evolvable Mach.*, vol. 10, no. 2, pp. 111–140, Jun. 2009.

[461] T. K. Paul and H. Iba, "Prediction of Cancer Class with Majority Voting Genetic Programming Classifier Using Gene Expression Data," *IEEE/ACM Trans. Comput.*

Biol. Bioinforma., vol. 6, no. 2, pp. 353–367, Apr. 2009.

[462] M. Zhang and P. Wong, "Genetic programming for medical classification: a program simplification approach," *Genet. Program. Evolvable Mach.*, vol. 9, no. 3, pp. 229–255, Sep. 2008.

[463] Y. Sun, A. K. C. Wong, M. Kamel, and S., "Classification of Imbalanced Data: a Review," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 23, no. 04, pp. 687–719, 2009.

[464] G. M. Weiss, "Mining with Rarity: A Unifying Framework," *SIGKDD Explor.*, vol. 6, no. 1, pp. 7–19, 2004.

[465] A. Fernández, S. García, J. Luengo, E. Bernadó-Mansilla, and F. Herrera, "Genetics-based machine learning for rule induction: State of the art, taxonomy, and comparative study," *IEEE Trans. Evol. Comput.*, vol. 14, no. 6, pp. 913–941, 2010.

[466] W. Koch *et al.*, "Diagnostic power of default mode network resting state fMRI in the detection of Alzheimer's disease," *Neurobiol. Aging*, vol. 33, no. 3, pp. 466–478, 2012.

[467] M. D. Greicius, G. Srivastava, A. L. Reiss, and V. Menon, "Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 13, pp. 4637–42, 2004.

[468] L. Q. Uddin *et al.*, "Network homogeneity reveals decreased integrity of default-mode network in ADHD," *J. Neurosci. Methods*, vol. 169, no. 1, pp. 249–254, 2008.

[469] E. B. Liddle *et al.*, "Task-related default mode network modulation and inhibitory control in ADHD: Effects of motivation and methylphenidate," *J. Child Psychol. Psychiatry Allied Discip.*, vol. 52, no. 7, pp. 761–771, 2011.

[470] Y. I. Sheline *et al.*, "The default mode network and self-referential processes in

depression," *Proc. Natl. Acad. Sci.*, vol. 106, no. 6, pp. 1942–1947, 2009.

[471] L. L. Zeng *et al.*, "Identifying major depression using whole-brain functional connectivity: A multivariate pattern analysis," *Brain*, vol. 135, no. 5, pp. 1498–1507, 2012.

[472] R. Bluhm *et al.*, "Resting state default-mode network connectivity in early depression using a seed region-of-interest analysis: Decreased connectivity with caudate nucleus," *Psychiatry Clin. Neurosci.*, vol. 63, no. 6, pp. 754–761, 2009.