2021-06-10

The relationship of genetics with cognitive and behavioral impairments in idiopathic Parkinson’s disease patients

Ramezani, Mehrafarin

UNIVERSITY OF CALGARY

The relationship of genetics with cognitive and behavioral impairments in idiopathic Parkinson's disease patients

by

Mehrafarin Ramezani

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN NEUROSCIENCE

CALGARY, ALBERTA

JUNE, 2021

To My Loving Parents

Acknowledgement

This thesis would not be possible without the help of my supervisor Dr. Oury Monchi. Thank you very much for offering me this unique opportunity. I appreciate your support, trust and generous guidance throughout this project. It was a great honour for me to be part of your team.

A special thanks is offered to my co-supervisor Dr. Gerald Pfeffer. Thank you for your advice and support during my project. Many thanks also for offering me to work in your laboratory and guiding me through my experiments.

I would like to offer my wholehearted gratitude to my excellent supervisory committee for your invaluable insight, feedback, support and guidance throughout this work, namely, Prof. Bruce Pike, Prof. Paul Arnold and last but not least Prof. Davide Martino.

I would like to express my appreciation to Dr. Zahinoor Ismail for offering me a chance to collaborate with his team. Thank you for your generous support and guidance throughout the mild behavioral impairment projects.

My sincere gratitude to Dr. Nils Forkert for offering me a collaboration opportunity. I appreciate your support, rigour and guidance through the machine learning project.

I would like to convey my warm thanks to Dr. Ziv Gan-Or and his team for collaborating on the genetic analysis.

A special thanks to Dr. Iris Kathol, the amazing laboratory manager of the PCAN lab. I could not imagine working with anyone better than you Iris. I am lost for words for your support and kindness throughout my project.

My warmest gratitude to Zainul Javer, the excellent research coordinator of CaPRI. Thank you for the diligence, patience and help you offered me throughout my projects.

I would like to express my appreciation to the wonderful research nurse of the PCAN lab, Tracy Hammer. Thank you for all you have taught me and for your excellent research skills.

A special thank you to the talented psychometrist of the PCAN lab, Mekale Kibreab. Thank you for all the hours of data preparation for the PCAN team.

My wholehearted gratitude to the former and current members of the Hotchkiss Brain Institute and to the PCAN Lab team members for your help, support and guidance: Dr. Eunjin Yoon,

Dr. Jean-Sebastien Provost, Dr. Hannes Almgren, Meng Wang, Lesley Towill, Pauline Mouches, Matthew Joel, Dr. Arash Aghamohammadi Sereshki, Kristina Martens, Jennifer Ruskey and all the people who directly and indirectly contributed to this work.

I would like to thank the Parkinson Association of Alberta for funding my project and the amazing team of PAA, in particular, in Calgary, which I had the honor of volunteering with.

My deepest gratitude goes to my lovely family, especially my kind and supportive parents and my amazing nieces. I am forever thankful for having such incredible parents, you can always brighten my gloomiest days. Thank you for being there for me, I love you so much.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADHD = Attention Deficit Hyperactivity Disorder

AD = Alzheimer's Disease

ALFF = Amplitude of Low-Frequency Fluctuations

ANCOVA = Analysis of Covariance

APOE = Apolipoprotein

BAI = Beck Anxiety Inventory

BAM = Binary Alignment/Map

BDI = Beck Depression Inventory

BDNF = Brain-Derived-Neurotrophic-Factor

BNT = Boston Naming Test

bp = Base pair

CamPaIGN = Cambridgeshire Parkinson's Incidence from GP to Neurologist

CDT = Clock Drawing Test

CHCR1 = Carbonyl Reductase 1

CNS = Central Nervous System

COMT = Catechol-O-Methyltransferase

CSF = Cerebrospinal Fluid

DAT1 = Dopamine Transporter

DLB = Dementia with Lewy Body

DNA = Deoxyribonucleic Acid

DRS-2 = Dementia Rating Scale-2

DV = Dependent Variable

EHI = Edinburgh Handedness Inventory

fMRI = Functional Magnetic Resonance Imaging

GATK = Genome Analysis Tool Kit

GBA1 = Glucocerebrosidase

GCA = Gaussian Classifier Atlas

GDS = Geriatric Depression Scale

GEE = Generalized Estimating Equation

GLM = General Linear Model

GWAS = Genome Wide Association Study

HC = Healthy Control

HVLT-R = Hopkins Verbal Learning Test-Revisited

HVOT = Hopkins Visual Organizational Test

indel = Insertion-deletion

iPD = Idiopathic Parkinson's Disease

IR-FSPGR = Inversion Recovery Prepped Fast Spoiled Gradient Echo

IV = Independent Variable

LB = Lewy Body

LD = Linkage Disequilibrium

LEDD = Levodopa Equivalency Daily Dosage

LONI = Laboratory of Neuro Imaging

LRRK2 = Leucine Rich Repeat Kinase 2

LTD = Long-Term Depression

MAOB = Mono Amine Oxidase B

MAPT = Microtubule Associated Protein Tau

MBI = Mild Behavioral Impairment

MBI-C = Mild Behavioral Impairment-Checklist

MCI = Mild Cognitive Impairment

MDS = Movement Disorder Society

Met = Methionine

MGB = Minor Grove Binder

miRNA = Micro-Ribonucleic Acid

ML = Machine Learning

MMSE = Mini Mental State Examination

MoCA = Montreal Cognitive Assessment

MRI = Magnetic Resonance Imaging

MSA = Multiple System Atrophy

M-W U = Mann-Whitney U test

NPS = Neuropsychiatric Symptoms

PARK2 = Parkin RBR E3 Ubiquitin Protein Ligase

PARK6 = PTEN Induced Kinase 1

PARK7 = Parkinsonism Associated Deglycase

PCA = Principal Component Analysis

PCR = Polymerase Chain Reaction

PD = Parkinson's Disease

PDD = Parkinson's Disease with Dementia

PD-MCI = Parkinson's Disease with Mild Cognitive Impairment

PPMI = Parkinson's Progression Marker Initiative

PTSD = Post Traumatic Stress Disorder

QUIP = Questionnaire of Impulsive/Compulsive Disorders

RBC = Red Blood Cells

RBD = REM Sleep Behavioral Disorder

RCFT = Rey Complex Figure Test

REM = Rapid Eye Movement

RF = Random Forest

ROI = Region of Interest

RPM = Revolutions Per Minute

SLC6A3 = Solute Carrier Family 6 Member 3

SNARE = NSF Attachment Protein Receptor

SNCA = Synuclein Alpha

SNP = Single Nucleotide Polymorphism

SNV = Single Nucleotide Variant

ssDNA = Single Stranded

SVR = Support Vector Regression

SVM = Support Vector Machine

STAI = State-Trait Anxiety Inventory

Taq Polymerase = Thermus Aquatics Polymerase

TRKB = Tyrosine Kinase Receptor B

UPDRS-III = Unified Parkinson's Disease Rating Scale Part III

UTR = Untranslated Region

Val = Valine

VCF = Variant Call Format

VNTR = Variable Number Tandem Repeat

WAIS-IV = Wechsler Adult Intelligence Scale-IV

WMS-IV = Wechsler Memory Scale-IV

Abstract

Parkinson's disease (PD) is currently characterized by cardinal motor symptoms of rigidity, tremor and bradykinesia. However, this disease is far from a mere movement disorder, and non-motor symptoms have a substantial adverse effect on the quality of life for PD patients. PD patients suffer from a broad range of non-motor symptoms but two of them are the focus of this thesis, neuropsychiatric and cognitive impairments. The true cause of these symptoms' manifestation is still unknown, but it is speculated that they might originate from the extensive neuronal damage due to PD. The investigation of genetic variants associated with these non-motor symptoms can provide valuable information on the possible causes of non-motor symptoms, their prevention, and even their treatment.

In this thesis, the association of some specified genetic variants were investigated with neuropsychiatric symptoms and cognitive impairments in idiopathic PD patients (iPD). In the ` part, mild behavioral impairment (MBI) in iPD patients was investigated using the mild behavioral impairment checklist (MBI-C). In chapter 3, the association of MBI and rs6265 in the Brain-Derived-Neurotrophic-Factor (*BDNF*) was studied and it was shown that the Met allele for this variant was linked to higher risk of MBI. It was observed that the Met allele was associated with specific neuropsychiatric symptoms related to emotional dysregulation and distorted thoughts. In chapter 4, the relationship between rs4680 in Catechol-O-methyltransferase (*COMT*) and rs28363170 in Solute carrier family 3 member 6 (*SLC6A3*) with MBI was explored in iPD patients. These two variants are both pertinent in the regulation of dopamine availability in the frontal lobe. No association was found for either of these variants with MBI in iPD patients.

In the second part, the association of a specific variant rs894280 in Synuclein-alpha (*SNCA*) and mild cognitive impairment (MCI) in iPD patients was explored. In chapter 5, a machine learning analysis was used to predict cognition in iPD patients. rs894280 was ranked as the second most important feature for prediction of cognition in PD patients. The post-hoc analysis demonstrated a connection between this variant and overall cognition, attention and visuo-spatial abilities in iPD patients. This variant was further investigated in chapter 6 using longitudinal data from the Parkinson's Progression Marker Initiative (PPMI) dataset. It was shown that rs894280 was linked to the cognitive status of drug-naïve iPD patients at baseline. This variant was associated with the rate of MCI conversion in iPD patients longitudinally, but no association was found between the variant and neuropsychiatric symptoms.

In conclusion, this thesis presents genetic association with two prominent non-motor symptoms in iPD patients. These findings can be used to assist identification of PD patients at risk of cognitive decline or neuropsychiatric impairments. Nonetheless, further studies should be conducted to elucidate and validate the results of this thesis further. Limitations of the present projects and a framework for future studies are discussed in the last chapter.

1 Chapter One: Introduction

1.1 Parkinson's Disease

Parkinson's disease (PD) is a chronic and progressive disease affecting primarily older adults. It is a common neurodegenerative disease affecting about 6% of the Canadian adult population over the age of 65 [1]. PD is classically diagnosed based on motor symptoms including tremor, rigidity, bradykinesia, postural instability, and abnormal gait. The first clinical description of the symptoms of PD was made by James Parkinson in 1817 [2, 3].

PD is accompanied by the loss of dopaminergic neurons in the substantia nigra, part of the basal ganglia located in the midbrain. However, the loss of dopaminergic neurons initiates decades prior to the manifestation of motor symptoms. By the time motor features of PD emerge, a large proportion of dopaminergic neurons are lost leading to a deficiency of dopamine in the substantia nigra. Post-mortem analysis of the substantia nigra of PD patients demonstrates pathological accumulation of deposits known as Lewy Bodies (LB) [4]. The spread of PD pathology is proposed to follow a specific spatial-temporal development (Fig.1.1). Based on the model by Braak et. al [5], the pathology appears initially in the vagal nerve nucleus in the brainstem and spreads upward through the brainstem into the medulla oblongata, and locus coeruleus. These two pathological stages are considered asymptomatic.

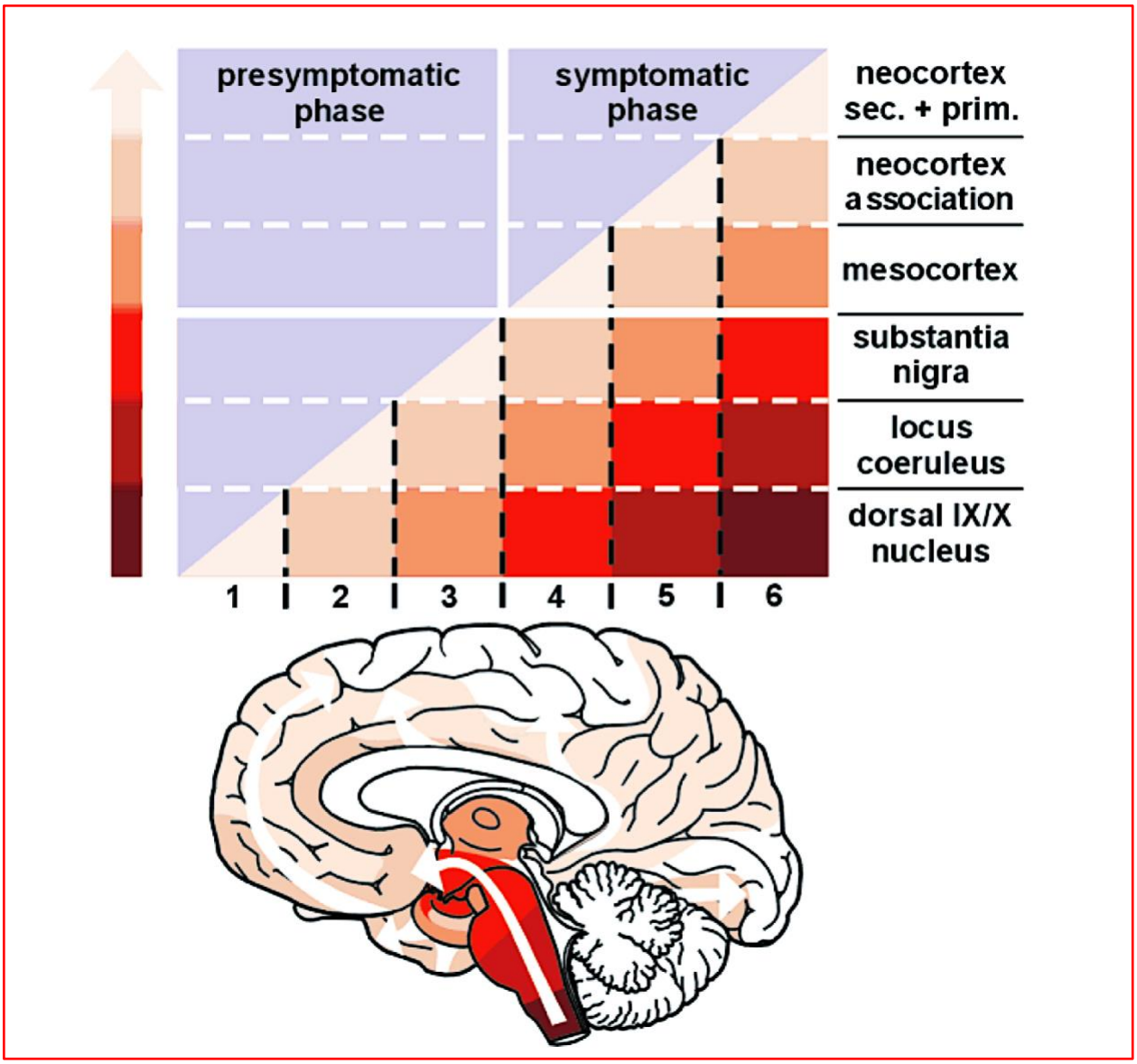

Figure 1-1. Figure 1. Schematic diagram of the Braak model of PD pathology progression.

Based on the Braak model, PD pathology initiates in the dorsal nucleus of the vagal nerve in the brainstem and follows a spatiotemporal pattern for spreading through the brain. In the top figure, the horizontal axis shows the six stages of PD pathology and the vertical axis demonstrates the dorso-lateral direction of pathology spread. (Figure obtained from Braak et al. paper[5], permission for use obtained)

The substantia nigra becomes affected in the third stage and the spread of LB significantly damages the neuromelanin projections of the substantia nigra. The spread of LB proceeds to the forebrain and temporal mesocortex, causing damage to the cholinergic neurons of the basal forebrain. During the next two stages, the pathologic process continues to damage the affected regions like the substantia nigra and temporal mesocortex more intensely and moving towards the outer layers of the cerebral cortex. According to the Braak model the clinical motor symptoms of PD can emerge during the last two stages. Also, abnormalities in limbic, autonomic and somatosensory functions can manifest in accordance with the criteria for iPD diagnosis at later stages of the Braak model [5].

1.2 Non-Motor Symptoms

Non-motor symptoms are fairly common in PD patients and might emerge decades prior to the classic motor symptoms [6]. Table 1.1 demonstrates an overview of the most reported non-motor symptoms in PD patients with their most frequent time course of manifestation. This thesis focuses on the cognitive impairment and neuropsychiatric symptoms of PD patients, and these non-motor symptoms will be discussed in greater detail in the following sections.

Table 1-1. A brief overview of non-motor symptoms in PD.

This table summarizes the most common non-motor symptoms of PD in relation to Braak's pathology model and the time course of their emergence compared to PD. This table is prepared based on the following studies [7-9].

Non-motor symptoms	Braak stage	Time of emergence
Constipation	I	Prodromal
Hyposmia	I	Prodromal
Rapid eye movement sleep behavioral disorder	II	Prodromal
Depression and other neuropsychiatric symptoms	III	Prodromal
Mild cognitive impairment	IV	Clinical symptoms
Orthostatic hypotension	V	Clinical symptoms
Hallucination and psychosis	V/VI	Clinical Symptoms
Dementia	VI	Clinical Symptoms

1.3 Non-Motor Symptoms: Cognitive Impairment

Cognitive impairment is one of the non-motor symptoms with a substantial negative impact on the quality of life of PD patients. PD patients are more susceptible to sustain deficits in their cognition compared to the general population. Moreover, PD patients are more likely to become demented [6] as the disease progresses. PD patients display cognitive impairments with a wide range of intensity. Mild cognitive impairment (MCI) is considered a defined initial stage of cognitive decline. MCI is described as neuropsychological deficits that do not interfere with daily life activities, but it can be perceived by the patients. This neuropsychological deficit can be quantitively evaluated by means of neuropsychological assessments [10]. A recent meta-analysis of 39 studies with 4011 PD patients reported that almost 28% of PD patients with normal cognition at the time of diagnosis develop MCI after three years [11]. Cognitive impairment has an inevitable negative impact on different aspects of PD patients' life and their families including, emotional, professional, and economic impacts. Also, cognitive decline is linked with higher likelihood of living in care centers, dementia and mortality [12-14].

The high prevalence of MCI in PD patients compelled the Movement Disorder Society (MDS) to create a taskforce in order to facilitate identification of PD patients who have MCI [10, 15].

A recent meta-analytical study suggested several risk factors for cognitive impairment in PD including hallucinations, older age, higher severity of motor symptoms, lower education level, and male sex, to name a few [6, 16]. Hallucinations are usually common non-motor symptoms of PD with minor hallucinations as the most prevalent form [8, 17]. Age appears to be one of the most influential risk factors for cognitive impairment in PD patients [8, 16]. The severity of motor symptoms is frequently reported as a risk factor for cognitive decline in PD [15, 18]. This might be justified by more severe damage in dopaminergic neurons resulting in both worse motor symptoms and vulnerability to cognitive decline [19].

Lower education level is another risk factor for cognitive impairment in PD [20, 21]. The effect of years of education could be explained by a possibly better protection of cognitive reserve in individuals with higher education [22]. Male sex is frequently reported as a strong risk factor for PD development at a rate of 3:2 compared to females and male sex demonstrated greater risk of cognitive impairment in PD [7, 21]. It should be noted that there are other risk factors reported by other studies and the above-mentioned risk factors are merely the most reliably reported ones.

1.4 Non-Motor Symptoms: Neuropsychiatric Symptoms

Neuropsychiatric symptoms (NPS) e.g., depression, anxiety, and apathy are another class of common non-motor symptoms in PD. Similar to cognitive impairments, NPS can precede PD diagnosis [6, 23] and have an immense negative impact on the quality of life of both PD patients and their families. The exact cause of these symptoms is not clear, and they sometimes are

perceived as a side effect diagnosis of PD. However, their frequent emergence prior to the diagnosis of PD rules out the mere diagnosis shock in PD patients.

The manifestation of NPS in PD patients prior to motor symptoms might be explained by the proposed model of Braak et al. on LB spread and pathology progression in the brain of PD patients [5]. The disruption of serotonergic and noradrenergic neurotransmitters might contribute to the NPS observed in PD patients [24, 25]. Some NPS could be the side-effects of pharmaceutical treatment of PD, for instance, impulse control disorder and drug-induced psychosis [26].

1.5 The Relationship of Genetics with Neuropsychiatric Symptoms and Cognitive Impairments in PD

The involvement of genetics in PD has been verified through numerous studies, and several genes have been identified in direct connection to hereditary forms of PD [27]. Several mutations in genes like *PARK2*, *PARK6* and *PARK7* can cause early onset recessive forms of PD [28]. Mutations in *SNCA*, *LRRK2*, and *PARK3* genes can cause autosomal dominant forms of PD [29-33]. Genome-wide-association-studies (GWAS) have facilitated identification of loci which exhibit an increased risk of iPD. The identified regions for iPD include Synuclein-Alpha (*SNCA*) and Leucine Rich Repeat Kinase 2 (*LRRK2)* which can be causal genes in some familial cases as mentioned above [27]. Genetic polymorphisms in these two regions have strong effect size and are risk factors for PD. Similarly, several other genes like *APOE*, *MAPT*, *GBA1* can be of substantial effect and are considered as risk factors for PD [27].

Mounting evidence supports a role for involvement of genetics in cognitive decline in iPD. Certain genetic variants seem to be associated with the progression of cognitive decline into dementia in

iPD patients. A hypothesis by Williams-Gray et al. suggests two distinct cognitive profiles for PD patients with cognitive impairments and associated genetic variants for each profile.

In the dual syndrome hypothesis, two main types of cognitive profiles in PD patients were described based on the cortical regions associated with the impairments (Fig 1.2). In the dysexecutive syndrome, deterioration of frontal lobe function in PD patients leads to deprived executive-function abilities e.g., problem solving, word finding. This type of cognitive impairment is more associated with genes involved in the modulation of dopamine availability in the frontal lobe such as *DAT1* and *COMT* [34, 35].

PD patients who suffer from the cognitive deficits which are derived from posterior cortical regions are more susceptible to dementia compared to PD patients with the dysexecutive syndrome [35]. According to Williams-Gray et al., poor performance in semantic fluency and copying of intersecting pentagons could be predictive of dementia in PD patients. To perform well in the above-mentioned tests, one requires well-functioning temporal and posterior cortical regions [35]. It is suggested that some genes e.g., *MAPT*, and *APOE* have an interaction with age-dependent cognitive impairments associated with temporal and posterior-cortical regions of the brain. Specific susceptibility variants in the genes which are associated with temporal and posterior-cortical regions can account for the cognitive decline with higher chance of dementia in PD patients.

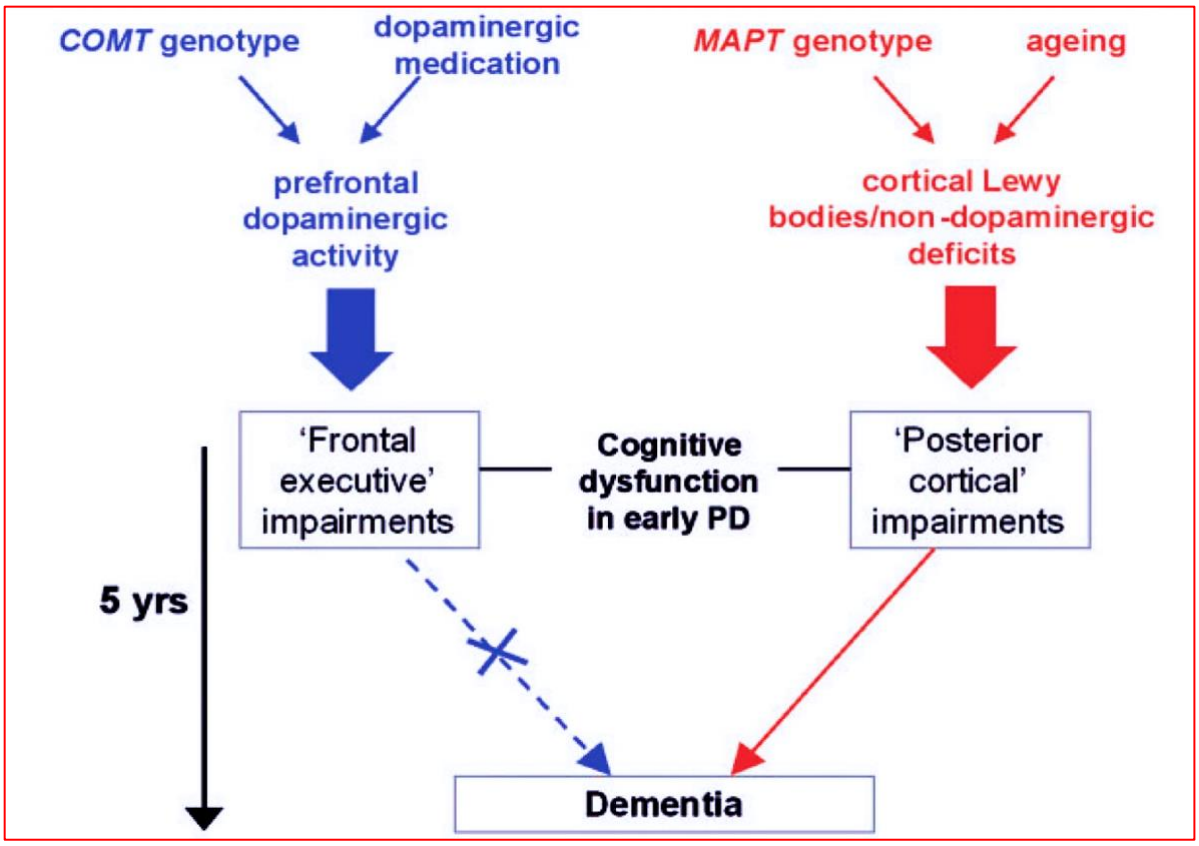

Figure 1-2. Schematic representation of the Dual syndrome hypothesis proposed by Williams-Grey et al.

In this model, Williams-Gray suggested that some genes are associated with cognitive impairments in PD patients. These genes can be divided into two main groups. The first group is associated with dopamine dysregulation in prefrontal regions such as DAT1 and COMT. The risk variants in these genes did not exhibit connections to dementing processes. The second group of genes such as MAPT, and APOE were linked to impairments in posterior-cortical regions. PD patients with the risk variants of these genes had higher likelihood of dementia. It should be noted that Williams-Gray et al. found other variables such as age or dopaminergic medications to have differential effects on risk of dementia in PD patients. Older age was reported to be linked to higher likelihood of dementia. Dopaminergic medication was found linked to the frontal executive impairments but not to dementia in PD. (Figure obtained from Williams-Gray et al. paper [35], permission for use obtained)

1.6 Apolipoprotein E

The apolipoprotein E (*APOE*) gene is located on the long arm of chromosome 19 q 13.23, at 44,905,796 to 44,909,393 bp according to GRCh38.p13 [36]. This gene encodes the apolipoprotein (ApoE) protein [37]. ApoE is part of a family of proteins called very low-density lipoproteins that regulate the homeostasis of lipids through modulation of lipid transport between cells or tissues and they are the main lipoproteins in the brain [37, 38].

Three major alleles are reported for *APOE*; *ε2*, *ε3*, and *ε4* [39]. These three alleles are derived by substitution of cysteine and arginine residues in the ApoE protein structure. The *ε*2 allele has cysteine residues in 112 and 158 residues of the protein, while in the *ε3* allele a cysteine residue is substituted with an arginine in the 158 position. The *ε4* allele differs from *ε*3 by substitution of a cysteine at position 112 with an arginine residue [40]. The *ε4* allele is well-recognized for its role in conferring an increased risk for Alzheimer disease (AD) up to 4 folds for heterozygous and up to 12 folds for homozygous individuals [41]. The *ε4* allele is also implicated in increased likelihoods for both, PD and cognitive decline in PD [42-45]. An increased susceptibility to PD with dementia has been reported in PD patients with *ε4* (OR 1.74 (with CI95% of 1.36-2.23), $p=1x\ 10^{-4}$) [46]. A genetic association study including 1079 PD participants discovered that *APOE ε4* had the most robust and extensive association with impairments of various cognitive measures. That study investigated the connection of two other genes (*MAPT* and *SNCA*) and their variants with cognitive performance. According to the study results *APOE ε4* was associated with worse performance in the following cognitive measures; HVLT-R total recall, HVLT-R delayed recall, HVLT-R recognition index, semantic verbal fluency, the letter-number sequencing and Trail making test B excluding the Trail A [42].

APOE is part of the dual syndrome hypothesis on progression of cognitive decline in PD. *APOE ε4* is suggested to be involved in the cognitive deficits that can be mapped to posterior-cortical regions of the brain [47]. Furthermore, studies reported a link between *APOE ε4*, and the overall likelihood of cognitive decline in both newly-diagnosed and mid-stage PD patients when using dementia rating scales [46, 48].

1.7 Brain-Derived-Neurotrophic-Factor

The Brain-Derived-Neurotrophic-Factor (*BDNF*) gene is located on the short arm (p) of chromosome 11 p 14.1, and at 27,654,893 to 27,722,030 bp according to GRCh38.p13 [36]. This gene encodes the BDNF protein, a prominent member of the neurotrophin family of proteins [49]. BDNF plays an essential role in the differentiation, protection and preservation of neurons through interaction with the tyrosine kinase receptor B (TrkB) [4, 50, 51]. This protein must go through post-translational modifications from a precursor form which has pre-pro domains to a fully functional (mature) form. These transformations result in the removal of the pre and the pro domains of the precursor form [52].

A well-characterized polymorphism in this gene is p.Val66Met (G758A, rs6265), in exon 11 of the gene. This variant substitutes a Valine (Val) residue at position 66 with a Methionine (Met) residue in the pro-domain of the BDNF protein. This substitution impacts the intracellular trafficking and activity-dependent secretion adversely leading to reduced availability of BDNF to CNS neurons [53-55]. This SNP has been shown to have a connection to both cognitive and psychiatric impairments in the geriatric population [56, 57].

A recent meta-analysis revealed an association between p.Val66Met and cognitive impairments in PD patients. The Meta-analysis consisted of six studies including 532 patients and 802 controls. Most of the participants were of European descent. It reported a significant relationship between the Met allele in PD patients and risk of cognitive impairments (p=0.029, OR (95%) = 1.06-2.75). However, this association was specific for the European descent population, p=0.003, OR (95%) = 1.47-6.34 [56]. At the protein level, the concentration of BDNF protein in serum demonstrated a positive correlation with executive-function and attention abilities in PD patients [58]. It should be mentioned that a link between the Met allele and better executive-function abilities specifically set-shifting was reported in PD patients [59].

There is a lack of consistent evidence about the role of this variant and NPS in PD patients. Nevertheless, there is evidence of involvement of this polymorphism and late-life depression in the general population. A meta-analysis on geriatric depression revealed an enhanced likelihood of depression in Met carriers (p=0.004, OR=1.48, 95% CI= 1.13-1.93) [60]. The role of BDNF in the relationship between anxiety and PD patients is not clear. However, a meta-analysis showed that BDNF protein levels are reduced in anxiety disorders e.g., post-traumatic stress disorder in the general population (PTSD) [61].

It should be mentioned that Cagni et al. reported a reverse relationship for BDNF and NPS in PD patients. They observed an association between the Val allele and an increase of depression/anxiety symptoms in PD patients [62]. This variant is reported in association with other psychiatric disturbances, but the current knowledge of its role in cognitive and psychiatric profiles of PD patients is still limited.

1.8 Catechol-O-Methyltransferase

The catechol-O-methyltransferase (*COMT*) gene is located on the long arm of chromosome 22 q 11.21 at 19941772 to 19969975 bp according to GRCh38.p13 [36]. This gene encodes an enzyme called catechol-o-methyltransferase which degrades catecholamines including dopamine. The prominent role of *COMT* in modulation of dopamine availability made it the focus of extensive studies. It has been reported in association with different neurological and psychiatric disorders [63-65].

One of the *COMT* polymorphisms with substantial effect on the enzyme function is p.Val158Met (rs4680). This missense polymorphism substitutes a valine residue at position 158 of the enzyme with a methionine residue. This substitution results in a reduced enzymatic activity of COMT up to approximately 40 % in vivo [66].

The p.Val158Met has a prominent role in prefrontal dopamine availability. In healthy controls the Met allele was found to be linked to better performance in the Wisconsin Card Sorting Task with regards to perseverative errors, working memory, and attentional control [67]. In addition, healthy Met carriers had enhanced physiological response to a simple cognitive task (two-back) during an fMRI study compared to non-carriers [67].

Williams-Gray et al. proposed that p.Val158Met could have a dual effect on the executive-function abilities of PD patients through the course of the disease (Fig 1.3). They proposed that COMT has an inverted U-shape relationship with the executive-function abilities in PD patients [35].

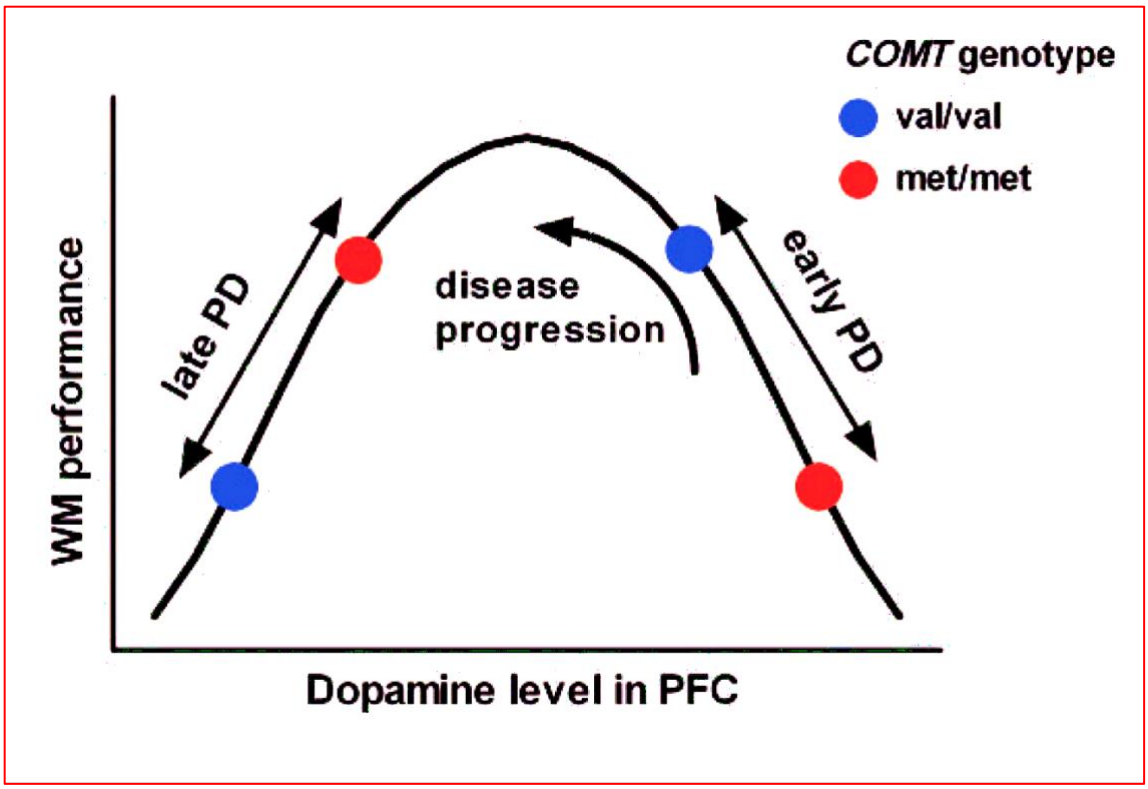

Figure 1-3. Schematic diagram of *COMT* alleles and their interaction with working memory throughout the progression of PD.

Based on Williams-Gray et al., PD patients with Val allele are more likely to demonstrate better performance in working-memory abilities compared to the Met carriers. However, as the disease progresses, the dopamine becomes scarcer. At this stage of PD, the Met carriers benefit from the lower enzymatic activity of COMT. The reduced activity of COMT, subsequently increases the availability of dopamine in the prefrontal regions. (Figure obtained from Williams-Gray et al. paper [35], permission for use obtained)

Based on the findings from the CAMPAIGN longitudinal study in PD patients, the Val carriers at early stages of the disease could display a superior cognitive performance compared to the Met carriers. However, as the disease progresses, the availability of dopamine becomes sparser in prefrontal regions. The Met carriers benefit from the lower activity of COMT enzyme by having higher dopamine availability in their prefrontal cortex at more advanced stages of PD [35].

1.9 Dopamine Transporter

The Solute Carrier family 6 member 3 (*SLC6A3*) commonly known as Dopamine Transporter (*DAT1*) is located on the short arm of chromosome 5p15.33 at 1392794 to 1445440 bp according to GRCh38.p13 [36]. It encodes the Dopamine transporter (DAT), a type of monoamine transporter protein located on the plasma membrane [68]. The DAT protein is exclusive to dopaminergic neurons with highest abundance in the striatum [68]. This protein is in charge of dopamine reuptake in dopaminergic neurons using a sodium/chloride gradient-dependent mechanism [68].

In agreement with its prominent function in modulation of dopamine availability, mutations in *DAT1* are linked to manifestation of PD-like symptoms in toddlers and familial forms of PD. Several mutations and polymorphisms are identified in *DAT1* in association with PD. A variable number of tandem repeat (VNTR) with polymorphism ID of rs28363170 has received attention in idiopathic PD [69, 70]. This VNTR of 40 base pair repeats is located at the 3' untranslated region (UTR) and have different alleles based on the length of VNTR in different populations. The most frequently reported alleles of this variant are 9 and 10 repeats (9-R, and 10-R, respectively).

This polymorphism can affect the abundance of the DAT protein in the putamen of non-PD individuals [71].

The 10-R allele was reported in connection of an increased expression of *DAT1* therefore this allele can modulate a higher rate of dopamine reuptake [72]. Furthermore, a meta-analytical study reported a protective effect for the 10-R allele in an Asian population but not for PD patients of European descent [70]. Also, PD patients with a 9-R allele exhibited reduced neural activity in prefrontal, premotor and caudate nucleus regions compared to 10-R patients during an executive-function task [73]. However, some evidence suggested a link between the 10-R allele and reduced reward reactivity in the orbitofrontal region of PD patients compared to HC [74]. *DAT1* might be involved in NPS in PD patients, for instance PD patients with the 9-R allele were shown to have a higher likelihood of psychosis [75].

1.10 Microtubule-Associated-Protein-Tau

The Microtubule-associated-protein-tau (*MAPT*) gene is located on the long arm of chromosome 17q 21.31 at 45,894,538 to 46,028,334 bp according to GRCh38.p13 [36]. *MAPT* is implicated in several neurodegenerative diseases like AD, and PD. It encodes for the Microtubule-associated-protein-tau which is the building block of tau neurofibrillary tangles, a common pathological inclusion body in AD and PD pathology [76]. This protein is primarily expressed in neurons and plays a substantial role in the cytoskeleton of neurons.

The *MAPT* gene has a prominent haplotype which appears to be derived from an inversion of approximately 970 Kbp region. This region is located inside a larger block spanning about 1.6 Mbp with linkage disequilibrium. H1 haplotype refers to the non-inverted region and it is the most

common haplotype globally. The other allele, H2 haplotype is the inverted region which was frequently found in up to 25% of people with an European descent [77]. These two haplotypes exhibited a robust connection to the risk of having PD.

The H1 haplotype was reported to be associated with an increased risk of PD, and H2 was found to be protective against the risk of PD [78, 79]. Based on a meta-analysis consisting of 43 studies, H2 haplotype might have a protective effect on PD (OR=0.76, 95% CI= 0.74-0.79, p=0.002) [79]. In a large-scale study of shared genetic associations between PD and AD in 89,904 cases one variant was found in Carbonyl Reductase 1 *(CHCR1)*, a gene adjacent to the *MAPT* locus. This variant displayed a strong Linkage Disequilibrium (LD) with the H1 haplotype [80].

The importance of *MAPT* in PD is not limited to the increased risk of PD and it has been found associated with cognitive impairments as well. H1 haplotype has been found to be linked to a higher rate of Parkinson's disease dementia (PDD) [81]. Nevertheless, a longitudinal study following PD patients for 10 years discovered that this relationship might be modulated with age in H1 haplotype carriers and not in H2 carriers [82].

fMRI studies reported significantly reduced neural activity in hippocampal, inferior temporal, fusiform and parahippocampal regions of H1 PD patients during a memory encoding task compared to non-carrier PD patients. Also, PD patients with the H1 haplotype had lower hippocampal activation during the memory encoding task [83]. However, several studies reported no association of *MAPT* H1 in cognitive impairment and dementia in PD [84]. Furthermore, some studies demonstrated a bias for association of *MAPT* H1 haplotype and early stages of PD [47, 85].

1.11 Synuclein-Alpha

The Synuclein-alpha (*SNCA*) gene is located on the long arm of chromosome 4 q 22.1 at 89,724,099 to 89,838,324 bp according to GRCh38.p13 [36]. The *SNCA* gene is a crucial gene in both, iPD and familial forms of PD. Point mutations and an increase in the *SNCA* gene copy number can create early onset autosomal dominant PD with moderate to severe cognitive complications [29, 86, 87]. However, the role of *SNCA* in PD is not limited to mendelian forms and several variants were found in connection with risk of PD and dementia in iPD [88-91].

This gene encodes the alpha-synuclein protein which was discovered initially as a component of amyloid-beta plaques in AD and later was identified as the main component of LBs in PD [4]. The true function of the alpha-synuclein protein is obscure at this stage, but based on its cellular location, it is speculated to be part of the NSF attachment protein receptor (SNARE). SNARE is a crucial protein complex in neurotransmitter release from the presynaptic neurons [92]. The term synucleinopathies refers to disorders like PD, DLB, AD, etc., in which the alpha-synuclein protein was found to be involved [93].

Nevertheless, the role of *SNCA* and its protein is not restricted to the risk of PD. Recent studies discovered a connection between variants in this gene and cognitive decline in iPD [88, 94]. In a thorough analysis of *SNCA* using 1,492 PD, 922 DLB and 971 HC participants a haplotype on intron 4 of *SNCA* was found associated with the risk of PDD [88]. Guella et al. included 43 *SNCA* variants in the analysis and found that two variants: rs10018362 and rs7689942 had strong connections to PDD after correction for age, sex, *APOE ε*4 dosage, and sample site [88].

In recent years, several variants in *SNCA* were reported in association with neuronal and cognitive changes in iPD. A Brazilian study reported a *SNCA* variant, rs2583988, linked to cognitive impairment in PD patients [94]. Similarly, evidence of *SNCA* effect on activity of brain networks was proposed recently through investigation of the effect of rs894278 on resting-state activity in both PD and HC groups [95]. This variant had an association with amplitude of low-frequency fluctuations (ALFF) in the fusiform region indicating a possible role in resting-state activity in iPD patients [95]. These studies although still preliminary, indicate a potential role for *SNCA* variants in iPD and encourage further investigations on this gene.

2 Chapter Two: Methods

In this chapter, the methods that were applied in this thesis will be briefly described. The methods were selected based on their availability, suitability and efficiency according to the requirements of each project.

2.1 Genetic Techniques

Genetic methods applied in this thesis were selected based on the study design, availability of laboratory facilities and efficiency of technique. These methods were employed in particular to aid the investigation of potential associations of variants in the above-mentioned genes and cognitive impairments or NPS in PD patients.

2.1.1 DNA Extraction

The DNA was extracted from a whole blood sample collected in a 10 ml lavender cap tube based on the following protocol.

The red blood cells (RBC) were lysed using 30 milliliters (ml) of RBC lysis solution in a 50 ml polypropylene canonical tube. The blood sample of one lavender cap tube was added to the RBC lysis solution and incubated on a shaker for 5-10 minutes. The solution was centrifuged at x4000 RPM for 2 minutes at room temperature. The supernatant was discarded, and the pellet was resuspended in the remaining volume by a brief vortex. To precipitate the protein content of the blood sample, 3.3 ml of protein precipitation solution was added to the resuspended pellet. 10 ml of RBC lysis solution was slowly added to the mixture and was vortexed intermittently for 30 seconds until red fragments of lysed cells were visible in the solution. The mixture was centrifuged

at x4000 RPM for 5 minutes. The supernatant was added to a clean 50 ml polypropylene canonical tube containing 10 ml of 2-propanol and inverted manually until the white strand of DNA emerged. The DNA pellet was precipitated using centrifugation at x4000 RPM for 3 minutes. The supernatant was discarded, and the DNA pellet was washed with 10 ml of 70% Ethanol. The resolved DNA mixture was centrifuged at x4000 RPM for 1 minute and the supernatant was discarded. The DNA pellet was air-dried for 10 minutes and was resolved in 1 ml of pure DNase free water. The concentration of DNA was measured using a Nanodrop spectrophotometer. All DNA samples were stored (short-term) at 4° C in a refrigerator.

2.1.2 TaqMan SNP Genotyping Assay

TaqMan Single Nucleotide Polymorphism (SNP) genotyping assay is an efficient and affordable genotyping technique. It was designed based on the 5' to 3' nuclease activity of DNA polymerases. This technique was originally invented based on the *Thermus aquaticus* (*Taq*) polymerase exonuclease activity by Holland et al in 1991 [96].

TaqMan SNPS genotyping assay is a mixture of forward and reverse primers specific to the target SNP. Each primer is conjugated with a TaqMan probe, a specific oligonucleotide with a fluorescent reporter dye at 5' and a quencher at 3' end adjacent to a minor grove binder (MGB). There are different combinations of fluorescent reporter dyes, but the most usual ones are VIC and FAM.

The TaqMan oligonucleotide is designed to be specific to the DNA sequence adjacent to the target SNP. The MGB molecule at 3' of TaqMan oligonucleotide increases the likelihood of hybridization between the oligonucleotide and the DNA template. The MGB molecule stabilizes hydrogen bonds between the oligonucleotide and the DNA template after they are hybridized.

The DNA polymerization begins after one of the primers and the TaqMan probe are hybridized with the DNA template. The AmpliTaq™ Gold DNA polymerase moves along the DNA for polymerization, and the fluorescent dye is removed by its exonuclease activity. The removal of fluorescent molecule from the primer creates a signal which can be detected. This fluorescent signal is an indication of a specific DNA template (specific Allele) [97].

2.1.3 Variable Number of Tandem Repeat Genotyping (VNTR)

Variable Number of Tandem Repeats (VNTR) are genomic elements that consist of short sequences of DNA consecutively repeated and located adjacent to each other. These genetic elements can be highly polymorphic and typically have more than two alleles among individuals of the same species [98]. These elements are spread throughout the genomic DNA. Some VNTRs which are located within the coding or regulatory regions can have a prominent role in genetic polymorphisms.

To identify different VNTR alleles, the genomic regions in which the VNTR are located are amplified using PCR primers. The PCR primers are designed to anneal to the 3' and 5' ends of the locus in genomic DNA containing the target VNTR. The PCR primers are extended using the genomic locus as the template and capture the DNA sequence and its length in the amplicon (final PCR product). The resulted amplicons are typically separated based on their length using agarose gel electrophoresis [98].

The gel electrophoresis is a routine method of choice for separation of charged molecules in laboratories. In the most basic form, an agarose gel of specific concentration is prepared. The gel is solidified into a porous structure which can slow down the movement of charged molecules once

it is exposed to an electric field. The amplicons have a net negative charge and move from the negative pole (cathode) to the positive pole (anode). Those amplicons which are larger in size, have slower migration to the anode. The difference in migration pace of amplicons facilitate separation of amplicons based on their size. Each participant's sample is run through the gel with a molecular marker to measure the approximate size of the amplicons. Different sizes of amplicons from the sample of a participant indicate a heterozygous genotype with respect to the target VNTR [98].

2.1.4 Exome Sequencing

Exome sequencing is a genomic technique which generates the sequence of the protein-coding regions (exome) of a genome [99]. This technique consists of two main steps; the first is the enrichment step, in which the target regions are selected for sequencing. The second step is the sequencing step, which generates the sequence of the selected regions using high throughput DNA sequencing methods [99].

The exome sequencing method which was applied in chapter 6 for PPMI data was Ilumina Nextra® Rapid Capture Expanded Exome (Illumina, Inc. SanDiego). This exome sequencing kit covers about 201,121 exons, UTRs and micro-RNA (miRNA) binding sites in the human genome. Comparing this library to Refseq showed 95.3% coverage of the human genome [100]. This method is briefly described in the following section.

The genomic DNA is prepared for the enrichment step using transposomes to cut it into fragments of few hundreds bp. The transposomes tagment the genomic DNA fragments by adding specific sequences (called adaptors) to both ends of them. The tagmented genomic fragments with adapter

sequence can hybridize with primer pairs specific to the adapter sequences and get amplified [100, 101]. The PCR products are denatured to become single-stranded DNA (ssDNA). The ssDNA fragments are enriched using biotinylated probes containing an oligonucleotide sequence complementary to regions of interest in the genome (Figure 2.1) [100, 101].

The ssDNA fragments are purified and sequenced in Ilumina HiSeq 2500 sequencing platforms (Illumina, Inc. SanDiego) [102]. The fragments are read using 2x100 bp paired-end read cycles. Using this method each fragment is read twice, one time the reading direction is from 3' to 5' and the next time, it is vice versa.

These sequences are stored in (. fastq) files in a database. The (. fastq) read files are aligned to the human genome (for instance, UCSC hg 19) using Burrows-Wheeler Aligner (BWA). The duplicates are removed, insertion-deletions (indels) were checked, and a base calibration was performed using the Piccard Software (picard.sourceforge.net/index.shtml) and the Genome Analysis Tool kit (GATK).

A variant calling, and the genotype likelihood are performed, and a Variant Calling File (VCF) is generated per participant. All participants VCF files are combined to generate a cohort VCF file using GATK tools [102]. Quality check can be performed with the PLINK toolset using variant call-rate, heterozygosity rate, gender check, etc. For further details please check the PPMI database or Ilumina website [101, 102].

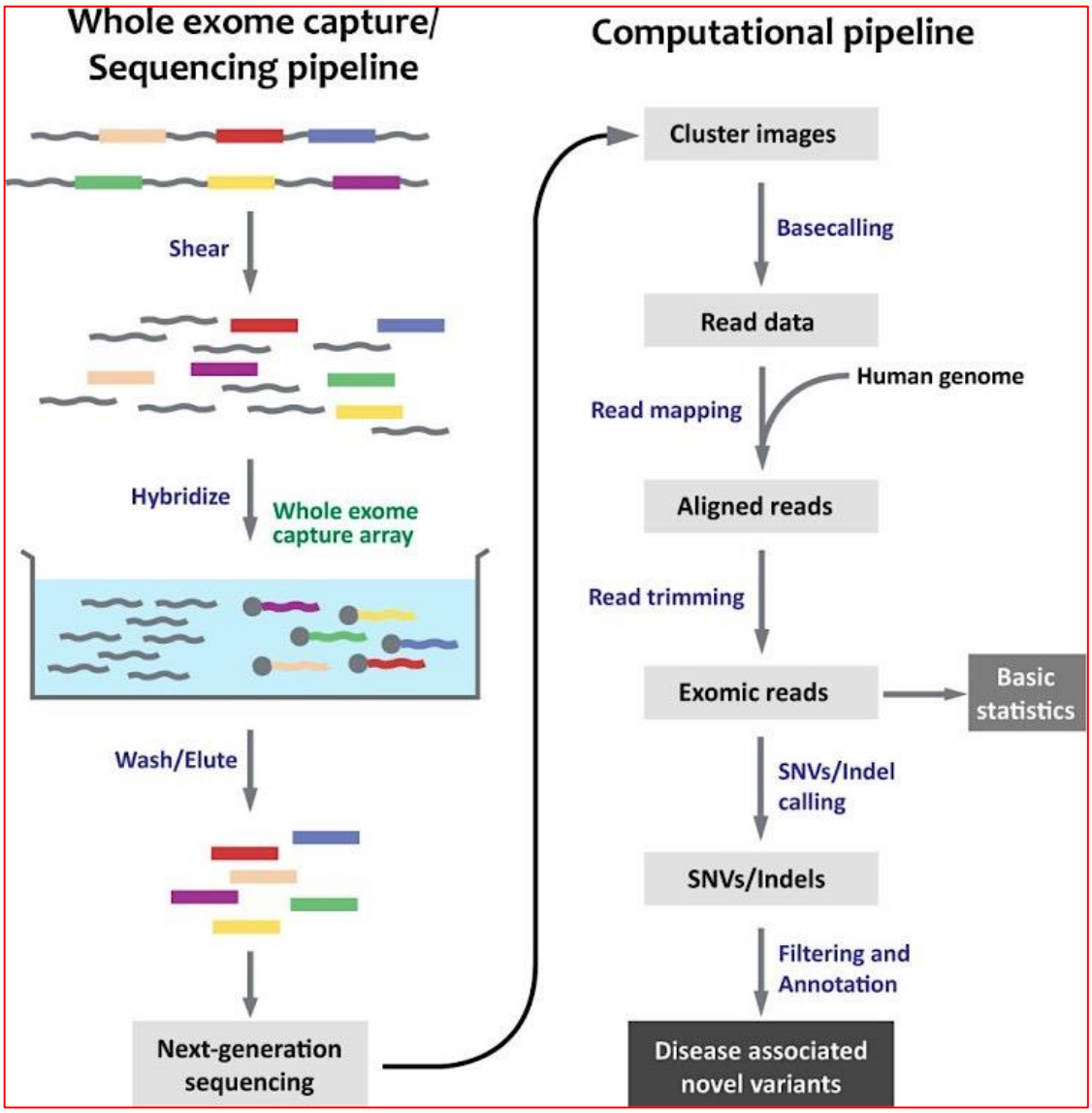

Figure 2-1. A schematic view of Exome sequencing pipeline.

The Genomic DNA is cut into fragments of few hundred bp, and during an in-solution enrichment step, the genomic fragments are hybridized with primer pairs. These hybrid oligonucleotides are sequenced and by use of bioinformatic tools, the exome sequence is reconstructed. Each oligonucleotide is sequenced, and oligonucleotides are aligned using the human genome as a reference. The duplicate reads are removed in a validation step. The exomic sequences are used to screen for indels and Single Nucleotide Variants (SNVs). (The image is adopted from the following paper [103], open access data, used in agreement with the creative commons attribution non-commercial license)

2.2 Neuropsychological and Neuropsychiatric Assessment

Neuropsychological assessment is defined as "the normatively informed application of performance-based assessment of various cognitive skills" [104]. The cognitive assessment is important in evaluating cognitive abilities of the examinee in an unbiased and standardized setting. The examinee's performance is compared to the normative data to assess their cognitive abilities. The normative data used for the cognitive assessment should be collected from a large sample reference group of the same age, sex, and years of education [104, 105]. The normative comparison enables the investigators to evaluate the cognitive performance of the examinee while considering these critical demographic factors which can influence cognitive performance [104].

The neuropsychological assessment usually includes more than one type of test per cognitive domain. The common cognitive domains are memory, processing speed, language, attention, visuo-spatial and executive-function [104].

In this thesis the following neuropsychological battery was applied to evaluate the cognitive abilities of the study participants (Table 2). This cognitive battery was designed to tap into the five main cognitive domains: attention, executive-function, language, memory and visuo-spatial function. Participants were diagnosed as MCI according to the guidelines suggested for the diagnosis of MCI in PD patients by the Movement Disorders Society (MDS) [15]. The participants were diagnosed as MCI if their performance in two cognitive measures was $\geq$ 1.5 standard deviations below the mean of the normative data. All cognitive assessments were performed by a senior psychometrist. The final scores were transformed to Z-scores and an average Z-score per cognitive domain was calculated. The global cognitive Z-score was calculated using a weighted average of the Z-scores of the examined domains.

Table 2.1. Cognitive battery applied to evaluate the cognitive abilities of the study participants.

Cognitive domain	Cognitive test/measure
Attention	Trail A Wechsler Memory Scale-IV (WMS-IV) Symbol Span Wechsler Adult Intelligence Scale-IV (WAIS-IV) Digit-Span Forward
Executive function	Brixton Clock Drawing Test (CDT) Command Dementia Rating Scale-2 (DRS-2) Initiation and Perseveration Hayling 2 Letter Fluency FAS Stroop Colour and Word Trail B WAIS-IV Digit Span Backward WAIS-IV Digit Span Sequencing
Language	Boston Naming Test (BNT) Category Fluency Animals Category Fluency Actions
Memory	Hopkins Verbal Learning Test (HVLT) Retention HVLT Recognition Rey Complex Figure Test (RCFT) Delay Recall WMS-IV Logical Memory Delayed Recall
Visuo-spatial	CDT Copy Hooper Visual Organizational Test (HVOT) RCFT Copy

Another method to evaluate the cognitive abilities of a participant is to use the Montreal Cognitive Assessment (MoCA) test. The MoCA is a brief test which can provide a general overview of cognitive abilities of the examinee. This test was administered by the psychometrist and the MoCA scores were corrected for age and education [106].

The NPS in this thesis were evaluated in accordance with the mild behavioral impairment (MBI) criteria in chapter three and four. MBI is a validated syndrome that is characterized by the emergence of persistent NPS in later life which indicate a potential risk of cognitive decline [107]. This syndrome is evaluated using the mild behavioral impairment checklist (MBI-C).

The MBI-C contains five major domains of behavioral symptoms: 1) impaired drive/motivation; 2) emotional dysregulation; 3) impulse dyscontrol; 4) social inappropriateness; and 5) abnormal thoughts/perception. This checklist is a 34-item questionnaire which explicitly captures symptoms emerging in later life, persisting for ≥6 months, and representing a meaningful change in the participant's behavior. Items were scored on a severity scale of 0-3; a cut-off score of ≥8 was used to classify a patient as MBI positive [108, 109]. The MBI-C was completed by a third party in close contact with the participant e.g., family member or caregiver

2.3 Structural MRI

Structural MRI is a non-invasive technique to examine the anatomy of the brain clinically and in research settings. Numerous studies reported structural changes in cortical and subcortical regions of PD patients compared to the general population [110-113]. There are various types of structural MRI including, T_1-weighted, T_2-weighted, and Fluid attenuated inversion recovery imaging.

T_1-weighted imaging is a basic technique to study the anatomy of the brain. It is based on the longitudinal relaxation (T_1) also known as spin-lattice relaxation. The T_1 relaxation is defined as the process in which net magnetization returns to each initial state which is parallel to the longitudinal component of net magnetization (B_0) [114]. The T_1 constant refers to a time constant through which 63 % of the exited spins return to the longitudinal component of net magnetization. Different brain tissues return to the B_0 at different rates. Cerebrospinal fluid (CSF) has the shortest recovery time while white matter displays the longest T_1. The differences in tissues' T_1 are used to differentiate anatomical structures of the brain.

Different software programs are available to process structural imaging and Freesurfer (http://surfer.nmr.mgh.harvard.edu/; version 6.0.0) is used in this thesis to measure cortical thickness and subcortical volumetry. Freesurfer software is publicly available and includes features enabling precise and automatic processing of structural data [115-118].

2.3.1 Cortical Thickness

Cortical thickness measurements aim to evaluate the thickness of cortical grey matter either locally (region of interest) or globally. Grey matter thickness changes with age, but some studies reported an accelerated cortical thinning in PD patients compared to age-matched controls in the general population [110, 119]. These cortical atrophies are associated with cognitive and psychiatric impairments which necessitates investigation of these changes in PD patients [110, 119, 120].

The automatic cortical thickness measurement with Freesurfer is discussed in brief below. First, structural images are required to be corrected for motion artifacts and field intensity inhomogeneity. These images are registered to the MNI-152 template using nine parameter affine transformation and the skull is removed. Next, Freesurfer segments the image and separates cortical grey matter from subcortical white matter through segmentation and reconstruction steps.

The segmentation process is comprised of several steps. First, different tissues are labeled based on their intensity information. Next, the whole volume is analyzed thoroughly to identify regions where more than one tissue type is located at proximity. A planar orientation is prepared for the voxels located at the borders where different types of tissues are occupying the same voxel. In-plane intensity inhomogeneity allows identification of voxels that have different intensities from their neighboring voxels, and the label for that individual voxel is reversed. The connected

components algorithm is used to fill out the footprint of white matter in the reconstructed grey matter. The boundaries of grey and white matter and the grey matter and pial membrane are reconstructed by tessellation of their surfaces. This can result in generating a polygonal mesh of the cortical surface made from approximately $150x10^{3}$ vertices. Freesurfer checks the topology of the resulting surface automatically. The cortical surface is evaluated and modeled using a surface deformation technique.

The cortical thickness is measured by calculation of the difference between the analogous vertices on the surface of the pial and white matter. Each vertex on the pial and the white matter are considered as a truncated pyramid which together they form a truncated tetrahedrane. The cortical volume is calculated as the volume of the truncated tetrahedrane. These cortical measurements are smoothed to diminish the regional variation [115, 116].

The statistical analysis is performed using a general linear model (GLM) in which confounding factors can be controlled for. The ROIs are identified and labeled using the Deskian-Killiany atlas and their average cortical thickness is extracted [121].

2.3.2 Subcortical Volumetry

This technique is performed using the Freesurfer platform and has some overlaps with the cortical thickness technique explained above. Freesurfer provides a highly automated pipeline for this technique and the following steps are performed to obtain subcortical volumetric measures [117, 118]. T_1-weighted images are motion-corrected and intensity-normalized. Images are registered to the MNI-152 common template using a nine-parameter affine transformation and the skull is removed.

The analysis proceeds to the segmentation step which comprises of several processes; first the subcortical regions are registered to the Gaussian Classifier Atlas (GCA), after which images are normalized and registered using canonical algorithms [117]. Neck removal follows this process and images are again registered with the skulls included in the image. Finally, the subcortical regions are labeled using the GCA atlas and their volumes are measured by Freesurfer. It is highly recommended to measure the intracranial volume for the subsequent normalization of subcortical volumetric measures. A total grey matter measure can be extracted which consists of the average surface-based grey matter volume and the voxel counts of subcortical grey matter [117, 118].

2.4 Machine Learning

This branch of artificial intelligence has the primary aim of developing algorithms which can detect the underlying relationship in a large and complex set of data accurately and efficiently. Machine learning (ML) applies various approaches to analyze a given dataset and these approaches can be divided into two major groups: supervised and unsupervised analyses. In the unsupervised models, there is no definite outcome, and the model is used to differentiate classes or clusters of data [122].

Supervised models are employed to predict a pre-determined outcome, and this requires training of the model on an example dataset, in which the outcome is predefined to learn to accurately predict the outcome. One of the most common applications of ML is to develop a predictive model. A predictive model is capable of making predictions about a very specific topic. To build the predictive model, a training dataset is typically used to train the algorithms and enables it to predict. In a supervised analysis, the choice of ML method for a predictive model depends on the nature of the target outcome. Continuous outcomes are processed using regression while categorical outcomes are analyzed using classification methods.

Feature selection is an optional step for ML data processing through which highly correlated and redundant input features are removed. Feature selection is essential to avoid unnecessary model complexity and overfitting issues in large datasets. Some of the most common feature selection algorithms are discussed below. These algorithms are selected based on their relevance to the methods used in this thesis.

One common feature selection algorithm is correlation-based feature selection which reduces the dimensionality of the dataset by removing the input features that are not highly predictive of the outcome variable [122]. Another common method is principal component analysis (PCA) which generates a set of new variables by combing the actual independent variables (IV) linearly. The new variables are created to maximize the amount of variance in themselves leading to fewer IVs in charge of substantial portion of variance.

The RELIEF algorithm is another common feature selection method in particular for classification analyses. In this method a given feature and its K nearest neighboring features are selected from one of the classes (called Hits) and M nearest features from the opposite class (called Misses). After selecting these features, the RELIEF algorithms analyze the features iteratively. By the end of the iteration process, the features are weighted and those features with greater weights gain more relevance in the analysis. This method is mostly employed when features display collinearity issues. A modified version of this feature selection algorithm is RRELIEF which can be used for regression problems.

In the following section, some of the most common supervised regression models for ML are summarized including linear regression, random forest, k-nearest neighbors and support vector regression. This section focuses on the support vector regression in more details, since it has been applied for the analysis of the project presented in chapter 5 of this thesis.

2.4.1 Linear Regression

Linear regression is one of the basic models to predict a continuous outcome in ML. Linear regression assumes that the relationship between the predictive and the outcome variables is linear. This model has different variations but two of the most common ones are simple and multiple linear regression. In simple linear regression, the outcome variable depends on changes in value of one independent variable (input). Multiple linear regression considers effects of several variables for prediction of the outcome.

2.4.2 Random Forest

The random forest (RF) technique is one of the decision tree methods. Decision tree is a common predictive modeling method. It is defined as a system of conditions and rules that can be graphed as a tree. The branches of the tree are constructed by optimization of the rules over the features that can optimally split the data according to the outcome variable. The trees contain several layers of branches and the smallest branches end up in leaves. The leaves of the tree demonstrate the probability based on the observations. In the process of building a decision tree, the tree structure is trained on choosing data points most similar to each other in the same node/leaf [122].

The random forest method differs from the simple decision tree by dividing the training dataset using sampling with replacement into groups. Each group is used to train a tree. The final result is provided by calculation of the average of each tree. In addition to this difference, RF uses a subset of available features in the formation of the decision tree to induce more randomness. This technique benefits the analysis by avoiding overfitting problems [122].

2.4.3 k-Nearest Neighbor

k-nearest neighbor is a non-parametric supervised algorithm. It utilizes the training dataset as a template and predicts a model based on the similarity between the training and test datasets. This algorithm uses the distance between each test sample and the training dataset and considers a weight for the neighboring observations. The average of the nearest neighbors is the predicted value for the test sample. For predictive purposes, the Euclidean distance is a common distance type, referring to the distance between the current test sample and the training datapoints. There are several limitations for this algorithm; it does not perform well with a high number of independent variables and is prone to error in small datasets.

2.4.4 Support Vector Regression

Support vector regression (SVR) is a derivative of the more known ML algorithm, support vector machine (SVM). SVM is a supervised classification method that detects a hyperplane (line/higher dimensional plane) which can separate different classes of data efficiently. To identify this line/hyperplane, datapoints might need to be transformed using mathematical functions called kernels. Common kernels are linear (the simplest method, suitable for large and sparse datasets), polynomial (used commonly with imaging data), or Gaussian (used mostly when no prior knowledge on the dataset is available).

SVR is, in principle, very similar to SVM classification models with slight differences in the adaption to a regression style problem [123]. More precisely, an SVR model is built based on only a subset of training data within the predefined margins that minimize the generalization error between the true and the predicted outcome.

Therefore, the data is first transformed into a higher dimensional space employing the polynomial kernel, thereby allowing linear models to fit the training data. The SVR model is less likely to overfit the data compared to other models, i.e., SVR is a model with adequate generalization capabilities and good prediction accuracy.

2.4.5 Model Evaluation

The model performance should be evaluated in order to validate its sensitivity and accuracy regarding the target outcome. The validation of a model is an important step through the model development to examine the reliability and reproducibility of the model. A test dataset is ideally required to evaluate the performance of a model [124]. However, in some cases there are no test datasets available for model evaluation. This issue can be resolved using cross-validation methods. Through cross-validation, the dataset is divided into training and test sets. For instance a fourfold cross-validation refers to dividing the dataset into four equal parts, in which one part is used for testing and the three other parts are used for training the model [124]. The nested leave-one-out cross-validation (an internal validation method) was applied in chapter 5, which is discussed briefly below.

The nested cross-validation achieves higher accuracy through splitting the validation dataset into smaller sets. This method is based on two loops, an inner loop which is nested within an outer loop [125]. The model score is maximized by fitting the model to the training set (inner loop). The generalization error (an index of the model generalization ability) is estimated through the outer loop which is divided into training and test sets for validation. As the test and training sets in this method are selected randomly the average error score obtained through this method is more reliable [125].

In the leave-p-out-cross-validation, a dataset with (n) number of datapoints is split to be used for both testing and training the model. This can be performed by using p number of data points which are randomly selected and used as a test set. At each validation test, p number of participants are used to test the model while n-p participants were used to train the model. The average metrics of these n models are usually used to report the model performance. In the leave-one-out cross-validation, p is equal to1 [124, 126].

Thesis Overview and Aims

The overall goal of the present thesis is to investigate the association between genetics and non-motor symptoms in iPD patients. This thesis specifically focuses on two of the non-motor symptoms, NPS and cognitive impairment.

The first part of the thesis includes two manuscripts exploring the relationship between certain genetic variants and NPS in iPD patients. In chapter 3, we investigated the association between the *BDNF* Val66Met and mild behavioral impairment (MBI) in iPD patients using the Mild Behavioral Impairment-Checklist (MBI-C) as the ascertainment tool. In chapter 4, we studied two well-characterized genetic variants in two dopamine modulating genes and their potential association with MBI in iPD patients. We aimed to investigate the possible link between *COMT* Val158Met and *DAT1* VNTR and MBI. In particular, we explored whether these variants are associated with mood impairments induced by dopamine dysregulation.

The second part of the thesis includes two manuscripts studying the association between cognitive impairment and a novel variant in the *SNCA* gene. In chapter 5, we aimed to predict global cognition in iPD patients using a combination of features including genetics, demographic, clinical characteristics, and structural imaging by means of machine learning. In a second step, we analyzed the most novel selected feature, rs894280 in a post-hoc analysis to study its relationship with each of the main cognitive domains. In chapter 6, we examined whether rs894280 demonstrates an association with cognitive impairment in PD patients longitudinally using PPMI data. In addition, we hypothesized a possible association between this variant and NPS in the PPMI dataset. This analysis was designed in order to replicate results, presented in chapter 5, in a larger

cohort. In addition, it was aimed at investigating whether rs894280 could be linked to cognitive and NPS symptoms both cross-sectionally and longitudinally.

In Chapter 7, the results of the studies are summarized. I describe how these findings can aid future research in iPD non-motor symptoms, specifically cognitive decline and NPS. I describe possible future investigations that can be conducted in order to elucidate the role of the abovementioned variants in non-motor symptoms of iPD patients. Finally, I discuss the implication of the main findings of this thesis in the conclusion.

Part one: Genetic Associations and Neuropsychiatric Symptoms in Idiopathic PD

3 Chapter Three

In this chapter, we studied the association of a common BDNF variant with NPS in PD patients. We studied whether this variant shows any association with the likelihood of NPS in general. We also investigated whether this variant is associated with specific domains of mood disorder in PD patients.

This chapter is reproduced from the paper published in Frontiers Neurology: Ramezani M, Ruskey JA, Martens K, Kibreab M, Javer Z, Kathol I, Hammer T, Cheetham J, Leveille E, Martino D, Sarna J, Gan-Or Z, Pfeffer G, Ismail Z, Monchi O. (2021) Association between BDNF Val66Met polymorphism and mild behavioral impairment in patients with Parkinson's disease. Frontiers in Neurology. Jan 21,10:587992.

Association between BDNF Val66Met polymorphism and mild behavioral impairment in patients with Parkinson's disease

Mehrafarin Ramezani[1], Jennifer A Ruskey[2,3], Kristina Martens[1], Mekale Kibreab[1], Zainul Javer[1], Iris Kathol[1], Tracy Hammer[1], Jenelle Cheetham[1], Etienne Leveille[2], Davide Martino[1], Justyna R Sarna[1], Ziv Gan-Or[2,3,4], Gerald Pfeffer[1], Zahinoor Ismail[1], Oury Monchi[1,3*]

[1]Hotchkiss Brain Institute, Department of Clinical Neuroscience, University of Calgary, Calgary, Alberta, Canada

[2] Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

[3] Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada

[4] Department of Human Genetics, McGill University, Montreal, Quebec, Canada

*** Correspondence:**

Corresponding Author

Oury.monchi@ucalgary.ca

Keywords: BDNF, Parkinson's disease, Mild Behavioural Impairment, Neuropsychiatric Symptoms (Min.5-Max. 8)

3.1 Abstract

Neuropsychiatric symptoms (NPS) are common in Parkinson's disease (PD) and have demonstrated an association with the p.Val66Met, a polymorphism in the *BDNF* gene. Mild behavioral impairment (MBI) is a validated syndrome describing emergent and persistent NPS in older adults as a marker of potential cognitive decline and dementia. This study investigated if PD patients with the Met allele were more likely to have MBI, and whether they had impairments in specific domains of MBI using the Mild Behavioral Impairment Checklist (MBI-C) as the MBI ascertainment tool. 146 PD patients were screened for neuropsychiatric and cognitive impairments with the MBI-C and the Montreal Cognitive Assessment (MoCA). All participants were genotyped for the *BDNF* p.Val66Met SNP using TaqMan Genotyping Assay. Statistical analysis was performed using multiple linear and logistic regression models. Met carriers had a two times higher likelihood of being MBI positive (MBI-C total score ≥8) than Val carriers. Met carriers had significantly higher MBI-C total scores, and significantly greater impairments in the mood/anxiety and the psychotic domains of MBI-C compared to Val carriers. These findings indicate that the *BDNF* Met allele is associated with a higher neuropsychiatric burden in PD.

3.2 Introduction

Neuropsychiatric symptoms (NPS) are common in Parkinson's disease (PD) patients. These NPS include depression, anxiety, psychosis, etc. which occur more frequently in PD patients than in the general population [6, 127]. NPS can be present at early stages of PD, and even precede the emergence of cardinal motor symptoms of PD [127]. They have a severe social and emotional impact on the quality of life in PD patients and their families/caregivers [128].

Mild behavioral impairment (MBI) is a validated syndrome characterized by the emergence of persistent NPS in older adults as an at-risk state for incident cognitive decline and for some, MBI is the index manifestation of dementia, emerging in advance of cognitive symptoms [129]. Early evidence in PD has linked MBI to altered corticostriatal connectivity, middle temporal lobe atrophy and cognitive impairment which suggest a higher risk of developing dementia [130, 131].

Brain-derived-neurotrophic-factor (BDNF) is a crucial protein in the central nervous system (CNS) with a substantial role in differentiation, survival and protection of CNS neurons [132]. Studies have investigated a potential role for p.Val66Met (G758A, rs6265), a single nucleotide polymorphism (SNP) in exon 11 of the *BDNF* gene, and PD [56, 57, 133]. The p.Val66Met SNP substitutes a Valine (Val) residue at position 66 with a Methionine (Met) residue in the pro-domain of the BDNF protein [132].The Met allele has been found to be associated with cognitive impairments in PD patients and late-life psychiatric symptoms in general population [56, 57, 134]. This substitution is not transferred to the final form of BDNF, however, this structural change in the BDNF protein precursor can significantly decrease the secretion of BDNF extracellularly and subsequently reduce its availability to the CNS neurons [132]. Recent evidence suggests a role of the Met allele in the induction of Long-Term Depression (LTD) in the brain, likely via altered interaction of BDNF pro-domain with sortilin receptors [54, 55]. The altered interaction might explain the connection of this polymorphism with NPS in the general population and in neurodegenerative diseases [60, 135]. Recent longitudinal data in Alzheimer Disease (AD) patients revealed strong evidence of Met association with depression [135]. Moreover, recent meta-analysis reported higher likelihood of mild cognitive impairment in PD patients with the Met allele [56].

These evidence suggest a link between the p.Val66Met polymorphism and NPS in PD patients, and therefore encourage an investigation of this relationship.

In this study, we tested whether the p.Val66Met SNP in PD patients is associated with MBI burden using the Mild Behavioral Impairment Checklist (MBI-C). Specifically, we hypothesized that PD patients with at least one Met allele (Met carriers) would have greater likelihood of having MBI and higher total MBI-C score than those who are Val homozygotes (Val group). Additionally, we hypothesized that Met carriers would have higher MBI-C domain scores compared to Val group.

3.3 Methods

3.3.1 Participants

146 PD patients at Hoehn & Yahr stages II-III were recruited. Patients had a confirmed diagnosis of idiopathic PD by a Movement Disorder Clinic neurologist, meeting the UK Brain Bank criteria for idiopathic PD. All patients were on prescribed dopaminergic medication and were responsive to it. Exclusion criteria were: 1) any neurological disorders other than PD; 2) alcohol dependency; 3) history or presence of a severe psychiatric disorder; and 4) cerebrovascular disorders. The severity of motor symptoms was assessed using the motor section of the Unified Parkinson's Disease Rating Scale (UPDRS-III). All participants provided written informed consent according to the declaration of Helsinki and the study was approved by the Conjoint Health Research Ethics Board (REB14-2463) at the University of Calgary.

3.3.2 Genotyping

A blood sample was collected from each participant and DNA was extracted using the MagMax DNA Multi-Sample Ultra 2.0 kit and the King Fisher Duo Prime Robot (Thermo Fisher Scientific). DNA samples were screened for the *BDNF* p.Val66Met SNP (rs6265) using TaqMan SNP Genotyping Assay C-11592758-10 on C-1000 Touch Thermal cycler (Bio-Rad). TaqMan assay reading was done on Applied Biosystems QuantStudio 7 Flex Real-Time PCR system (Fisher Scientific) according to the manufacturer's instructions. The TaqMan assay results were analyzed using the Bio-Rad CFX Maestro software.

3.3.3 Neuropsychiatric and cognitive assessment

NPS in all participants were evaluated using the MBI-C [107, 136]. The MBI-C contains 34 questions to cover the five domains of MBI including: 1) impaired drive/motivation (apathy); 2) emotional dysregulation (mood and anxiety symptoms); 3) impulse dyscontrol (agitation, aggression, abnormal reinforcement and reward salience); 4) social inappropriateness (impaired social cognition); and 5) abnormal thoughts/perception (psychotic symptoms). This checklist is completed by each patient's caregiver/close family member. Consistent with the MBI criteria, symptoms should have lasted for at least six months and present a meaningful change of behavior from longstanding patterns. An MBI-C total score cut point of ≥8 was used to classify a patient as MBI case positive [130, 137, 138]. All participants completed the Montreal Cognitive Assessment (MoCA) for a brief cognitive assessment and completed a questionnaire on their demographics, and daily activity level.

3.3.4 Statistical analysis

Statistical analyses of continuous variables were performed using either the student T-test or Mann-Whitney (M-W) U test based on the data normality. The Fisher Exact and Chi-square tests were used to test the categorical variables. Logistic regression was used to test the relationship between the MBI positive condition (the categorical dependent variable) and the two BDNF genotype groups, including any independent variables that were significantly different between the two conditions.

MBI-C total score (the continuous dependent variable) was compared between the two groups using a multiple linear regression model after checking for the multiple linear regression model assumptions. The independent variables that were correlated with MBI-C total score were included in the regression model. Values of $p<0.05$ were considered significant for single tests and significance of 0.01 was used to test MBI-C domains using Bonferroni correction. The same analysis was used to study association of p. Val66Met and MBI-C domain scores. All statistical tests were performed using IBM SPSS Statistics for Mac v. 26 (IBM Corp., Armonk, N.Y., USA). Power analysis was performed using G-Power software 3.1.9.6 [139].

3.4 Results

3.4.1 Demographics of participants

The demographic and clinical characteristics of the participants are summarized in Table 4.1.

10 patients were identified as outliers based on values that were more than 3 standard deviation away from the mean of each allelic group for the following variables; age, education, UPDRS, LEDD, disease duration, MoCA and MBI-C total score.

Among the 136 remaining PD participants, Val homozygous patients (GG) represented the majority of the cohort (*n*=90). Because of the low number of homozygous Met/Met patients (*n*=4), all Met carriers were pooled as one group. 46 patients were heterozygous or homozygous for the Met allele (GA, AA) with a frequency of 0.18 which was in accordance with Hardy-Weinberg equilibrium. The two groups had no significant differences in any of their demographic, or clinical characteristics, ethnicity and weekly exercise level (Table 4.1).

Table 4-1. Demographic and Clinical Characteristics of PD Participants.

Outliers were identified based on 3 standard deviations away from the mean values of demographic and clinical characteristics, ten participants were removed as outliers.

Characteristics **Mean, SD (Min-Max)**	**Val carriers (GG)** **n=90**	**Met carriers (GA, AA)** **n=46 (TT=4)**	**p-value**
Age	69.2, ± 8.1 (47-86)	66.7, ± 7.8 (48-79)	0.09[a]
Sex (female percentage)	36%	48 %	0.20[b]
Education (year)	14.8, ± 2.8 (8-21)	14.87, ± 2.5 (9-19)	0.89[a]
LEDD	809.7, ± 401.6 (200-1925)	822.3, ± 373.3(225-1675)	0.86[a]
Disease Duration (year)	5.71, ± 4.4 (0.2-16.1)	5.57, ± 3.9 (0.2-18.2)	0.86[a]
UPDRS-III	18.2, ± 10.0 (0-50)	20.3, ± 11.2 (0-49)	0.27[a]
MoCA	25.3, ± 4.0 (13-30)	25.9, ± 3.2 (18-30)	0.42[a]
Handedness **Right-handed** **Left-handed** **Ambidextrous** **NA**	 84% 12% 2% 1%	 87% 6% 2% 4%	0.49[c]
Ethnicity % **Caucasian** **Other** **NA**	 86.7% 7.8% 5.5%	 91.3% 4.0% 4.0%	0.7[c]
Exercise (hours per week)	5.8 ± 4.8 (0-28) [d]	6.2 ± 4.5 (0-20)	0.8[c]

Abbreviations: Val= Valine, Met= Methionine, SD =Standard Deviation, Min= Minimum, Max= Maximum, LEDD= Levodopa Equivalent Daily Dosage, UPDRS-III= Unified Parkinson's Disease Rating Scale part III, MoCA= Montreal Cognitive Assessment

a Student t test

b Fisher Exact test (2-sided)

c Chi-square test (2-sided)

d data was not available for two of the participants in Val group

3.4.2 Association of Val66Met and MBI-C score

Met carriers were twice as likely to be MBI positive than the Val carriers, 39% of Met carriers were MBI positive whereas in the Val group only 20 % of patients were MBI positive. The Met group had a significantly greater mean value for MBI-C total score than Val carriers (7.39 versus 4.06, respectively). Two factors were included in the multiple logistic regression as independent variables in addition to the *BDNF* groups based on significant differences between the two groups in the Mann-Whitney U test; MoCA and UPDRS-III (*M-W U*= 1230.5, p=0.005, and *M-W U*=2441.5, p=0.002, respectively). The logistic regression analysis revealed a significant contribution of the Met allele for the likelihood of being MBI positive (*OR*= 2.88, *CI 95%* = 1.22- 6.78, p=0.02) (Table 4.2).

Table 4-2. Multiple logistic regression analysis of BDNF p.Val66Met and MBI positive likelihood.

Covariate	Estimate	S E	Wald's chi square	p-value*	OR	95% CI
***BDNF* Met allele**	1.06	0.44	5.82	0.02[a]*	2.88	1.22-6.78
MoCA	-0.14	0.06	5.96	0.02[a]*	0.87	0.78-0.97
UPDRS-III	0.05	0.02	4.68	0.03[a]*	1.05	1.00-1.09
Constant	0.10	1.60	0.004	0.95[a]	NA	NA

Abbreviations: Val= Valine, Met= Methionine, MBI-C= Mild Behavioral Impairment Checklist, SE= Standard Error, OR= Odds Ratio, CI= Confidence interval, NA= Not Applicable

[a] = Multiple logistic regression model, N=136, Nagelkerke pseudo R^2= 0.22, Hosmer and Lemeshow goodness-of-fit test p=0.85 (df=8), correct cases overall percentage= 74.3 %

Most Val carriers had either zero or very low MBI-C total score (Figure 4.1).

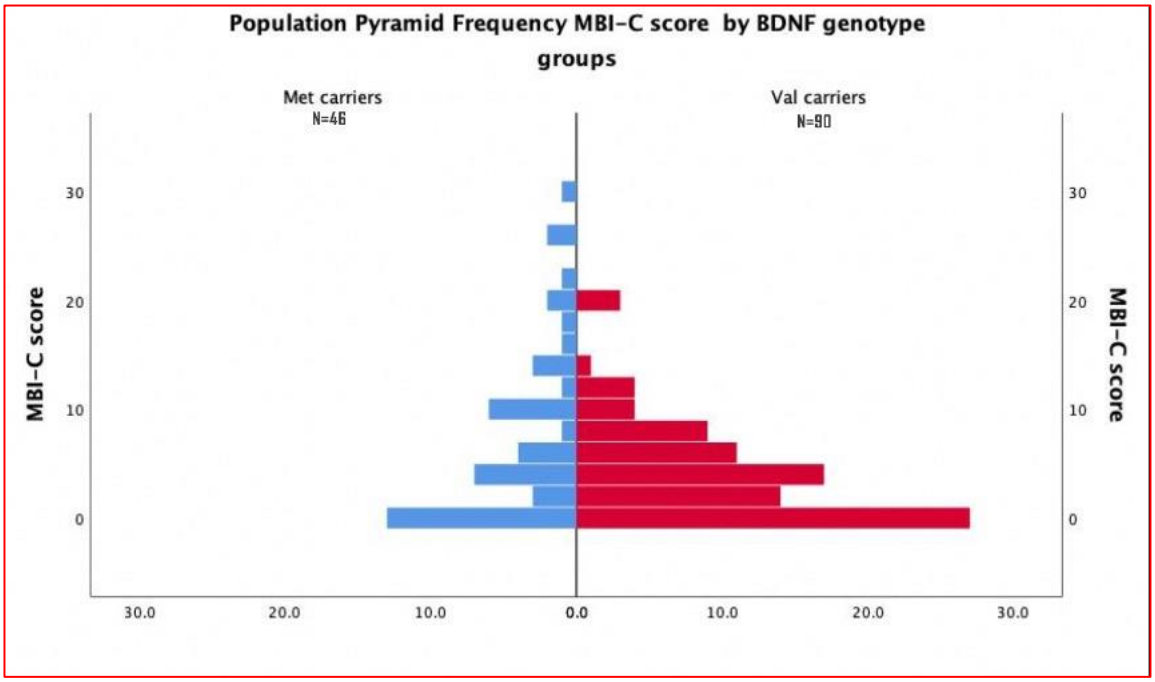

Figure 4-1. The population pyramid frequency of MBI-C total score by BDNF genotype.

MBI-C total and MoCA scores were negatively correlated (*Pearson's r* = (134) 0.17, *p*=0.04). Also, UPDRS-III and MBI-C total scores had a positive correlation (*Pearson's r* = (134) 0.23, *p*=0.007). These factors were included in the multiple linear regression model. The difference between the MBI-C total score between the two allelic groups was statistically significant when controlling for MoCA and UPDRS-III scores in the regression model (r^2= 0.13, *Beta*= 0.25, *F* (3, 135) = 6.36, *p*= 0.013, Cohen's f^2 = 0.15). A power analysis was conducted, which revealed that our samples size of 136, would yield a power of 0.99 assuming type-I error rate of 0.05.

Association results of *BDNF* alleles with MBI-C domain scores are shown in Table 4.3.

Table 4-3. MBI-C total and domain scores.

MoCA and UPDRS scores were used in multiple linear regression model. In total, 10 PD participants were identified as outliers and removed from the analysis.

MBI-C scores Mean, SD (Min-Max)	**Val carriers (GG) N=90**	**Met carriers (GA, AA) N=46**	**r^2, Beta (CI 95%)**	**p-value***
MBI-C total	4.06, ± 4.50 (0-20)	7.39, ± 8.05 (0-29)	0.13, 0.25 (1.17-5.35)	0.003**
Drive/Motivation	1.10, ± 1.72 (0-8)	1.43, ± 2.02 (0-8)	0.06, 0.07 (-0.38-0.92)	0.41
Mood/Anxiety	1.48, ± 2.21 (0-13)	2.87, ± 3.36 (0-12)	0.10, 0.24 (0.46-2.35)	0.004**
Impulse Dyscontrol	0.79, ± 1.43 (0-8)	1.15, ± 2.80 (0-13)	0.03, 0.10 (-0.28-1.16)	0.23
Social Inappropriateness	0.16, ± 0.62 (0-4)	0.43, ± 1.41 (0-7)	0.04, 0.14 (-0.07-0.63)	0.11
Abnormal Thoughts/Perception	0.53, ± 1.09 (0-4)	1.50, ± 2.54 (0-11)	0.12, 0.23 (0.26-1.48)	0.006**

Abbreviations: Val= Valine, Met= Methionine, MBI-C= Mild Behavioral Impairment Checklist, SD= Standard Deviation, Min= Minimum, Max= Maximum, MoCA= Montreal Cognitive Assessment, CI 95%= 95% Confidence Interval

* Analysis was performed using multiple linear regression model including MoCA and UPDRS scores in the model. MoCA and MBI-C total scores were negatively correlated r_p= 0.17, p = 0.043 while UPDRS and MBI-C total scores were positively correlated r_p= 0.23, p = 0.007.

** p-value was set to < 0.01 to correct for multiple tests, Bonferroni correction

Patients with the Met allele had significantly higher MBI-C scores for the mood/anxiety (r^2= 0.10, *Beta*= 0.24, *p*=0.004) and the psychosis domains (r^2= 0.12, *Beta*= 0.23, *p*=0.006) when controlling for MoCA and UPDRS-III scores (Table 4.3).

We performed an extra analysis in order to confirm that MBI classification results were derived from *BDNF* alleles and not driven by a few participants with marginally higher or lower MBI-C total score than the cut-off value (≥8). All participants with an MBI-C total score of 7 and 8 were excluded from the sample, and analysis was repeated. In total, 10 patients were removed; only one Met-carrier had the score of 7. In the Val group, three patients had the score of 7 and six patients had the score of 8 for MBI-C.

Similar to the main analysis, two factors were included in the multiple logistic regression as independent variables in addition to the *BDNF* groups, based on their significant differences between the two groups in the Mann-Whitney U test; MoCA and UPDRS-III (*M-W U*= 1024.5, *p*=0.02, and *M-W U*=1914.5, *p*=0.007, respectively). Results revealed a significant contribution of the Met allele for the likelihood of being MBI positive (*OR*= 4.38, *CI 95% =* 1.72- 11.14, *p*=0.002) (Table 4.4).

Table 4-4. Multiple logistic regression analysis of BDNF p.Val66Met and MBI positive likelihood after removing participants with total MBI-C score of 7 and 8.

10 participants were removed, including only one Met carrier with the score of 7 (n=126).

Covariate	Estimate	S E	Wald's chi square	p-value*	OR	95% CI
***BDNF* Met allele**	1.48	0.48	9.54	0.002[a*]	4.37	1.71-11.14
MoCA	-0.14	0.06	4.84	0.03[a]	0.87	0.77-0.98
UPDRS-III	0.04	0.02	3.76	0.053[a]	1.04	1.00-1.09
Constant	-0.71	1.71	0.19	0.67[a]	NA	NA

Abbreviations: Val= Valine, Met= Methionine, MBI-C= Mild Behavioral Impairment Checklist, SE= Standard Error, OR= Odds Ratio, CI= Confidence interval, NA= Not Applicable

[a]= Multiple logistic regression model, *N*=126, Nagelkerke pseudo R^2= 0.24, Hosmer and Lemeshow goodness-of-fit test *p*=0.31 (df=8), correct cases overall percentage= 77.8 %

* p-value was set to <0.01 corrected for multiple tests, Bonferroni correction

Each MBI-C domain score was compared between the two groups using Mann-Whitney U test (Table 5). The results were similar to the whole cohort analysis, Met carriers had significantly higher MBI-C total score than the Val group. Also, Met carriers had significantly higher score for the mood/anxiety domain when compared to the Val group and a trend for higher psychosis score (Table 4.5).

Table 4-5. MBI-C total and domain scores after removing participants with total MBI-C score of 7 and 8.

10 participants were removed, including only one Met carrier with the score of 7 (n=126).

MBI-C scores Mean, SD (Min-Max)	**Val carriers (GG) N=81**	**Met carriers (GA, AA) N=45**	**p-value***
MBI-C total	3.65, ± 4.56 (0-20)	7.40, ± 8.14 (0-29)	0.03**
Drive/Motivation	1.00, ± 1.72 (0-8)	1.40, ± 2.03 (0-8)	0.26
Mood/Anxiety	1.35, ± 2.21 (0-13)	2.84, ± 3.40 (0-12)	0.008**
Impulse Dyscontrol	0.64, ± 1.34 (0-8)	1.18, ± 2.83 (0-13)	0.87
Social Inappropriateness	0.10, ± 0.46 (0-3)	0.44, ± 1.42 (0-7)	0.09
Abnormal Thoughts/Perception	0.57, ± 1.13 (0-4)	1.53, ± 2.56 (0-11)	0.02

Abbreviations: Val= Valine, Met= Methionine, MBI-C = Mild Behavioral Impairment Checklist, MoCA= Montreal Cognitive Assessment, SD =Standard Deviation, Min= Minimum, Max= Maximum

* All the analysis was done by Mann-Whitney U test because the data normality.

** p-value was set to < 0.01 to correct for multiple tests, Bonferroni correction

3.5 Discussion

To our knowledge, the present study is the first to explore the association of the *BDNF* p.Val66Met SNP and MBI in patients with PD. Patients with at least one Met allele had significantly higher MBI-C total score and significantly higher scores in the emotional dysregulation and the abnormal thoughts/perception domains. Furthermore, PD patients with at least one Met allele had a significantly higher prevalence of MBI than patients in the Val group using MBI-C as the case ascertainment instrument. Our findings implicate the *BDNF* p.Val66Met SNP in the pathogenesis of MBI in PD patients, and suggests this variant as a genetic risk factor for MBI in PD with a medium effect size (Cohen's f^2= 0.15). These findings are consistent with the evidence in the AD, which imply that the Met allele can be a risk factor for incident cognitive decline and dementia in PD [129, 135, 140].

An increasing body of evidence suggests a link between the development of NPS and cognitive decline in different types of dementia [130, 141-143]. Studies demonstrating that *BDNF* p.Val66Met SNP is found to be associated with both NPS and cognitive impairments in AD are consistent with a biological understanding of NPS [144-148], and previous evidence linking *BDNF* and NPS [57, 135]. The presence of Amyloid-β pathology in PD patients together with Lewy-body pathology might be a possible explanation of similar NPS profile in AD and PD patients. However, it should be mentioned that different studies have teased apart how the psychiatric profile of PD and AD patients are different [149, 150].

Patients who experience NPS in the early stages of PD show an increased risk of cognitive decline [127, 130], which is consistent with the findings in non-PD dementia [143, 151-155]. A recent study reported that PD patients with a variety of NPS e.g. depression, apathy, hallucinations, displayed impairments in at least one of the main cognitive domains (executive-function, language, memory, attention and visuospatial) and in global cognition [130]. These findings hint at the importance of early diagnosis of sustained NPS as markers of cognitive decline in PD patients, in order to identify patients at risk of incident cognitive decline and dementia.

A recent meta-analysis reported an association between the *BDNF* Met allele and cognitive impairments in PD for 532 patients and 802 controls (p=0.003). However, the cognitive impairments were found to be more specific to the Caucasian populations [56]. Several studies suggested that p.Val66Met SNP might have an association with depression, particularly geriatric depression [60, 156, 157]. Nonetheless, this association might differ based on a variety of factors e.g. the origin of the study population, sex and the fundamental issue of whether depression is chronic, recurrent, or of later life onset as most depression rating scales

do not differentiate [57, 152]. The association of *BDNF* Met allele and geriatric depression was investigated in a meta-analysis including 523 cases and 1,220 controls (age ≥ 60 years)[60]. An association between the Met allele and an increased risk for late-life depression was reported (p=0.004)[57]. However, one study reported that the Val alleles is associated with anxiety and depression using the Beck Depression Inventory (BDI) and Beck Anxiety Inventory (BAI) in 104 PD patients. Nevertheless, their population structure was greatly different than the one in our study. 17 % of their participants had early onset PD with a positive family history [62], while in our sample there were only two participants (1.4%) with early onset PD and all of the participants had a confirmed diagnosis of idiopathic PD. Other reasons explaining the different results between the Cagni et al., 2013 study and ours are linked to measurement differences of NPS. The BDI and BAI are self-report measures assessing the presence of mood and anxiety symptoms over the last 2 and 4 weeks respectively. In contrast, the MBI-C measures later life emergent and persistent (for at least 6 months) NPS, identified by a reliable informant. These are quite different approaches to measurement of symptoms, with the MBI-C developed explicitly to capture later life emergent symptoms that either serve as risk factors for cognitive decline and dementia, or more likely represent early manifestations of dementia.

Furthermore, BDNF involvement in depression/anxiety disorders has been confirmed through measuring peripheral BDNF levels as well [158, 159]. A meta-analysis reported a strong evidence of an association between depression and a decrease in BDNF levels ($p< 6.8 \times 10^{-8}$) [158]. These meta-analyses highlight the crucial role of BDNF in depressive disorders and specifically the impact of p.Val66Met SNP on geriatric depression. The Met allele exhibits LTD properties, reduces the neural plasticity, and can also substantially affect the docking of

BDNF secretory vesicles into the cellular membrane and decrease its release into the synaptic cleft [132]. The BDNF pro-domain with a Met residue is shown to have an independent function by induction of LTD, reducing spine density and neuronal plasticity. These molecular changes are linked to depression and anxiety disorders in both animal models and clinical studies [132, 160, 161]. These findings are in agreement with our results that PD patients with at least one Met allele are more susceptible to impairments in the affective/mood dysregulation and abnormal thoughts/perception domains of MBI-C.

We found a strong association between abnormal thoughts/perception in PD patients and Met allele in our cohort. Abnormal thoughts/perception represent psychotic symptoms, specifically hallucinations and delusions, which are associated with impairments in global cognition [127, 162]. An abrupt visual memory function is suggested as a potential cause of visual hallucinations. Since BDNF plays a prominent role in the molecular mechanisms of memory in hippocampus, this indicates a possible role for BDNF in the development of such NPS[127, 132]. Meta-analytical results of p.Val66Met SNP and psychotic disorders e.g. schizophrenia are inconclusive at the moment [57, 163]. However, the Met allele was found to be linked to higher susceptibility to hippocampal volume loss and deteriorated memory abilities in bipolar patients [164]. These findings are in agreement with our results that PD patients with at least one Met allele are more susceptible to symptoms in the affective/mood dysregulation and abnormal thoughts/perception domains.

This study has some limitations that need to be considered. Although the findings of this study showed a fair level of robustness, the sample size is relatively small. The results of this study need to be replicated in a larger sample with an age-matched control group.

Nevertheless, the post-hoc power analysis indicates sufficient power in our cohort to detect the true effect of *BDNF* p.Val66Met. A full cognitive assessment of participants would benefit the exploration of p.Val66Met SNP impact on the aging brain, specifically in PD. Antidepressant medications was not considered. It has been shown that antidepressant medications can elevate peripheral BDNF and improve reversible NPS [158].

In conclusion, we observed an association between the *BDNF* p.Val66Met SNP and susceptibility for the development of late-life behavioral changes in PD patients. PD patients with at least one Met allele had a higher likelihood of MBI compared to non-carriers. Moreover, PD patients with one Met allele had a greater tendency to exhibit mood and anxiety symptoms as well as psychotic symptoms compared to the Val carriers. These findings indicate a potential role for the *BDNF* p.Val66Met SNP in late-life psychiatric impairment, subsequent cognitive decline and dementia in PD patients.

3.6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

3.7 Author Contributions

MR; study design, data analysis, interpretation, manuscript writing and revision

JR; data collection

KM; data collection

MK; data collection

ZJ: data collection

IK; data collection

TH; data collection

JC; data collection

EL; data collection

DM; data collection

JR; data collection

ZGO; data collection

GP; data collection

ZI; study design, analysis, manuscript revision

OM; study design, data analysis, supervision, manuscript revision

3.8 Funding

This work was funded by a project grant from the Canadian Institutes of Health Research (CIHR) (PJT-166123), the Tourmaline Oil Chair in Parkinson's Disease, the Canada Research Chair in non-motor symptoms of Parkinson’s disease to OM. MR received a scholarship from the Parkinson Association of Alberta.

3.9 Acknowledgments

The authors would like to thank all the PD patients and their families who were willing to participate and provide information for this study.

3.10 Data Availability Statement

The datasets for this study can be accessed upon contact with the corresponding author.

4 Chapter Four

In this chapter, we studied the association of two common variants (rs4680 and rs28363170) in the genes involved in dopamine modulation with NPS in PD patients. We tested the hypothesis that Met allele carriers of rs4680 and 9-R carries of rs28363170 were more susceptible to behavioral symptoms specifically to emotional dysregulation (mood and anxiety).

This chapter is reproduced from a manuscript which has been submitted for publication at the time of the thesis defense.

Lack of association between the genes involved in dopamine availability and the risk of mild behavioral impairment in Parkinson's disease patients

Mehrafarin Ramezani[1,2], Jennifer A Ruskey[3,4], Kristina Martens[1,2], Zainul Javer[1,2], Tracy Hammer[1,2], Mekale Kibreab[1,2], Iris Kathol[1,2], Nadia Maarouf[1,2], Etienne Leveille[3], Davide Martino[1,2], Justyna R Sarna[1,2], Gerald Pfeffer[1,2,5], Zahinoor Ismail[1,2,6,7,8], Oury Monchi[1,2,4,9*]

1-Department of Clinical Neurosciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

2- Hotchkiss Brain Institute (HBI), Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

3- Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

4-Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada

5-Alberta Children's Hospital Research Institute, Department of Medical Genetics, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; Hotchkiss Brain Institute

6-Department of Psychiatry, University of Calgary, Calgary, Alberta, Canada

7-Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada

8-O'Brien Institute for Public Health, Cumming School of Medicine, Calgary, Alberta, Canada

9-Department of Radiology, University of Calgary, Calgary, Alberta, Canada

*** Correspondence:**

Corresponding Author

Oury.monchi@ucalgary.ca

4.1 Abstract

Parkinson's disease (PD) is one of the most common neurological disorders in older adults in which patients suffer from the loss of dopaminergic neurons. PD patients can display neuropsychiatric symptoms (NPS) in addition to motor symptoms. The catechol-O-methyltransferase (*COMT*) contains the rs4680 polymorphism. The two common alleles of rs4680, Val and Met, can modify the rate of dopamine degradation. The rs28363170 in Solute Carrier Family 6 member 3 (*SLC6A3*) is a variable number of tandem repeat (VNTR) polymorphism, with two common alleles 9-R and 10-R. These alleles affect the rate of transportation of dopamine. The mentioned variants were reported in association with NPS and cognitive changes in PD. In this study, we tested the hypothesis that Met allele carriers and 9-R carries are more susceptible to NPS, both globally and with respect to symptoms of emotional dysregulation, using the Mild Behavioral Impairment-Checklist (MBI-C) as the screening tool. 177 PD patients were recruited and screened for cognitive deficits, and NPS using the Montreal Cognitive Assessment and MBI-C. No associations were found between either variant and the risk of NPS or emotional dysregulation in PD patients using a regression model. The current results might indicate a lack of connection between these variants and the risk of NPS in PD patients.

Keywords: Parkinson's disease, cognitive impairment, catechol-o-methyltransferase, solute carrier family 6 member 3, mild behavioral impairment, dopamine transporter

4.2 Introduction

Parkinson's disease (PD) is one of the most common neurological disorders in the geriatric population. PD patients suffer from motor disabilities, but they can also manifest a wide range of non-motor symptoms including autonomic dysfunction, cognitive and neuropsychiatric symptoms (NPS) [6, 119, 127, 165]. The substantial loss of dopaminergic neurons in the substantia nigra of PD patients generates a low dopamine state. Dopaminergic medications are commonly prescribed to compensate for the reduced dopamine availability in the brain of PD patients [166]. However, dopamine availability in the brain depends on several biological factors. Two prominent internal factors are the catechol-O-methyltransferase (COMT) enzyme and the dopamine transporter protein.

The COMT enzyme is involved in the degradation of catecholamines (e.g. dopamine) in the brain by addition of a methyl group to them [64]. This enzyme has two common isoforms, Val and Met isoforms. These two isoforms are results of a single nucleotide polymorphism (SNP) in *COMT* known as rs4680 (p.Val158Met, G>A). The Val isoform has the normal activity level while the Met isoform displays up to 40% reduced activity *in vivo,* which can consequently increase the dopamine availability in the synaptic cleft [66].

The dopamine transporter protein can influence the availability of dopamine through reuptake and terminates its action in the synaptic cleft [167]. The gene encoding this transporter protein is the Solute Carrier Family 6 member 3 (*SLC6A3*). A common variant in a form of variable number of tandem repeat (VNTR) located in the 3' untranslated region of this gene affects the activity of this transporter protein. This VNTR (rs28363170), has two common alleles: 9 and 10 repeats (9-R and 10-R, respectively) [168]. The 10-R allele is associated with higher expression of the SLC6A3 [169]. The rs28363170 was reported to be linked to better cognitive ability in PD patients [73].

Both rs4680 and rs28363170 are associated with psychiatric disorders such as bipolar disorder, depression, and attention deficit hyperactivity disorder (ADHD) in the general population [170]. The relationship between these variants and NPS in PD patients has not been investigated. The rs28363170 has been reported in association with psychosis in PD [171]. Although the association of rs4680 with cognitive impairment in PD has been reported in several studies, little is known of its potential association with NPS in PD patients [172-174].

Mild behavioral impairment (MBI) is a validated syndrome that is characterized by the emergence of persistent NPS in later life which indicate a potential risk of cognitive decline [107]. In some, MBI can be an early clinical manifestation of neurodegeneration, and preceding cognitive symptoms [175]. This syndrome contains five major domains of behavioral symptoms: 1) impaired drive/motivation (apathy); 2) emotional dysregulation (mood and anxiety symptoms); 3) impulse dyscontrol (agitation, aggression, abnormal reinforcement and reward salience); 4) social inappropriateness (impaired social cognition); and 5) abnormal thoughts/perception (psychotic symptoms). Studies have demonstrated an association between MBI and cognitive impairment in PD and the general population [120, 176, 177].

In this study, we explored the association between the rs4680 and rs28363170 variants and NPS in PD patients using the Mild Behavioral Impairment-checklist (MBI-C) [107] as the measurement tool. We hypothesized that PD patients with the Met allele for rs4680 would be more likely to have MBI as well as symptoms of emotional dysregulation. Similarly, we hypothesized that PD patients with the 9-R allele for rs28363170 would be more likely to have MBI as well as symptoms of emotional dysregulation.

4.3 Methods

4.3.1 Participants

177 PD participants were recruited at the University of Calgary, Calgary, Alberta, Canada as part of the PD-MCI study. All patients had a diagnosis of idiopathic PD by a neurologist and met the UK Brain Bank criteria [178]. Participants did not have any other neurological disorders.

The severity of motor symptoms was assessed using the Hoehn & Yahr scale [179] by an experienced research nurse based on each participant's medical records and the notes of the assigned neurologist of the participant. Disease duration, and levodopa equivalency daily dosage (LEDD) for each patient who had initiated the dopaminergic therapy were calculated by a research nurse. A list of psychiatric and mono amine oxidase B (MAOB) medications taken by the participants was also collected.

All participants provided written consent in accordance with the declaration of Helsinki. This study was approved by the Conjoint Health Research Ethics Board (REB14-2463) at the University of Calgary, Calgary, AB, Canada.

4.3.2 Genotyping

All participants had a blood draw, and their DNA sample was extracted using an isopropanol-based method. The rs4680 isoform was genotyped using a TaqMan genotyping assay with the ID of C-25746809-50 on a 1000-Touch Thermal cycler. The reading of TaqMan genotyping assay was performed on an Applied Biosystem Quantstudio Flex 7 Real-Time PCR system (Fisher Scientific) following the manufacturer's instructions. The analysis of TaqMan genotyping assay was done using Bio-Rad CFX Maestro software.

The rs28363170 polymorphism was analyzed using the following forward and reverse primers to amplify the region surrounding this variant:

F-TGTGGTGTAGGGAACGGCCTGAG

R-CTTCCTGGAGGTCACGGCTCAAGG.

The PCR protocol was 30 seconds at 95°C, 36 cycles of (95°C for 15 seconds, 60°C for 30 seconds, 70°C for 60 seconds), 68°C for 5 minutes and 4°C for hold on a C-1000 Touch Thermal Cycler (Biorad). The primer sequences and protocol were described previously [180].

4.3.3 Neuropsychiatric and Neuropsychological Assessment

All patients were assessed for NPS using the MBI-C [107], a general psychopathology scale developed to capture NPS in accordance with the MBI criteria, and which we have implemented previously in a PD population [120, 176, 181]. The MBI-C is a 34-item questionnaire which explicitly captures symptoms emerging in later life, persisting for ≥6 months, and representing a meaningful change in the participant's behavior. Items were scored on severity from 0-3; a cut-off score of ≥8 was used to classify a patient as MBI positive [108, 109]. The MBI-C was completed by a third party in close contact with the participant e.g., close relative, family member.

All participants received a brief cognitive assessment using the Montreal Cognitive Assessment (MoCA) administered by a psychometrist or a trained study staff [106]. Patients also completed demographic questionnaires.

4.3.4 Statistical Analysis

The between group differences for the demographic and clinical measures were tested using either F-test, student t-test or Kruskal-Wallis test based on the normality of the data. The correlation of the MBI-C total score with demographic and clinical measures were examined using appropriate correlation tests according to the normality of the variables. The association of each variant with MBI likelihood was tested using a multiple logistic regression model.

The relationship of MBI-C total score (dependent variable) and either of the variants was investigated using ANCOVA. The covariates were selected using the demographic and clinical measures correlated with MBI-C total score or any variables which were significantly different between the groups when the participants were divided according to the genotype of the target variant. The same analysis was performed to study the relationship between these variants and the emotional dysregulation domain of MBI-C. All statistical analysis were performed using IBM SPSS Statistics for Mac v. 27 (IBM Corp., Armonk, N.Y., USA).

4.4 Results

4.4.1 Demographics of Participants

The PD participants were divided into three groups based on their genotype for rs4680 (GG, GA and AA) to compare the effect of genotype on MBI likelihood. The same procedure was repeated for rs28363170, in which the participants were divided into two groups (10-R/10-R, and 9-R carriers) because of the low number of homozygous 9-R PD patients.

The demographic and clinical characteristics of PD patients are summarized in Table 5-1. The outliers were identified by evaluating the following variables in each group per variant; age, years

of education, years of disease duration, LEDD and MoCA score. In total, 7 participants were identified as outliers based on 3 standard deviations distance from the mean of each group for the controlled variables.

The 170 PD patients had the following genotypes for rs4680: GG = 43, GA = 82 and AA = 40. Three outliers had GG genotype and the rest of the outliers were heterozygous. The frequency of the minor allele (A) was 0.48.

Table 5-6. Demographic and clinical measures for PD patients (N_T = 170), when patients were divided based on their rs4680 and rs28363170 genotypes.

Demographic / Clinical measures **Mean, SD, (Min-Mix)**	***COMT* (rs4680)**				***SLC6A3 (*rs28363170)**		
	GG (N=41)	GA (N=89)	AA (N=40)	sig*	10/10 (N=)	10/9 & 9/9 (N=)	sig*
Age	70.19,± 8.19(47 - 82)	67.90, ± 8.40 (42 – 86)	66.87, ± 9.39 (46 - 88)	0.20[a]	69.50, ± 8.05 (47 - 88)	66.71, ± 9.08 (42 - 86)	0.04[*d]
Education	14.46, ± 2.65 (9 - 20)	15.29, ± 2.56 (9 - 21)	13.95, ± 2.52 (8 - 20)	0.02[*a]	15.11, ± 2.61 (9 - 21)	14.38, ± 2.61 (8 - 20)	0.07[d]
Sex (Female %)	39.0 %	44.9 %	25.0 %	0.10[b]	32.6 %	46.2 %	0.08[b]
Disease Duration	6.01, ± 3.97 (1.78 – 15.50)	5.53, ± 4.03 (0.69 – 17.25)	7.11, ± 5.63 (0.75 – 23.17)	0.46[c]	6.20, ± 4.88 (0.72 - 23.17)	6.13, ± 4.26 (0.69 – 19.58)	0.78[e]
LEDD	833.0,± 329.4 (300 - 1862)	721.3, ± 374.6 (0 - 1800)	985.4, ± 503.7 (225 - 2075)	0.003[*c]	765.8, ± 376.9 (0 - 2075)	873.8, ± 457.8 (200 - 1995)	0.09[e]
H & Y **Stage I** **Stage II** **Stage III** **Stage IV**	7.3 % 78.0 % 9.8 % 4.9 %	5.6 % 82.0 % 10.1 % 2.2 %	10.0 % 67.5 % 22.5 % 0 %	0.30[b]	7.6 % 76.1 % 13.0 % 3.3 %	6.4 % 79.5 % 12.8 % 1.3 %	0.84[b]
Handedness **Right-handed** **Left-handed** **Ambidextrous**	92.7 % 4.9 % 2.4 %	84.3 % 9.0 % 6.7 %	95.0 % 2.5 % 2.5 %	0.39[b]	88.0 % 8.7 % 3.3 %	89.7 % 3.8 % 6.4 %	0.30[b]
MoCA	25.51, ± 3.39 (15 - 30)	25.62, ± 3.76 (14 - 31)	25.50, ± 3.63 (19 - 30)	0.94[c]	25.14, ± 3.60 (14 - 31)	26.06, ± 3.62 (14 - 30)	0.02[*e]
MAOB inhibitors (on %)	9.8 %	27.0%	27.5%	0.07[b]	25.5 %	20.5%	0.58[b]
Psychiatric medications (on %)	31.7%	37.1%	55%	0.07[b]	41.3%	38.5%	0.76[b]
Family history of PD (positive%) [**]	17.1	14.6%	15.0%	0.99	15.2%	15.4%	0.89
Exercise (hours per week) [***]	5.51, ± 3.24, (0 - 14)	6.12, ± 4.15, (0 - 20)	5.06, ± 4.33, (0 - 15)	0.37	5.31, ± 3.72 ,(0 - 17)	6.18, ± 4.48, (0-20)	0.17[d]

Abbreviations: Min = Minimum, Max = Maximum, LEDD = Levodopa Equivalency Daily Dosage, H and Y = Hoehn & Yahr, MoCA = Montreal Cognitive Assessment, MAOB inhibitor = Monoamine Oxidase B inhibitor, COMT = Catechol-O-Methyltransferase, SLC6A3 = Solute Carrier family 6 member3

a F-test

b Fisher-test

c Kruskal-Wallis test

d t-test

e Mann-Whitney U test

* significance set to < 0.05

** data was unavailable for 28 participants due to unknown family history or missing data

*** data was not available for 10 participants due to missing data

The distribution of rs28363170 alleles were as follows; 10-R/10-R = 92, 10-R/9-R = 67, 9-R/9-R = 11. Three outlier had 10-R/10-R genotype and the other 4 outliers were heterozygous. The frequency of the minor allele (9-R) was 0.26.

The demographic characteristics of the participants were evaluated based on their rs4680 groups. Years of education and LEDD were significantly different between the three groups ($p = 0.02$, and $p = 0.003$, respectively). A trend was found for the MAOB inhibitor and the psychiatric medications between the three groups ($p = 0.07$, $p = 0.07$, respectively).

The demographic characteristics of the participants were also evaluated based on their rs28363170 groups. It was observed that age was significantly different between the two groups ($p = 0.04$). A trend was observed for years of education between the two groups ($p=0.07$).

The MoCA score was significantly different between the two groups $p = 0.02$, with 10-R/10-R PD patients having a lower average MoCA score compared to 9-R carriers.

The demographic and clinical measures for this cohort are detailed in Table 5-1.

4.4.2 Association of p.Val158Met with MBI

The association of p.Val158Met with MBI likelihood was explored including age, years of education, LEDD, MoCA score, and the MAOB inhibitor in the multiple logistic regression model.

The covariates were selected based on the demonstration of a difference between the three groups or a significant correlation with the MBI-C total score. MoCA score was negatively correlated with the MBI-C total score (R^2 = 0.04, p = 0.01). The MAOB inhibitor use was positively associated with the MBI-C total score (Pearson chi-square = 38.91, p = 0.02, respectively).

Table 5-7. Association of rs4680 in *COMT* gene with MBI likelihood using multiple logistic regression. The following variables were included as covariates: age, education, LEDD, MoCA, and MAOB inhibitors.

Covariate	**Estimate**	**S E**	**Wald's chi square**	**p-value***	**OR**	**95% CI**
***COMT* rs4680**	–	–	3.08	0.21		
AA > GG	0.98	0.56	3.08	0.08	2.66	0.89 – 7.92
GA > GG	0.60	0.51	1.39	0.24	1.83	0.67 – 4.99
GA > AA	-0.37	0.45	0.69	0.41	0.69	0.28 – 1.67
Constant	3.27	2.62	1.56	0.21	NA	NA

Abbreviations; *COMT*= Catechol-O-Methyltransferase, S.E.= Square root error, OR=Odds ratio, 95% CI= 95% Confidence Interval

Nagelkerke R^2 = 0.14, Hosmer and Lemeshow test = 4.51, df = 8, p = 0.81, the correct overall percentage = 74.7

* the significance level was set to <0.05

The p.Val158Met was not associated with MBI likelihood, Wald = 3.08 (df = 2), p = 0.21. The pair-wise comparisons of the three groups did not reveal any associations between MBI-C likelihood and any of the groups (Table 4-2). The frequency of MBI positive patients in the three groups was the following: AA = 37.5%, GA = 27.0%, and GG = 17.0% (Figure 5-1 A).

The association of rs4680 and MBI-C total score was tested using ANCOVA with age, years of education, LEDD, MoCA score and MAOB inhibitor as covariates. No association was observed between this variant and MBI-C total; F (2, 169) = 0.66, p = 0.52.

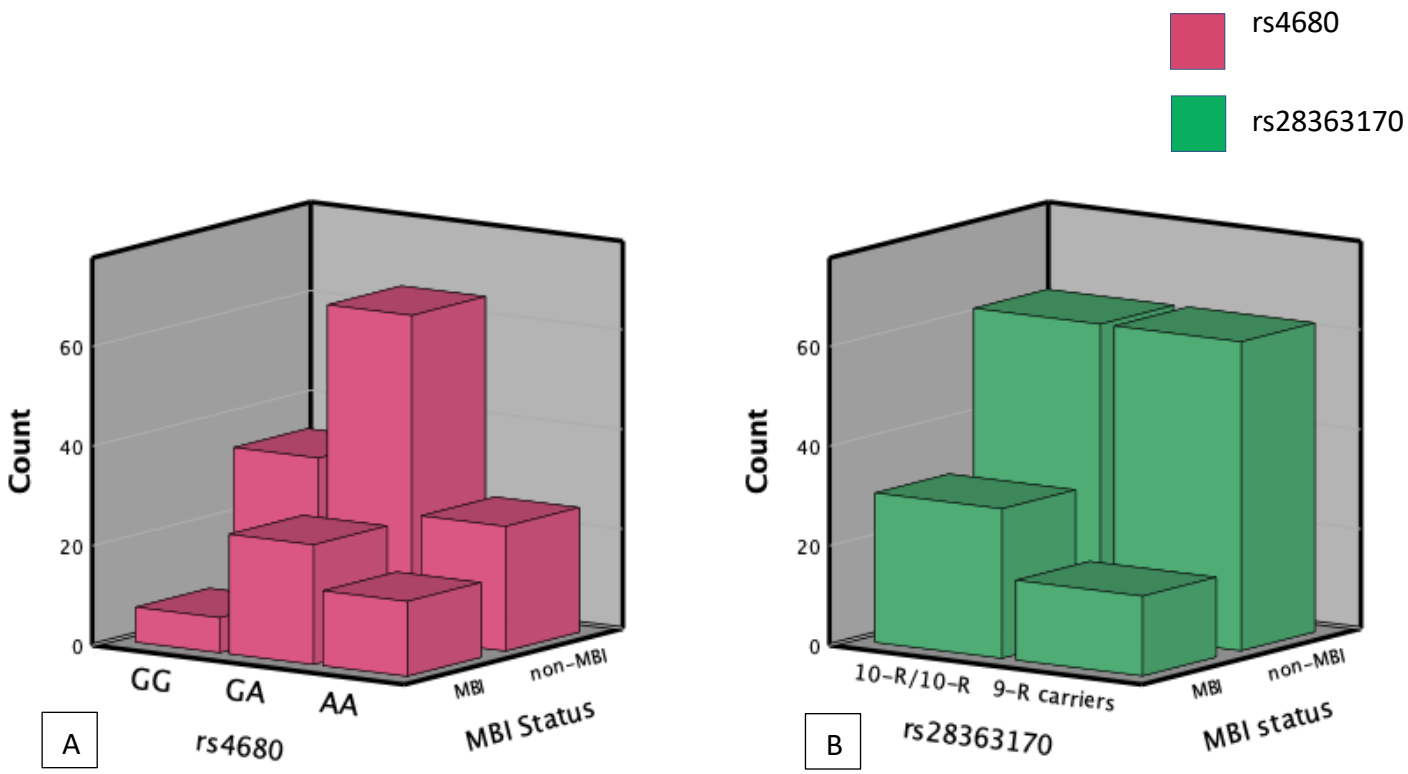

Figure 5-6. Schematic view of MBI likelihood in PD patients when the cohort was divided based on the rs4680 and rs28363170 genotypes.

A) Column chart of participants divided into 3 groups according to rs4680 genotype. Met/Met carriers have the highest ratio of MBI patients compared to the other groups (AA = 37.5% versus GA = 27.0% and GG = 17.0%, respectively). **B)** Column chart of participants divided into 2 groups according to rs28363170 genotype. PD patients with 10-R/10-R genotype had higher rate of MBI compared to 9-R carriers but did not reach significance (32.6 % 10-R/10-R homozygous versus 20.5 % for 9-R carriers).

The relationship between this variant and the emotional dysregulation domain of MBI-C was explored and no association was found ($F(2,169) = 0.64$, $p = 0.53$). Details for this analysis can be found in Table 5-3.

The analysis was repeated using COMT genotype to divide the participants into two groups (GG vs GA+AA and AA vs GG+GA). The same procedure as described above was performed to check whether this SNP has any association with MBI risk.

Table 5-8. The results of association analysis between rs4680 in *COMT* gene and MBI-C total score, and MBI-C emotional dysregulation domain scores. The following variables were included as covariates: age, education, LEDD, MoCA, and MAOB inhibitors.

Covariate	**MBI-C total score**		**MBI-C emotional dysregulation domain score**	
	Mean Difference	p-value*	Mean Difference	p-value*
***COMT* rs4680**	NA	0.52	NA	0.53
AA > GG	1.59	0.25	0.73	0.26
GA > GG	0.85	0.47	0.41	0.46
GA > AA	-0.70	0.56	-0.32	0.58

Abbreviations; COMT = Catechol-O-Methyltransferase, MBI-C = Mild Behavioral Impairment Checklist

* The significance level was set to <0.05

No significant association was found between rs4680 and the risk of MBI for either of these analyses, Wald $_{\text{GG+GA versus AA}}$ = 1.83 (df = 1), p = 0.18, and Wald $_{\text{GG versus GA+AA}}$ = 2.37 (df = 1), p = 0.12. No significant differences were found for either of these two grouping (Table S5-1).

4.4.3 Association of rs28363170 with MBI

The association of rs28363170 with MBI likelihood was tested using multiple logistic regression with age, MoCA score, and MAOB inhibitor use as covariates. MoCA and MAOB inhibitor use displayed an association with MBI-C total score as mentioned before. The rs28363170 VNTR did not show an association with MBI likelihood, Wald = 2.04, df = 1, p = 0.15.

The frequency of MBI positive patients in 9-R group was 20.5 %, versus 32.6 % for homozygous 10-R group (Figure 5-1.B).

The association of this variant with MBI-C total score was examined using ANCOVA with age, MoCA score and MAOB inhibitor use as covariates. No association was found between this variant and the MBI-C total score F (1, 169) = 0.05, p = 0.82. A lack of association was observed for this variant and the total score of the emotional dysregulation domain of MBI-C, F (1, 169) = 0.04, p = 0.85 (Table 5-4).

Table 5-9. The results of association analysis between rs28363170 and MBI-C total score, and MBI-C emotional dysregulation domain scores. The following variables were included as covariates: age, MoCA, and MAOB inhibitors.

Covariate	**MBI-C total score**		**MBI-C emotional dysregulation domain score**	
	Mean Difference	p-value*	Mean Difference	p-value*
***SLC6A3* rs28363170**	0.44	0.82	0.17	0.85

Abbreviations; SLC6A3 = Solute Carrier Family 6 member 3, MBI-C = Mild Behavioral Impairment Checklist

* The significance level was set to <0.05

4.5 Discussion

In this study, we investigated whether two common variants involved in dopamine availability were linked to NPS in PD patients. We did not find any associations for either the Met allele of rs4680 or the 9-R allele of rs28363170 with either global NPS burden (i.e., MBI) or symptoms of emotional dysregulation. To the best of our knowledge, this is the first study to investigate the relationships between these two common variants involved in dopamine availability in frontal regions and the risk of NPS in PD patients using the MBI-C.

The rs4680 variant has a prominent role in dopamine availability and is reported in association with various psychiatric disorders in the general population *e.g.*, schizophrenia, ADHD, aggressive behavior and depression [63, 65, 182]. A well-known model for the effect of this genetic variant on mood was proposed by Goldman et. al [183].

They suggested that its effect could be described through the warrior/worrier model in addictive disorders. They observed that substance abusers with the Val allele were susceptible to dyscontrol or impulsivity. On the other hand, the Met allele was observed in higher frequencies in substance abusers who suffered from anxiety disorders [183]. This model proposed a better anxiety resilience for the Val carriers versus better overall cognition in the Met carriers [183]. The substitution of valine residue with methionine affects the enzymatic activity considerably [64]. The Goldman model seems to be in agreement with the impact of this variant on the enzymatic activity of the COMT enzyme. *In vivo* studies revealed that the Met allele can decrease the enzymatic activity up to 40 % [66]. Several studies reported the effect of rs4680 in PD patients, and these findings were in agreement with the Goldman model [172-174].

It has been proposed that rs4680 might have a dual effect on cognition in PD patients [34]. It has been suggested the Met allele might benefit PD patients at the more advanced stages of the disease. The inverted U-shape model proposed by Williams-Gray et al. described the effect of this variant with more advantages for the Val carriers at the early stages of the disease when the dopamine availability is not too scarce. But as the disease progresses the higher activity of COMT enzyme in Val/Val carriers diminishes the dopamine level in the frontal regions to a higher degree compared to the Met carriers. This could hinder the dopamine dependent activity of the frontal regions in the Val/Val carriers to a higher extent compared to non-carriers [34]. To date there is no evidence suggesting the same effect proposed by the Williams-Gray model or the Goldman model for this variant regarding NPS.

The possible link between this variant and behavioral aspects of PD was not addressed in these studies. A meta-analysis in the general population reported a pleiotropic effect for this variant regarding the cognitive and emotional paradigms processing.

According to this meta-analysis, the Met carriers demonstrated better cognitive abilities while the Val carriers displayed better emotional processing abilities [184].

Considering the potential effect of the disease duration on the effect of COMT in PD patients, the lack of association in the current study might be due to heterogeneity of patients regarding the disease stage. The disease duration for PD patients in this cohort was between 6 to 7 years on average but the range was from less than one year to more than 23 years. Also, the Hoehn & Yahr stage of participants varied from stage 1 to 4 (23 % versus 10%, respectively). The difference in Hoehn and Yahr stage can be a potential source of heterogeneity in this cohort [185]. Furthermore, considering the inverted U-shape effect of this enzyme on the dopamine availability in the brain, the broad range of disease stages in this cohort might explain the negative results in this study.

No association was observed between the rs28363170 polymorphism and the risk of NPS in PD patients in this cohort. The 10-R allele of rs28363170 is linked to an increased expression of *SLC6A3* [169]. This variant was reported in association of cognitive decline and NPS in PD patients [73, 74, 171]. However, there are contradictory results from different studies which reported a lack of association between this variant and cognitive impairments in PD. Also, studies reported a lack of a relationship between this variant and impulsive/compulsive disorders, depression, and psychosis in PD patients [186-189].

PD patients who suffer from NPS are at greater risk of cognitive decline [6, 110, 119, 120]. We have previously shown that PD patients with MBI demonstrate cognitive impairments in attention, memory, executive-function and visuo-spatial abilities [120]. Moreover, it was shown that PD patients with depression display a similar pattern of cortical atrophy as PD patients with cognitive decline over time [110, 119]. A functional MRI (fMRI) study showed that PD patients with the 9-R allele displayed a diminished neural activity, similar to PD patients with cognitive impairment.

Based on the findings discussed above, a potential link between rs28363170 and the risk of behavioral symptoms in PD patients could be suggested [73]. Nevertheless, the current findings do not indicate any associations between this variant and the likelihood of having MBI in PD patients.

It should be mentioned that 10-R/10-R patients had lower MoCA scores compared to 9-R carriers in this study. This result was unlike the previous findings reporting that 9-R carriers were at higher risk of cognitive decline [73]. The difference in MoCA score of 10-R homozygous patients might be attributed to the significant age difference between the two groups of rs28363170 in this study. Although MoCA score is corrected for age and education the 10-R/10-R group had a significantly higher average age compared to 9-R carriers.

The lack of association of rs28363170 with MBI in this study, might be explained by the Williams-Gray et al. model. According to this model the genes which are involved in dopamine modulation do not propel irreversible cognitive impairments. These cognitive impairments are associated with temporal, parietal and more posterior regions of the brain and are linked to higher risk of dementia in PD [35]. In other words, cognitive impairments that are associated with anomalies in the frontal region (fronto-striatal dopamine) were not found to be linked to the onset of dementia in PD patients [35]. This could suggest that despite the evidence of an association between the 9-R allele and cognitive impairments in PD patients, this relationship might not indicate further risks regarding NPS, more severe cognitive impairments and dementia [35, 73].

This study had several important limitations which need to be addressed. This study had a relatively small sample size for an association analysis. It would be useful to include an age and sex matched control group to compare the possible effects of these variants between PD and controls.

It should be mentioned that these variants might have different mechanisms for their association in the psychiatric disorders versus the neurodegenerative diseases. It needs to be mentioned that these negative results do not rule out the possibility of undetected NPS that were below the syndromic threshold of MBI. A comprehensive cognitive assessment along with MBI data might assist elucidating the relationship of these variants with cognition and behavior in PD in more detail. Also, given the potential inverted U-shape effect of COMT, it might be helpful to recruit the participants with similar disease duration and at the similar Hoehn &Yahr stages to have a more uniform cohort.

In conclusion, we did not observe any association between either rs4680 or rs28363170, and the likelihood of MBI in this cohort. We did not find any associations between the above-mentioned variants and the emotional dysregulation in PD patients. These findings might indicate that these variants are not involved in the disruption of neural circuits in charge of NPS in PD.

Data Availability

The data is available upon request. Please contact the corresponding author at oury.monchi@ucalgary.ca.

Acknowledgement

The authors would like to thank all PD patients and their families who were willing to participate and provide information for this study.

This work was funded by a project grant from the Canadian Institutes of Health Research (CIHR) (PJT-166123), the Tourmaline Oil Chair in Parkinson's Disease, the Canada Research Chair in non-motor symptoms of Parkinson's disease to OM. MR received a scholarship from the Parkinson Association of Alberta (PAA). This work was financially supported by Parkinson Canada, the Michael J. Fox Foundation, the Canadian Consortium on Neurodegeneration in Aging (CCNA) and the Canada First Research Excellence Fund (CFREF), awarded to McGill University for the Healthy Brains for Healthy Lives (HBHL) program. Infrastructure for work performed in the GP lab was supported by a grant from the Canada Foundation for Innovation John R Evans Leaders Fund.

Competing Interests

The authors declare no competing interests.

Table S5-1 The results of association analysis between rs4680 in *COMT* and MBI-C total, and MBI-C emotional dysregulation domain scores when the *COMT* genotype was used to divide the participants into two groups.

The following variables were included as covariates: age, education, LEDD, MoCA, and MAOB inhibitors.

Covariate	**MBI-C total score**		**MBI-C emotional dysregulation domain score**	
	Mean Difference	p-value*	Mean Difference	p-value*
***COMT* GG versus GA+AA**	-1.29	0.32	-0.54	0.32
***COMT* GG+GA versus AA**	-1.8	0.36	-0.85	0.39

Abbreviations; COMT = Catechol-O-Methyltransferase, MBI-C = Mild Behavioral Impairment Checklist

*The significance level was set to <0.05

Part two: Genetic Association and Cognitive Impairment in Idiopathic PD

5 Chapter Five

In this chapter we aimed to predict cognition Z-scores in iPD patients through machine learning using a combination of genetic, demographic, clinical measures, and structural neuroimaging. In the second step, we used post-hoc analysis to investigate the effect of the most novel selected feature of the machine learning analysis and explored its association with different cognitive domains.

This chapter is reproduced from a paper published in Scientific Reports:

Ramezani M*, Mouches P, Yoon EJ*, Rajashekar D*, Ruskey JA, Leveille E, Martens K, Kibreab M, Hammer T*, Kathol I*, Maarouf N*, Sarna J, Martino D, Pfeffer G, Gan-Or Z, Forkert ND M,onchi O. (2021) Investigating the relationship between the SNCA gene and cognitive abilities in idiopathic Parkinson's disease using machine learning. Scientific Reports. 11(1): 4917.

Investigating the relationship between the *SNCA* gene and cognitive abilities in idiopathic Parkinson's disease using machine learning

Mehrafarin Ramezani[1,2], Pauline Mouches[2,3,4], Eunjin Yoon[1,2], Deepthi Rajashekar[2,3,4], Jennifer A. Ruskey[5,6], Etienne Leveille[5], Kristina Martens[2], Mekale Kibreab[1,2], Tracy Hammer[1,2], Iris Kathol[1,2], Nadia Maarouf[1,2], Justyna Sarna[1,2], Davide Martino[1,2], Gerald Pfeffer[1,2,7], Ziv Gan-Or[5,6,8], Nils D. Forkert [1,2,4], Oury Monchi[1,2,4,6*]

Authors Affiliations

1-Department of Clinical Neurosciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.

2- Hotchkiss Brain Institute (HBI), Cummings School of Medicine, University of Calgary, Calgary, Alberta, Canada

3- Biomedical Engineering Graduate Program, University of Calgary, Calgary, Canada.

4-Radiology, University of Calgary, Calgary, Alberta, Canada

5- Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

6-Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada

7-Alberta Children's Hospital Research Institute, Department of Medical Genetics, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; Hotchkiss Brain Institute

8-Department of Human Genetics, McGill University, Montreal, Quebec, Canada

[*]**Corresponding Author** oury.monchi@ucalgary.ca

5.1 Abstract

Cognitive impairments are prevalent in Parkinson's disease (PD), but the underlying mechanisms of their development are unknown. In this study, we aimed to predict global cognition (GC) in PD with machine learning (ML) using structural neuroimaging, genetics and clinical and demographic characteristics. As a post-hoc analysis, we aimed to explore the connection between novel selected features and GC more precisely and to investigate whether this relationship is specific to GC or is driven by specific cognitive domains. 101 idiopathic PD patients had a cognitive assessment, structural MRI and blood draw. ML was performed on 102 input features including demographics, cortical thickness and subcortical measures, and several genetic variants (*APOE*, *MAPT*, *SNCA*, etc.). Using the combination of RRELIEFF and Support Vector Regression, 11 features were found to be predictive of GC including sex, rs894280, Edinburgh Handedness Inventory, UPDRS-III, education, five cortical thickness measures (R-parahippocampal, L-entorhinal, R-rostral anterior cingulate, L-middle temporal, and R-transverse temporal), and R-caudate volume. The rs894280 of *SNCA* gene was selected as the most novel finding of ML. Post-hoc analysis revealed a robust association between rs894280 and greater GC, attention, and visuospatial abilities. This variant indicates a potential role for the *SNCA* gene in cognitive impairments of idiopathic PD.

Keywords: Parkinson, cognition, machine learning, SNCA, mild cognitive impairment, dementia

5.2 Introduction

Parkinson's disease (PD) is traditionally associated with motor symptoms. Non-motor deficits have received less attention but are also quite common and can even precede motor symptoms. PD patients can display a wide range of cognitive deficits of various intensity from Mild Cognitive Impairment (MCI: neuropsychological deficits that do not interfere with daily life activities) to dementia [190]. PD patients with cognitive deficits are more likely to develop Parkinson's disease dementia (PDD) as the disease progresses, compared to patients without such deficits [14]. To address the high prevalence and the inevitable negative impact of cognitive deficits, the Movement Disorders Society (MDS) prepared a task force to assist identification of MCI in PD patients [10, 15]. Nevertheless, the underlying mechanisms of PD-MCI development are still obscure and further investigation is required to unravel its nature. At the cellular level, PD is characterized by presence of Lewy Body (LB) deposits in the substantia nigra [191]. LB mainly consists of alpha-synuclein [192], a protein that is encoded by the *SNCA* gene [193]. *SNCA* is a prominent potential genetic marker in PD due to its involvement in familial PD through point mutations or gene dosage effect [29, 86, 194]. *SNCA* is also implicated in a class of disorders called synucleinopathies, which have LB pathology in common, e.g. PD (with or without dementia), dementia with LB (DLB), Multiple System Atrophy (MSA), idiopathic REM Behavioral Disorder (RBD). PDD and DLB share numerous similarities contributing to the challenge of distinguishing them from each other [88, 195]. The substantial involvement of *SNCA* in PDD and DLB pathology through LB and the similarity of the symptomatology, such as the presence of dementia in these two diseases, indicate a potential role for the *SNCA* gene in cognitive decline of idiopathic PD patients and necessitates further investigation. In recent years, several variants of *SNCA* have been discovered in connection with cognitive impairments or dementia in PD [88, 94]. Exploring possible genetic variants

associated with cognitive impairments in PD could improve the understanding of the primary biological mechanisms of PD-MCI and identification of patients at risk of cognitive decline. However, given the complexity of PD, further investigation is required to unravel the true involvement of the known and novel genetic variants linked to cognitive deficits in PD using advanced techniques [196].

In addition to genetic risk factors, structural neuroimaging has been used in PD participants to identify image-based biomarkers of cognitive decline [113, 197-200]. A growing body of evidence supports the use of structural neuroimaging as a biomarker for PD-MCI identification. One example is thinning of frontal and temporal cortices, which has been associated with cognitive decline in PD [197, 200, 201]. Also, corticometric and volumetric analyses have shown a reduced volume in frontal and limbic regions in PD-MCI compared to PD-nonMCI patients[113, 119, 202].

In this study, we aimed to predict global cognition in PD participants using a machine learning approach including genetic, structural neuroimaging, clinical, and demographic data as input. In a second step, using post-hoc analyses, we aimed to identify the relationship between the most novel genetic features and the cognitive profile of PD patients in more depth. Furthermore, we aimed to investigate whether this relationship is specific to global cognition or driven by specific cognitive domains including executive function, attention, visuospatial abilities, memory, and language.

5.3 Results

5.3.1 Participants

The demographic and clinical characteristics of 101 PD participants included in this study are summarized in Table 6.3. The mean age of participants was 70 years with a mean disease duration of 6 years. Thirty-two precents of the participants were female (n=32). Ninety participants were right-handed (89%), six participants were left-handed (6%), and the remaining five participants were identified as ambidextrous (5%). The majority of the participants were of European descent (84%).

Table 5-1. The demographic and clinical characteristics.

Demographic/clinical characteristics	**Mean, SD (Min, Max)**
Age (years)	70, ± 7.3 (53, 84)
Education (years)	14.4, ± 2.9 (8, 21)
Disease Duration (years)	5.7, ± 4.0 (1, 18)
LEDD (mg/d)	869.0, ± 457.6 (200, 2650)
UPDRS-III	19.46, ± 9.41 (4, 53)
EHI	68.9, ± 41.6 (-100, 100)
Sex (female %)	32%
Ethnicity **European Descent** **Others** **NA**	 84% 10% 6%
Global cognition Z-score	-0.27, ± 0.6 (-2, 1)
Attention Z-score	-0.24, ± 0.6 (-1, 1)
Executive function	-0.43, ± 0.8 (-3, 1)
Language	-0.13, ± 0.8 (-2, 1)
Memory	-0.15, ± 0.8 (-2, 1)
Visuo-spatial	-0.43, ± 0.9 (-3, 1)

Abbreviations; LEDD= Levodopa Equivalency Daily dosage, UPDRS-III= Unified Parkinson's disease Rating Scale part III, EHI= Edinburg Handedness Inventory, NA= Not available

5.3.2 Machine Learning Analysis

The global cognition scores were predicted using the 102 features employing the machine learning framework including feature ranking and a support vector regression model. The best model performance predicting global cognition was achieved when including the 11 top-ranked features (Table 6.4). This resulted in a correlation coefficient of 0.54 and mean absolute error of 0.39. The selected features were (in order of descending importance): sex, rs894280, EHI, UPDRS-III, education, five measures of cortical thickness (right parahippocampal cortex, left entorhinal cortex, right rostral anterior cingulate cortex, left middle temporal cortex, and right transverse temporal cortex), and right caudate volume.

Table 5-2. Machine learning analysis results.

Features are presented in the right column based on their importance in the machine learning model. The features are presented in a descending order (the most important to the least important).

Cognitive domain	**Correlation Coefficient, R^2**	**Features**
Global cognition	0.54, 0.29	Sex rs894280 EHI score UPDRS-III Education R parahippocampal-thickness L entorhinal-thickness R rostralanteriorcingulate-thickness L middletemporal-thickness R transeversetemporal-thickness R caudate-volume

Abbreviations: UPDRS-III= Unified Parkinson's disease Rating Scale part III, EHI= Edinburgh Handedness Inventory

The high-ranked demographic features were further analyzed individually using Pearson correlation or chi-square test. Correlation analysis of sex and EHI score with global cognition were found to be non-significant (χ^2= (1, n=101) 75.59 p=0.46 and R^2=0.005, p=0.50, respectively). Also, UPDRS-III and years of education exhibited significant correlation with global cognition. The UPDRS-III score had a negative correlation while years of education had a positive correlation with global cognition (R^2= - 0.26, $p < 0.001$ and R^2= 0.07, p= 0.007, respectively). The structural measures were tested for correlation with global cognition and two measures showed significant correlation with global cognition: R rostral-anterior cingulate-thickness, and L middle temporal-thickness (R^2= 0.08, p=0.004, R^2=0.04, p=0.03, respectively)

5.3.3 Association of the SNCA Variant rs894280 and Global Cognition

Based on the machine learning results, post-hoc analyses were performed to study the association of the alleles of the novel variant rs894280 with global cognition and specific cognitive domains.

Out of 101 participants, 33 had CC genotype, 48 had CT genotype, and 20 had TT genotype. Based on the preliminary analysis of this SNP, participants with the T allele were pooled in one group, resulting in dividing the participants in two allelic groups [203]. The demographic and clinical characteristics were not significantly different between the two allelic groups (Table 6.5).

This variant showed a significant association with global cognition when controlling for UPDRS-III, education, R rostral-anterior cingulate and L middle temporal thickness measures (F (5,95) = 12.17, p<0.001, R^2= 0.35). Based on the ANCOVA results, global cognition was significantly different between the CC and CT/TT groups, F (5,95) = 4.20, p=0.04.

Table 5-3. Demographic and clinical characteristics of the two groups of rs894280 variant.

Demographic/clinical characteristics **Mean, SD, (Min, Max)**	**CC n=33**	**CT/TT n=68**	**Significance**
Age (years)	71.5, ± 7.3	69.2, ± 7.3	0.14[a]
Education (years)	14.6, ± 3.2	14.3, ± 2.3	0.66[a]
Disease Duration (years)	5.6, ± 4.2	5.7, ± 4.0	0.56[b]
LEDD (mg/d)	767.7, ± 300.9	918.2, ± 511.7	0.31[b]
UPDRS-III	18.14, ± 7.82	20.10, ± 10.09	0.44[b]
EHI	69, ± 40	69, ± 40	0.83[b]
Sex (female%)	33%	31%	0.82[c]
Ethnicity			0.67[d]
European descent	88%	82%	
Others	9%	10%	
NA	3%	7%	

Abbreviations; SD= Standard Deviation, LEDD= Levodopa Equivalency Daily dosage, UPDRS-III= Unified Parkinson's disease Rating Scale part III, EHI= Edinburgh Handedness Inventory, NA= Not available

[a] student-t test

[b] Mann-Whitney U test

[c] Fisher exact test

[d] Chi-square test

According to ANCOVA analysis, each participant's Z-score of global cognition increased by 0.25 (95% CI= 0.01– 0.42) when the participant had CC genotype (i.e. reference sequence) for rs894280. The calculated Hedges' effect size for rs894280 is 0.4, which represents a medium effect size based on 95% CI.

Association analysis of rs894280 and each cognitive domain was performed using ANCOVA including any demographic/clinical factor with a significant correlation with the domain of interest as covariates. UPDRS-III, education, R rostral-anterior cingulate and L middle temporal thickness measures were included for all domains, except for the visuo-spatial domain for which the EHI

score was also added to the model (R^2=0.05, p=0.02). Significance level was set to 0.01 using Bonferroni correction to correct for multiple tests.

A significant association was found between rs894280 and the attention domain score (B= -0.34, p=0.003) with the CC group displaying better attention abilities. Moreover, this variant showed an association with the visuo-spatial domain score (B=-0.51, p=0.005). Similar to results from global cognition and attention, PD patients with CC genotype had better visuo-spatial abilities compared to the other group (CT/TT). The results of association analysis are summarized in Table 6.6. A trend was observed for the memory domain (p=0.02) while executive function and language did not show any association with this variant. PD participants homozygous for C allele of rs894280 displayed superior attention and visuo-spatial abilities compared to participants who had one or more T alleles.

Table 5-4. Results of ANCOVA analysis of rs894280, and global cognition and the five cognitive domains.

Cognitive domain Mean, SD, (Min-Max)	**CC N=33**	**CT/TT N=68**	**CI95% rs894280**	**Significance**
Global cognition	0, ± 0.5, (-1, 1)	-0.4, ± 0.6, (-2, 1)	0.006 – 0.42[a]	0.04[*]
Attention	0, ± 0.5, (-1, 1)	-0.4, ± 0.6, (-1, 1)	0.12 – 0.56[a]	0.003[**]
Executive function	-0.3, ± 0.7, (-2, 1)	-0.5 ± 0.9, (-3, 1)	-0.10 – 0.55[a]	0.19
Language	0, ± 0.7, (-1, 1)	-0.2, ± 0.8, (-2, 1)	-0.16 – 0.44[a]	0.36
Memory	0.2, ± 0.7, (-1, 1)	-0.3, ± 0.8, (-2, 1)	0.07 – 0.70[a]	0.02
Visuo-spatial	-0.1, ± 0.6, (-2, 1)	-0.6, ± 1, (-3, 1)	0.16 – 0.87[b]	0.005[**]

SD= Standard Deviation, CI 95%= 95% confidence intervals

[a] UPDRS-III, Education, R rostral-anterior cingulate and L middle temporal thickness measures were entered in the model

[b] UPDRS-III, Education, R rostral-anterior cingulate and L middle temporal thickness measures and EHI score were entered in the model

[*] significance level was set to 0.05

[**] significance level was set to 0.01, Bonferroni correction

5.4 Discussion

In this study, we used machine learning to predict global cognition in PD patients and post-hoc analysis to investigate the *SNCA* rs894280 variant as a feature associated with cognitive deficits in PD. Using the RRELIEFF feature selection algorithm and SVR, eleven features were selected as the best predictor of global cognition Z-score in this cohort; sex, rs894280, EHI, UPDRS-III, education, five measures of cortical thickness (right parahippocampal cortex, left entorhinal cortex, right rostral anterior cingulate cortex, left middle temporal cortex, and right transverse temporal cortex), and right caudate volume. The selection of features indicate that these variables are informative for prediction of the global cognition score but the direction for each single feature cannot be easily determined based on the machine learning model. Consistent with the machine learning results, which revealed rs894280 as the only genetic factor informative for PD cognition prediction, further analysis was performed on the association of this variant and global cognition. The results suggest that this variant is associated with differences in global cognition, as well as attention and visuo-spatial domains in our cohort, with a medium effect size (Hedges' $g = 0.4$).

The RRELIEFF approach was used to remove redundant and non-informative features and to select the optimal subset of features. This feature selection method has been used for optimal selection of genetic features in previous studies [204-207]. The SVR was implemented to build the model based on the selected features in order to evade the collinearity issue of the features. The SVR model has been used previously to model PD diagnosis and progression [208, 209] but the current combination used in this study has not been applied before specifically in investigation of cognitive deficits in PD with similar set of inputs.

Except for rs894280, all other features used in the optimal regression model have been reported in different studies to be associated with cognitive decline in PD patients. There is a substantial body of evidence on the role of sex in cognitive decline in PD, with male patients showing greater risk of cognitive impairments [35, 210, 211]. However, it also needs to be mentioned that several other studies reported no evidence of the impact of sex in cognitive decline in PD [15, 21]. The reason for these conflicting results remains speculative but could be related to the sample size. EHI is a well-known screening tool to determine handedness [212]. The correlation between the dominant hand and the side of motor symptoms onset has been suggested by several studies [213-215]. This relationship might extend to the cognitive impairments in PD as some studies suggest [216, 217].

The severity of motor symptoms has been suggested to be one of the strongest risk factors for cognitive deficits in idiopathic PD [15, 18, 218, 219]. UPDRS-III is one of the most widely used screening tools for the severity of motor symptoms in PD and can accurately and efficiently note the presence and progression of those symptoms [220]. In line with the findings of this study, other studies have also found evidence for a connection between the severity of motor symptoms and emergence of cognitive decline [221, 222]. The number of years of education was also found to be predictive of global cognition in our cohort. Similarly, a large body of evidence reported a negative correlation between higher education level and the likelihood of cognitive impairments in PD patients [15, 21, 35]. It has been suggested that education has a role in preserving the cognitive reserve in PD patients at risk of cognitive decline [22].

The structural biomarkers of cognitive decline in PD have been investigated extensively using different techniques [21, 223, 224]. The right parahippocampal gyrus, the top anatomical feature identified in this study, has been reported as one of the main brain regions showing significant Dopamine receptor (D2) binding reduction in PD patients [225].

Another machine learning study reported the parahippocampal region as a top feature showing the highest correlation with the motor score in PD [226]. Similarly, this region was also identified as one of the top features in Alzheimer's disease (AD). These findings suggest a more general function of the parahippocampal region in neurodegenerative diseases given its prominent role in memory [227]. The entorhinal cortex was previously reported as one of the main brain regions allowing a fine distinction between PD-MCI and PD-nonMCI patients [228]. The right entorhinal volume was observed to be positively correlated with memory abilities in early drug-naïve PD-MCI patients [228]. Additionally, cortical thinning of the entorhinal region was found to correlate significantly with memory impairments in PD patients [229]. The anterior cingulate cortex is another ROI associated with cognitive impairment in PD. A large body of evidence indicates a link between PD-MCI cognitive status and the anterior cingulate [165, 225, 230, 231]. These findings are in accordance with the results of the present study and indicate a potential relationship between this region and cognition in PD. The last anatomical feature predictive of global cognition was the right caudate volume. The caudate nucleus is one of the chief regions in PD pathology and extensive loss of neurons in this nucleus was reported in association with cognitive impairment and dementia in PD [73, 232-234].

We used a combination of known genetic risk factors (H1 *MAPT, ε*4 *APOE, COMT p.Val158Met, DAT1 VNTR, BDNF p.Val66Met*) and novel genetic variants to predict the global cognition in this cohort [35, 73, 88]. The rs894280 was selected as the novel finding for the post-hoc analysis because of its importance in the machine learning model. Ranked as the second top feature, this variant could present a meaningful role in prediction of global cognition in this cohort. This finding was in contrast with the known genetic risk factors included in this study which were not selected by the machine learning model. This variant is an intronic polymorphism located on the 5' region

of *SNCA* gene and was initially reported in association with dementia with Lewy bodies (DLB) [235]. The role of *SNCA* gene mutations in familial PD has been known for decades. However, new data suggest a role for this gene in cognitive deficits and dementia in idiopathic PD [88],[94, 174, 236]. A recent study indicated association between several *SNCA* variants and worse performance in Mini Mental state examination (MMSE) in PD patients[174]. Similar association was reported on the association of several *SNCA* variants and PDD[88]. A microsatellite (Rep1) is located on the 3' region of the *SNCA* gene and has two common alleles (short repeat and long repeat). The long repeat allele of Rep1 seems to increase *SNCA* transcription and was reported linked to lower MMSE scores in PD patients[236]. On the other hand, the 5' region of the *SNCA* gene was considered as a haplotype specific for DLB and not PD. This evidence was further supported by another study investigating the *SNCA* role in both DLB and PD [88]. Both PDD and DLB are classified as synucleinopathies and share substantial similarities in symptoms and pathology, to the point that the exact differentiation of these two disorders clinically and pathologically are still a matter of debate [93]. The rs894280 has been reported in association with both DLB and PDD and this might suggest a more general role for this variant in LB pathology.

Furthermore, rs894280 is in linkage disequilibrium (LD) with rs1348224 with comparable odds ratio (D'=1.0, R^2=1.0). The rs1348224 variant was previously reported in association with PDD, surviving multiple testing in a sample of 1492 PD patients [88]. Moreover, a strong correlation was reported between rs894280 and the Hopkins Verbal Learning Test-Revised (HVLT-R) total recall in PD patients (p=6.1 x 10^{-4}), and it displayed the strongest relationship with cognitive abilities out of 39 *SNCA* variants included in the study. However, this association did not survive after correction for multiple comparisons. This could indicate a role for rs894280 in PD cognitive abilities, especially in the memory domain [237]. A Brazilian study found cognitive impairments

in PD patients carrying T allele of rs2583988 of *SNCA*. The rs2583988 is in a strong LD with rs894280 (D'=1.0, R^2=0.40 in European descendent populations), which further indicates a possible role for rs894280 in cognitive decline of PD patients [94].

Deficits in the attention, visuo-spatial, and memory domains are frequently reported in PD-MCI patients [10, 238, 239]. Association of rs894280 with impairments in these domains in idiopathic PD patients may indicate a role for this variant in the development of such deficits. Specifically, this SNP might be connected to visuo-spatial abilities given that attention measures used in this study have a prominent visuo-spatial component. Studies have shown that attention measures with a visual component can tap on to visuo-spatial abilities [240]. Out of three attention measures used in this study, two of them; Trail A and Symbol Span have the required component to engage both visuo-spatial and attention suggesting a potential role for this SNP in connection to visuo-spatial abilities.

We did not observe any association between executive function and language domains and rs894280 in our cohort. A possible explanation for this could be that executive function impairments involve the frontal-striatal areas while most of the cognitive deficits identified in this study are focused in more medial temporal lobe and posterior cortical regions [35]. Although deficits in the language domain are reported in association with dementia in PD [35], this SNP did not show any link to language abilities in this cohort.

This study had some limitations that should be mentioned. We used a machine learning approach in this study in an effort to capture the underlying complex patterns of the various input features and focused on the most unique and relevant features for further investigation. These findings are preliminary and need replication in larger cohorts.

The present cohort size was small for a genetic analysis, but our results displayed a fair level of robustness, in a cohort that is extensively phenotyped and well-characterized. These results need to be replicated in a larger cohort with higher number of genetic variants to avoid missing effect of other potential risk variants before a definite conclusion can be inferred on this specific variant and cognitive impairment in PD.

In conclusion, using machine learning, we found that rs894280 in *SNCA* was one of the top features predictive of cognition in PD patients. Further analysis in the same cohort revealed association of this variant (CC genotype) with attention and visuo-spatial abilities in PD patients with a trend in the same direction for memory abilities. These results indicate a potential involvement of *SNCA* variant rs894280 in the cognitive deficits and even dementia in idiopathic PD patients.

5.5 Methods

5.5.1 Participants

101 PD patients at Hoehn and Yahr stages II-III were recruited. All patients had a confirmed diagnosis of idiopathic PD by a movement disorder clinic neurologist, meeting the UK brain bank criteria for idiopathic PD. All participants were responsive to dopaminergic medications and took their usual dosage of medications during all study visits. None of the participants were asked to modify their medications for this study. Exclusion criteria were: 1) any neurological disorder other than PD, 2) alcohol dependency, 3) history or presence of a severe psychiatric disorder, and 4) cerebrovascular disorders. The severity of motor symptoms was assessed by a trained professional using the motor section of the Unified Parkinson's Disease Rating Scale (UPDRS-III).

Levodopa Equivalency Daily Dosage (LEDD) and disease duration of all participants were calculated by a research nurse.

All participants provided written informed consent according to the declaration of Helsinki and the study was approved by the Conjoint Health Research Ethics Board (REB14-2463) at the University of Calgary, AB, Canada. All methods were carried out in accordance with the relevant guidelines and regulations.

5.5.2 Genotyping

A blood sample was collected from each participant and DNA was extracted using an isopropanol-based protocol. DNA samples were screened for several Single Nucleotide Polymorphisms (SNP) using TaqMan genotyping assays on a C-1000 Touch Thermal Cycler. The list of all SNPs and TaqMan assays investigated in this study are shown in Table 6.1. TaqMan assay reading was done using Applied Biosystem Quantstudio Flex 7 Real-Time PCR system (Fisher Scientific) according to the manufacturer's instructions. TaqMan (assays) results were analyzed using Bio-Rad CFX Maestro software. The 40 bp Variable Tandem Repeats (VNTR) located on the 3' region of Solute Carrier 6 family 3 *(SLC6A3)* was amplified using PCR (30 seconds at 95°C , 36 cycles of (95°C for 15 seconds, 60°C for 30 seconds, 70°C for 60 seconds), 68°C for 5 minutes and 4°C for hold) on a C-1000 Touch Thermal Cycler (Biorad), using the primers and protocol described previously [180]. PCR products were mixed with loading dye and loaded on 2% agarose gel containing gelstar and run at 120 V for 30 minutes followed by 60 minutes at 100 V. A 100 bp DNA ladder (Biohelix, DM 001-R500F, FroggaBio) was loaded on each gel to determine the molecular size of PCR products. The length of PCR products was captured using the Chemidoc Imaging System (Biorad).

Table 5-5. The list of the TaqMan genotyping assays and the primers used for genotyping.

Gene	Variant	TaqMan genotyping Assay
SNCA	rs7689942	C-28994912-10
	rs894280	C-8933128-10
APOE	rs7412	C-904973-10
	rs429358	C-3084793-20
COMT	rs4680	C-25746809-50
MAPT	rs393152	C-2265265-10
BDNF	rs6265	C-11592758_10
SLC6A3*	VNTR 9/10 R	F-TGTGGTGTAGGGAACGGCCTGAG R-CTTCCTGGAGGTCACGGCTCAAGG

**SLC6A3* VNTR was genotyped using custom made primers. For details, please refer to methods.

5.5.3 Neuropsychological Assessment

All participants completed a comprehensive cognitive assessment. The cognitive battery applied in this study consists of tests and measures covering five cognitive domains: executive function, memory, attention, visuo-spatial, and language. The full list of cognitive tests is shown in Table 6.2. All tests were scored by a trained psychometrist. The cognitive tests were first scored using the test makers manual, which details specific parameters to evaluate an examinee's performance. The total raw score is then converted to a standardized score by comparing the examinee's score to other healthy controls matched for age. All neuropsychological tests that were corrected for age, and most were corrected for years of education, and sex.

The measures corresponding to the same cognitive domain were averaged to obtain the average Z-score for each of the cognitive domains. The global cognition Z-score was calculated by averaging all cognitive domains' Z-scores. The Edinburgh Handedness Inventory (EHI) was administered to the participants and scored to identify each participant's dominant hand.

Table 5-6. The neuropsychiatric battery.

Cognitive domain	Cognitive test/measure
Attention	Trail A Wechsler Memory Scale (WMS) symbol span Digit-span Forward
Executive function	Hayling 2 Brixton Trail B Stroop Colour/Word Clock drawing command
Language	Animals Actions Boston Naming test (BNT)
Memory	Rey Complex Figure Test (RCFT) delay recall Hopkins Verbal Learning Test (HVLT) Retention HVLT Recognition Logical memory delayed recall
Visuo-spatial	RCFT copy Hooper Visual Organizational Test

5.5.4 MRI Acquisition

Each participant had an MRI scan within two weeks of the neuropsychological assessment using the GE Discovery 750 3T MRI at the Seaman Family Imaging Centre at the University of Calgary, Calgary, Alberta, Canada. A high-resolution T1-weighted 3D inversion recovery prepared fast spoiled gradient recalled (IR-FSPGR) sequence was acquired for each participant (repetition time = 7.176 ms, echo time = 2.252 ms, flip angle = 10°, acquisition matrix = 256×256, voxel size = 1×1×1 mm^3, 172 slices).

5.5.5 Cortical Thickness and Subcortical Volume

Freesurfer (http://surfer.nmr.mgh.harvard.edu/; version 6.0.0) was used to perform cortical thickness and subcortical volumetric analyses. The analysis was performed following the procedure detailed in prior publications [115, 116, 241]. Cortical segmentation was performed

automatically and upon visual inspection, appropriate manual adjustments were made. The manual editing was carefully performed in accordance with the Freesurfer manual in several steps (https://surfer.nmr.mgh.harvard.edu/fswiki/Edits). The errors resulting from imperfect intensity normalization were corrected by inserting control points and where appropriate, extraneous tissue were removed from the brain volumes (wm.mgz and brainmask.mgz). A total of 29 segmentation results were manually edited. Cortical thickness was computed for 68 Regions-of-Interest (ROIs) using the Deskian-Killiany brain atlas in Freesurfer [121]. Furthermore, the mean cortical thickness was calculated for each hemisphere.

Subcortical volumetric measures were computed from eight regions per hemisphere including caudate nucleus, putamen, pallidum, nucleus accumbens, hippocampus, amygdala, thalamus, and ventral diencephalon, plus the brain stem. All subcortical volumetric measures were corrected for the intracranial volume.

5.5.6 Machine Learning

The input features used for machine learning were sex, age, EHI score, years of education, years of disease duration, LEDD, UPDRS-III, rs6265, rs7689942, rs894280, rs7412, rs393152, rs429358, rs4680, and *SLC6A3* VNTR. Furthermore, the following imaging measures were included: cortical thickness measures (68 ROIs), subcortical volumetric measures adjusted for the intracranial volume (17 ROIs), and 2 measures of global thickness. In total, 102 measures were available per participant for the machine learning analysis to predict the outcome variable, the global cognition Z-score.

The machine learning analysis consisted of two main steps: (1) feature ranking and selection, and (2) regression analysis. The feature ranking was employed to rank the 102 features (8 genetic, 87 neuroimaging, 3 clinical, and 4 demographic) based on their contribution to the outcome variable and to select the most efficient combination of features that can predict the outcome variable of the regression problem. Reducing the number of features is essential for improving model performance by eliminating features that are redundant and non-informative. In this study, the RRELIEFF feature selection algorithm was used for this purpose [242]. In the next step, the machine learning regression modelling was performed based on the ranked features using a Support Vector Regression (SVR) model with a polynomial kernel. The SVR is, in principle, very similar to the support vector machine classification models with slight differences for the adaption to a regression style problem [123]. More precisely, an SVR model is built based on only a subset of training data within the predefined margins that minimize the generalization error. Therefore, the data is first transformed into a higher dimensional space employing the polynomial kernel, thereby allowing linear models to fit the training data. The SVR model was used in this study for regression modeling as it is less likely to overfit the data compared to other models, i.e. SVR is a model with adequate generalization capabilities and good prediction accuracy.

The least informative feature was iteratively removed from the set of ranked features until only two features were left for model training to identify the optimal subset of features. The model performance was evaluated for each iteration using the root mean squared error comparing the predicted with true observations. Finally, the model with the optimal feature subset was further evaluated using additional metrics including the coefficient of determination (R^2) and the correlation value. The coefficient of determination quantifies the amount of variance in the outcome variable that is explained by the selected features in the model. A nested leave-one-out

method of the cross-validation was employed through the feature selection and regression in which the number of model validation was set to N where N is equal to the number of participants in the sample. At each validation test, one participant is used to test the model while N-1 participants were used to train the model. All metrics reported for machine learning were attained by averaging the metrics of these N models. This method was used to overcome the small sample size and to prevent double-dipping.

5.5.7 Statistical Analysis

Statistical analyses of continuous variables were performed using either a student-t test or Mann-Whitney U test based on the data normality. The Fisher exact test was used for categorical variables.

The post-hoc statistical analysis was designed in compliance with the machine learning results. Pearson correlation test was used to select independent factors correlated with the target feature. ANCOVA was used to explore the allelic group differences in the rs894280 variant of the *SCNA* gene with regards to global cognition. Demographic and clinical factors that were significantly correlated with a cognitive domain score of (attention, language, etc.) were included as covariates in the ANCOVA to control for them. A value of $p < 0.05$ was considered significant for the single tests, and Bonferroni correction was used to correct for multiple testing. The chi-square test was used to explore association of rs894280 with other genetic variants available in this cohort. All statistical tests were performed using IBM Corp. Released 2019. IBM SPSS Statistics for Mcintosh, Version 26.0. Armonk, NY: IBM Corp.

5.6 Acknowledgments

This work was funded by a project grant from the Canadian Institutes of Health Research (CIHR) (PJT-166123), the Tourmaline Oil Chair in Parkinson's Disease, the Canada Research Chair in non-motor symptoms of Parkinson's disease to OM. MR received a scholarship from the Parkinson Association of Alberta. NDF is supported by the Canada Research Chair program and the River Fund at Calgary Foundation. This work was financially supported by Parkinson Canada, the Michael J. Fox Foundation, the Canadian Consortium on Neurodegeneration in Aging (CCNA) and the Canada First Research Excellence Fund (CFREF), awarded to McGill University for the Healthy Brains for Healthy Lives (HBHL) program. ZGO is supported by the Fonds de recherche du Québec-Santé (FRQS) Chercheurs-boursiers award and is a Parkinson's Disease Canada New Investigator awardee. Infrastructure for work performed in the GP lab was supported by a grant from the Canada Foundation for Innovation John R Evans Leaders Fund.

5.7 Author Contributions

MR has prepared the data, performed the statistical analysis, interpreted the results, wrote and edited the manuscript. PM and DR performed the machine learning analysis and helped with Methods section.PM also helped with manuscript edition. EJY prepared the structural neuroimaging data and edited the manuscript. JAR, EL and KM prepared the genetic data. MK administered, scored and prepared the neuropsychological data. TH calculated and collected the clinical data. IK and NM recruited the participants and collected clinical data. JS and DM were involved in data collection and manuscript edition. GP and ZGO were involved in genetic data collection. NDF supervised the machine learning analysis and manuscript writing and editing. OM was involved in study design, supervision of the project and manuscript writing and edition.

5.8 Additional Information

Ethics Declaration

5.9 Competing Interests

The authors declare no competing interests.

The data is available upon request. Please contact the corresponding author at oury.monchi@ucalgary.ca

6 Chapter Six

In this chapter, we aimed to test the hypothesis whether rs894280 in SNCA was associated with longitudinal cognitive changes in PD patients. Also, we hypothesized that this variant was associated with NPS in PD patients cross-sectionally and longitudinally.

This chapter is reproduced from a manuscript that has been submitted to Parkinson's Disease Journal under the same title.

SNCA variant shows association with cognitive decline in Parkinson's disease patients in the PPMI dataset longitudinally

Mehrafarin Ramezani [a,b], Matthew M. Joel [b,c], Gerald Peffer [a,b,d], Oury Monchi [*a,b,e]

Author Affiliations

a-Department of Clinical Neurosciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

b-Hotchkiss Brain Institute (HBI), Cummings School of Medicine, University of Calgary, Calgary, Alberta, Canada

c-Department of Neurosciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

d-Alberta Children's Hospital Research Institute, Department of Medical Genetics, Cumming

e-Radiology, University of Calgary, Calgary, Alberta, Canada

*Corresponding Author oury.monchi@ucalgary.ca , Address: HSC-2915, 3330 Hospital DR, NW, Calgary, AB, Canada T2N 4N1 telephone: +1- 403.220.8048

Running Title: SNCA variant and Cognition in PD

6.1 Abstract

BACKGROUND: Neuropsychiatric symptoms (NPS) and cognitive impairments are common non-motor symptoms of Parkinson's disease (PD). The *SNCA* is implicated in familial PD and dementia in idiopathic PD (iPD). A single nucleotide polymorphism (SNP) in 5'untranslated region (UTR) of *SNCA* (rs894280) demonstrated an association with cognition in iPD.

OBJECTIVE: We hypothesized that rs894280 is associated with cognitive impairments and NPS in iPD patients both cross-sectionally and longitudinally.

METHODS: To test this hypothesis, 285 participants with available data for genetic, demographic, NPS and cognitive measures from the Parkinson's Progression Marker Initiative (PPMI) were selected. The participants had data at baseline and 5-year follow-up.

RESULTS: The rs894280 was associated with cognitive status at baseline and the change of cognitive status of iPD patients in a five-year follow-up when controlling for age, sex, education, ethnicity, unified UPDRS-III and LEDD. CT-carriers were twice more likely to develop MCI after 5 years compared to non-carries. No association was found between rs894280 and NPS.

CONCLUSIONS: These findings suggest a role for rs894280 or *SNCA* 5' UTR in the cognitive changes of iPD patients longitudinally.

Keywords: Parkinson's disease, cognitive impairments, SNCA, Neuropsychiatric symptoms, PPMI

6.2 Introduction

Parkinson's disease (PD) is a common neurological disorder in senior adults. PD patients suffer from a broad range of non-motor symptoms e.g., autonomic dysfunction, cognitive and/or neuropsychiatric symptoms (NPS). These non-motor symptoms have a negative impact on the quality of life of both PD patients and their family/caregiver causing social, emotional and economic issues [12, 14]. It needs to be noted that PD patients with advanced cognitive impairments have increased susceptibility to dementia and even mortality [11, 14]. The non-motor symptoms sometimes precede the onset of motor symptoms of PD and their early detection provides clinicians with valuable time to slow down or prevent the onset of more advanced symptoms [6].

PD patients demonstrate a high prevalence of cognitive impairments in different cognitive domains (attention, executive-function, language, memory, and visuospatial function) and in their overall cognitive profile compared to healthy controls [6, 113, 119, 232]. Moreover, PD patients are highly susceptible to suffer from a range of NPS e.g., depression, anxiety, and apathy etc. [6, 110, 120, 127].

The Synuclein Alpha gene (*SNCA*) was the first gene discovered in the familial forms of PD [86, 87]. Individuals with certain *SNCA* mutations or who carry multiple copies of this gene develop a familial form of PD, accompanied with early age at onset compared to non-carriers [29, 86, 87]. It should be noted that PD patients with familial forms of the disease caused by *SNCA* usually suffer from moderate to severe cognitive impairments [243].

Other evidence for the involvement of *SNCA* in PD was the discovery of alpha-synuclein protein as the building block of Lewy bodies, the main pathologic inclusion body in PD [4]. Additionally, several *SNCA* variants were reported in association with dementia in idiopathic PD (iPD) [88]. Recently we reported an association between rs894280 in *SNCA*, and impairment in overall cognitive performance, attention, and visuo-spatial domains in iPD patients [244]. These findings highlight the importance of this gene in relation to cognitive changes in iPD patients [244].

In this study we tested two hypotheses using data from the PPMI dataset [245]. We hypothesized that rs894280 is associated with cognitive decline and NPS in PD, both cross-sectionally and longitudinally.

6.3 Methods

6.3.1 Participants Data

The data used in this study were obtained from the PPMI dataset at www.ppmi-info.org. PPMI is a multicenter comprehensive and longitudinal research study that aims to identify early markers of PD. This study has several major and minor funding agencies including Michael J. Fox foundation, Abbvie, Allergan, Amathus Therapeutics, Avid Radiopharmaceudicals, Bial Biotech, Biogen, BioLengend, Bristol-Myers Squibb, Calico, Celgene, Denali Therapeutics,4D pharma plc, GE Healthcare, Genentech, GlaxoSmithKline, Golub Capital, Handl Therapeutics, Insitro, Janssen Neuroscience, Lilly, Lundbeck, Merck, Meso Scale Discovery, Neurocrine Biosciences, Pfizer, Piramal Imaging, Prevail Therapeutics, Roche, Sanofi Genzyme, Servier, Takeda, Teva, UCB, Verily, and Voyager [102, 245, 246].

The data for the study were downloaded from the PPMI website using curated available data (last updated on the April 2020). All the participants identifiers were removed in the PPMI dataset. A subset of data including 285 PD patients was selected for the analysis based on the availability of the data in two time points, at baseline and five-year follow-up (T_5). The participants were included for the analysis based on the following criteria. They should not have any other neurological disorders other than PD. They need to have an available neuropsychological evaluation with a determined Mild Cognitive Impairment Status (MCI-status). Finally, the selected participants should have available genetic data to be included in the analysis.

The demographics of the participants (age, sex, years of education, family history of PD and ethnicity) and their clinical measures (Unified Parkinson's disease rating scale-part III (UPDRS-III) score, Hoehn & Yahr stage, levodopa daily dosage equivalency (LEDD)) were collected for the analysis.

Each PPMI recruiting center received written consent form from all the participants in accordance with the declaration of Helsinki. Each center also received ethics approval from their local ethics board. For further details on participants' recruitment, please visit the PPMI website at www.ppmi-info.org/data, additional information can be found in the reference article [245]. For up-to-date information about the PPMI study please visit the PPMI website at www.ppmi-info.org. This study was approved by the Conjoint Health Research Ethics Board (REB14-2463) at the University of Calgary.

6.3.2 Neuropsychological and Neuropsychiatric Assessment

All participants recruited in the PPMI study underwent a thorough neuropsychiatric and neuropsychological assessment at each time point. The neuropsychological and neuropsychiatric tests/questionnaires used for the screening are shown in Table 1. The participant's performance for each test was scored according to the instructions. The MCI status was determined from the neuropsychological assessment. The participants who scored at least 1.5 standard deviations below the standardized mean in at least two cognitive tests were classified as MCI. For further details please refer to PPMI data manual available on the PPMI website.

The neuropsychiatric assessment was done at the same time point as the neuropsychological evaluation. The participants were screened for depression, impulsive/compulsive disorders, and anxiety (Table 7-1). All the neuropsychiatric questionnaires were scored according to the instructions. Depression was screened using the Geriatric Depression Scale (GDS) in which a score of > 5 were indicated as depressed. Impulsive/compulsive impairments were assessed using Questionnaire for Impulsive/Compulsive disorders (QUIP-Short). This questionnaire collected data on the following behaviors: buying, eating, gambling, hobbies, punding, sex, and walking/driving. If one or more behaviors received a positive answer, the PD patient was considered to have an impulsive/compulsive disorder. Anxiety was assessed using the Trait score of State-Trait Anxiety Inventory questionnaire. The Trait section of this questionnaire contains questions on the general feeling of anxiety [247]. All the tests were scored by the PPMI team according to the guidelines.

Table 7-7. The list of neuropsychological and neuropsychiatric tests/measures extracted from the PPMI dataset for use in the analysis.

Type of assessment	Name of the test/measure
Neuropsychological	Semantic Fluency (Animal, Vegetables, Fruits)
	Letter Number Sequencing
	Hopkins Verbal Learning Test (HVLT)
	Benton Judgment of Line Orientation (JoLO)
	Symbol Digit Modalities
	Montreal Cognitive Assessment (MoCA)
Neuropsychiatric	Geriatric Depression Scale (GDS)
	State-Trait Anxiety Inventory
	Questionnaire for Impulsive-Compulsive disorders (QUIP-short)

6.3.3 Genetic data

The genetic analysis was performed using DNA samples extracted from the whole blood for each participant. The DNA samples were used to obtain genetic data of more than 93% of each participant's genome using >340 $x10^3$ probes for exome sequencing. The genetic data were collected using the Ilumina Nextera Rapid Capture Expanded Exome Kit (Ilumina, Inc., San Diego, California, USA). Samples were divided into groups of 12 and sequenced on an Ilumina HiSeq 2500 sequencing platform (Ilumina, Inc., San Diego, California, USA). The sequencing was performed using 2x100 bp paired-end read cycles. The pair-end reads were aligned to UCSC hg19 as the human genome reference.

The Genome Analysis Tool Kit (GATK) was used to check the quality for reads per each participant's Binary Alignment/Map (BAM) file. GATK haplotype caller was utilized to generate variant calling and genotype frequencies for each participant. The Cohort's combined genotyping

was generated from each participant's Variant Call Format (VCF) file using the GATK tools: Combine GVCF and Genotype GVCFs.

The quality control measures were performed using variant call rate, heterozygosity rate, etc. using the PLINK toolset. This information was obtained from the PPMI website where further details can be found [102].

The *SNCA* SNP, rs894280 was accessed from the whole exome VCF file corresponding to chromosome 4 from the genetic data repository of PPMI [102]. The VCF was filtered with the bcftools software (version 1.9, http://samtools.github.io/bcftools/bcftools.html) using the view command for the target SNP (rs894280), by providing the appropriate genetic position at chr4:89839732.

The command used for data filter using VCF files is: "bcftools view -r -.vcf.gz"

6.3.4 Missing Data

Some of the participants had missing values for Hoehn & Yahr stage and UPDRS-III score. To resolve this issue the Last Observation Carried Forward method was used [248]. The missing values of Hoehn & Yahr scores, were replaced by the scores chronologically closest to the missing value. The Hoehn & Yahr scores were measured in PPMI in both off and on modes. In total 15 participants had missing values for T_5 Hoehn & Yahr scores. The same procedure was performed for missing UPDRS-III scores for the on mode (n= 13, all from the T_5). Only one person had missing QUIP data for T_5, the missing value was replaced by the data from T_4.

6.3.5 Statistical Analysis

The demographic and clinical measures were statistically analyzed between the three possible rs894280 genotypes using ANOVA, or Kruskal-Wallis tests based on the data's normality. Categorical measures were analyzed with Fisher test.

Longitudinal analysis was performed with General Estimating Equations (GEE), using the binary logistic model. The MCI-status was the binary dependent variable (DV) in the logistic model and rs894280 genotypes were the main independent categorical variables (IV). The same procedure was performed for neuropsychiatric measures except for the STAI test which had total score available. To analyze the STAI-trait data, GEE used a linear regression model, using STAI score as the DV and rs894280 groups as the main IV.

The cross-sectional analyses for baseline and T_5 were done using binary logistic regression with MCI-status as the DV and the rs894280 genotypes as the IV. The same procedure was performed for neuropsychiatric measures for cross-sectional analysis except for the STAI-trait test, which was analyzed using a linear regression model, with STAI as the DV and rs894280 groups as the main IV.

6.4 Results

6.4.1 Participants

The demographic and the clinical characteristics of iPD participants for this study are summarized in Table 2. Based on the analysis, the participants had an average age of 60.7 years, and an average MoCA score of 27. The participants had an average disease duration of 6.6 years at baseline, and 94 % of them were at stage I or II of Hoehn and Yahr at the baseline visit (Table 7-2).

Table 7-8. The demographic and clinical characteristics of PD participants at baseline.

Demographic and Clinical characteristics	Mean, SD, (Min - Max)
Age at baseline	60.72, ± 9.75, (34 – 85)
Education (years)	15.62, ± 2.93, (5 – 26)
<13 years	16%
13-23 years	83%
>23 years	0.7%
Sex (Female%)	66 %
Disease Duration (years)	6.60, ± 6.63, (0.7 – 35.8)
UPDRS-III	19.87, ± 8.43 (4 – 47)
Hoehn and Yahr score (%)	
Stage I	50.2%
Stage II	49.1%
Stage III	0.7%
MoCA	27.21, ± 2.26, (17 – 30)
Ethnicity	
European Decent	93.3 %
African American	1.4 %
South East Asian	1.1 %
Others	4.2 %

Abbreviations: SD = standard deviation, Min = Minimum, Max = Maximum, UPDRS-III = Unified Parkinson's Disease Rating Scale part-III, MoCA = Montreal Cognitive Assessment

The demographic and clinical measures for each group are summarized in Table 7-3. The participants had the following genotype frequencies: CC=64, CT=147, TT= 74.

Table 7-9. The demographic and clinical measures for each rs894280 genotype group.

Demographic and Clinical characteristics	**CC N= 64** **Mean, SD (Min - Max)**	**CT N= 147** **Mean, SD (Min - Max)**	**TT N= 74** **Mean, SD (Min - Max)**
Age at baseline	62.3, ± 8.4 (45.2 – 75.2)	60.6, ± 9.9 (33.5 – 81.8)	59.4, ±10.3 (33.7 – 84.9)
Education (years)	15.8, ± 3.5 (5 – 3424)	15.7, ± 2.6 (5 – 22)	15.6, ± 3.1 (8 – 26)
<13 years	20.3 %	12.2%	20.3%
13-23 years	78.1%	87.8%	78.4%
>23 years	1.6%	NA	1.4%
Sex (Female%)	70 %	63.9 %	67.6%
Disease Duration (years)	6.8, ± 6.2 (1.0 – 31.9)	6.4, ±6.7 (0.9– 35.8)	6.9, ±6.8 (0.7 – 29.8)
UPDRS-III	19.6, ± 7.5 (6 – 41)	19.3, ± 7.7 (4 – 40)	21.2, ± 10.4 (6 – 47)
Hoehn & Yahr score (%)			
Stage I	51.6%	51.7%	45.9%
Stage II	48.4%	46.9%	54.1%
Stage III	NA	1.4%	NA
MoCA	27.3, ± 2.5 (20 – 30)	27.2, ± 2.2 (17 – 30)	27.3, ± 2.3 (21 – 30)
Ethnicity			
European Decent	95.3 %	95.2%	87.8%
African American	NA	1.4 %	2.7%
South East Asian	NA	0.7%	2.7%
Others	4.7%	2.7%	6.8%

Abbreviations: SD = standard deviations, Min = Minimum, Max = Maximum, UPDRS-III = Unified Parkinson's Disease Rating Scale part-III, MoCA = Montreal Cognitive Assessment

The allele frequencies were C=0.48 and T=0.52. No deviations from the Hardy-Weinberg were detected for these alleles.

6.4.2 Association of rs894280 with Cognitive Status in iPD Patients

The percentage of iPD patients with MCI diagnosis based on the neuropsychological assessment is summarized in Table 7-4.

Table 7-10. PD-MCI percentage per genetic group at each time point.

Time point	CC N = 64	CT N =147	TT N = 74
MCI % Baseline	23.4%	8.2 %	20.3 %
MCI % T_5	18.8 % (-4.6 %)	14.3 % (6.1 %)	23.0% (2.7%)

The CC group had higher percentage of MCI patients compared to the CT and TT groups at baseline. However, the CC carriers showed a negative MCI progression rate after five years of follow-up, with some of the CC-MCI patients reverting back to normal cognition after five years. The CT group had the lowest percentage of MCI patients at the baseline and demonstrated the highest MCI progression rate compared to the CC and TT groups after the five-year follow-up.

It was found that rs894280 was associated with MCI status of PD patients over the 5-year follow-up (Fig. 7-1). This association held when controlling for age, sex, years of education, ethnicity, UPDRS-III score, and LEDD (Wald Chi-square = 6.42 (df = 2), significance = 0.04). The φ effect size measure showed a medium effect size (φ = 0.4) for the longitudinal association of rs894280 and MCI status change of PD patients in the PPMI dataset.

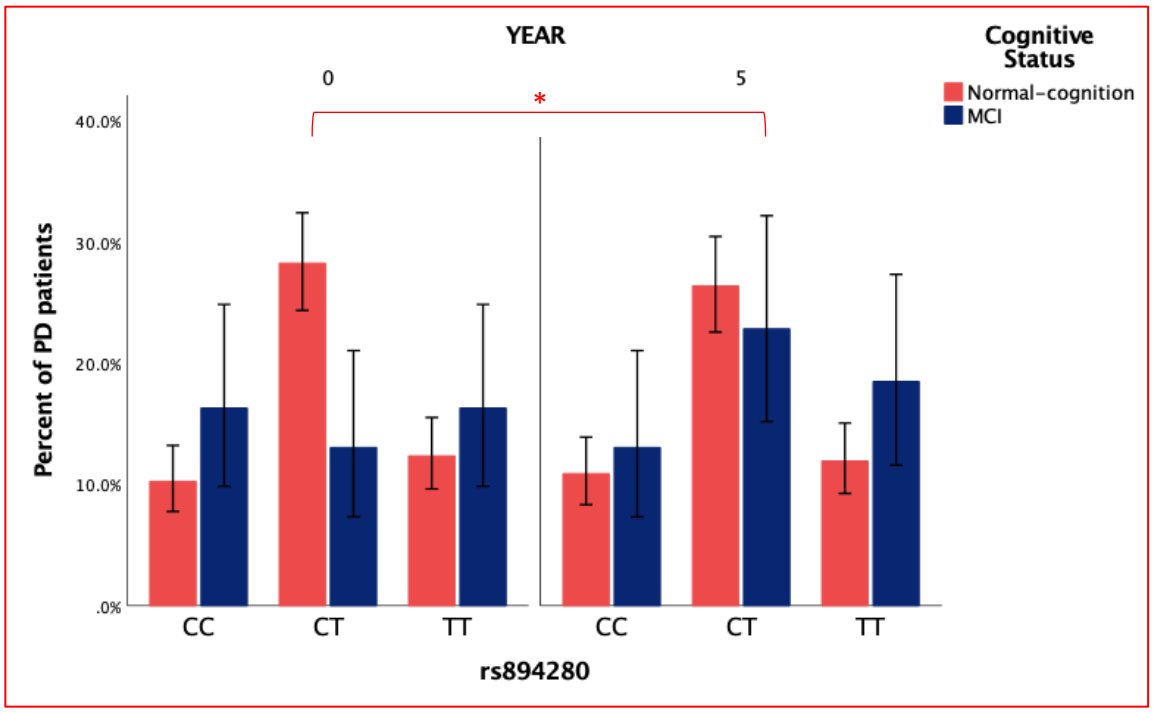

Figure 7-3. The change in cognitive status of PD patients based on their rs894280 genotype over five-year follow-up in PPMI dataset (N = 285).

The CC-carriers showed a significant increase for conversion to MCI longitudinally (Wald chi-square = 6.42 (df = 2), significance = 0.04) when controlling for the age, sex, years of education, ethnicity, UPDRS-III score and LEDD. The φ effect size measure showed a medium effect size of 0.4 for the longitudinal association of rs894280 and the MCI conversion of PD patients in PPMI dataset. The error bars represent 95% confidence interval (CI).

Post-hoc analysis did not find any significant differences between the three groups when controlling for age, sex, years of education, ethnicity, UPDRS-III score, and LEDD. However, the PD patients with CT genotype demonstrated a trend for higher likelihood of becoming MCI than the CC and TT groups. The significance did not survive the Bonferroni correction for either of the pairwise comparisons (p = 0.03, p = 0.04, respectively). The PD patients in the CT group were almost twice as likely to develop MCI compared to the other two groups after the five-year follow-up. The PD patients with CT genotype had greater likelihood of becoming MCI compared to the CC and TT groups with odds ratio of 2.04 and 1.94, respectively (Table 7-5). No differences were identified between the CC and TT groups longitudinally (Table 7-5).

Table 7-11. Longitudinal association of rs894280 with cognitive status in PPMI dataset over 5 years. The following covariates were controlled in the GEE analysis; age, sex, years of education, ethnicity, UPDRS-III score and LEDD.

Group	Beta	Wald χ^2	Wald χ^2 - 95% CI	sig.**	Exp(B)	Exp (B) - 95% CI
intercept	0.83	0.39	-1.8 – 3.42	0.53	-	-
CC>TT	-0.09	0.06	-0.82 – 0.64	0.81	1.09	0.53 – 2.27
CT>CC	0.75	4.81	0.08 – 1.42	0.03	2.05	1.08 – 4.31
CT>TT	0.66	4.06	0.02 – 1.30	0.04	1.94	1.02 – 3.68

Abbreviations: sig = significance, Exp (B) = Exponential Beta, CI = confidence interval

Association of rs894280 with MCI-status in PPMI dataset in general; Wald chi-square test= 6.42 (df = 2), significance= 0.04*

** significance level was set to 0.02 for pair-wise comparisons, Bonferroni correction

The cross-sectional association analysis demonstrated that rs894280 was linked to the MCI- status of PD patients at baseline (Wald chi-square test = 7.58, (df = 2) significance = 0.02), when controlling for age, sex, years of education, ethnicity, UPDRS-III score and LEDD (Table 7-6). The effect size for this association demonstrated the same medium effect size with (φ = 0.4) as the longitudinal analysis.

Table 7-12. Association of rs894280 with cognitive status of PD patient in PPMI dataset at baseline. The following covariates were controlled in the binary logistic regression analysis; age, sex, years of education, ethnicity, UPDRS-III score and LEDD

Group	Beta	Wald chi-square	sig.	Exp(B)	Exp (B) - 95% CI
intercept	-0.61	0.12	0.73	-	-
CC>TT	0.37	0.69	0.41	0.69	0.29 – 1.66
CT>CC	-1.18	7.22	0.007**	0.31	0.13 – 0.73
CT>TT	-0.81	3.34	0.07	0.45	0.19 – 1.06

Abbreviations: sig = significance, Exp (B) = Exponential Beta, CI = confidence interval

Association of rs894280 with MCI-status in PPMI dataset at baseline in general; Wald chi-square test (2)7.58 significance=0.02*, Nagelkerke R^2 = 0.17, Hosmer and Lemeshow test goodness of fit-test p = 0.43 (df = 8), correct cases overall percentage = 86.0%

** significance level was set to 0.02 for pair-wise comparisons, Bonferroni correction

PD patients with the CT genotype showed a lower likelihood of being MCI when compared to the CC group (Wald chi-square = 7.22, df = 2, significance = 0.007) at the baseline. The PD patients with CT genotype did not show any statistical differences when they were compared to the TT group (Wald chi-square = 3.33, df = 2, significance = 0.07) (Fig. 7-2).

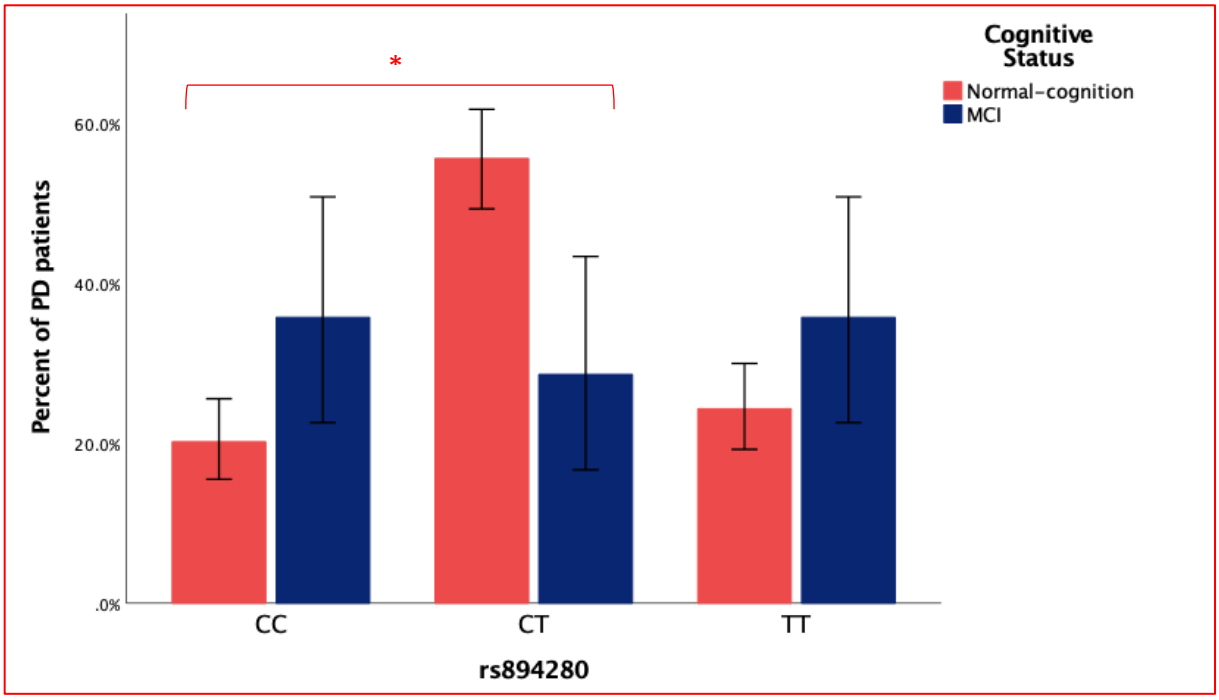

Figure 7-4. The cognitive status of PD patients based on their rs894280 genotype at baseline.

This variant demonstrated an association with the cognitive status of PD patients at baseline (Wald chi-square test =7.58, (df = 2), significance = 0.02). The PD patients with CT genotype showed lower likelihood of being MCI when compared to the CC group (Wald chi-square = 7.22, (df=2), p = 0.007). The effect size for this association demonstrated a medium effect size with (φ= 0.4). The error bars represent 95% confidence interval (CI).

The rs894280 was not linked to MCI- status of PD patients at T_5, (Wald chi-square = 7.26 (df = 2), significance = 0.32; Table 7-7).

Table 7-13. Association of rs894280 with cognitive status of PD patient in PPMI dataset at T5 (five years follow-up). age, sex, years of education, ethnicity, UPDRS-III score and LEDD.

Group	Beta	Wald chi-square	sig.*	Exp(B)	Exp (B) - 95% CI
intercept	-2.58	2.31	0.13	-	-
CC>TT	-0.12	0.07	0.79	1.13	0.37 – 2.17
CT>CC	-0.43	1.06	0.30	0.65	0.28 – 1.48
CT>TT	-0.55	2.00	0.16	0.58	0.27 – 1.24

Abbreviations: sig= significance, Exp (B)= Exponential Beta, CI= confidence interval

Association of rs894280 with MCI-status in PPMI dataset at T_5 in general; Wald chi-square test (2)2.28 significance=0.32, Nagelkerke R^2 = 0.16, Hosmer and Lemeshow test goodness of fit-test p=0.10 (df=8), correct cases overall percentage = 83.2%

* significance level was set to 0.02 for pair-wise comparisons, Bonferroni correction

6.4.3 Association of rs894280 with NPS in iPD Patients

The longitudinal analysis of rs894280 with NPS (depression, impulsive/compulsive, and anxiety) demonstrated no relationship between them over the five-year follow-up (Table 7-8). No association was found between this variant and the risk of developing depression or anxiety over five years (Wald chi-square test = 0.82, (df = 2), significance = 0.66, Wald chi-square test = 0.89 (df = 2), significance = 0.64, respectively).

No association was found between this variant and the risk of developing impulsive/compulsive disorders in this cohort after five-year follow-up (Wald chi-square test = 0.23, (df = 2), significance = 0.89).

Table 7-14. Results of association analysis of rs894280 and GDS, QUIP and STAI tests over 5 years follow-up. The following covariates were included in the analysis age, sex, years of education, ethnicity, UPDRS-III score and LEDD.

Test	Group	Beta	Wald chi-square	sig.*	Exp(B)	Exp (B) - 95% CI
GDS[a]	intercept	2.02	2.33	0.13	-	-
	CC>TT	- 0.26	0.40	0.53	0.77	0.35 – 1.72
	CT>CC	-0.04	0.02	0.89	0.96	0.50 – 1.83
	CT>TT	-0.30	0.81	0.37	0.74	0.38 – 1.43
QUIP[a]	intercept	-0.24	0.07	0.80	-	-
	CC>TT	- 0.13	0.19	0.66	0.88	0.49 – 1.58
	CT>CC	0.02	0.008	0.93	0.98	0.61 – 1.72
	CT>TT	-0.11	0.16	0.69	0.90	0.53 – 1.52
STAI-trait[b]	intercept	39.23	60.25	<0.001	-	-
	CC>TT	-1.12	0.53	0.47	NA	NA
	CT>CC	1.02	0.70	0.40	NA	NA
	CT>TT	-0.09	0.005	0.94	NA	NA

Abbreviations: sig = significance, Exp (B) = Exponential Beta, CI = confidence interval, NA = Not Applicable, GDS = Geriatric Depression Scale, QUIP = Questionnaire of Impulsive/Compulsive disorders, STAI = State-Trait Anxiety Inventory

Association of rs894280 with GDS in PPMI dataset longitudinally; Wald chi-square test=0.82, (df = 2) significance = 0.66

Association of rs894280 with QUIP in PPMI dataset longitudinally; Wald chi-square test = 0.22 (df = 2) significance = 0.89

Association of rs894280 with STAI-trait in PPMI dataset longitudinally; Wald chi-square test = 0.79 (df = 2) significance = 0.67

a = binary logistic model was used for GEE analysis

b = linear regression model was used for GEE analysis

* significance level was set to 0.02, Bonferroni correction

No association was observed for rs894280 and any of the neuropsychiatric measures at baseline or T_5 likewise. These results are listed in Table S7-1.

This variant had no association with depression measured by GDS (Wald chi-square test = 0.17 (df = 2), significance = 0.98) or the impulsive/compulsive impairments using QUIP at the baseline (Wald chi-square test (df = 2) = 0.70, significance = 0.70). No association was observed for anxiety using STAI-trait score at the baseline (F (8, 284) = 0.13, significance = 0.88).

No relationship was observed between rs894280 and any of the neuropsychiatric measures after five years, including depression using GDS (Wald chi-square test= 0.85, (df = 2) significance = 0.65), the impulsive/compulsive impairments using QUIP at T_5 (Wald chi-square test = 0.27, (df = 2) significance = 0.86), or anxiety using STAI-trait at T5 (F (8, 284) = 1,03, significance = 0.36).

6.4.4 Discussion

In this study, we hypothesized that rs894280 in *SNCA* was associated with cognition of iPD patients cross-sectionally and longitudinally, using PPMI data. Based on the association analysis, rs894280 was found to be associated with the rate of MCI conversion in iPD patients who were followed up for 5 years. The cross-sectional analysis demonstrated an association between the rs894280 and MCI status of iPD patients at the baseline but not at the T_5. The observed effect size for this variant was (φ= 0.4), which indicates a medium effect size. We hypothesized that rs894280 should be associated with NPS in iPD patients of PPMI study. However, no association was found for depression, anxiety, or impulsive/compulsive disorders in these iPD patients cross-sectionally or over the five-year follow-up. These results indicate a potential role for this *SNCA* SNP in the pathogenesis of cognitive impairment in iPD patients and possibly dementia. Based on the current results there is no evidence for the involvement of this variant with neuropsychiatric changes.

The role of *SNCA* in familial cases of PD is well-established [29, 86] and in recent years several variants in *SNCA* have been reported in association with cognitive impairment in iPD patients [88, 94, 174, 236]. The extensive LB pathology in PD patients and recent findings on the relationship of *SNCA* variants with cognitive impairments in iPD signify the importance of this locus to cognitive changes in iPD patients. This SNP is located on the 5' untranslated region (UTR) of the *SNCA* and was first reported in association with dementia with LB (DLB) [235].

DLB is very similar in both symptoms and pathology to PD with dementia (PDD), and they are both synucleinopathies [93]. The shared LB pathology between these two disorders, together with other similar symptoms indicate a potential role for *SNCA* in a common pathologic process. A large-scale study including 1,492 PD patients and 922 DLB patients discovered a SNP in 5' UTR of *SNCA* in nearly full Linkage Disequilibrium (LD) with rs894280, which was linked to PDD. Based on their findings, the 5' region of the *SNCA* contains a risk haplotype for DLB [88].

In our analysis, rs894280 demonstrated an association with the rate of conversion to MCI status in iPD patients over a five-year follow-up. This finding suggests a role for this SNP in shaping the cognitive profile of iPD patients, but pair-wise comparisons of genetic groups did not yield any robust relationship between the rs894280 and cognitive status (table 5). Based on the current results PD patients with heterozygous genotype might have a higher likelihood of developing MCI compared to the homozygous groups. Interestingly, both homozygous groups displayed very similar odds for MCI conversion over five-year span (2.04 vs 1.94). This result can seem to contradict the results of the cross-sectional analysis at the baseline, which demonstrated that PD patients with the CT genotype were less likely to be MCI. Also, PD patients with the CT genotype showed a lower likelihood of being MCI than the CC group at the baseline. Also, the CT group had a trend for a lower likelihood to be MCI when they were compared to the TT group at the baseline.

According to the baseline results, the PD patients carrying the CC genotype demonstrated considerably higher likelihood of being MCI at baseline but the rate of conversion reverses at the T_5. On the contrary, both the CT and TT groups showed relatively lower numbers of MCI patients compared to the CC group at the baseline (table 4).

However, both of the CT and TT groups had positive MCI conversion rate in the five-year span. While preliminary, these results suggest that the CC group might have a lower likelihood of cognitive decline over time.

There are a number of possible reasons for the above-mentioned results. The reduced number of PD-MCI cases in the CC group might be attributed to dopaminergic medications. The effect of the dopaminergic medications on cognitive abilities in PD patients is still not clear [249, 250]. These medications might reverse some mild cognitive impairments derived from low dopamine availability in PD patients [251]. It has been proposed that low dopamine availability in the frontostriatal circuits can contribute to impairment of the cognitive abilities which are linked to this region such as executive-function [251]. In addition, longitudinal studies demonstrated that cognitive impairments derived from insufficiency of the dopamine in frontostriatal regions are reported to be not linked to the development of dementia in PD patients [35].

The results mentioned above support our speculation that the CC group might be more susceptible to the cognitive impairments which are attributed to the low dopamine availability. These impairments might not exacerbate into more severe impairments and more importantly dementia [35]. It can be speculated that the CT and TT groups might be more susceptible to cognitive impairments in temporal-parietal regions which are linked to more severe irreversible cognitive decline and even dementia [16].

These findings were similar to the well-known Catechol O-methyltransferase (*COMT*) p.Val158Met SNP (rs4680), which is proposed to have a dual effect on the cognitive profile of PD patients [34]. This SNP has a prominent effect in the degradation of the dopamine in frontal regions of the brain.

The PD patients with the Val allele display better working memory than the Met carriers at early stages of PD. But as PD progresses and the dopamine becomes scarcer, the Met carriers benefit from the lower activity of COMT enzyme [34].

No association was found between rs894280 and NPS in PD patients from the PPMI cohort at a given time point or longitudinally. These results were in contrast with our hypothesis that rs894280 should be linked to NPS in iPD patients. Familial PD caused by *SNCA* mutations or an increased *SNCA* gene copy number was reported to be associated with a range of NPS e.g., depression, hallucination etc. [30, 194, 252]. Also, studies suggest a link between NPS in PD and a higher likelihood of cognitive impairments [110, 111, 120]. However, no association was found between this variant and depression, anxiety or impulsive/compulsive disorders in this cohort.

In this study, the PPMI data was used to investigate the association of rs894280 with cognition in iPD participants. The PPMI study aimed to recruit drug-naïve PD patients in order to explore and discover early and novel biomarkers of PD [102, 245]. Although cognitive impairments can precede the PD diagnosis, the severity of cognitive impairments usually increases with time. The disease duration of the PPMI participants included in the current study was about 6.6 years at baseline compared to 5.7 years in our previous cross-sectional study of this SNP [244]. The prior study included 101 iPD patients of PD-MCI cohort at University of Calgary, in which rs894280 demonstrated an association with the overall cognitive performance of the patients [244]. It was also observed that rs894280 was related to attention, and visuo-spatial abilities in iPD patients [244]. This association demonstrated an advantage for CC carriers while in the current study CC carriers have higher likelihood of MCI at baseline compared to non-carriers [244].

The discrepancies between these two studies can be explained considering that in the previous analysis, PD patients already received dopaminergic medications before enrolling in the study and they were responsive to it. The *SNCA* gene was reported in connection with levodopa response in PD patients and the alpha-synuclein protein can interact with dopamine transporter [253, 254]. Another important difference between the two studies is the distribution male and female participants. The percentage of females in the previous study was about 30% while in the PPMI cohort 70% of the participants were female. This can introduce large differences in the average cognitive performance of the patients. Recent evidence suggests a difference between male and female patients regarding the susceptibility to cognitive impairment [7, 255, 256]. It has been reported that the male PD patients have higher likelihood of executive-function and processing speed impairments compared to the female PD patients [210]. Considering the substantial sex difference between the two cohorts, the average cognitive performance can be considerably different.

The last difference to be mentioned is the average age of the participants between the two studies. In the previous study, participants had an average age of ~ 70 years old with disease duration of 6.6 years. In the current cohort the average age is about 9 years less than PD-MCI cohort while the disease duration is still comparable to the PD-MCI cohort (6.6 versus 5.7). This difference in demographics might be due to recruitment of PD patients below the age of 50 years old unlike the PD-MCI cohort [244].

Several limitations need to be mentioned. PD patients in PPMI are drug-naïve at baseline (recruitment time) and their disease duration are relatively short at the time of recruitment. Severe cognitive impairments are infrequent in early PD but as the disease progresses the likelihood of more severe and irreversible cognitive impairments escalate.

The findings of the current study require follow-up investigations in later time points to reproduce, validate, and further study the longitudinal association of this variant with cognition. Another limitation was the lack of availability of standardized scores for cognitive and neuropsychiatric measures. A common scale for all the administered tests/ measures can assist direct comparison, and a more thorough analysis of cognitive domains and it can improve statistical power of the analysis.

In conclusion, we used publicly available data from the PPMI study to investigate whether rs894280 in *SNCA* was associated with cognitive and neuropsychiatric changes in PD patients at a given time point and over time. We found this variant was linked to the cognitive impairments of iPD patients over time. The cognitive status of iPD patients was associated with this SNP at baseline but not after the five year follow-up. PD patients with the CT genotype had a higher likelihood to become MCI after five years compared to the non-carriers. No association was found between this variant and any NPS longitudinally or cross-sectionally. These findings suggest the involvement of rs894280 or other variants in the close proximity of it with cognitive impairments in iPD patients.

6.5 Data Availability

All the data is used in this study belong to third party. Interested researchers can apply to PPMI database to access the data at the following link: https://www.ppmi-info.org/access-data-specimens/download-data/

6.6 Acknowledgement

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including AbbVie, Allergran, Amathus Therapeutics, Avid Radiopharmaceuticals, Biogen, BioLegend, Bristol, Myers Squibb, Celgene, Denali Therapeutics, GE Healthcare, Genentech, GlaxoSmithKline, Golub Capital, Handl Therapeutics, Insitro, Janssen Neuroscience, Eli Lilly and Company, Lundbeck, Merck, Meso Scale Discovery, Pfizer, Piramal, Prevail Therapeutics, Roche, Sanofi Genzyme, Servier, Takeda, Teva Pharmaceuticals, UCB, Verily and Voyager Therapeutics.

This work was funded by a project grant from the Canadian Institutes of Health Research (CIHR) (PJT-166123), the Tourmaline Oil Chair in Parkinson's Disease, the Canada Research Chair in non-motor symptoms of Parkinson's disease to OM. MR received a scholarship from the Parkinson Association of Alberta (PAA). MJ received a scholarship from Alberta Graduate Excellence Scholarship "Katharine Sarah Melinda Mei-Ling Thomas Rare Diseases/Biomedical Engineering Research Scholarship". GP received Canada Foundation for Innovation John R Evans Leaders Fund.

The Authors would like to thank Miss Meng Wang for her statistical advice in this project.

6.7 Conflict of Interest

The authors declare no competing interests.

Table S7-1. Association analysis of neuropsychiatric measures with rs894280 genotype at baseline and T5.

Time point	Test	Group	Beta	Wald chi-square	sig.*	Exp(B)	Exp (B) - 95% CI
baseline	*GDS*[a]	intercept	0.03	0	1.00	-	-
		CC>TT	0.08	0.02	0.88	1.08	0.38 – 3.06
		CT>CC	0.10	0.05	0.83	1.10	0.45 – 2.70
		CT>TT	0.18	0.16	0.69	1.19	0.51 – 2.79
	QUIP[a]	intercept	-0.43	0.10	0.76	-	-
		CC>TT	0.35	0.69	0.41	1.41	0.62 – 3.24
		CT>CC	-0.20	0.31	0.58	0.71	0.31 – 1.61
		CT>TT	0.15	0.18	0.67	1.17	0.57 – 2.38
	STAI-trait[b]	CC>TT	-0.11	0.13	0.90	NA	NA
		CT>CC	0.31	0.24	0.81	NA	NA
		CT>TT	0.75	0.57	0.57	NA	NA
T_5	*GDS*[a]	intercept	-2.93	3.43	0.06	-	-
		CC>TT	0.35	0.59	0.44	1.42	0.58 – 3.49
		CT>CC	-0.02	0.003	0.96	0.98	0.46 – 2.09
		CT>TT	0.33	0.74	0.39	1.39	0.66 – 2.95
	QUIP[a]	intercept	1.40	1.00	0.32	-	-
		CC>TT	-0.08	0.04	0.85	0.93	0.42 – 2.07
		CT>CC	0.18	0.24	0.62	1.19	0.59 – 2.42
		CT>TT	0.10	0.09	0.76	1.11	0.57 – 2.15
	STAI-trait[b]	CC>TT	-1.38	1.50	0.14	NA	NA
		CT>CC	1.75	1.21	0.23	NA	NA
		CT>TT	-0.72	0.49	0.63	NA	NA

Abbreviations: sig = significance, Exp (B) = Exponential Beta, CI = confidence interval, GDS = Geriatric Depression Scale, QUIP = Questionnaire for Impulsive-Compulsive disorders, STAI = State-Trait Anxiety Inventory, NA = Not Applicable

Association of rs894280 with GDS in PPMI dataset at baseline; Wald chi-square test (df = 2) = 0.17, significance = 0.98, Nagelkerke R^2 = 0.08, Hosmer and Lemeshow test goodness of fit-test p = 1.00 (df=8), correct cases overall percentage = 86.3 %

Association of rs894280 with QUIP in PPMI dataset at baseline; Wald chi-square test (df = 2) = 0.70, significance = 0.70, Nagelkerke R^2 = 0.01, Hosmer and Lemeshow test goodness of fit-test p = 0.54 (df = 8), correct cases overall percentage = 78.2%

Association of rs894280 with STAI-trait in PPMI dataset at baseline; F (8, 284) = 0.13, significance = 0.88

Association of rs894280 with GDS in PPMI dataset at T5; Wald chi-square test (df = 2) = 0.85, significance = 0.65, Nagelkerke R^2 = 0.12, Hosmer and Lemeshow test goodness of fit-test p = 0.61 (df = 8), correct cases overall percentage = 81.1%

Association of rs894280 with QUIP in PPMI dataset at T_5; Wald chi-square test (df = 2) =0.27, significance = 0.86, Nagelkerke R^2 = 0.06, Hosmer and Lemeshow test goodness of fit-test p = 0.24 (df = 8), correct cases overall percentage = 73.9 %

Association of rs894280 with STAI-trait in PPMI dataset at T5; F (8,284) = 1,03, significance = 0.36

a = binary logistic model

b = linear regression model

* significance level was set to 0.02, Bonferroni correction

7 Chapter Seven: Discussion

In this chapter, the main findings of this thesis are summarized and discussed, followed by limitations and possible future directions. The main findings are discussed in the order of the chapters that they were first presented in.

In the first part of the thesis, the association of several common variants with NPS was investigated. In chapter three, it was reported that a common variant in *BDNF* p.Val66Met (rs6265) was linked to NPS in PD patients using MBI-C as the screening tool. It was shown that a positive correlation between MBI-C and UPDRS-III scores was observed. Also, a negative correlation was observed between the MBI-C total and MoCA scores for PD patients in that study. Considering these two factors in the analysis, it was shown that p.Val66Met could be linked to a higher likelihood of adverse changes in the emotional regulation domain which includes depressive and anxious symptoms. Moreover, this variant demonstrated an association with abnormal thoughts/perception domain which includes psychotic symptoms in PD patients. These results were in line with the previous findings in a meta-analysis on the association of Met allele of this variant with worse cognition in PD [56].

Moreover, a relationship between the severity of motor symptoms evaluated by UPDRS-III scores and the likelihood of NPS was observed. Some studies suggested an association between the motor severity and NPS in PD patients [176, 257, 258]. A recent study reported an association between the severity of NPS in PD patients and the alteration of corticostriatal network connectivity. Also, a negative relationship between NPS and severity of motor symptoms was reported in that study [176].

Additionally, the correlation between MoCA score and MBI-C suggests a relationship between NPS and the cognitive deficits. This relationship between NPS and cognition in PD was reported by several studies [6, 120, 259].

The possibility of a link between cognitive decline and NPS in PD patients were suggested in different studies [6, 120, 131, 259]. The current results from chapter three, indicate the same relationship between NPS and a variant in PD patients which has been previously reported in a meta-analysis in an association with cognitive decline [56]. One of the advantages of using MBI-C in this study was that MBI-C includes the most common geriatric NPS in one questionnaire. This provided us with the opportunity to investigate most of the common NPS in PD patients in a unified manner. Additionally, the length of NPS to be included as positive in MBI-C is six months which can eliminate inclusion of those NPS which are transient in nature.

In chapter four, the relationship between two common variants in the genes which are involved in the dopamine modulation and NPS was investigated in 170 iPD patients using MBI-C as the ascertainment tool. rs4680 in the *COMT* gene was found to be not associated with either the risk of NPS or the emotional dysregulation domain of MBI-C. Several factors were controlled for in the analysis including age, years of education, LEDD, MoCA score, and the MAOB inhibitor. Similar to the project presented in chapter three, a negative correlation was observed between the MoCA and the MBI-C total scores. Moreover, the use of MAOB inhibitors in patients was checked and it was observed that there was an association between the use of MAOB inhibitors and higher MBI-C scores. This association might be because of the common practice of prescribing the MAOB inhibitors in order to address NPS in PD [260].

The other studied variant was rs28363170 in *SLC6A3*, the gene encoding the dopamine transporter protein. The 10-R allele for this variant was reported in connection with higher abundance of dopamine transporter protein *in vivo* [169]. The enhanced expression of SLC6A3 leads to a higher rate of dopamine reuptake from the synaptic cleft space. We showed that there was no association between this variant and the likelihood of NPS in iPD patients in our study. Similar to the *COMT* analysis several factors were controlled for including age, MAOB inhibitor use and MoCA score. Also, we did not find any associations between this variant and the emotional dysregulation domain score in iPD patients. These results might indicate a lack of involvement for these variants in the general pathology of non-motor symptoms in PD patients.

Previous findings demonstrated a robust association between severity of NPS with cognitive impairments in PD patients [120, 259]. Both rs4680 and rs28363170, were reported in connection with cognitive impairments in PD patients [73, 172]. However, it was suggested that those genes, which are directly involved in the regulation of the dopamine level in the brain might not be involved in the pathological process of dementia in PD patients [35]. The impairments in frontostriatal regions can be derived from a diminished level of dopamine and affect cognitive abilities which are strongly correlated to it. The frontostriatal region has a substantial role in executive-function and attention abilities. It was proposed that PD patients with cognitive impairments linked to frontostriatal activity are at lower risk of dementia compared to patients with impairments linked to more posterior regions [35]. However, a recent study reported that an altered corticostriatal connectivity was linked to the severity of NPS in PD patients [176]. This apparent discrepancy might be due to damage of none-dopaminergic circuits in this network which do not depend on the function of the abovementioned variants [261].

The overall results of the first part of this thesis indicate a possible link between the likelihood of cognitive decline and the risk of NPS in PD. rs6265 in *BDNF* was reported in association with cognitive impairment [59, 174]. The Val allele for this variant was linked to better cognition in the meta-analysis nonetheless, this relationship might be more pronounced in patients of European descent [56] like the majority of our study participants. On the other hand, investigation of the relationship of two common variants, rs4680 and rs28363170 with risk of NPS in PD patients did not yield any significant associations for either of the variants. Both of these variants are linked to frontostriatal dopamine availability. There were reports of their association with cognitive impairment in PD as mentioned above. Nevertheless, these variants did not demonstrate any relationships with cognitive changes longitudinally in the CamPaIGN study [35, 262].

The lack of studies investigating the relationship of rs6265 in *BDNF*, and cognition of PD patients longitudinally complicates drawing any conclusions of the potential lasting relationship between this variant and cognitive decline. Considering the evidence for a protective role of BDNF in the brain and in particular dopaminergic neurons, this variant might play a role as a precursor for neuronal damage, and even their loss [133]. This might generate a hotspot for further destruction of neurons by unknown factors e.g., biological or environmental. Furthermore, there has been evidence of involvement of this variant and LTD induction which can be a potential issue in PD patients [54, 55]. Taken all these findings together, this variant might be involved in creating a condition which can facilitate non-motor symptoms in PD patients.

Chapter five presented the results of a ML analysis in a sample of 101 PD patients for prediction of Z-score of overall cognitive abilities using RRELIEF feature selection and SVR. The ML analysis yielded 11 features including rs894280, with overall correlation coefficient of 0.54 and mean absolute error of 0.39 for the predictive model.

It was found that rs894280 in *SNCA* was linked to cognition in PD patients when controlling for UPDRS-III scores, years of education, R rostral-anterior cingulate thickness, and L middle temporal thickness. The post-hoc analysis revealed that this variant was linked to better abilities in attention, and visuo-spatial domains. Also, a trend was observed for this variant and the memory domain. In recent years several studies reported discovery of variants in *SNCA* which are linked to differences in neuronal activity and cognition of iPD patients [94, 95]. The rs894280 SNP was previously reported in connection with DLB and PDD [88, 235]. DLB and PDD share a substantial similarity in symptoms and PD-MCI can be considered as part of the continuum to PDD. These similarities together with the present results could indicate a possible role for this SNP in the underlying shared pathology of these two disorders. The possible link between the severity of motor symptoms which is often evaluated by the UPDRS-III score and cognitive impairment in PD was reported by previous studies [221, 222]. This association might indicate the extent of the neuronal damage in PD patients. Previous studies reported a negative correlation between the density of neurons in substantia nigra and UPDRS-III score [263, 264]. According to the study by Greffard et al. each point in UPDRS-III score can be equivalent to loss of 25 neurons/mm^3 [263]. The years of education was reported as a potential protective factor against cognitive impairment in PD [21, 22]. The anterior cingulate cortex and middle temporal regions are both reported in association with cognitive impairment and the middle temporal regions are suggested to be specifically linked to dementia in PD [35, 165, 225, 230, 231]. The more posterior-temporal regions are part of a model proposed by Williams-Gray et al. of the dementing process. According to this model the impairments in these regions are associated with severe cognitive impairments and dementia in PD patients [35].

Based on the current analysis, rs894280 had an association of medium effect size with cognition of PD patients. In particular, PD patients with the CC genotype demonstrated healthier cognitive abilities generally, and better performance in attention, visuo-spatial function and memory. On the other hand, PD patients with at least one T allele were at greater loss for the above-mentioned cognitive abilities.

In chapter six, an analysis was performed in PD patients using the data from the PPMI study to replicate and validate the connection of rs894280 with cognition of PD patients as presented in chapter five. Also, we investigated whether the association between this variant and cognition could be detected over time. The second aim was to explore the potential relationship between this variant and NPS.

It was found that rs894280 was linked to the cognitive status (MCI versus non-MCI) of PD patients at baseline. Moreover, it was observed that this polymorphism was associated with the cognitive status change in iPD patients longitudinally. No association was found for this polymorphism with available neuropsychiatric data on depression, anxiety and or impulsive/compulsive impairments.

These findings were in agreement with previous results presented in chapter five. We have reported an association between the CC genotype of this polymorphism with the general cognitive performance, attention, and visuo-spatial function. A trend was also found in favor of PD patients with CC genotype and memory abilities. Nevertheless, the replication of the precise association results on this polymorphism with attention and visuo-spatial function abilities similar to the findings of chapter five was not possible due to unavailability of the domain scores. However, the analysis detected a higher rate of MCI likelihood in PD patients with the CC genotype compared to patients with other genotypes at baseline. This result might seem contrary to the greater performance of CC carriers reported in chapter five.

This disparity might be due to several fundamental differences between the PPMI study cohort and the PD-MCI study cohort at the University of Calgary. The first difference could be due to the small size of the PD-MCI cohort compared to the PPMI Cohort. Nevertheless, the vigorous phenotyping in the PD-MCI cohort, the consistency of methods throughout the study, and the application of a strict method of correction for multiple comparisons seem to reduce the risk of this issue. The second difference might be the state of PD patients in the PPMI study. PD patients in the PPMI study were recruited at the initial diagnosis state while still being drug-naïve (before the start of dopaminergic therapy). The difference between PPMI and PD-MCI results might be explained as an effect of the dopaminergic medications on PD patients. Specifically, there has been evidence of interaction between *SNCA,* and Levodopa medication based on ML analysis [253].

Another possible explanation for the different results between the two cohorts could be that the CC genotype might be associated with more reversible cognitive impairments. There have been reports of restoration and improvement of cognitive functions in PD patients after dopaminergic therapy [265-267]. However, the later topic of efficacy of dopaminergic therapy in restoration of cognitive function is still a matter of debate [268-270]. Last but not least is the substantial difference between the percentage of male and female participants between the PPMI and PD-MCI study. The percentage of female and male participants is exactly reversed between the two studies. This can induce substantial differences between the two studies. Based on a recent study, male PD patients are at higher risk of cognitive impairments in the executive-function and in processing speed abilities compared to female PD patients with similar demographics [210]. Furthermore, female PD patients have lower risk of cognitive impairment compared to male patients over time (hazard ratio 0.59 [0.48, 0.73], $P = 4.6E-7$, adjusted for the years of education) [255].

Other factors that might contribute to the higher reversion rate in CC carriers might be the quality of care, socio-economic status, and other factors [271].

Overall, rs894280 seems to be a promising variant in the identification of PD patients with an elevated risk of cognitive decline. Although we found promising evidence on the association of this variant with cognitive changes in PD patients, further investigations are required to elucidate this relationship in more detail.

7.1 Limitations

There are several limitations that should be considered for the first part of this thesis. One of the limitations of chapter three and four studies, was the lack of a control group. *BDNF* was reported in several studies in association with depression and specifically geriatric depression, as well as in the general population [60]. The investigation of rs6265 association with NPS in PD patients will be more thorough if NPS can be studied in an age-sex matched control group.

The application of MBI-C might have some drawback. The MBI-C questionnaire was designed for the general senior population [107]. Although there is a substantial overlap between PD patients and the general senior population regarding NPS, there are some considerable differences between the two groups. One example could be the prevalence of impulsive/compulsive disorders in PD patients which can emerge after taking dopaminergic medication. Although the abnormal thoughts/perception domain of MBI-C includes several important symptoms, the presence of visual hallucinations which is a serious issue in PD patients is not addressed. Moreover, there are no results from a large-scale study to validate the use of MBI-C in PD populations and to validate the cutoff scores used for MBI positive identification.

A comprehensive cognitive assessment for each of the chapters' results alongside with the MBI assessment would be beneficial. It has been demonstrated that a relationship exists between cognitive impairment and NPS in PD patients. Including a comprehensive cognitive assessment can improve the generalizability of the findings to PD patients outside these studies as well. q

The most important limitation for the second part of this thesis is the novelty of the results for rs894280 in the *SNCA* gene. The lack of previous data on the exact relationship between this variant and cognition of PD patients restricts any comprehensive conclusions. Further investigations in independent cohorts and population of PD patients could be helpful in explaining the true nature of the association between this variant and the cognitive changes in PD. Furthermore, a thorough exploration of the 5' region of the *SNCA* along with a complete cognitive and neuropsychiatric assessment are required. Such investigations could validate the reported association for this variant. Also, they could potentially identify other closely related variants which actually modulate the relationship with cognition in 5' UTR region of *SNCA.*

Similar to the limitations of part one, the lack of an age-sex matched control group complicates drawing conclusions for this variant. Therefore, future studies might consider including a control group to explore the relationship of rs894280 with cognition in the general aging population. Lack of longitudinal data for more than five years, was another limitation for part two. It would be considerably helpful to study the relationship of rs894280 with cognition and behavior in PD patients over a longer period of time. It has been argued that PD patients have a higher likelihood of dementia as the disease progresses [272].

Last but not least, is the potential issues of the candidate genes approach (CGA) in this thesis. PD is a very complex and heterogenous disease and using CGA helped with statistical power of the presented projects. Focusing on a candidate gene can introduce problems in studying the actual

relationship between the symptom in question and the candidate gene [273]. Also using CGA can readily ignore the population stratification. Nevertheless, this was partially addressed by controlling for ethnicity of the study participants.

7.2 Future Directions

The importance of early detection and treatment of cognitive impairments in PD has been the focus of many research groups. NPS and their relationship with cognitive impairment in iPD patients have gained more importance in recent years. My colleagues and I explored several variants in association with NPS and cognitive impairments. Given the nature of the genetic studies and how various populations might display very different results regarding associative analysis, the current findings should be considered with caution. However, these findings, if replicated in other PD cohorts, can be applied in the clinics for a straightforward and prompt identification of PD patients at risk of cognitive impairment or NPS.

The ideal study design for investigation of the above-mentioned variants would be a longitudinal study of more than 10 years follow-up and multi-centric using the exome sequencing technique which covers coding regions, of most of the known and regulatory non-coding regions proximate to the genes. The suggested 10-year of follow-up is because of the evidence that the likelihood of PDD increases and can be observed in up to 80% of PD patients after 20 years of disease diagnosis [272]. The 10-year follow-up might provide ample chance of studying and identifying PD patients who would develop PDD. The use of multiple recruitment centers across the country, the continent or globally could incorporate the potential genetic stratifications [56]. The use of genetic techniques focusing on variant genotyping could provide high accuracy but the complicated

structure of the *SNCA* gene makes the exome sequencing a more appropriate option for this gene. The non-coding regions of SNCA might have a substantial role for this gene in iPD patients [88]. Ideally this technique could be applied to study other genes mentioned in this thesis (*BDNF*, *COMT* and *SLC6A3*). However, for the purpose of this thesis the TaqMan genotyping and VNTR techniques could readily provide the data needed for this purpose at a much lower cost. To increase the power of analysis including a control group can be useful [274].

Several steps can be taken to shed more light on the findings presented in this thesis. The BDNF protein was reported as a potential protective factor for dopaminergic neurons in the substantia nigra [275]. The association of BDNF protein and dopaminergic neuron density can be explored using neuromelanin-sensitive MRI technique [276, 277]. Using neuromelanin-sensitive MRI, we hypothesize that iPD patients with the Val allele would have greater melanin density in their substantia nigra and possibly higher density of the dopaminergic neurons [264, 275]. This could be attributed to the higher level of functional BDNF in the Val carriers.

To study the possible link between BDNF and dopaminergic neurons, and given the dominant model for *BDNF*, OR = 2.0, α= 5%, MAF = 0.1, 264 participants (case-control ratio of 1:1) are required to achieve statistical power of 80% [274]. However, the detected OR for rs6265 and MBI likelihood was 2.8 which lowers the required sample size to achieve the sufficient statistical power.

The other future plan to follow is to investigate whether rs894280 in *SNCA* has any association with NPS in iPD patients. To explore such association, assuming a dominant model for this variant similar to rs6265, OR=2.3, α= 5%, MAF = 0.5, less than 160 participants are required to achieve 80% power [274].

The interaction of rs6265 and rs894280 might facilitate the calculation of a risk score for these two variants combined [278]. However, to be able to compute the potential interaction between these two variants an at least four-fold larger sample size is required [278]. Considering that the dominant model of rs6265 might require a sample size of 264 participants, a four-fold larger sample size yields to 1,056. The four-fold increase in sample size is the result of 2x2 interaction model. If rs894280 interaction with rs6265 is explored, a 3x2 or even 3x3 larger sample size will be required, six-fold and nine-fold, respectively [274, 278].

It needs to be noted that recent studies demonstrated an association between NPS and cognitive impairments in AD and PD patients [120, 154]. Considering these findings, the possibility of identification of PD patients at risk of NPS might facilitate early intervention to prevent or slow down the evolution of these symptoms. These interventions might be change of lifestyle, receiving appropriate medical treatments (such as prescribing drugs improving mood or even cognition), or cognitive training [271, 279].

In conclusion, my colleagues and I have found a common polymorphism p.Val66Met in *BDNF* gene associated with NPS. We reported lack of association for two common variants in genes modulating dopamine availability, rs4680 in *COMT*, and rs28363170 in *SLC6A3* gene, respectively. On the other hand, the lack of association for rs4680 or rs28363170 variants in our cohort can suggest that these variants are not involved in the pathology of NPS in iPD patients. These findings indicate a potential role for the *BDNF* gene in the pathology of NPS in iPD patients while rs4680 and rs28363170 did not display such a relationship.

A machine learning analysis revealed a substantial contribution of rs894280, a variant in *SNCA*, to cognitive deficits of iPD patients. My colleagues and I observed that rs894280 was associated with overall cognition, attention and visuospatial abilities in iPD patients. We later demonstrated

that this association could be replicated for cognitive status of iPD patients at baseline. We additionally discovered an association between this variant and the evolution of cognitive status in iPD patients. These findings indicate a potential role for rs894280 in cognition of PD patients. This variant might serve as a proxy for identification of PD patients at risk of cognitive impairments in the early stages of PD.

8 References

[1] Wong SL, Gilmour H, Ramage-Morin PL (2014) Parkinson's disease- Prevalence, diagnosis and impact. *Health Reports* **25**, 10-14.

[2] Parkinson J (1817) Essay on the Shaking Palsy. *London: Sherwwod, Neely and Jones*.

[3] Manyam BV (1990) Paralysis agitans and levodopa in "Ayurveda": ancient Indian medical treatise. *Mov Disord* **5**, 47-48.

[4] Spillantini MG, Schmidt ML, Lee VM-Y, Trojanowski JQ, Goedert M, Jakes R (1997) alpha-Synuclein in Lewy bodies. *Nature* **388**, 839-840.

[5] Braak H, Del Tredici K, Bratzke H, Hamm-Clement J, Sandmann-Keil D, Rub U (2002) Staging of the intracerebral inclusion body pathology associated with idiopathic Parkinson's disease (preclinical and clinical stages). *J Neurol* **249 Suppl 3**, III/1-5.

[6] Weintraub D, Simuni T, Caspell-Garcia C, Coffey C, Lasch S, Siderowf A, Aarsland D, Barone P, Burn D, Chahine LM, Eberling J, Espay AJ, Foster ED, Leverenz JB, Litvan I, Richard I, Troyer MD, Hawkins KA, Parkinson's Progression Markers I (2015) Cognitive performance and neuropsychiatric symptoms in early, untreated Parkinson's disease. *Mov Disord* **30**, 919-927.

[7] Kalia LV, Lang AE (2015) Parkinson's disease. *The Lancet* **386**, 896-912.

[8] Marinus J, Zhu K, Marras C, Aarsland D, van Hilten JJ (2018) Risk factors for non-motor symptoms in Parkinson's disease. *The Lancet Neurology* **17**, 559-568.

[9] Mantri S, Morley JF, Siderowf AD (2019) The importance of preclinical diagnostics in Parkinson disease. *Parkinsonism Relat Disord* **64**, 20-28.

[10] Litvan I, Aarsland D, Adler CH, Goldman JG, Kulisevsky J, Mollenhauer B, Rodriguez-Oroz MC, Troster AI, Weintraub D (2011) MDS Task Force on mild cognitive impairment in Parkinson's disease: critical review of PD-MCI. *Mov Disord* **26**, 1814-1824.

[11] Saredakis D, Collins-Praino LE, Gutteridge DS, Stephan BCM, Keage HAD (2019) Conversion to MCI and dementia in Parkinson's disease: a systematic review and meta-analysis. *Parkinsonism Relat Disord* **65**, 20-31.

[12] Oosterveld LP, Allen JC, Jr., Reinoso G, Seah SH, Tay KY, Au WL, Tan LC (2015) Prognostic factors for early mortality in Parkinson's disease. *Parkinsonism Relat Disord* **21**, 226-230.

[13] Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *J Intern Med* **256**, 183-194.

[14] Galtier I, Nieto A, Lorenzo JN, Barroso J (2016) Mild cognitive impairment in Parkinson's disease: Diagnosis and progression to dementia. *Journal of clinical and Experimental Neuropsychology* **38**, 40-50.

[15] Litvan I, Goldman JG, Troster AI, Schmand BA, Weintraub D, Petersen RC, Mollenhauer B, Adler CH, Marder K, Williams-Gray CH, Aarsland D, Kulisevsky J, Rodriguez-Oroz MC, Burn DJ, Barker RA, Emre M (2012) Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Mov Disord* **27**, 349-356.

[16] Williams-Gray CH, Mason SL, Evans JR, Foltynie T, Brayne C, Robbins TW, Barker RA (2013) The CamPaIGN study of Parkinson's disease: 10-year outlook in an incident population-based cohort. *J Neurol Neurosurg Psychiatry* **84**, 1258-1264.

[17] Barrett MJ, Smolkin ME, Flanigan JL, Shah BB, Harrison MB, Sperling SA (2017) Characteristics, correlates, and assessment of psychosis in Parkinson disease without dementia. *Parkinsonism Relat Disord* **43**, 56-60.

[18] Verbaan D, Marinus J, Visser M, van Rooden SM, Stiggelbout AM, Middelkoop HA, van Hilten JJ (2007) Cognitive impairment in Parkinson's disease. *J Neurol Neurosurg Psychiatry* **78**, 1182-1187.

[19] Alves G, Larsen JP, Emre M, Wentzel-Larsen T, Aarsland D (2006) Changes in motor subtype and risk for incident dementia in Parkinson's disease. *Mov Disord* **21**, 1123-1130.

[20] Muslimović D, Schmand B, Speelman JD, de Haan RJ (2007) Course of cognitive decline in Parkinson's disease: a meta-analysis. *Journal of the International Neuropsychological Society* **13**, 920-932.

[21] Cammisuli DM, Cammisuli SM, Fusi J, Franzoni F, Pruneti C (2019) Parkinson's Disease-Mild Cognitive Impairment (PD-MCI): A Useful Summary of Update Knowledge. *Front Aging Neurosci* **11**, 303.

[22] Armstrong MJ, Naglie G, Duff-Canning S, Meaney C, Gill D, Eslinger PJ, Zadikoff C, Mapstone M, Chou KL, Persad C, Litvan I, Mast BT, Fox S, Tang-Wai DF, Marras C (2012) Roles of Education and IQ in Cognitive Reserve in Parkinson's Disease-Mild Cognitive Impairment. *Dement Geriatr Cogn Disord Extra* **2**, 343-352.

[23] Gallagher DA, Schrag A- (2012) Psychosis, apathy, depression and anxiety in Parkinson's disease. *Neurobiol Dis* **46**, 581-589.

[24] Remy P, Doder M, Lees A, Turjanski N, Brooks D (2005) Depression in Parkinson's disease: loss of dopamine and noradrenaline innervation in the limbic system. *Brain* **128**, 1314-1322.

[25] Storch A, Schneider CB, Wolz M, Stürwald Y, Angelika Nebe PO, Andreas Mahler, Gerd Fuchs, Wolfgang H Jost, K Ray Chaudhuri, Rainer Koch, Heinz Reichmann, Georg Ebersbach (2013) Nonmotor fluctuations in Parkinson disease: severity and correlation with motor complications. *Neurology* **80**, 800-809.

[26] Cooney JW, Stacy M (2016) Neuropsychiatric Issues in Parkinson's Disease. *Curr Neurol Neurosci Rep* **16**, 49.

[27] Klein C, Westenberger A (2012) Genetics of Parkinson's disease. *Cold Spring Harb Perspect Med* **2**, a008888.

[28] Gasser T (2005) Genetics of Parkinson's Disease. *Mov Disord* **18**, 363-369.

[29] Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, Hulihan M, Peuralinna T, Dutra A, Nussbaum R, Lincoln S, Crawley A, Hanson M, Maraganore D, Adler C, Cookson MR, Muenter M, Baptista M, Miller D, Blancato J, Hardy J, Gwinn-Hardy K (2003) alpha-Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841.

[30] Zarranz JZ, Alegre J, Go´mez-Esteban JC, Lezcano E, Ros R, Ampuero I, Lı´dice Vidal P, 2 Janet Hoenicka, PhD,2 Olga Rodriguez, MD,3, Begon˜a Atare´s M, 4 Vero´nica Llorens, MD,5 Estrella Gomez Tortosa, MD, PhD,2,6, Teodoro del Ser M, PhD,7 David G. Mun˜oz, MD, PhD,2 and Justo G. de Yebenes, (2004) The New Mutation, E46K, of -Synuclein Causes Parkinson and Lewy Body Dementia. *Ann Neurol* **55**.

[31] Gasser T, Müller-Myhsok B, Wszolek ZK, al. e (1998) A susceptibility locus for Parkinson's disease maps to chromosome 2p13. *Nat Genet* **18**, 262-265.

[32] Zimprich A, Biskup S, Leitner P, al. e (2004) Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **44**, 601-607.

[33] Paisan-Ruiz C, Jain S, Evans EW, al. e (2004) Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron* **44**.

[34] Williams-Gray CH, Hampshire A, Robbins TW, Owen AM, Barker RA (2007) Catechol O-methyltransferase Val158Met genotype influences frontoparietal activity during planning in patients with Parkinson's disease. *J Neurosci* **27**, 4832-4838.

[35] Williams-Gray CH, Evans JR, Goris A, Foltynie T, Ban M, Robbins TW, Brayne C, Kolachana BS, Weinberger DR, Sawcer SJ, Barker RA (2009) The distinct cognitive syndromes of Parkinson's disease: 5 year follow-up of the CamPaIGN cohort. *Brain* **132**, 2958-2969.

[36] (2020), ed. NCBI.

[37] Mahley RW (1988) Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science* **240**, 622-630.

[38] Rebeck GW (2017) The role of APOE on lipid homeostasis and inflammation in normal brains. *J Lipid Res* **58**, 1493-1499.

[39] Zannis VI, Breslow JL, G Utermann RWM, K H Weisgraber, R J Havel, J L Goldstein, M S Brown, G Schonfeld, W R Hazzard, C Blum (1982) Proposed nomenclature of apoE isoproteins, apoE genotypes, and phenotypes. *J Lipid Res* **6**, 911-914.

[40] Weisgraber KH, Rall SC, Mahley RW (1981) Human E apoprotein heterogeneity. Cysteine-arginine interchanges in the amino acid sequence of the apo-E isoforms. . *J Biol Chem* **256**, 9077-9083.

[41] Spinney L (2014) Alzheimer's disease: The forgetting gene. *Nature* **510**, 26-28.

[42] Mata IF, Leverenz JB, Weintraub D, Trojanowski JQ, Hurtig HI, Van Deerlin VM, Ritz B, Rausch R, Rhodes SL, Factor SA, Wood-Siverio C, Quinn JF, Chung KA, Peterson AL, Espay AJ, Revilla FJ, Devoto J, Hu SC, Cholerton BA, Wan JY, Montine TJ, Edwards KL, Zabetian CP (2014) APOE, MAPT, and SNCA genes and cognitive performance in Parkinson disease. *JAMA Neurol* **71**, 1405-1412.

[43] Gao J, Huang X, Park Y, Liu R, Hollenbeck A, Schatzkin A, Mailman RB, Chen H (2011) Apolipoprotein E genotypes and the risk of Parkinson disease. *Neurobiol Aging* **32**, 2106 e2101-2106.

[44] Zareparsi S, Camicioli R, Sexton G, Bird TSP, Kaye J, Nutt J, Payami H. (2002) Age at onset of Parkinson disease and apolipoprotein E genotypes. *Am J Med Genet B Neuropsychiatr Genet* **107**, 156-161.

[45] Pankratz N, Byder L, Halter C, Rudolph A, Clifford W Shults PMC, Tatiana Foroud, William C Nichols (2006) Presence of an APOE4 allele results in significantly earlier onset of Parkinson's disease and a higher risk with dementia. *Mov Disord* **21**, 45-49.

[46] Williams-Gray CH, Goris A, Saiki M, Thomas Foltynie DASC, Stephen J Sawcer, Roger A Barker (2009) Apolipoprotein E genotype as a risk factor for susceptibility to and dementia in Parkinson's disease. *J Neurol* **256**, 493-498.

[47] Collins LM, Williams-Gray CH (2016) The Genetic Basis of Cognitive Impairment and Dementia in Parkinson's Disease. *Front Psychiatry* **7**, 89.

[48] Morley JF, Xie SX, Hurtig HI, Stern MB, Colcher A, Horn S, Dahodwala N, Duda JE, Weintraub D, Chen-Plotkin AS, Van Deerlin V, Falcone D, Siderowf A (2012) Genetic influences on cognitive decline in Parkinson's disease. *Mov Disord* **27**, 512-618.

[49] Hofer M, Pagliusi SR, Hohn A, Leibrock J, Barde YA (1990) Regional distribution of brain-derived neurotrophic factor mRNA in the adult mouse brain. *EMBO J* **9**, 2459-2464.

[50] Ernfors P, Lee KF, Jaenisch R (1994) Mice lacking brain-derived neurotrophic factor develop with sensory deficits. *Nature* **368**, 147-150.

[51] Acheson A, Conover JC, Fandl JP, DeChiara TM, Russell M, Thadani A, Squinto SP, Yancopoulos GD, Lindsay RM (1995) A BDNF autocrine loop in adult sensory neurons prevents cell death. *Nature* **374**, 450-453.

[52] Blum R, Konnerth A (2005) Neurotrophin-mediated rapid signaling in the CNS, mechanisms and functions. *Physiology (Bethesda)* **20**, 70-78.

[53] Chen ZY, Patel PD, Sant G, Meng CX, Teng KK, Hempstead BL, Lee FS (2004) Variant brain-derived neurotrophic factor (BDNF) (Met66) alters the intracellular trafficking and activity dependent secretion of wild-type BDNF in neurosecretory cells and cortical neurons. *J Neurosci* **24**, 4401-4411.

[54] Hempstead BL (2015) Brain-Derived Neurotrophic Factro: three ligands, many actions. . *Transactions of the American Clinical and Climatological Association.* **126**, 9-19.

[55] Kowianski P, Lietzau G, Czuba E, Waskow M, Steliga A, Morys J (2018) BDNF: A Key Factor with Multipotent Impact on Brain Signaling and Synaptic Plasticity. *Cell Mol Neurobiol* **38**, 579-593.

[56] Wang Q, Liu J, Guo Y, Dong G, Zou W, Chen Z (2019) Association between BDNF G196A (Val66Met) polymorphism and cognitive impairment in patients with Parkinson's disease: a meta-analysis. *Braz J Med Biol Res* **52**, e8443.

[57] Tsai SJ (2018) Critical Issues in BDNF Val66Met Genetic Studies of Neuropsychiatric Disorders. *Front Mol Neurosci* **11**, 156.

[58] Costa A, Peppe A, Carlesimo GA, Zabberoni S, Scalici F, Caltagirone C, Angelucci F (2015) Brain-derived neurotrophic factor serum levels correlate with cognitive performance in

Parkinson's disease patients with mild cognitive impairment. *Front Behav Neurosci* **9**, 253.

[59] van der Kolk NM, Speelman AD, van Nimwegen M, Kessels RP, IntHout J, Hakobjan M, Munneke M, Bloem BR, van de Warrenburg BP (2015) BDNF polymorphism associates with decline in set shifting in Parkinson's disease. *Neurobiol Aging* **36**, 1605 e1601-1606.

[60] Pei Y, Smith AK, Wang Y, Pan Y, Yang J, Chen Q, Pan W, Bao F, Zhao L, Tie C, Wang Y, Wang J, Zhen W, Zhou J, Ma X (2012) The brain-derived neurotrophic-factor (BDNF) val66met polymorphism is associated with geriatric depression: a meta-analysis. *Am J Med Genet B Neuropsychiatr Genet* **159B**, 560-566.

[61] Suliman S, Hemmings SM, Seedat S (2013) Brain-Derived Neurotrophic Factor (BDNF) protein levels in anxiety disorders: systematic review and meta-regression analysis. *Front Integr Neurosci* **7**, 55.

[62] Cagni FC, Campelo C, Coimbra DG, Barbosa MR, Junior LGO, Neto ABS, Ribeiro AM, Junior COG, Gomes de Andrade T, Silva RH (2017) Association of BDNF Val66MET Polymorphism With Parkinson's Disease and Depression and Anxiety Symptoms. *J Neuropsychiatry Clin Neurosci* **29**, 142-147.

[63] Antypa N, Drago A, Serrettia A (2013) The role of COMT gene variants in depression: Bridging neuropsychological, behavioral and clinical phenotypes. *Neuroscience & Biobehavioral Reviews* **37**, 1597-1610.

[64] Lachman HM, Papolos DF, Saito T, Yu YM, Szumlanski CL, Weinshilboum RM (1996) Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics* **6**, 243-250.

[65] Qayyum A, Zai CC, Hirata Y, Tiwari AK, Cheema S, Nowrouzi B, Beitchman JH, Kennedy JL (2015) The Role of the Catechol-o-methyltransferase (COMT) Gene Val158Met in Aggressive Behavior, A Review of Genetic Studies. *curr Neurolpharmacol* **13**, 802-814.

[66] Chen J, Lipska BK, Halim N, Ma QD, Matsumoto M, S. Melhem BSK, T.M. Hyde, M.M. Herman, J. Apud, M.F. Egan, , Kleinman JE, Weinberger DR (2004) Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): effects on mRNA, protein, and enzyme activity in postmortem human brain. *Am. J. Hum. Genet.* **75**, 807-821.

[67] Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mazzanti CM, Straub RE, Goldman D, Weinberger DR (2001) Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc Natl Acad Sci U S A* **98**, 6917-6922.

[68] Piccini PP (2003) Dopamine transporter: basic aspects and neuroimaging. *Mov Disord* **18 Suppl 7**, S3-8.

[69] Kim JW, Kim DH, Kim SH, J.K. C (2000) Association of the dopamine transporter gene with Parkinson's disease in Korean patients. *J Korean Med Sci* **15**, 449-451.

[70] Zhai D, Li b, S., Zhao Y, Linc Z (2014) SLC6A3 is a risk factor for Parkinson's disease: A meta-analysis of sixteen years' studies. *Neurosci Lett* **564**, 99-104.

[71] Heinz A, Goldman D, Jones DW, Palmour R, Hommer D, Gorey JG, Lee KS, Linnoila M, Weinberger DR (2000) Genotype influences in vivo dopamine transporter availability in human striatum. *Neuropsychopharmacology* **22**, 133-139.

[72] Mill J, Asherson P, Browes C, D'Souza U, Craig I (2002) Expression of the dopamine transporter gene is regulated by the 3' UTR VNTR: evidence from brain and lymphocytes using quantitative RT-PCR. *Am J Med Genet* **114**, 975-979.

[73] Habak C, Noreau A, Nagano-Saito A, Mejia-Constain B, Degroot C, Strafella AP, Chouinard S, Lafontaine AL, Rouleau GA, Monchi O (2014) Dopamine transporter SLC6A3 genotype affects cortico-striatal activity of set-shifts in Parkinson's disease. *Brain* **137**, 3025-3035.

[74] du Plessis S, Bekker M, Buckle C, Vink M, Seedat S, Bardien S, Carr J, Abrahams S (2020) Association Between a Variable Number Tandem Repeat Polymorphism Within the DAT1 Gene and the Mesolimbic Pathway in Parkinson's Disease. *Front Neurol* **11**, 982.

[75] Kaiser R, Hofer A, Grapengiesser A, Gasser T, Kupsch T, Roots I, Brockmöller J (2003) L -dopa-induced adverse effects in PD and dopamine transporter gene polymorphism. *Neurology* **60**, 1750-1755.

[76] Compta Y, Parkkinen L, O'Sullivan SS, Vandrovcova J, Holton JL, Collins C, Lashley T, Kallis C, Williams DR, de Silva R, al. e (2011) Lewy- and Alzheimer-type pathologies in Parkinson's disease dementia: which is more important? *Brain* **134**, 1493-1505.

[77] Zody MC, Jiang Z, Fung HC, Antonacci F, Hillier LW, Cardone MF, Graves TA, Kidd JM, Cheng Z, Abouelleil A, al. e (2008) Evolutionary toggling of the MAPT 17q21.31 inversion region. *Nat. Genet.* **40**, 1076–1083.

[78] Zabetian CP, Hutter CM, Factor SA, Nutt JG, Higgins DS, Griffith A, Roberts JW, Leis BC, Kay DM, Yearout D, Montimurro JS, Edwards KL, Samii A, Payami H (2010) Association Analysis of MAPT H1 Haplotype and Subhaplotypes in Parkinson's Disease. *Ann Neurol* **62**, 137-144.

[79] Zhang CC, Zhu JX, Wan Y, Tan L, Wang HF, Yu JT, Tan L (2017) Meta-analysis of the association between variants in MAPT and neurodegenerative diseases. *Oncotarget* **8**, 44994-45007.

[80] Desikan RS, Schork AJ, Wang Y, Witoelar A, Sharma M, McEvoy LK, Holland D, Brewer JB, Chen CH, Thompson WK, Harold D, Williams J, Owen MJ, O'Donovan MC, Pericak-Vance MA, Mayeux R, Haines JL, Farrer LA, Schellenberg GD, Heutink P, Singleton AB, Brice A, Wood NW, Hardy J, Martinez M, Choi SH, DeStefano A, Ikram MA, Bis JC, Smith A, Fitzpatrick AL, Launer L, van Duijn C, Seshadri S, Ulstein ID, Aarsland D, Fladby T, Djurovic S, Hyman BT, Snaedal J, Stefansson H, Stefansson K, Gasser T, Andreassen OA,

Dale AM, Adni AGC, Investigators I (2015) Genetic overlap between Alzheimer's disease and Parkinson's disease at the MAPT locus. *Mol Psychiatry* **20**, 1588-1595.

[81] Seto-Salvia N, Clarimo J, Pagonabarraga J, Pascual-Sedano B, Campolongo A, Combarros O, MD PJIM, MD, PhD; Daniel Regan˜ a, BS;, Merce Martı´nez-Corral MMM, MD; Daniel Alcolea, MD; Marc Sua´ rez-Calvet, MD;, Laura Molina-Porcel MOD, BS; Teresa Go´mez-Isla, MD, PhD; Rafael Blesa, MD, PhD;, Alberto Lleo´ M, PhD; Jaime Kulisevsky, (2011) Dementia risk in Parkinson disease- disentangling the role of MAPT haplotypes. *Arch Neurol* **68**, 359-364.

[82] Evans JR, Mason SL, Williams-Gray CH, Foltynie T, Brayne C, Robbins TW, Barker RA (2011) The natural history of treated Parkinson's disease in an incident, community based cohort. *J. Neurol. Neurosurg. Psychiatry* **82**, 1112e1118.

[83] Winder-Rhodes SE, Hampshire A, Rowe JB, Peelle JE, Robbins TW, Owen AM, Barker RA (2015) Association between MAPT haplotype and memory function in patients with Parkinson's disease and healthy aging individuals. *Neurobiol Aging* **36**, 1519-1528.

[84] Huertas I, Jesus S, Garcia-Gomez FJ, Lojo JA, Bernal-Bernal I, Bonilla-Toribio M, Martin-Rodriguez JF, Garcia-Solis D, Gomez-Garre P, Mir P (2017) Genetic factors influencing frontostriatal dysfunction and the development of dementia in Parkinson's disease. *PLoS One* **12**, e0175560.

[85] Fagan ES, Pihlstrom L (2017) Genetic risk factors for cognitive decline in Parkinson's disease: a review of the literature. *Eur J Neurol* **24**, 561-e520.

[86] Polymeropoulos MH, Lavedan† C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R, Stenroos ES, Chandrasekharappa S, Athanassiadou A, Papapetropoulos T, Johnson WG, Lazzarini AM, Duvoisin RC, Di Iorio G, Golbe LI, Nussbaum RL (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045-2047.

[87] Hattori N, Kitada T, Matsumine H, Asakawa S, Yamamura Y, Yoshino H, Kobayashi T, Yokochi M, Wang M, Yoritaka A, Kondo T, Kuzuhara S, Nakamura S, Shimizu N, Mizuno Y (1998) Molecular genetic analysis of a novel parkin gene in Japanese families with autosomal recessive juvenile parkinsonism: evidence for variable homozygous deletions in the parkin gene in affected individuals. . *Ann Neurol* **44**, 935-941.

[88] Guella I, Evans DM, Szu-Tu C, Nosova E, Bortnick SF, Group SCS, Goldman JG, Dalrymple-Alford JC, Geurtsen GJ, Litvan I, Ross OA, Middleton LT, Parkkinen L, Farrer MJ (2016) alpha-synuclein genetic variability: A biomarker for dementia in Parkinson disease. *Ann Neurol* **79**, 991-999.

[89] Oczkowska A, Kozubski W, Lianeri M, Dorszewska J (2013) Mutations in PRKN and SNCA Genes Important for the Progress of Parkinson's Disease *Current Genomics* **14**, 502-517.

[90] Rajput A, Vilariño-Güell C, Rajput ML, Ross OA, Soto-Ortolaza AI, Lincoln SJ, Cobb SA, Heckman MG, Farrer MJ, Rajput A (2009) Alpha-synuclein polymorphisms are associated with Parkinson's disease in a Saskatchewan population. *Mov Disord* **24**, 2411-2414.

[91] Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, DeStefano AL, Kara E, Bras J, Sharma M, Schulte C, Keller MF, Arepalli S, Letson C, Edsall C, Stefansson H, Liu X, Pliner H, Lee JH, Cheng R, International Parkinson's Disease Genomics C, Parkinson's Study Group Parkinson's Research: The Organized GI, andMe, GenePd, NeuroGenetics Research C, Hussman Institute of Human G, Ashkenazi Jewish Dataset I, Cohorts for H, Aging Research in Genetic E, North American Brain Expression C, United Kingdom Brain Expression C, Greek Parkinson's Disease C, Alzheimer Genetic Analysis G, Ikram MA, Ioannidis JP, Hadjigeorgiou GM, Bis JC, Martinez M, Perlmutter JS, Goate A, Marder K, Fiske B, Sutherland M, Xiromerisiou G, Myers RH, Clark LN, Stefansson K, Hardy JA, Heutink P, Chen H, Wood NW, Houlden H, Payami H, Brice A, Scott WK, Gasser T, Bertram L, Eriksson N, Foroud T, Singleton AB (2014) Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* **46**, 989-993.

[92] Burré J, Sharma M, Tsetsenis T, Buchman V, Etherton M, Südhof TC (2010) α-Synuclein Promotes SNARE-Complex Assembly in vivo and in vitro. *Science* **329**, 1663-1667.

[93] Goldman JG, Williams-Gray C, Barker RA, Duda JE, Galvin JE (2014) The spectrum of cognitive impairment in Lewy body diseases. *Mov Disord* **29**, 608-621.

[94] Campelo CLC, Cagni FC, de Siqueira Figueredo D, Oliveira LG, Jr., Silva-Neto AB, Macedo PT, Santos JR, Izidio GS, Ribeiro AM, de Andrade TG, de Oliveira Godeiro C, Jr., Silva RH (2017) Variants in SNCA Gene Are Associated with Parkinson's Disease Risk and Cognitive Symptoms in a Brazilian Sample. *Front Aging Neurosci* **9**, 198.

[95] Zhang K, Tang Y, Meng L, Zhu L, Zhou X, Zhao Y, Yan X, Tang B, Guo J (2019) The Effects of SNCA rs894278 on Resting-State Brain Activity in Parkinson's Disease. *Front Neurosci* **13**, 47.

[96] Holland PM, Abramson RD, Watson R, Gelfand DH (1991) Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc Natl Acad Sci U S A* **88**, 7276–7280.

[97] Scientific TF, How TaqMan Assay Works,Thermo Fisher Scientific,

[98] Marwal A, Sahu AK, Gaur RK (2014) Moleculat Markers: tool for genetic analysis In *Animal Biotechnology; Models in Discovery and Translation*, Verma AS, Singh A, eds. Academic Press, pp. 289-305.

[99] Choi M, Scholl UI, Ji W., Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloğlu A, Ozen S., Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton RP (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A* **106**, 19096-19101.

[100] Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, McCarroll SA, Visscher PM (2009) Finding the missing heritability of complex diseases. *Nature* **4618**, 747-753.

[101] Ilumina I, https://www.illumina.com,

[102] PPMI, https://www.ppmi-info.org,

[103] Goh G, Choi M (2012) Application of whole exome sequencing to identify disease-causing variants in inherited human diseases. *Genomics Inform.*, 214-219.

[104] Harvey PD (2012) Clinical applications of neuropsychological assessment. *Dialogues Clin Neurosci* **14**, 91-99.

[105] Anastasi A, Urbina S (1997) *Psychological testing* Prentice-Hall, Upper Saddle River, NJ.

[106] Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal Cognitive Assessment, MoCA- A Brief ScreeningTool For Mild Cognitive Impairment. *Journal of American Geriatrics Society* **53**, 695-699.

[107] Ismail Z, Aguera-Ortiz L, Brodaty H, Cieslak A, Cummings J, Fischer CE, Gauthier S, Geda YE, Herrmann N, Kanji J, Lanctot KL, Miller DS, Mortby ME, Onyike CU, Rosenberg PB, Smith EE, Smith GS, Sultzer DL, Lyketsos C, Research NPSPIAotISotAAs, Treatment (2017) The Mild Behavioral Impairment Checklist (MBI-C): A Rating Scale for Neuropsychiatric Symptoms in Pre-Dementia Populations. *J Alzheimers Dis* **56**, 929-938.

[108] Creese B, Brooker H, Ismail Z, Wesnes KA, Hampshire A, Khan Z, Megalogeni M, Corbett A, Aarsland D, Ballard C (2019) Mild Behavioral Impairment as a Marker of Cognitive Decline in Cognitively Normal Older Adults. *Am J Geriatr Psychiatry* **27**, 823-834.

[109] Creese B, Arathimos R, Brooker H, Aarsland D, Corbett A, Lewis C, Ballard C, Ismail Z (2021) Genetic risk for Alzheimer's disease, cognition, and mild behavioral impairment in healthy older adults. *Alzheimers Dement (Amst)* **13**, e12164.

[110] Hanganu A, Bruneau MA, Degroot C, Bedetti C, Mejia-Constain B, Lafontaine AL, Chouinard S, Monchi O (2017) Depressive symptoms in Parkinson's disease correlate with cortical atrophy over time. *Brain Cogn* **111**, 127-133.

[111] Hanganu A, Monchi O (2018) Interaction Between Neuropsychiatric Symptoms and Cognitive Performance in Parkinson's Disease: What Do Clinical and Neuroimaging Studies Tell Us? *Curr Neurol Neurosci Rep* **18**, 91.

[112] Burton EJ, McKeith IG, Burn DJ, Williams ED, O'Brien JT (2004) Cerebral atrophy in Parkinson's disease with and without dementia: a comparison with Alzheimer's disease, dementia with Lewy bodies and controls. *Brain* **127**, 791-800.

[113] Nagano-Saito A, Washimi Y, Arahata Y, Kachi T, Lerch JP, Evans AC, Dagher A, Ito K (2005) Cerebral atrophy and its relation to cognitive impairment in Parkinson disease. *Neurology* **64**, 224-229.

[114] Bloch F (1946) Nuclear induction *Phys Rev* **70**, 460-474.

[115] Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179-194.

[116] Fischl B, Serenob MI, Dalea AM (1999) Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *Neuroimage* **9**, 195-207.

[117] Fischl B, Salat DH, Busa E, Albert MS, Dieterich M, Haselgrove C, van der Kouwe AJ, Killiany RJ, Kennedy D, Klaveness S, Albert Montillo, Nikos Makris, Rosen B, Dale AM (2002) Whole Brain Segmentation- Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuroimage* **33**, 341-355.

[118] Fischl B, Salat DH, van der Kouwe AJ, Makris N, Segonne F, Quinn BT, Dale AM (2004) Sequence-independent segmentation of magnetic resonance images. *Neuroimage* **23 Suppl 1**, S69-84.

[119] Hanganu A, Bedetti C, Jubault T, Gagnon JF, Mejia-Constain B, Degroot C, Lafontaine AL, Chouinard S, Monchi O (2013) Mild cognitive impairment in patients with Parkinson's disease is associated with increased cortical degeneration. *Mov Disord* **28**, 1360-1369.

[120] Yoon EJ, Ismail Z, Hanganu A, Kibreab M, Hammer T, Cheetham J, Kathol I, Sarna JR, Martino D, Furtado S, Monchi O (2019) Mild behavioral impairment is linked to worse cognition and brain atrophy in Parkinson disease. *Neurology* **93**, e766-e777.

[121] Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MT, Killiany RJ (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968-980.

[122] Awad M, Khanna R (2015) *Efficient Learning Machines. Theories, concepts, and applications for engineers and system designers*, Apress open.

[123] Vapnik V, Golowich SE, Smola AJ (1997) Support vector method for function approximation, regression estimation and signal processing. *Advances in Neural Information Processing Systems* **9**, 281-287.

[124] Lo Vercio L, Amador K, Bannister JJ, Crites S, Gutierrez A, MacDonald ME, Moore J, Mouches P, Rajasheka D, Schimert S, Subbanna N, Tuladhar A, Wang N, Wilms M, Winder A, Forkert ND (2020) Supervised machine learning tools: a tutorial for clinicians. *J Neural Eng*.

[125] Maleki F, Muthukrishnan N, Ovens K, Reinhold C, Forghani R (2020) Machine Learning Algorithm Validation: From Essentials to Advanced Applications and Implications for Regulatory Certification and Deployment. *Neuroimaging Clin N Am* **30**, 433-445.

[126] Molinaro AM, Simon R, Pfeiffer RM (2005) Prediction error estimation: a comparison of resampling methods. *Bioinformatics* **21**, 3301-3307.

[127] Gallagher DA, Schrag A (2012) Psychosis, apathy, depression and anxiety in Parkinson's disease. *Neurobiol Dis* **46**, 581-589.

[128] Martinez-Martin P, Rodriguez-Blazqueza C, JoãoForjaz M, Frades-Payo B, Agüera-Ortizc L, al. e (2015) Neuropsychiatric symptoms and caregiver's burden in Parkinson's disease. *Parkinsonims and related disorders* **21**, 629-634.

[129] Ismail Z, Smith EE, Geda Y, Sultzer D, Brodaty H, Smith G, Agüera-Ortiz L, Sweet R, Miller D, Lyketsos CG (2016) Neuropsychiatric symptoms as early manifestations of emergent dementia: provisional diagnostic criteria for mild behavioral impairment. *Alzheimer's & Dementia* **12**, 195-202.

[130] Yoon E, Ismail Z, Hanganu A, Kibreab M, Hammer T, Cheetham J, Kathol I, Sarna JR, Martino D, Furtado S, Monchi O (2019) Mild Behavioral Impairment is linked to worse cognition and brain atrophy in Parkinson's disease. *Neurology* **93**, e766-e777.

[131] Lang S, Yoon EJ, Kibreab M, Kathol I, Cheetham J, Hammer T, Sarna J, Ismail Z, Monchi O (2020) Mild behavioral impairment in Parkinson's disease is associated with altered corticostriatal connectivity. *NeuroImage: Clinical*, 102252.

[132] Hempstead BL (2015) Brain-Derived Neurotrophic Factro: three ligands, many actions. *TRANSACTIONS OF THE AMERICAN CLINICAL AND CLIMATOLOGICAL ASSOCIATION* **126**, 9-19.

[133] Mogi M, Togaria A, Kondo T, Mizuno Y, Komure O, Kuno S, Ichinose H, Nagatsud T (1999) Brain-derived growth factor and nerve growth factor concentrations are decreased in the substantia nigra in Parkinson's disease

[134] Altmann V, Schumacher-Schuh AF, Rieck M, Callegari-Jacques SM, Rieder CR, Hutz MH (2016) Val66Met BDNF polymorphism is associated with Parkinson's disease cognitive impairment. *Neurosci Lett* **615**, 88-91.

[135] Bessi V, Mazzeo S, Bagnoli S, Padiglioni S, Carraro M, Piaceri I, Bracco L, Sorbi S, Nacmias B (2019) The implication of BDNF Val66Met polymorphism in progression from subjective cognitive decline to mild cognitive impairment and Alzheimer's disease: a 9-year follow-up study. *Eur Arch Psychiatry Clin Neurosci*.

[136] Creese B, Griffiths A, Brooker H, Corbett A, Aarsland D, Ballard C, Ismail Z (2020) Profile of Mild Behavioral Impairment and Factor Structure of the Mild Behavioral Impairment Checklist in Cognitively Normal Older Adults. *International Psychogeriatrics* **32**, 705-717.

[137] Mallo SC, Ismail Z, Pereiro AX, Facal D, Lojo-Seoane C, al. e (2019) Assessing mild behavioral impairment with the mild behavioral impairment checklist in people with subjective cognitive decline. *International Psychogeriatrics* **31**, 231-239.

[138] Mallo SC, Ismail Z, Pereiro AX, Facal D, Lojo-Seoane C, Campos-Magdaleno M, Juncos-Rabadán O (2017) Assessing mild behavioral impairment in people with subjective

cognitive complaints (SCCS) with the mild behavioral impairment checklist (MBI-C): a pilot study. *the journal of Alzheimer Association* **13**, 364-365.

[139] Faul F, Erdfelder E, Lang A, Buchner A (2007) G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research* **39**, 175-191.

[140] VanDam D, Vermeiren Y, Dekker AD, Naudé PJW, DeDeyn PP (2016) Neuropsychiatric Disturbances in Alzheimer's Disease- What Have We Learned from Neuropathological Studies? *Curr Alzheimer Res.* **13**, 1145-1164.

[141] Chiu PY, Tsai CT, Chen PK, Chen WJ, Lai TJ (2016) Neuropsychiatric Symptoms in Parkinson's Disease Dementia Are More Similar to Alzheimer's Disease than Dementia with Lewy Bodies: A Case-Control Study. *PLoS One* **11**, e0153989.

[142] López-Pousa S, Vilalta-Franch J, Garre-Olmo J, Pons S, Cucurella MG (2007) Characterisation and prevalence of the psychological and behavioural symptoms in patients with dementia. *Europe PMC* **45**, 683-688.

[143] Taragano FE, Allegri RF, Heisecke SL, Martelli MI, Feldman ML, Sánchez V, García VA, Tufro G, Castro DM, Leguizamón PP (2018) Risk of Conversion to Dementia in a Mild Behavioral Impairment Group Compared to a Psychiatric Group and to a Mild Cognitive Impairment Group. *Journal of Alzheimer's Disease* **62**, 227-238.

[144] Naude J, Gill S, Hu S, McGirr A, Forkert N, Monchi O, Stys P, Smith EE, Ismail Z (2020) Plasma Neurofilament Light: a marker of cognitive decline in Mild Behavioural Impairment. *Journal of Alzheimer's disease* **76**, 1017-1027.

[145] Andrews SJ, Ismail Z, Anstey KJ, Mortby ME (2018) Association of Alzheimer's genetic loci with mild behavioral impairment. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **177**, 727-735.

[146] Gill S, Mouches P, Hu S, Rajashekar D, MacMaster FP, Smith EE, Forkert ND, Ismail Z, Initiative AsDN (2020) Using Machine Learning to Predict Dementia from Neuropsychiatric Symptom and Neuroimaging Data. *Journal of Alzheimer's Disease* **75**, 277-288.

[147] Creese B, Brooker H, Aarsland D, Corbett A, Ballard C, Ismail Z (2020) Genetic risk for Alzheimer disease, cognition and Mild Behavioral Impairment in healthy older adults. *medRxiv*, 2020.2005.2013.20100800.

[148] Lussier FZ, Pascoal TA, Chamoun M, Therriault J, Tissot C, Savard M, Kang MS, Mathotaarachchi S, Benedet AL, Parsons M, Qureshi MNI, Thomas ÉM, Shin M, Dion L-A, Massarweh G, Soucy J-P, Tsai I-H, Vitali P, Ismail Z, Rosa-Neto P, Gauthier S (2020) Mild behavioral impairment is associated with β-amyloid but not tau or neurodegeneration in cognitively intact elderly individuals. *Alzheimer's & Dementia* **16**, 192-199.

[149] Aarsland D, Cummings JL, Larsen JP (2001) Neuropsychiatric difference between Parkinson's disease with dementia and Alzheimer's disease. *Geriatric Psychiatry* **16**, 184-189.

[150] Noe E, Marder K, Bell KL, D.M. J, Manly JJ, Yaakov Stern Y (2004) Comparison of Dementia With Lewy Bodies to Alzheimer's Disease and Parkinson's Disease With Dementia. . *Mov Disord.*, 60-67.

[151] Matsuoka T, Ismail Z, Narumoto J (2019) Prevalence of Mild Behavioral Impairment and Risk of Dementia in a Psychiatric Outpatient Clinic. *Journal of Alzheimer's Disease* **70**, 505-513.

[152] Ismail Z, Gatchel J, Bateman DR, Barcelos-Ferreira R, Chantillon M, Jaeger J, Donovan NJ, Mortby ME (2018) Affective and emotional dysregulation as pre-dementia risk markers: exploring the mild behavioral impairment symptoms of depression, anxiety, irritability, and euphoria. *International Psychogeriatrics* **30**, 185-196.

[153] Ismail Z, McGirr A, Gill S, Hu S, Forkert ND, Smith EE (2020) Mild Behavioral Impairment and Subjective Cognitive Decline predict Mild Cognitive Impairment. *medRxiv*, 2020.2005.2024.20112284.

[154] Creese B, Brooker H, Ismail Z, Wesnes KA, Hampshire A, Khan Z, Megalogeni M, Corbett A, Aarsland D, Ballard C (2019) Mild Behavioral Impairment as a Marker of Cognitive

Decline in Cognitively Normal Older Adults. *The American Journal of Geriatric Psychiatry* **27**, 823-834.

[155] Mortby ME, Black SE, Gauthier S, Miller DS, Porsteinsson A, Smith EE, Ismail Z (2018) Dementia clinical trial implications of Mild Behavioral Impairment. *International Psychogeriatrics* **30**, 171-175.

[156] Taylor WD, Züchner S, McQuoid DR, Steffens DC, Speer MC, Krishnan KRR (2007) Allelic Differences in the Brain-Derived Neurotrophic Factor Val66Met Polymorphism in Late-Life Depression. *The American Journal of Geriatric Psychiatry* **15**, 850-857.

[157] Hwang JP, Tsai SJ, Hong CJ, Yang CH, Lirng JF, Yang YM (2006) The Val66Met polymorphism of the brain-derived neurotrophic-factor gene is associated with geriatric depression. *Neurobiol Aging* **27**, 1834-1837.

[158] Sen S, Duman R, Sanacora G (2008) Serum brain-derived neurotrophic factor, depression, and antidepressant medications: meta-analyses and implications. *Biol Psychiatry* **64**, 527-532.

[159] Suliman S, Hemmings, S. M., Seedat, S. (2013) Brain-Derived Neurotrophic Factor (BDNF) protein levels in anxiety disorders: systematic review and meta-regression analysis. *Front Integr Neurosci* **7**, 55.

[160] De Vincenti AP, Rios AS, Paratcha G, Ledda F (2019) Mechanisms That Modulate and Diversify BDNF Functions: Implications for Hippocampal Synaptic Plasticity. *Front Cell Neurosci* **13**, 135.

[161] Rex CS, Lauterborn JC, Lin CY, Kramar EA, Rogers GA, Gall CM, Lynch G (2006) Restoration of long-term potentiation in middle-aged hippocampus after induction of brain-derived neurotrophic factor. *J Neurophysiol* **96**, 677-685.

[162] Aarsland D, Marsh L, Schrag A (2009) Neuropsychiatric symptoms in Parkinson's disease. *Mov Disord* **24**, 2175-2186.

[163] Goren JL (2016) Brain-derived neurotrophic factor and schizophrenia. *Ment Health Clin* **6**, 285-288.

[164] Cao B, Bauer IE, Sharma AN, Mwangi B, Frazier T, Lavagnino L, Zunta-Soares GB, Walss-Bass C, Glahn DC, Kapczinski F, Nielsen DA, Soares JC (2016) Reduced hippocampus volume and memory performance in bipolar disorder patients carrying the BDNF val66met met allele. *J Affect Disord* **198**, 198-205.

[165] Monchi O, Stoessl AJ (2012) Imaging neural correlates of mild cognitive impairment in Parkinson’s disease. *Lancet Neurol* **11**, 653-655.

[166] Fahn S, Oakes D, Shoulson I, Kieburtz K, Rudolph A, Lang A, Olanow CW, Tanner C, Marek K, Group PS (2004) Levodopa and the progression of Parkinson's disease. *N Engl J Med* **351**, 2498-2508.

[167] Overk CR, Mufson EJ (2010) Dopamine Transporter: Aging and Parkinson's Disease In *Encyclopedia of Movement Disorders* Academic Press, pp. 330-332.

[168] Vandenbergh DJ, Persico AM, Hawkins AL, Griffin CA, Li X, Jabs EW (1992) Human dopamine transporter gene (DAT1) maps to chromosome 5p15.3 and displays a VNTR. *Genomics* **14**, 1104-1106.

[169] Fuke S, Suo S, Takahashi N, Koike H, Sasagawa N, Ishiura S (2001) The VNTR polymorphism of the human dopamine transporter (DAT1) gene affects gene expression. *the pharmacogenomics journal* **1**, 152-156.

[170] Vaughan RA, Foster JD (2013) Mechanisms of dopamine transporter regulation in normal and disease states. *Trends Pharmacol Sci.* **34**.

[171] Kaiser R, Hofer A, Grapengiesser A, Gasser T, Kupsch A, Roots I, Brockmöller J (2003) L - dopa-induced adverse effects in PD and dopamine transporter gene polymorphism. *Neurology* **11**, 1750-1755.

[172] Bialecka M, Kurzawski M, Roszmann A, Robowski P, Sitek EJ, Honczarenko K, Gorzkowska A, Budrewicz S, Mak M, Jarosz M, Golab-Janowska M, Koziorowska-Gawron

E, Drozdzik M, Slawek J (2012) Association of COMT, MTHFR, and SLC19A1(RFC-1) polymorphisms with homocysteine blood levels and cognitive impairment in Parkinson's disease. *Pharmacogenet Genomics* **22**, 716-724.

[173] Sampedro F, Marín-Lahoz J, Martínez-Horta S, Pagonabarraga J, Kulisevsky J (2020) Reduced gray matter volume in cognitively preserved COMT 158 Val/Val Parkinson's disease patients and its association with cognitive decline. *Brain Imaging Behav* **14**, 321-328.

[174] Senkevich K, Miliukhina I, Shadenkov V, Gracheva E, Maria B, Kulabukhova D, Emelyanov A, Pchelina S (2018) Associations of genetic variants in COMT, BDNF, SNCA, MAPT genes with cognitive impairment in Parkinson's disease. (P6.090). *Neurology* **90**.

[175] Ismail Z, McGirr A, Gill S, Hu S, Forkert ND, Smith EE (2021) Mild Behavioral Impairment and Subjective Cognitive Decline Predict Cognitive and Functional Decline. *Journal of Alzheimer's disease*, 459-469.

[176] Lang S, Yoon EJ, Kibreab M, Kathol I, Cheetham J, Hammer T, Sarna J, Ismail Z, Monchi O (2020) Mild behavioral impairment in Parkinson's disease is associated with altered corticostriatal connectivity. *Neuroimage Clin* **26**, 102252.

[177] Rouse HH, Small BJ, Schinka JA, Loewenstein DA, Duara R, Potter H (2020) Mild behavioral impairment as a predictor of cognitive functioning in older adults. *Int Psychogeriatr* **33**, 285-293.

[178] Hughes AJ, Daniel SE, Kilford L, Lees AJ (1992) Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* **55**, 181-184.

[179] Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, Giladi N, Holloway RG, Moore CG, Wenning GK, Yahr MD, Seidl L, Movement Disorder Society Task Force on Rating Scales for Parkinson's D (2004) Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord* **19**, 1020-1028.

[180] Shinohara M, Mizushima H, Hirano M, Shioe K, Nakazawa M, Hiejima Y, Ono Y, Kanba S (2004) Eating disorders with binge-eating behaviour are associated with the s allele of the 3'-UTR VNTR polymorphism of the dopamine transporter gene. *J Psychiatry Neurosci* **29**, 134-137.

[181] Ramezani M, Ruskey JA, Martens K, Kibreab M, Javer Z, Kathol I, Hammer T, Cheetham J, Leveille E, Martino D, Sarna JR, Gan-Or Z, Pfeffer G, Ismail Z, Monchi O (2020) Association Between BDNF Val66Met Polymorphism and Mild Behavioral Impairment in Patients With Parkinson's Disease. *Front Neurol* **11**, 587992.

[182] Strous RD, Nolan KA, Lapidus R, Diaz L, Saito T, Lachman HM (2003) Aggressive behavior in schizophrenia is associated with the low enzyme activity COMT polymorphism: a replication study. *Am J Med Genet B Neuropsychiatr Genet* **120B**, 29-34.

[183] Goldman D, Oroszi G, Ducci F (2005) The genetics of addictions: uncovering the genes. *Nature Reviews Genetics* **6**, 521-532.

[184] Mier D, Kirsch P, Meyer-Lindenberg A (2009) Neural substrates of pleiotropic action of genetic variation in COMT: a meta-analysis. *Mol. Psychiatry* **15**, 918-927.

[185] Poewe W (2006) The natural history of Parkinson's disease. *J Neurol* **253 Suppl 7**, VII2-6.

[186] Cormier F, Muellner J, Corvol JC (2013) Genetics of impulse control disorders in Parkinson's disease. *J Neural Transm (Vienna)* **120**, 665-671.

[187] Geissler JM, International Parkinson Disease Genomics Consortium members, Romanos M, Gerlach M, Berg D, Schulte C (2017) No genetic association between attention-deficit/hyperactivity disorder (ADHD) and Parkinson's disease in nine ADHD candidate SNPs. *Atten Defic Hyperact Disord* **9**, 121-127.

[188] Kraemmer J, Smith K, Weintraub D, Guillemot V, Nalls MA, Cormier-Dequaire F, Moszer I, Corvol JC (2016) Clinical-genetic model predicts incident impulse control disorders in Parkinson's disease. *J Neurol Neurosurg Psychiatry* **87**, 1106–1111.

[189] González-Giraldo Y, Camargo A, López-León S, Forero DA (2015) No Association of BDNF, COMT, MAOA, SLC6A3, and SLC6A4 Genes and Depressive Symptoms in a Sample of Healthy Colombian Subjects. *Depression research and treatment* **2015**.

[190] Aarsland D, Creese B, Politis M, Chaudhuri KR, Ffytche DH, Weintraub D, Ballard C (2017) Cognitive decline in Parkinson disease. *Nat Rev Neurol* **13**, 217-231.

[191] Moore DJ, West AB, Dawson VL, Dawson TM (2005) Molecular pathophysiology of Parkinson's disease. *Annu Rev Neurosci* **28**, 57-87.

[192] Spillantini MG, Schmidt ML, Lee VM-Y, Trojanowski JQ, Jakes R, Goedert M (1997) Alpha-synuclein in Lewy bodies. *Nature* **38**, 839-840.

[193] Uéda K, Fukushima H, Masliah E, Xia Y, Iwai A, Yoshimoto M, O'tero DA, Kondo J, Ihara Y, Saitoh T (1993) Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America.* **90**, 11282-11286.

[194] Ross OA, Braithwaite AT, Skipper LM, Kachergus J, Hulihan MM, Middleton FA, Nishioka K, Fuchs J, Gasser T, Maraganore DM, Adler CH, Larvor L, Chartier-Harlin MC, Nilsson C, Langston JW, Gwinn K, Hattori N, Farrer MJ (2008) Genomic investigation of alpha-synuclein multiplication and parkinsonism. *Ann Neurol* **63**, 743-750.

[195] Atik A, Stewart T, Zhang J (2016) Alpha-synuclein as a biomarker for Parkinson's disease *Brain Pathol* **26**, 410-418.

[196] Libbrecht MW, Noble WS (2015) Machine learning applications in genetics and genomics. *Nat Rev Genet* **16**, 321-332.

[197] Fuchs J, Nilsson C, Kachergus J, Munz M, Larsson EM, Schule B, Langston JW, Middleton FA, Ross OA, Hulihan M, Gasser T, Farrer MJ (2007) Phenotypic variation in a large Swedish pedigree due to SNCA duplication and triplication. *Neurology* **68**, 916-922.

[198] Beyer MK, Janvin CC, Larsen JP, Aarsland D (2007) Amagnetic resonance imaging study of patients with Parkinson's disease with mild cognitive impairment and dementia using voxelbased morphometry. *Journal of neurology* **78**, 254-259.

[199] Weintraub D, Doshi J, Koka D, Davatzikos C, Siderowf AD, Duda JE, Clark CM (2011) Neurodegeneration across stages of cognitive decline in Parkinson disease. *Archives of Neurology* **68**, 1562-1568.

[200] Weintraub D, Dietz N, Duda JE, Wolk DA, Doshi J, Xie SX, Siderowf A (2012) Alzheimer's disease pattern of brain atrophy predicts cognitive decline in Parkinson's disease. *Brain* **135**, 170-180.

[201] Somme J. H. G-EJC, Molano A., Tijero B., Lezcano E., Zarranz J.J. (2011) Initial neuropsychological impairments in patients with the E46K mutation of the a-synuclein gene (PARK 1). *J. Neurol. Sci.* **310**.

[202] Camicioli R, Gee M, Bouchard TP, Fisher NJ, Hanstock CC, Emery DJ, Martin WRW (2009) Voxel-based morphometry reveals extra-nigral atrophy patterns associated with dopamine refractory cognitive and motor impairment in Parkinsonism. *Parkinsonism & Related Disorders* **15**, 187-195.

[203] Ramezani M, Provost JS, Kathol I, Kibreab M, Yoon EJ, Cheetham J, Lang ST, Hammer T, Robichaud A, Alrazi T, Ruskey JA, Gan-Or Z, Martino D, Sarna JR, Monchi O (2019) Synuclein-alpha´s variant, rs 894280 shows association with memory and attention in PD patients. *MDS international Congress* **1725**.

[204] Urbanowicz RJ, Meeker M, La Cava W, Olson RS, Moorea JH (2019) Relief-Based Feature Selection: Introduction and Review. *J Biomed Inform*.

[205] Beretta L, Santaniello A (2011) Implementing relieff filters to extract meaningful features from genetic lifetime datasets. *Journal of biomedical informatics* **44**, 361-369.

[206] Eppstein MJ, Payne JL, White BC, Moore JH (2007) Genomic mining for complex disease traits with random chemistry. *Genetic Programming and Evolvable Machines* **8**, 395-411.

[207] McKinney BA, White BC, Grill DE, Li PW, Kennedy RB, Poland GA, Oberg AL (2013) Reliefseq: a gene-wise adaptive-k nearest-neighbor feature selection tool for finding gene-gene interactions and main effects in mrna-seq gene expression data. . *PLoS One* **8**, e81527.

[208] Lei H, Huang Z, Han T, Luo Q, Cai Y, Liu G, Leib B (2018) Joint regression and classification via relational regularization for Parkinson's disease diagnosis. *Technol Health Care* **26**, 19-30.

[209] Nilashi M, Ibrahim O, Ahani A (2016) Accuracy Improvement for Predicting Parkinson's Disease Progression. *nature* **6**.

[210] Reekes TH, Higginson CI, Ledbetter CR, Sathivadivel N, Zweig RM, Disbrow EA (2020) Sex specific cognitive differences in Parkinson disease. *NPJ Parkinsons Dis* **6**, 7.

[211] Cholerton B, Johnson CO, Fish B, Quinn JF, Chung KA, Peterson-Hiller AL, Rosenthal LS, Dawson TM, Albert MS, Hu SC, Mata IF, Leverenz JB, Poston KL, Montine TJ, Zabetian CP, Edwards KL (2018) Sex differences in progression to mild cognitive impairment and dementia in Parkinson's disease. *Parkinsonism Relat Disord* **50**, 29-36.

[212] Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9**, 97-113.

[213] Yust-Katz S, Tesler D, Treves TA, Melamed E, Djaldetti R (2008) Handedness as a predictor of side of onset of Parkinson's disease. *Parkinsonism Relat Disord* **14**, 633-635.

[214] van der Hoorn A, Burger H, Leenders KL, de Jong BM (2012) Handedness correlates with the dominant Parkinson side: a systematic review and meta-analysis. *Mov Disord* **27**, 206-210.

[215] Prasad S, Saini J, Yadav R, Pal PK (2018) Motor asymmetry and neuromelanin imaging: Concordance in Parkinson's disease. *Parkinsonism Relat Disord* **53**, 28-32.

[216] Claassen DO, McDonell KE, Donahue M, Rawal S, Wylie SA, Neimat JS, Kang H, Hedera P, Zald D, Landman B, Dawant B, Rane S (2016) Cortical asymmetry in Parkinson's disease: early susceptibility of the left hemisphere. *Brain Behav* **6**, e00573.

[217] Modestino EJ, Amenechi C, Reinhofer A, O'Toole P (2016) Side-of-onset of Parkinson's disease in relation to neuropsychological measures. *Brain Behav* **19**, e00590.

[218] Tomer R LB, Weiner WJ. (1993) Side of onset of motor symptoms influences cognition in Parkinson's disease. *Ann Neurol.* **34**, 579-584.

[219] Aarsland D. TE, Larsen, J.P.; et al (1996) Frequency of Dementia in Parkinson Disease. *Arch Neurol* **53**, 538-542.

[220] Bloem BR, Marinus J, Almeida Q, Dibble L, Nieuwboer A, Post B, Ruzicka E, Goetz C, Stebbins G, Martinez-Martin P, Schrag A, Committee MDSRS (2016) Measurement instruments to assess posture, gait, and balance in P arkinson's disease: Critique and recommendations. *Mov Disord* **31**, 1342-1355.

[221] Karr JE, Graham RB, Hofer SB, Muniz-Terrera G (2019) When does cognitive decline begin? A systematic review of change point studies on accelerated decline in cognitive and neurological outcomes preceding mild cognitive impairment, dementia, and death. *Psychol Aging* **33**, 195-218.

[222] Lin YS, Lee WJ, Wang SJ, Fuh JL (2018) Levels of plasma neurofilament light chain and cognitive function in patients with Alzheimer or Parkinson disease. *Sci Rep* **8**.

[223] Monchi O, Hanganu A, Bellec P (2016) Markers of cognitive decline in PD: The case for heterogeneity. *Parkinsonism Relat Disord* **24**, 8-14.

[224] Mateos-Perez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, Evans AC (2018) Structural neuroimaging as clinical predictor: A review of machine learning applications. *Neuroimage Clin* **20**, 506-522.

[225] Christopher L, Duff-Canning S, Koshimori Y, Segura B, Boileau I, Chen R, Lang AE, Houle S, Rusjan P, Strafella AP (2015) Salience network and parahippocampal dopamine dysfunction in memory-impaired Parkinson disease. *Ann Neurol* **77**, 269-280.

[226] Nguyen AA, Maia PD, Gao X, P FD, Raj A (2020) Dynamical Role of Pivotal Brain Regions in Parkinson Symptomatology Uncovered with Deep Learning. *Brain Sci* **10**.

[227] Zhang D, Wang, Y., Zhou, L., Yuan, H., Shen, D., Initiative, A.D.N., et al. (2011) Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* **55**, 856-867.

[228] Jia X, Wang Z, Yang T, Li Y, Shuai L, Wu GG, Jiang T, Liang P (2019) Entorhinal Cortex Atrophy in Early, Drug-naive Parkinson's Disease with Mild Cognitive Impairment. *Aging Dis* **10**, 1221-1232.

[229] Pagonabarraga J, Corcuera-Solano I, Vives-Gilabert Y, Llebaria G, García-Sánchez C, Pascual-Sedano B, Delfino M, Kulisevsky J, Gómez-Ansón B (2013) Pattern of regional cortical thinning associated with cognitive deterioration in Parkinson's disease. *PLoS One* **8**, e54980.

[230] Ekman U, Eriksson J, Forsgren L, Jakobson S, Riklund K, Nyberg L (2012) Functional brain activity and presynaptic dopamine uptake in patients with Parkinson's disease and mild cognitive impairment: a cross-sectional study. *Lancet Neurol* **8**, 679-687.

[231] Katako A, Shelton P, Goertzen AL, Levin D, Bybel B, Aljuaid M, Yoon HJ, Kang DY, Kim SM, Lee CS, Ko JH (2018) Machine learning identified an Alzheimer's disease-related FDG-PET pattern which is also expressed in Lewy body dementia and Parkinson's disease dementia. *Sci Rep* **8**, 13236.

[232] Monchi O, Petrides M, Doyon J, Postuma RB, Worsley K, Dagher A (2004) Neural bases of set-shifting deficits in Parkinson's disease. *J Neurosci* **24**, 702-710.

[233] Halliday GM, Leverenz JB, Schneider JS, Adler CH (2014) The Neurobiological Basis of Cognitive Impairment in Parkinson'S Disease. *Mov Disord* **29**, 634-650.

[234] Palumbo B, Fravolini ML, Buresta T, Pompili F, Forini N, Nigro P, Calabresi P, Tambasco N (2014) Diagnostic accuracy of Parkinson disease by support vector machine (SVM) analysis of 123I-FP-CIT brain SPECT data: implications of putaminal findings and age. *Medicine (Baltimore)* **93**, e228.

[235] Bras J, Guerreiro R, Darwent L, Parkkinen L, Ansorge O, Escott-Price V, Hernandez DG, Nalls MA, Clark LN, Honig LS, Marder K, Van Der Flier WM, Lemstra A, Scheltens P, Rogaeva E, St George-Hyslop P, Londos E, Zetterberg H, Ortega-Cubero S, Pastor P, Ferman TJ, Graff-Radford NR, Ross OA, Barber I, Braae A, Brown K, Morgan K, Maetzler W, Berg D, Troakes C, Al-Sarraj S, Lashley T, Compta Y, Revesz T, Lees A, Cairns N, Halliday GM, Mann D, Pickering-Brown S, Dickson DW, Singleton A, Hardy J (2014) Genetic analysis implicates APOE, SNCA and suggests lysosomal dysfunction in the etiology of dementia with Lewy bodies. *Hum Mol Genet* **23**, 6139-6146.

[236] Ng ASL, Tan YJ, Zhao Y, Saffari SE, Lu Z, Ng EYL, Ng SYE, Chia NSY, Setiawan F, Xu Z, Tay KY, Au WL, Tan LCS, Tan EK (2019) SNCA Rep1 promoter variability influences cognition in Parkinson's disease. *Mov Disord* **34**, 1232-1236.

[237] Mata IF, Johnson CO, Leverenz JB, Weintraub D, Trojanowski JQ, Van Deerlin VM, Ritz B, Rausch R, Factor SA, Wood-Siverio C, Quinn JF, Chung KA, Peterson-Hiller AL, Espay AJ, Revilla FJ, Devoto J, Yearout D, Hu SC, Cholerton BA, Montine TJ, Edwards KL, Zabetian CP (2017) Large-scale exploratory genetic analysis of cognitive impairment in Parkinson's disease. *Neurobiol Aging* **56**, 211 e211-211 e217.

[238] Pfeiffer HC, Lokkegaard A, Zoetmulder M, Friberg L, Werdelin L (2014) Cognitive impairment in early-stage non-demented Parkinson's disease patients. *Acta Neurol Scand* **129**, 307-318.

[239] Cammisuli DM, Crowe S (2018) Spatial disorientation and executive dysfunction in elderly nondemented patients with Parkinson's disease. *Neuropsychiatr Dis Treat* **14**, 2531-2539.

[240] Donolato E, Giofre D, Mammarella IC (2017) Differences in Verbal and Visuospatial Forward and Backward Order Recall: A Review of the Literature. *Front Psychol* **8**, 663.

[241] Yoon EJ, Ismail Z, Hanganu A, Kibreab M, Hammer T, Cheetham J, Kathol I, Sarna JR, Martino D, Furtado S, Monchi O (2019) Mild behavioral impairment is linked to worse cognition and brain atrophy in Parkinson disease. *Neurology* **98**.

[242] Kira K, Rendell LA (1992) The Feature Selection Problem: Traditional Methods and a New Algorithm. *AAAI-92 Proceedings*.

[243] Hernandez DG, Reed X, Singleton AB (2016) Genetics in Parkinson disease: Mendelian versus non-Mendelian inheritance. *J Neurochem* **139 Suppl 1**, 59-74.

[244] Ramezani M, Mouches P, Yoon E, Rajashekar D, Ruskey JA, Leveille E, Martens K, Kibreab M, Hammer T, Kathol I, Maarouf N, Sarna J, Martino D, Pfeffer G, Gan-Or Z, Forkert ND, Monchi O (2021) Investigating the relationship between the SNCA gene and cognitive abilities in idiopathic Parkinson's disease using machine learning. *Scientific Reports* **11**.

[245] Parkinson Progression Marker I (2011) The Parkinson Progression Marker Initiative (PPMI). *Prog Neurobiol* **95**, 629-635.

[246] The Lancet N (2010) Biomarker promise for Parkinson's disease. *The Lancet Neurology* **9**.

[247] Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA (1983) *Manual for the State-Trait Anxiety Inventory*, Consulting Psychologists Press, Palo Alto, CA.

[248] Jorgensen AW, Lundstrom LH, Wetterslev J, Astrup A, Gotzsche PC (2014) Comparison of results from different imputation techniques for missing data from an anti-obesity drug trial. *PLoS One* **9**, e111964.

[249] Chaudhri KR, Schapira AH (2009) Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment. *Lancet Neurol* **8**, 464-474.

[250] Schapira AH, Chaudhuri KR, Jenner P (2018) Non-motor Features of Parkinson's disease. *nature reviews neuroscience* **q8**, 435-450.

[251] Poletti M, Bonuccelli U (2013) Acute and chronic cognitive effects of levodopa and dopamine agonists on patients with Parkinson's disease: a review. *Ther Adv Psychopharmacol* **3**, 101-113.

[252] Konno T, Ross OA, Puschmann A, Dennis W. Dickson daZKW, * (2016) Autosomal dominant Parkinson's disease caused by SNCA duplications. *Parkinsonism Relat Disord* **22**, S1-S6.

[253] Guin D, Mishra MK, Talwar P, Rawat C, Kushwaha SS, Kukreti S, Kukreti R (2017) A systematic review and integrative approach to decode the common molecular link between levodopa response and Parkinson's disease. *BMC Med Genomics* **10**, 56.

[254] Bernal-Conde LD, Ramos-Acevedo R, Reyes-Hernandez MA, Balbuena-Olvera AJ, Morales-Moreno ID, Arguero-Sanchez R, Schule B, Guerra-Crespo M (2019) Alpha-Synuclein Physiology and Pathology: A Perspective on Cellular Structures and Organelles. *Front Neurosci* **13**, 1399.

[255] Iwaki H, Blauwendraat C, Leonard HL, Makarious MB, Kim JJ, Liu G, Maple-Grodem J, Corvol JC, Pihlstrom L, van Nimwegen M, Smolensky L, Amondikar N, Hutten SJ, Frasier M, Nguyen KH, Rick J, Eberly S, Faghri F, Auinger P, Scott KM, Wijeyekoon R, Van Deerlin VM, Hernandez DG, Gibbs RJ, Day-Williams AG, Brice A, Alves G, Noyce AJ, Tysnes OB, Evans JR, Breen DP, Estrada K, Wegel CE, Danjou F, Simon DK, Andreassen OA, Ravina B, Toft M, Heutink P, Bloem BR, Weintraub D, Barker RA, Williams-Gray CH, van de Warrenburg BP, Van Hilten JJ, Scherzer CR, Singleton AB, Nalls MA (2021) Differences in the Presentation and Progression of Parkinson's Disease by Sex. *Mov Disord* **36**, 106-117.

[256] Miller IN, Cronin-Golomb A (2010) GENDER DIFFERENCES IN PARKINSON'S DISEASE: CLINICAL CHARACTERISTICS AND COGNITION. *Mov Disord* **25**, 2695-2703.

[257] Zahodne LB, Marsiske M, Okun MS, Rodriguez RL, Malaty I, Bowers D (2012) Mood and motor trajectories in Parkinson's disease: Multivariate latent growth curve modeling. *Neuropsychology* **26**, 71-80.

[258] Burn DJ, Landau S, Hindle JV, Samuel M, Wilson KC, Hurt CS, Brown RG, Group P-PS (2012) Parkinson's disease motor subtypes and mood. *Mov Disord* **27**, 379-386.

[259] Pirogovsky-Turk E MR, Filoteo JV, Litvan I, Song DD, Lessig SL, et al. (2017) Neuropsychiatric Predictors of Cognitive Decline in Parkinson Disease: A Longitudinal Study. *Am J Geriatr Psychiatry* **25**, 279-289.

[260] Weintraub D, Stern MB (2005) Psychiatric Complications in Parkinson Disease. *The American Journal of Geriatric Psychiatry* **13**, 844-851.

[261] Parent A, Hazrati L (1995) Functional anatomy of the basal ganglia.The cortico-basal ganglia-thalamo-corticalloop. *Brain Research Reviews* **20**, 91-127.

[262] Williams-Gray CH, Foltynie T, Brayne CE, Robbins TW, Barker RA (2007) Evolution of cognitive dysfunction in an incident Parkinson's disease cohort. *Brain* **130**, 1787-1798.

[263] Greffard S, Verny M, Bonnet AM, Beinis JY, Gallinari C, Meaume S, Piette F, Hauw JJ, Duyckaerts C (2006) Motor Score of the Unified Parkinson Disease Rating Scale as a Good Predictor of Lewy Body–Associated Neuronal Loss in the Substantia Nigra. *JAMA Neurol* **63**, 584-588.

[264] Gaurav R, Yahia-Cherif L, Pyatigorskaya N, Mangone G, Biondetti E, Valabregue R, Ewenczyk C, Hutchison RM, Cedarbaum JM, Corvol JC, Vidailhet M, Lehericy S (2021) Longitudinal Changes in Neuromelanin MRI Signal in Parkinson's Disease: A Progression Marker. *Mov Disord*.

[265] Kulisevsky J, García-Sánchez C, Berthier ML, Barbanoj M, Pascual-Sedano B, Gironell A, Estévez-González A (2000) Chronic effects of dopaminergic replacement on cognitive function in Parkinson's disease: a two-year follow-up study of previously untreated patients. *Mov Disord* **15**, 613-626.

[266] Ikeda M, Kataoka H, Ueno S (2017) Can levodopa prevent cognitive decline in patients with Parkinson's disease? *Am J Neurodegener Dis* **6**, 9-14.

[267] McGuigan S, Zhou SH, Brosnan MB, Thyagarajan D, Bellgrove MA, Chong TTJ (2019) Dopamine restores cognitive motivation in Parkinson's disease. *Brain* **142**, 719-732.

[268] MacDonald PA, MacDonald AA, Seergobin KN, Tamjeedi R, Ganjavi H, Provost JS, Monchi O (2011) The effect of dopamine therapy on ventral and dorsal striatum-mediated cognition in Parkinson's disease: support from functional MRI. *Brain* **134**, 1447-1463.

[269] MacDonald AA, Seergobin KN, Tamjeedi R, Owen AM, Provost JS, Monchi O, Ganjavi H, MacDonald PA (2014) Examining dorsal striatum in cognitive effort using Parkinson's disease and fMRI. *Ann Clin Transl Neurol* **1**, 390-400.

[270] Poletti M, Bonuccelli U (2013) Acute and chronic cognitive effects of levodopa and dopamine agonists on patients with Parkinson's disease: a review. *Ther Adv Psychopharmacol* **3**, 101-113.

[271] Shimada H, Doi T, Lee S, Makizako H (2019) Reversible predictors of reversion from mild cognitive impairment to normal cognition: a 4-year longitudinal study. *Alzheimer's research and therapy* **24**.

[272] Aarsland D, Kurz MW (2010) The epidemiology of dementia associated with Parkinson disease. *J Neurol Sci* **289**, 18–22.

[273] Pierce SE, Booms A, Prahl J, van der Schans EJC, Tyson T, Coetzee GA (2020) Post-GWAS knowledge gap: the how, where, and when. *NPJ Parkinsons Dis* **6**.

[274] Hong EP, Park JW (2012) Sample size and statistical power calculation in genetic association studies. *Genomics Inform* **10**, 117-122.

[275] Howells DW, Porritt MJ, Wong JY, Batchelor PE, Kalnins R, Hughes AJ, Donnan GA (2000) Reduced BDNF mRNA expression in the Parkinson's disease substantia nigra. *Exp Neurol* **166**, 127-135.

[276] Cassidy CM, Zucca FA, Girgis RR, Baker SC, Weinstein JJ, Sharp ME, Bellei C, Valmadre A, Vanegas N, Kegeles LS, Brucato G, Kang UJ, Sulzer D, Zecca L, Abi-Dargham A, Horga G

(2019) Neuromelanin-sensitive MRI as a noninvasive proxy measure of dopamine function in the human brain. *Proc Natl Acad Sci U S A* **116**, 5108-5117.

[277] Biondetti E, Gaurav R, Yahia-Cherif L, Mangone G, Pyatigorskaya N, Valabregue R, Ewenczyk C, Hutchison M, Francois C, Corvol J, Vidailhet M, Lehericy S (2020) Spatiotemporal changes in substantia nigra neuromelanin content in Parkinson's disease. *Brain* **143**, 2575-2770.

[278] Leon AC, Heo M (2009) Sample Sizes Required to Detect Interactions between Two Binary Fixed-Effects in a Mixed-Effects Linear Regression Model. *Comput Stat Data Anal* **53**, 603-608.

[279] Angelucci F, Peppe A, Carlesimo GA, Serafini F, Zabberoni S, Barban F, Shofany J, Caltagirone C, Costa A (2015) A pilot study on the effect of cognitive training on BDNF serum levels in individuals with Parkinson's disease. *Front Hum Neurosci* **9**, 130.

9 Appendix

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Mar 15, 2021

This Agreement between Miss. Mehrafarin Ramezani ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4946861362749
License date	Nov 12, 2020
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Journal of Neurology
Licensed Content Title	Staging of the intracerebral inclusion body pathology associated with idiopathic Parkinson's disease (preclinical and clinical stages)
Licensed Content Author	Heiko Braak et al
Licensed Content Date	Jan 1, 2002
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic

Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation/distribution	30 - 99
Author of this Springer Nature content	no
Title	PhD dissertation
Institution name	University of Calgary
Expected presentation date	Mar 2021
Portions	Figure 3 on page 4
Requestor Location	Miss. Mehrafarin Ramezani 2954A-HSC, 3330 Hospital Drive, NW Calgary, AB T2N4N1 Canada Attn: University of Calgary
Total	0.00 CAD

Terms and Conditions

Springer Nature Customer Service Centre GmbH
Terms and Conditions

This agreement sets out the terms and conditions of the licence (the **Licence**) between you and **Springer Nature Customer Service Centre GmbH** (the **Licensor**). By clicking 'accept' and completing the transaction for the material (**Licensed Material**), you also confirm your acceptance of these terms and conditions.

1. Grant of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, world-wide licence to reproduce the Licensed Material for the purpose specified in your order only. Licences are granted for the specific use requested in the order and for no other use, subject to the conditions below.

1. 2. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

1. 3. If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Scope of Licence

2. 1. You may only use the Licensed Content in the manner and to the extent permitted by these Ts&Cs and any applicable laws.

2. 2. A separate licence may be required for any additional use of the Licensed Material, e.g. where a licence has been purchased for print only use, separate permission must be obtained for electronic re-use. Similarly, a licence is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence. Any content owned by third parties are expressly excluded from the licence.

2. 3. Similarly, rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

2. 4. Where permission has been granted **free of charge** for material in print, permission may also be granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

2. 5. An alternative scope of licence may apply to signatories of the STM Permissions Guidelines, as amended from time to time.

3. Duration of Licence

3. 1. A licence for is valid from the date of purchase ('Licence Date') at the end of the relevant period in the below table:

Scope of Licence	Duration of Licence
Post on a website	12 months
Presentations	12 months
Books and journals	Lifetime of the edition in the language purchased

4. Acknowledgement

4. 1. The Licensor's permission must be acknowledged next to the Licenced Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

5. Restrictions on use

5. 1. Use of the Licensed Material may be permitted for incidental promotional use and minor editing privileges e.g. minor adaptations of single figures, changes of format, colour and/or style where the adaptation is credited as set out in Appendix 1 below. Any other changes including but not limited to, cropping, adapting, omitting material that affect the meaning, intention or moral rights of the author are strictly prohibited.

5. 2. You must not use any Licensed Material as part of any design or trademark.

5. 3. Licensed Material may be used in Open Access Publications (OAP) before publication by Springer Nature, but any Licensed Material must be removed from OAP sites prior to final publication.

6. Ownership of Rights

6. 1. Licensed Material remains the property of either Licensor or the relevant third party and any rights not explicitly granted herein are expressly reserved.

7. Warranty

IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND
WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

8. Limitations

8. 1. ***BOOKS ONLY:*** Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

9. Termination and Cancellation

9. 1. Licences will expire after the period shown in Clause 3 (above).

9. 2. Licensee reserves the right to terminate the Licence in the event that payment is not received in full or if there has been a breach of this agreement by you.

Appendix 1 — Acknowledgements:

For Journal Content:
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g.

Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

For Adaptations/Translations:
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT]** (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

For Book content:
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

Other Conditions:

Version 1.2

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Mar 15, 2021

This Agreement between Miss. Mehrafarin Ramezani ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	4946880761904
License date	Nov 13, 2020
Licensed content publisher	Oxford University Press
Licensed content publication	Brain
Licensed content title	The distinct cognitive syndromes of Parkinson's disease: 5 year follow-up of the CamPaIGN cohort
Licensed content author	Williams-Gray, Caroline H.; Evans, Jonathan R.
Licensed content date	Oct 7, 2009
Type of Use	Thesis/Dissertation
Institution name	

Title of your work	PhD dissertation
Publisher of your work	University of Calgary
Expected publication date	Mar 2021
Permissions cost	0.00 CAD
Value added tax	0.00 CAD
Total	0.00 CAD
Title	PhD dissertation
Institution name	University of Calgary
Expected presentation date	Mar 2021
Portions	Fig2 on page 2964 Fig3 on page 2966
Requestor Location	Miss. Mehrafarin Ramezani 2954A-HSC, 3330 Hospital Drive, NW Calgary, AB T2N4N1 Canada Attn: University of Calgary
Publisher Tax ID	GB125506730
Total	0.00 CAD

Terms and Conditions

STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL

1. Use of the material is restricted to the type of use specified in your order details.

2. This permission covers the use of the material in the English language in the following territory: world. If you have requested additional permission to translate this material, the terms and conditions of this reuse will be set out in clause 12.

3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.

4. No alteration, omission or addition is made to the material without our written consent. Permission must be re-cleared with Oxford University Press if/when you decide to reprint.

5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society must be included in the credit line.

6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from www.oxfordjournals.org Should there be a problem clearing these rights, please contact journals.permissions@oup.com

7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.

8. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment

terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.

9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission.

10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employs and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

12. Other Terms and Conditions:

v1.4

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.