UNIVERSITY OF MOLISE
CAMPOBASSO

PhD School in Translational and Clinical Medicine

XXXII Cycle

Department of Medicine and Health Science

S.S.D. MED/50

Final Dissertation

"Investigating the relation of candidate genetic markers with Parkinson Disease, related endophenotypes and L-Dopa induced dyskinesia in an Italian cohort"

Tutor:
Prof. Daniela CARNEVALE
Dept. of Angio-Cardio-Neurology
IRCCS Neuromed
"Sapienza" University of Rome

Supervisor:
Dr Teresa ESPOSITO
Institute of Genetics and Biophysics
CNR, Naples

Coordinator:
Prof. Marco SARCHIAPONE
Dept. of Medicine and Health Science
University of Molise

Supervisor:
Dr Alessandro GIALLUISI
Dept. of Epidemiology and Prevention
IRCCS Neuromed

PhD Student:
Dr. Alfonsina TIROZZI
Matr. 157764

Academic Year 2018/2019

INDEX

ABSTRACT

INTRODUZIONE: il Morbo di Parkinson è una patologia neurodegenerativa caratterizzata da numerosi sintomi motori e non motori, che di solito vengono stimati attraverso diverse scale. La formazione di aggregati tossici della proteina α-sinucleina (codificata dal gene *SNCA*) è stata proposta come uno dei principali meccanismi molecolari alla base del Parkinson, e sembra che tale meccanismo dipenda anche dai livelli di espressione del gene *SNCA*. L'attuale trattamento è solo sintomatico e la Levodopa (L-Dopa) rimane il farmaco migliore, nonostante crei, in alcuni casi, gravi effetti collaterali, come movimenti involontari chiamati discinesie (indotte da L-Dopa, o LID). Poiché il Parkinson mostra un'estrema eterogeneità genetica, che è anche influenzata dal background genetico di ciascun soggetto, sono stati condotti studi sulle popolazioni di differenti etnie, in particolare per le varianti di suscettibilità più studiate. Tuttavia, pochi studi si sono focalizzati su fenotipi di tipo continuo correlati al Parkinson quali scale di sintomi neurologici, cognitivi e clinici, anche noti come endofenotipi. Analogamente, la genetica delle LID è in gran parte poco chiara e solo alcune varianti sono state testate in relazione al loro rischio incidente.

OBIETTIVI E METODI: Abbiamo studiato due varianti di suscettibilità del gene *SNCA* - rs356219 e D4S3481 - associate al livello di espressione del gene e al rischio Parkinson, in una coorte italiana di 472 pazienti e 518 controlli. Prima abbiamo testato la potenziale influenza di queste varianti sul rischio prevalente, attraverso test di associazione caso-controllo aggiustati per sesso. Quindi abbiamo testato, nei soggetti affetti, associazioni con scale motorie (UPDRS), cognitive (MoCA) e non motorie (NMS), e con l'età di insorgenza della patologia (AAO), che ne rappresentano un altro importante endofenotipo. Infine, abbiamo testato l'influenza di rs356219 e D4S3481 sul rischio di insorgenza di LID, attraverso regressioni di Cox (follow-up totale 17.434 persone-mese, tempo di follow-up mediano 49 mesi). Queste analisi sono state aggiustate tenendo in considerazione diverse covariate quali età, sesso, terapia con L-Dopa (stato ON/OFF e dosaggio) e ulteriori scale di stadiazione della malattia.

RISULTATI: Abbiamo riscontrato due associazioni nominalmente significative del microsatellite D4S3481, una per l'allele 261 con una minore età di insorgenza della malattia (β (SE) = -2.02 (1.00); p = 0.045) - trend non confermato per l'allele di rischio putativo 263 - e l'altra con il rischio incidente di LID, in cui i portatori dell'allele 263 mostrano un rischio ridotto di complicanze motorie (HR [CI] = 0,56 [0,32; 0,98], p = 0,04). Tali associazioni non risultano significative dopo correzione per test multipli. Non sono state osservate altre associazioni significative per nessuno dei modelli genetici alternativi testati.

DISCUSSIONE: Per la prima volta viene riportata un'associazione della variante D4S3481-261 bp con una minore età di insorgenza della malattia, e un effetto protettivo della nota variante di rischio D4S3481-263 bp contro l'insorgenza di LID, con i portatori dell'allele 263 che mostrano la metà del rischio rispetto ai non portatori. Nonostante l'assenza di una significatività statistica dopo correzione per test multipli, questo risultato potrebbe avere un impatto importante nella gestione del trattamento del PD e pertanto è necessario confermarlo in coorti Parkinson indipendenti di grandi dimensioni.

ABSTRACT

BACKGROUND: Parkinson's disease (PD) is a neurodegenerative disorder characterized by several motor and non-motor symptoms, which are usually evaluated trough different scales (see below). Toxic aggregates of α-synuclein (encoded by the *SNCA* gene) have been proposed as one of the main molecular mechanisms at the basis of PD, which seem to depend also on the levels of expression of the gene. Although current PD treatment is only symptomatic, Levodopa (L-Dopa) remains the therapeutic gold standard for PD, which however creates in some cases severe side effects like involuntary movements called L-Dopa induced Dyskinesias (LIDs). Since PD shows an extreme genetic heterogeneity, which is also influenced by different genetic backgrounds, population-specific studies are warranted, also for known PD susceptibility variants. Similarly, the genetic of LIDs is largely unclear, and only a few variants in candidate PD genes have been tested with relation to LID risk.

OBJECTIVE & METHODS: Here, we investigated two candidate *SNCA* susceptibility variants - rs356219 and D4S3481 - which have been linked with the level of expression of the gene and have been consistently associated with PD risk, in an Italian cohort (472 patients and 518 controls). First we tested the potential influence of these variants on PD prevalent risk, through crude case-control association tests adjusted for sex. Then we tested, within PD cases, associations with scales assessing motor (UPDRS), cognitive (MoCA) and non-motor symptoms (NMS), and on PD age-at-onset (AAO), which represent powerful PD endophenotypes. Finally, we tested the influence of rs356219 and D4S3481 on the incident risk of LIDs, through multivariable Cox PH regressions (total follow-up 17,434 person-months, median follow-up time 49 months). These analyses were adjusted for an extended panel of covariates which may influence the outcome, including age, sex, L-Dopa therapy (status and dosage), and additional PD staging scales, where appropriate.

RESULTS: We observed a nominally significant association of D4S3481 with incident risk of LIDs, where carriers of the 263 (putative risk) allele showed a decreased risk (HR [CI] = 0.56 [0.32; 0.98], p = 0.04) of motor complications. Another nominally significant association was observed with AAO for D4S3481-261 bp allele vs 259 bp allele carriers (β (SE) = -2.02 (1.00); p = 0.045) in a pseudo-additive model, where however we did not observe any evidence of association for 263 vs 259 bp allele carriers. Both these associations did not survive correction for multiple testing. No other significant associations were observed for any of the alternative genetic models tested, neither in the case-control test nor in the analysis of continuous PD endophenotypes.

DISCUSSION: Here we report for the first time an association of D4S3481-261 bp variant with earlier age at PD onset, and a protective effect of the known PD risk D4S3481-263 bp variant against motor side-effects of L-Dopa treatment, with 263 carriers showing half the risk of non-carriers. Since this aspect has never been investigated before for D4S3481 and we observed only a nominally significant association, further studies in large independent PD cohorts are warranted to clarify the potential influence of this marker on LID susceptibility.

Chapter 1

INTRODUCTION

Parkinson Disease (PD)

Idiopathic Parkinson's Disease (PD) is the second most common neurodegenerative disorder after Alzheimer's disease (1). Despite almost 200 years since James Parkinson first described the disease, the exact mechanisms underlying this condition remain unclear (2).

PD is a progressive disorder characterized by dopaminergic cell degeneration in the substantia nigra pars compacta and is associated with intracytoplasmic Lewy body inclusions (3).

PD affects about 1% of people above 60 years of age and 4% of adults over 80 years (4), with increased prevalence in advancing age (5).

PD is characterized by several motor symptoms such as resting tremor, rigidity, bradykinesia and postural instability (6), and also non-motor symptoms (NMSs) such as depression, dementia, rapid eye movement, sleep behaviour disorder and anosmia, among others (7). Motor symptoms result from the degeneration of dopaminergic neurons in the midbrain substantia nigra, whereas NMSs are thought to result from the dysfunction of the serotonergic, cholinergic, and catecholaminergic systems (8). Based on these clinical and pathological findings, PD is recognized as a disease involving multiple systems and neurotransmitters (9).

In spite of the increasing knowledge of PD mechanisms, so far only symptomatic treatments have been discovered, either through pharmacological therapy or electrostimulation (7). Among pharmacological treatments, the most used active principle is Levodopa (L-Dopa; l-3,4-dihydroxyphenylalanine), a metabolic precursor of dopamine which is considered a gold standard in the field. L-Dopa is actively absorbed in the upper small intestine, and transported across the intestinal mucosa and blood–brain barrier (BBB). Once absorbed, it is converted into dopamine by aromatic amino acid decarboxylase (AADC) and metabolized to 3-O-methyldopa (3-OMD) by catechol-O-methyltransferase (COMT). Inhibitors of AADC (carbidopa or benserazide) and COMT are co-administered with L-Dopa to suppress the peripheral degradation of dopamine. This is done in order to reduce the exogenous dose of L-Dopa by maximizing the amount of the substance transported across the BBB, and to reduce adverse effects of peripheral dopamine, such as nausea and hypotension (10). Unfortunately, long-term L-Dopa treatment and over-dosage cause important side effects like L-Dopa induced Dyskinesias (LIDs). This motor complication - characterized by involuntary movements throughout the body - represent an important source of disability and notably worsens patients' quality of life.

Since this dissertation mainly focuses on investigating the genetic underpinnings of PD, of related neurological and clinical endophenotypes, and of side effects of its pharmacological treatment (i.e LIDs), we briefly review these aspects below.

Genetics of PD

PD tends to recur in families and is moderately heritable, with about 60% of its variance being explained by genetic factors (11), and is characterized by a complex architecture, with a number of genetic and environmental factors influencing susceptibility to the disease (12). It shows an extreme genetic heterogeneity, with 10% of PD cases having Mendelian inheritance (13,14). The genes which have been most robustly implicated in Mendelian forms of PD include:

SNCA (4q22.1; *α-synuclein*), encoding α-synuclein, a neuronal protein that plays several roles in synaptic activity, such as regulation of synaptic vesicle, trafficking and subsequent neurotransmitter release. Mutations have emerged as a rare, but important cause of PD with high penetrance (15). Since most of this dissertation focuses on the analysis of *SNCA* variants, this gene is reviewed more in details below.

*LRRK2 (*12q12*; Leucine Rich Repeat Kinase 2)*, encoding a leucine rich repeat kinase 2 containing multiple functional domains. *LRRK2* has been implicated in several autosomal dominant forms of PD, where several mutations have been identified (reviewed (16)), which make a large contribution towards both sporadic and familial forms of PD (17). Different studies have repeatedly shown linkage of PD risk to *LRRK2*, and a meta-analysis indicated *LRRK2* as one of the most important genomic loci influencing PD risk (18).

PARK2 (Parkin, 6q26*)*, encodes Parkin, an E3 ubiquitin ligase protein, but the mechanism of its pathogenicity remains unclear. Point mutations in this gene are mostly transmitted from common founders (19). These mutations are involved in development of Parkinson's disease probably by a loss-of-function mechanism (20). Patients with Parkinson's disease and Parkin mutations have a mean age at onset of 32 years in the Caucasian population (21). Hence, Parkin mutations are the most common cause of early-onset Parkinson's disease, occurring in up to 50% of those with age at onset under 25 years (and only 3%–7% in those with age at onset 30–45 years) (15).

ATP13A2 (Cation-transporting ATPase 13A2, 1p36*)*, encodes an ATPase that plays a role in intracellular cation homeostasis and in the maintenance of neuronal integrity (22). It is required for a proper lysosomal and mitochondrial maintenance (23,24), where it regulates the autophagy-lysosome pathway through the control of *SYT11* expression, both at the transcriptional and at the

post-translational levels (25). Mutations in *ATP13A2* cause also autosomal recessive parkinsonism with a complex phenotype (15).

PINK1 (*Serine/threonine-protein kinase*, 1p36.12) codes for a serin/threonine kinase localized to the mitochondria. Mutations in PINK1 are a rare cause of early-onset PD, accounting only for 2%–4% of early-onset cases in Caucasian populations (26,27) and 4%–9% in Asian populations (28,29). The penetrance for homozygous and compound heterozygous mutation carriers seems to be 100% but the specific mechanism of pathogenicity in PD is unclear and require further investigations (15).

DJ-1 (*Protein/nucleic acid deglycase DJ-1*, 1p36.23), also known as *PARK7* since it encodes Parkinson disease protein 7. Mutations in *DJ-1* cause autosomal recessive PD. Its product inhibits the aggregation of α-synuclein via its chaperone activity, (30,31) acting as a redox-sensitive chaperone protein and as a sensor for oxidative stress (15).

VPS35 (*VPS35 endosomal protein sorting factor-like*, 16q11.2) this gene belongs to a group of vacuolar protein sorting (VPS) genes. The encoded protein is a component of a large multimeric complex, termed the retromer complex, involved in retrograde transport of proteins from endosomes to the trans-Golgi network. Mutations in this gene cause an autosomal dominant, adult-onset form of the disorder. It is phenotypically similar to idiopathic PD (32).

DNAJC13 (*DnaJ heat shock protein family (Hsp40) member C13*, 3q22.1) is involved in membrane trafficking through early endosomes. In fact, it is implicated in the transport and recycling of transferrin and in the transport and degradation of endosomal growth factors from early endosome to late endosome (33). A novel mutation in this gene (p.Asn855Ser) was found to segregate with PD (34).

GBA (*Glucosylceramidase Beta*, 1q22) (35) encodes the lysosomal glucocerebrosidase enzyme, which cleaves the β-glucosyl. Proposed gain-of-function mechanisms include facilitation of α-synuclein accumulation perhaps loss-of-function mechanisms include substrate accumulation (35).

These genes are extensively reviewed in (13,36,37). Although other chromosomal loci - including *PARK3*, *PARK10*, *PARK11* and others (13) - have been identified, and these regions might contain further genes for typical, late-onset PD (13), we do not review them here since these have been not robustly implicated in PD as the candidate loci mentioned above. In these and other genes, rare mutations with both dominant (12,14) or recessive inheritance modes (38,39) have been identified, often through genome-wide linkage studies followed by targeted genotyping (e.g. 14) or, more

recently, through Next Generation Sequencing (NGS) studies (e.g. (40,41)). In addition to rare mutations, also common susceptibility variants like Single Nucleotide Polymorphisms (SNPs) have been detected within these genes, e.g. in *LRRK2* and *SNCA* (13). However, the genetic variants identified so far – be they common or rare - explain only a minor part of PD heritability (34), and for a large majority of PD cases the genetic diagnosis remains unresolved. The issue of missing heritability has been tackled through different approaches, including Genome Wide Association Scans (GWAS) to identify common variants with moderate/weak effect sizes on PD susceptibility (e.g. (42)), and NGS (mostly Whole Exome Sequencing) studies to identify rare causative mutations (e.g. (12,14,43–47)). Moreover, the genetic architecture and the mutational spectrum of PD can vary based on the ethnic and genetic background of the population (46,48) hence population-specific genetic studies are warranted (as in (43,46)).

Large-scale genomic studies carried out so far have scarcely investigated inter-individual variation in PD endophenotypes like neurological scales (12,14,42–47,49).

A GWAS study of age-at-onset in 25,568 PD cases reported two genome-wide significant associations within *SNCA* and *TMEM175* (50), while other preliminary GWAS of cognitive performance and motor symptoms progression are ongoing (51,52). Other SNP-based genomic studies tested associations of Polygenic Risk Scores (PRS) for PD with alpha-synuclein levels in the cerebrospinal fluid, age-at-onset of the disease, motor/cognitive symptoms and PD status (as reviewed in (53), detecting significant associations with PD risk (54), earlier PD onset (54,55), and faster motor and cognitive decline (56). With regard to large scale Next Generation Sequencing (NGS) studies, several Whole Exome Sequencing (WES) but no Whole Genome Sequencing (WGS) studies have been carried out so far on PD (12,14,18,44–47,57).These mostly focused on PD case-control analysis, but failed to find robust statistical evidence of association, probably due to the small sample size - compared to the huge genetic heterogeneity of the disease - and to the difficulty in recruiting proper neurological controls (i.e. people free of disease at an advanced age). Among these, our group attempted to identify genetic variants associated with continuous scales associated with PD (or *PD endophenotypes*, see below), assessing motor, cognitive and non-motor PD symptoms, but found no statistically significant associations (57).

On the other hand, association with specific scales related to PD has been more often tested for genetic variants in candidate PD susceptibility genes. Loss-of-function *GBA* mutations have been associated with a distinct cognitive profile characterized by greater impairment in working memory/executive function and visuospatial abilities in PD patients (58). PD cases carrying variants in *PARK16* - another gene implicated in PD (59) - exhibited greater motor progression

after 5 years of disease compared with non-carriers, based on assessment through Hoehn & Yahr (HY) staging scale, UPDRS motor score and UPDRS sub-scores (see below for details on these scales) (60). The common variant rs356182 in *SNCA* has been associated with a more tremor-predominant phenotype and predicted a slower rate of motor progression (61), while rs11931074 showed an association with worse motor symptoms (62). PD patients carrying rare variants in the *APP*, *PSEN1*, *PSEN2*, and *GRN* genes exhibit lower cognitive tests scores than non-carriers, regardless of age at PD diagnosis, age at evaluation, *APOE* status or recruitment site (63).

One of the most investigate genes in relation to PD endophenotypes is by far *SNCA*, the first PD locus identified (64). Since this dissertation focuses on the investigation of SNCA variants, we review below this gene and its implication in PD.

SNCA

SNCA (4q22.1) was the first gene identified as associated with idiopathic PD (38,64). Linkage analysis study of a large Italian kindred with autosomal dominant PD form revealed a locus at 4q22.1-q22.3 associated to the disease (64). This was further refined through the identification of a causative mutation in the *SNCA* gene (Ala53Thr), in the same Italian pedigree and in three unrelated dominant families of Greek origin (65). Since then, several studies have examined *SNCA* in relation to PD risk and its endophenotypes (reviewed in (66)).

SNCA (**Figure 1.1**) gene encodes for alpha-synuclein (α-syn) protein, a member of the synuclein family, which also includes beta- and gamma-synuclein. Synucleins are abundantly expressed in the brain, and alpha- and beta-synuclein inhibit phospholipase D2 selectively (67). α-syn plays a fundamental role in the molecular pathogenesis of PD, forming toxic oligomers and aggregates within neurons (68), acting in a prion-like manner. These aggregations ultimately result in Lewy bodies, which represent the histopathological hallmark lesions of PD (69). Similarly, *SNCA* has been implicated in another neurological disorder highly comorbid with PD, with a partly shared etiopathological mechanism, namely Dementia with Lewy Bodies (70). α-syn peptides are also a major component of amyloid plaques in the brains of patients with Alzheimer's Disease (71). In physiological conditions, neuronal α-syn protein plays several roles in synaptic activity, such as regulation of synaptic vesicle trafficking and subsequent neurotransmitter release (72). It also participates as a monomer in synaptic vesicle exocytosis by enhancing vesicle priming, fusion and dilation of exocytotic fusion pores (73). Mechanistically, α-syn acts by increasing local Ca^{2+} release from micro-domains, which is essential for the enhancement of ATP-induced exocytosis (73). It also acts as a molecular chaperone in its multimeric membrane-bound state, assisting in

the folding of synaptic fusion components called SNAREs (Soluble NSF Attachment Protein Receptors) at presynaptic plasma membrane, in conjunction with cysteine string protein-alpha (74). This chaperone activity is important to sustain normal SNARE-complex assembly during aging (74). *SNCA* plays also a role in the regulation of the dopamine neurotransmission in association with the dopamine transporter (*DAT1*) and thereby modulating its activity (73).

α−synuclein is currently seen as one of the most promising targets of disease-modifying therapies for PD (37), which is why investigating in detail the genetic risk/protection conferred by its genetic variants is more and more important. Studies carried out so far supported an influence of polymorphisms in multiple regions of *SNCA* gene, such as the promoter (5') region (REP1-SNCA), 3′ end (e.g., rs11931074 and rs356219), 3′ untranslated regions (e.g., rs356165), and introns (e.g., rs7684318, rs894278, and rs276990) (as reviewed in (75)). Among these variants, increasing attention have received specific variants which have been reported to alter *SNCA* gene expression levels (75–78), which is considered one of the main mechanisms through which α-syn causes PD (79–82).

One of these variants is represented by D4S3481 (commonly known as REP1), a complex polymorphic microsatellite (dinucleotide) repeat located ~10 kb upstream of the translation start site of *SNCA* (83). A microsatellite is a tract of repetitive DNA in which certain DNA motifs (ranging in length from one to ten base pairs) are repeated, typically 5 to 50 times. Several small studies have suggested that certain alleles of a dinucleotide repeat sequence (REP1) of the *SNCA* promoter might be associated with the risk of developing PD (84). SNCA-REP1 is essentially triallelic (259, 261, and 263 bp in length) (85). The 259/259 bp genotype has been associated with a decreased levels of expression of α-syn in the blood, compared to genotypes 261/261, 259/261, and 259/263 (78). A down-regulation effect of the 259 bp variant on *SNCA* gene expression has been supported also by functional analyses (86–89).

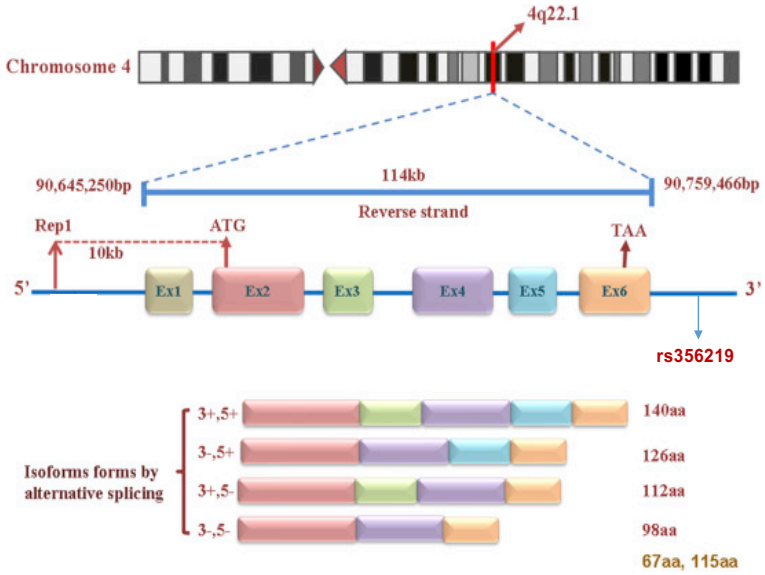
A meta-analysis of association studies showed higher frequency of 263 bp allele in cases compared to controls (90). Conversely, the 259 bp allele was found to be associated with a decreased risk of PD, while no effect was observed for the 261 bp allele (85). These alleles have been also associated with continuous PD-related traits, although not always consistently. PD patients carrying at least one 263bp allele in SNCA-REP1 exhibited four-fold higher odds of fast disease progression compared to non-carriers (91), and the 263 allele was also associated with a worse cognitive outcome in PD. (85). Conversely, REP1-259 allele was also associated with the development of worse motor outcomes (92). As opposed to these lines of evidence, other studies have reported an association of REP1-263 allele with better motor and cognitive outcomes or no association, as in

Ritz et al, 2012. In the latter study, authors observed no association between SNCA-REP1-259 allele and motor symptom progression under a dominant genetic model, although the risk was in the expected (‘‘protective’’) direction (91).

Another variant which has been robustly associated with changes in the level of expression of *SNCA* is the single-nucleotide polymorphism (SNP) rs356219, which lays in the 3′ region of the gene (**Figure 1.1**). This SNP is probably the most investigated common variant in *SNCA*, and it stands out as a consistent risk factor for PD in several studies (as reviewed in (75)). Moreover, this variant has showed a significant effect on *SNCA* mRNA levels in the substantia nigra and in the cerebellum (78), and was shown to affect also the blood plasma levels of α-syn (93). This evidence is in line with independent transcriptomic analyses which revealed a positive association between the count of the rs356219-G allele and the level of expression of specific *SNCA* isoforms, assessed through quantitative Polymerase Chain Reaction (qPCR), RNA-sequencing (RNA-seq) and cap analysis of gene expression (CAGE-seq) in postmortem frontal cortex tissues of neurologically healthy subjects (94).

In a two-tiered analysis of 1,956 patients with PD and 2,112 controls on 15 candidate SNPs within *SNCA*, rs356219 showed the most significant association among all variants tested, which was larger than and independent of the REP1 marker (95). Author suggested that this effect on increased PD susceptibility might by mediated by an upregulation of *SNCA* expression in a dose-dependent manner (95). In a later meta-analysis of 18 PD case-control observational studies focused on rs356219, (86,96–98), a significant association with PD risk was found in Caucasian populations, showing an increased risk by ~26% and ~38% in the dominant and recessive models, respectively (96). This finding was later supported in a PD case-control GWAS, where rs356219 was detected as a genome-wide significant hit (99), and in candidate variant studies of different genetic ancestries, such as Chinese Han (100) and South-Americans (96). In the latter study, rs356219-G allele was associated with an increased risk for cognitive impairment in Brazilian PD patients (96). Of note, rs356219 was also shown to significantly contribute to other PD continuous endophenotypes, such as an earlier age at onset of the disease (101).

Figure 1.1: SNCA gene structure and protein isoforms generated by alternative splicing.



Modified by: The link between the SNCA gene and parkinsonism Wei Xu et al. 2015

PD endophenotypes

Endophenotypes are measurable components (e.g., neurophysiological, biochemical, neuroanatomical, cognitive or neuropsychological) that exist between the behavioural symptoms of a disease and a distal genotype (102).

The purpose of the endophenotype concept is to divide symptoms and signs of a disease into more stable phenotypes with a clear genetic connection. The rationale at its basis is that a smaller number of genes will be associated with a less complex phenotype than a complex disorder, increasing the power to detect genetic associations with the endophenotype and, indirectly, with the disease of interest.

An ideal endophenotype should meet the following criteria:

- association with the disease in the population;
- heritability;
- primary state-independence (i.e. it should be seen in individuals with and without the active illness/diagnosis).
- co-segregation with the disease within families;

- familial clustering (i.e. it should present in both patients and their unaffected relatives at a higher rate than in the general population or, alternatively, should show intermediate values between probands and the general population, in case of continuous traits).

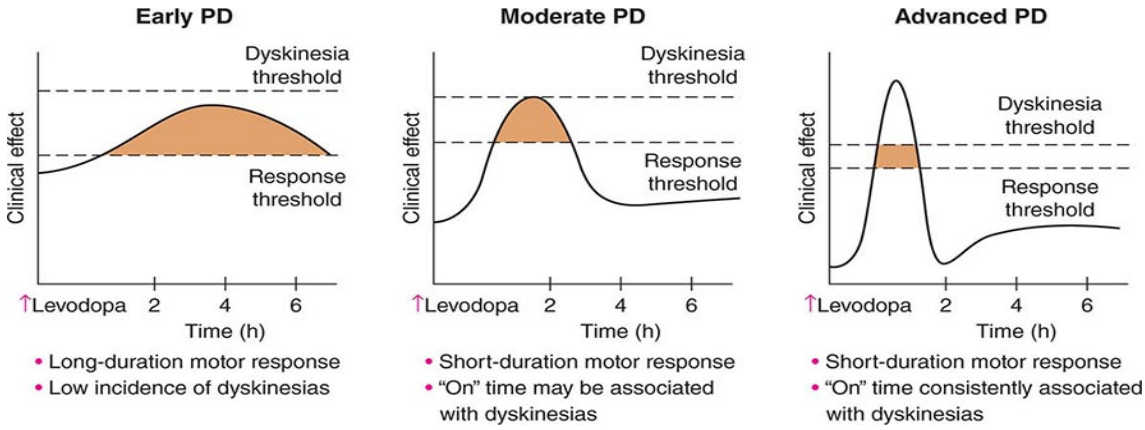
In PD, several continuous phenotypes (also known as "traits") can be considered as endophenotypes, e.g. motor symptoms, cognitive performance, depression, and age-at-onset. Currently, these components are notably under investigated in relation with PD patients' genetic profile (57). For specific PD candidate genes like *SNCA*, the influence of common genetic variants within or close to the gene is related to different aspects of PD phenotypic spectrum, as we briefly reviewed above. More in general, the study of common variants may provide valuable insights into the mechanisms underlying heterogeneity in PD (103).

Dyskinesias

Dyskinesias represent "*Abnormal involuntary movements attributed to pathologic state of one or more parts of the striate body and characterized by insuppressible, stereotyped, automatic movements that cease only during sleep.*" (The American Heritage® Stedman's Medical Dictionary). Schoenecker recorded the first clear description of clinical dyskinesia in 1957. The term "tardive dyskinesia" was coined to indicate abnormal movements induced by neuroleptics (104). In PD, dyskinesia was recognized with the advent of L-Dopa, and since then "levodopa-induced dyskinesias" (LIDs) (**Figure 1.2**) has become one of the major clinical limitations of the long-term treatment of PD. By the late 1970s, several classifications of LIDs have been proposed based on the type of movements, the timing of L-Dopa dosage and combinations of the two factors:

1. Peak-dose dyskinesia: dyskinesia noted at the peak clinical benefit of L-Dopa;
2. Dystonia-improvement-dystonia (105), or diphasic dyskinesia (106): appearing at beginning and at the end of each L-Dopa dose;
3. OFF dystonia: dystonia occurring early in the morning, when the effect of previous night's dose of L-Dopa had completely worn off.

Figure1.2: Changes in motor response associated with chronic levodopa treatment.



Levodopa-induced motor complications. Here we report a schematic illustration of the gradual shortening of the duration of a beneficial motor response to L-Dopa treatment (wearing off), and the appearance of dyskinesias as this time range ("on" time) shortens. Image courtesy of Harrison's Neurology in Clinical Medicine, 3rd Edition. C. Warren Olanow Image Anthony H.V. Schapira

LIDs comprise a variety of phenomena, the most common of which are chorea, choreo-athetosis, and dystonia. Chorea is the most common form of LID and it is most commonly associated with peak dose dyskinesia (107). Dystonia is the second most common form of LID, while ballism is characterized by abnormal choreic movements of the proximal parts of the limbs causing flinging movements, which can be unilateral or bilateral (107). Myoclonus, a sudden brief shock-like involuntary movement, is rarely classified as a part of LIDs (107). Other LID movements include respiratory dyskinesia (108,109), ocular dyskinesia (110), restlessness/hyperactivity, akathisia and enhanced tremor (107). The rate of LID development ranges between 3 and 94% among PD patients, depending on different factors which mainly include PD age-at-onset, disease duration, severity, and duration of L-Dopa therapy (3) (see *Risk factors of LIDs* subsection below for details).

Hypothesized mechanisms of LIDs

The aetiology of LIDs is largely unknown yet. With the reduction of dopamine in PD patients, it is believed that hypersensitization of the dopamine receptor contributes to the development of LIDs (111). The short half-life of L-Dopa and pulsatile release of dopamine, once the buffering capacity of dopamine transporter has been lost, is considered to be one of the major mechanism

generating LIDs (112). The use of extended-release carbidopa/levodopa and continuous intrajejunal infusion of carbidopa/levodopa intestinal gel has been reported to improve motor symptoms and motor fluctuations, without aggravating dyskinesia when compared to standard L-Dopa (113–115). Recently, it has been reported that carbidopa/levodopa intestinal gel infusion cause dyskinesias, including diphasic dyskinesia (116).

Compared to L-Dopa, dopamine agonists cause less dyskinesia, given that they have longer half-life (117–119). The use of a dopamine agonist in early stage PD patients to delay the use of L-Dopa is considered to be clinically effective, and to successfully postpone the occurrence of LIDs (120).

Risk factors of LIDs

In addition to the use of L-dopa rather than other dopamine agonists for the treatment of PD (see Sharma et al., 2010 for a review (121)), many other risk factors have been associated with the onset of LIDs. Some of them are modifiable, like L-Dopa dose and body weight, while others are non-modifiable, like age, sex, PD age at onset, duration of disease, clinical subtype, disease progression, disease severity, and genetic factors (which we review in the next subsection) (121).

PD age at onset (AAO) represents one of the main risk factors for dyskinesias (122). The younger is AAO, the more likely is the development of LIDs (123–125). A 5-year follow-up study of PD patients showed a prevalence of LIDs up to 50% at age 40–59, and 16% after 70 years (126). Another study found that after 5 years of L-Dopa treatment, the rates of dyskinesia in patients with PD onset at 40–49, 50–59, 60–69, and 70–79 years were 70%, 42%, 33%, and 24%, respectively (123). Furthermore, patients with AAO < 40 years (young-onset PD) had a higher incidence of LIDs than those with late-onset PD (AAO ≥ 50 years) (125). In line with this evidence, patients with longer duration of disease - which is connected with AAO - are more likely to develop LIDs (127). Of note, age per se has been detected as a risk factor in a single cross-sectional study, which reported a positive correlation between patients' age and time from onset to development of motor complications (128).

Sex represents another important risk factor for LIDs, with women showing greater incidence of dyskinesias than men (124,129). Moreover, women develop dyskinesias earlier in relation to time of L-Dopa administration, compared to men (130). This may be due to the fact that women have less "genetic protection" related to lower expression of dopamine receptor *DRD2*, which seems to exert a protective role against dyskinesia in men (131). An alternative explanation may be the

higher bio-availability of L-dopa in women, due to their lower body weight (121). Of note, the higher LID risk conferred by sex was not confirmed in another study (96), and in a multivariate analysis including additional risk factors (AAO and L-Dopa dosage) (132).

As mentioned above, low body weight and a resulting higher bioavailability of L-Dopa is a prominent risk factor for LIDs (124,132,133), which may be also easily explainable from a biological point of view. Indeed, several studies have proposed that the increased risk of motor complications in PD patients with lower body weight may be due to elevated peripheral L-Dopa levels in these patients ((133); Group 1996).

Clinical subtypes of PD are also an important risk factor for LIDs. An observational study of 144 L-Dopa-treated patients showed that the tremor dominant subgroup had lower rates of dyskinesia (29%) compared to the bradykinesia dominant subgroup (69%) (Friedman 1985). Similarly, in another study, resting tremor subtype was associated with lower risk of developing LIDs than other initial manifestations (134). Of note, resting tremor subtype is considered to be independent on all the other known risk factors for LIDs, for the occurrence of motor complications (135).

In PD patients in the early stage of the disease (HY score 1), the time from the beginning of L-Dopa treatment to the occurrence of dyskinesias was 66 months, while, in "stage 3" patients (i.e. with HY score 3), it was only 24 months (136). Similarly, a recent analysis of Chinese PD patients revealed a positive association of prevalent LID risk with low UPDRS-III and high HY scores in ON-state (i.e. under L-Dopa treatment), which indicated severity of motor symptoms and progression of the disease, in addition to early AAO, long disease duration, female sex, and high L-Dopa equivalent dose (137). Of note, the emergence of dyskinesia had no association with the initiation time of L-Dopa (137). A community-based study of L-Dopa-related motor complications in PD found that the overall dose of L-Dopa was the most important predictor of motor fluctuation, with the dose and treatment having the strongest impact on LID prevalence (138). The recommended initial dose - less than 400 mg per day - helps to reduce the risk of motor complications (124,139). These studies suggest that L-Dopa dosage may be more important than the duration of treatment.

In other words, the higher the dose, the greater the risk of dyskinesia (134).

In addition to classical risk factors, more recently functional imaging has also been used to find predictors of LIDs (reviewed in (140)). E.g., a research showed that pre-synaptic dopamine deficiency assessed through PET scanning in 127 drug-naive de novo patients with PD predicted the risk of LIDs (141). Given the focus of the present dissertation, here we do not extensively review these works.

Genetics of LIDs

The development of LIDs reflects a profound reorganization of the neural circuit and balance between different pathways in the basal ganglia (142). LIDs are determined in part by genetic factors with multiple polymorphisms in various candidate genes. PD patients show a remarkable heterogeneity in their response to L-Dopa and this likely suggests that there is a certain genetic predisposition. However, if and how the inherited predisposition to PD affects the development of LIDs is currently an unanswered and largely under-investigated issue, both in candidate gene studies and in genome or exome-wide studies with no a priori hypotheses. We briefly review below the different genes which have been studied in relation to LID onset, and the genetic influences identified so far.

Dopamine receptors

Dopamine exerts its physiological action through the activation of dopamine receptors (DRD1–DRD5), which can be divided into D1-like receptors (DRD1 and DRD5), and D2-like receptors (DRD2, DRD3, and DRD4) (143). Normally, dopamine triggers an excitatory response on direct pathways through D1-like receptors, and an inhibitory response on the indirect pathway through D2-like receptors (144). In PD, which is characterized by the loss of dopamine, usually underactivity of direct pathways and hyperactivity of indirect pathways is observed (142).

The *DRD2* gene (*dopamine receptor D2*, 11q22-23) is one of the most investigated with reference to LID risk (see below). It encodes for a transmembrane G protein coupled receptor which activates intracellular signalling by the inhibition of cAMP synthesis (145). Oliveri et al were the first to study an intronic short tandem repeat (CA_n-STR) with four common alleles (13, 14, 15, and 16 CA repeats) in this gene, reporting a higher frequency of the 13 and 14 alleles in non-dyskinetic compared to dyskinetic PD patients (146). Another study reported a similar protective effect in males but not in female PD patients (129). Strong et al, found that the 14 allele and/or the 14/15 genotype was a risk factor for dyskinesia, in partial contrast with the above mentioned studies (147). The impact of another polymorphism in the *DRD2/ANKK1* locus, rs1800497 (or Taq1A, coding for Glu713Lys change in the protein) has been found to influence the risk of developing 'wearing off' motor fluctuations in PD (148). An influence of other variants in the *DRD2/ANKK1* region - including 141CIns/Del (rs1799732), rs2283265, rs1076560, C957T (rs6277), rs1800497 and rs2734849 - on LIDs was also reported in a recent study (1). Similarly, Kusters et al found three *DRD2*-haploblocks to be associated with dyskinesia in about 60% of the studied patients carrying one to three risk haplotypes (149). After combining "risk haplotypes" into a *DRD2*

genotypic risk score, they observed this was associated with an increased risk of dyskinesias and with their severity (149).

The *DRD3* gene (*dopamine receptor D3,* 3q13.3) - encoding for a receptor with an activity mediated by G proteins which inhibit adenylyl cyclase - has been reported to be overexpressed in experimental primate animal models presenting with LIDs (150). The overexpression was in accordance with the severity of LIDs and was prominent in the D1 expressing neurons (151). Similarly, the *DRD3* rs6280-A allele, encoding a p.S9G substitution which confers a high binding affinity to dopamine, has been associated with tardive dyskinesia, that can be attributed to dopamine-receptor hypersensitivity (152–154). According to this, rs6280 was associated with the presence of diphasic dyskinesia (i.e. taking place at the beginning and/or end of dose), after adjusting for gender, age at PD onset, Hoehn & Yahr stage and duration of L-Dopa treatment, (155). A recent study supported this association, with patients carrying the rs6280-A allele showing an increased risk of LIDs (111).

Some studies have explored the possible role of *DRD1* (*dopamine receptor D1*) variants on LIDs development in PD, but the reported results were not consistent (111,149). Notably, a growing body of biochemical and biophysical studies show that dopamine receptors can form homomeric and heteromeric complexes (156), hence it may be hypothesized that synergistic interactions between different receptors may induce LID in PD (157).

Other receptors

The *adenosine A2A receptor* (*Adora2A*) gene (22q11.23), encodes a receptor binding to G proteins which is highly expressed in the striatum of the brain, where it indirectly competes with *DRD2*, regulating neurotransmission (158). A recent study showed that Adora2A receptors are highly expressed in the basal ganglia of PD patients (159), especially in the striatum of PD patients who had developed dyskinesia (160).

Adora2A polymorphisms located in intron 1 of the gene - like rs2298383 and rs3761422 - were recently associated with LID events in PD patients (161). Previously, an association of the rs2298383 polymorphism with LID risk was revealed (161), although this association warrants further investigations (161).

It has been suggested that some of the changes in opioid transmission are directly implicated in LIDs (162). Opioids are co-transmitters in both the direct and the indirect basal ganglia pathways, where they regulate dopamine function, and basal ganglia have one of the highest levels of

endogenous opioids and opioid peptide receptors in the brain (163). Importantly, a Positron Emission Tomography (PET)-scan study revealed that PD dyskinetic patients had lower opioid binding in striatum and thalamus (162). Among opioid receptors, μ (mu) receptors received the main attention with reference to LIDs. Indeed, in the human *mu opioid receptor* (*MOR*) gene, the SNP rs1799971 has been associated with earlier development of dyskinesia in L-Dopa–treated PD patients (147), and has been found to increase binding affinity and functional activity of the endogenous opioid peptide, endorphin (164). Interestingly, receptor-specific opioid antagonists used in primate models have also been observed to affect LIDs (165,166).

N-methyl-D-aspartate ionotropic glutamate receptor (NMDAR) is a ligand-gated ion channel that responds to the neurotransmitters glutamate and NMDA. Dyskinesia, partly involves also changes in glutamatergic receptors in the striatum (142). This hypothesis is supported by evidence that amantadine - a NMDA receptor antagonist widely used in PD patients - reduces LIDs (142,167). The predominant inhibitory mechanism results from the increasing rate of channel closed states (ref). Interestingly, susceptibility to LIDs was recently associated with two *GRIN2A* (*glutamate ionotropic receptor NMDA type subunit 2A*) variants, rs7192557 and rs8057394, which had been previously associated with the age of dyskinesia onset in Huntington's Disease, suggesting that these movement complications may arise from the same neuronal pathways (79).

Enzymes involved in dopamine metabolism

Catechol-O-methyltransferase (*COMT,* 22q11.21) is an enzyme that inactivates catechols and degrades catecholamine neurotransmitters, including dopamine (168). It is implicated in the metabolism of L-Dopa, producing 3-O-methyldopa (3-OMD), which antagonizes L-Dopa's therapeutic action. COMT inhibitors, e.g. Tolcapone and Entacapone, reduce the conversion of L-Dopa to 3-OMD and thus improve its bioavailability in the brain (169,170). A common polymorphism in exon 4 of the *COMT* gene, rs4680, causes a Valine to Methionine substitution in the protein (Val108/158Met, depending on the *COMT* isoform). This results in altered activity of the enzyme: high activity in Val/Val, intermediate activity in Val/Met, and low activity in Met/Met genotype. Patients with the Met/Met (i.e. rs4680-A/A) genotype have been documented to experience more frequently severe dyskinesias and other motor fluctuations (171), and especially LIDs (172). Moreover, the doses of L-Dopa treatment for PD patients have previously been found to be influenced by specific *COMT* haplotypes (173). However, other studies have failed to confirm these associations (174–176).

Monoamine oxidase (MAO) is an enzyme regulating the metabolism of neurotransmitters

including, among others, norepinephrine, dopamine, and serotonin. Two distinct forms of the enzyme exist, encoded by *MAOA* (*monoamine oxidase A*; Xp11.3) and *MAOB* (*monoamine oxidase B*; Xp11.3). A recent study found that patients carrying *MAOB* rs1799836-A allele and -AA genotype suffered more frequently from LIDs (172), but no other studies have supported these associations (177).

Dopamine transporters

DAT (*dopamine transporter*; 5p.15.32) encodes a product which is fundamental for transporting dopamine across the plasma membrane. According to Sossi et al, greater DAT levels are directly associated with lower dopamine turnover and lower changes in synaptic dopamine concentration in PD patients (178). In this gene, a statistically significant association between the C allele of the intronic SNP rs393795 and longer time to LID has been found, which was hypothesized to be due to an altered rate of dopamine reuptake in the synapse (179). Furthermore, the nine copy allele of the 40-bp Variable Number Tandem Repeat (VNTR) polymorphism rs28363170 significantly predicted the occurrence of dyskinesia in a retrospective study on L-Dopa treated PD patients (180).

Other pathways and PD genes

The human *BDNF* (*brain-derived neurotrophic factor*; 11p14.1) gene encodes a precursor protein, proBDNF, which is then cleaved to the mature 14-kDa form (mBDNF) by protease tissue plasminogen activator (tPA)-mediated activation of plasmin (181). BDNF exerts multiple biological functions in the central nervous system, and its expression is decreased in PD (182). PD patients with Val66Met polymorphism (rs6265) in the 5'-pro-BDNF sequence had a significantly higher risk of developing dyskinesias earlier in the course of treatment with dopaminergic agents (183). A recent study has also found an association of the minor (A) allele with dyskinesia risk after dopaminergic treatment (184). Recently, the possible role of BDNF in levodopa motor complications was also highlighted in experimental animal models. E.g., rats that over-expressed BDNF were more prone to develop LIDs (157).

The *leucine-rich repeat kinase 2* (*LRRK2*; 12q12) is one of the genes most robustly implicated in PD aetiology (17), and has been also associated with LID onset in some studies. In a North African cohort, the prevalence of LIDs was significantly higher in carriers of the known PD-causative mutation G2019S, compared to non-carriers (185). However, a study in the Israeli population did

not replicate this association (186). A recent study showed that LRRK2 phosphorylation levels directly correlate with LID onset, and inhibition of LRRK2 induced a significant increase in the dyskinetic score in L-DOPA treated parkinsonian rat animal models (187).

As *LRRK2*, also *SNCA* has been robustly implicated in PD aetiology and progression (188), but has so far been mostly neglected with regard to motor complications connected to the treatment of the disease, in spite of some interesting findings. A heterozygous autosomal dominant point mutation in *SNCA* (c.158C>A; p.A53E in transcripts NM_000345.3, NM_001146054.1, NC_000004.11) was revealed in two Finnish PD patients, a mother and her daughter, characterized by severe bradykinesia, very little tremor and early onset of LIDs (189). No cognitive decline or dysautonomic features have emerged in these patients during more than 5 years of follow-up. In a recent study, C. elegans model overexpressing human α-synuclein was exposed to L-Dopa in continuous and alternating fashions (190). Chronic exposure to the drug led to hyperactivity of the animal model without meaningful increase in motor activity, and to an increase in peripheral clustering and expression of dopamine receptors in motor neurons. Both of these changes were significantly higher in alternating, compared to continuous, exposure to L-Dopa (190). More recently, Corrado et al (149) investigated the influence of the D4S3481-263 bp allele on the incident risk of LIDs, in an longitudinal cohort of Italian PD patients, reporting no significant differences between 263 allele carriers vs non carriers. These lines of evidence warrant further investigation of *SNCA* influence on LID onset, in addition to PD risk and endophenotypes.

Chapter 2

AIMS

To sum up, most of the reported heritability of Parkinson Disease is largely unknown, and its genetic bases remain unclear. This is likely due to the notable genetic heterogeneity of the disease, and to the relatively low power of genetic studies carried out so far. Moreover, PD endophenotypes, such as scales assessing motor, cognitive and other non-motor symptoms, have been largely under-investigated, due to the difficulties to collect PD cohorts with complete and detailed phenotypic assessment. Using such continuous scales to investigate PD genetics may provide powerful tools to identify PD susceptibility variants. Similarly, the genetic of LIDs is largely unclear, with different single variant associations reported, which have not been replicated yet. Therefore, further studies in independent cohorts are needed to clarify the genetic underpinnings of PD, its endophenotypes and genetic influences on side effects of L-Dopa therapy. To investigate these aspects, we adopted a multi-faceted and comprehensive approach (resumed below).

First, we investigated in an Italian PD cohort (N=470) collected at IRCCS Neuromed, the *SNCA* gene in order to:

1. clarify the role of two of the most investigated PD susceptibility variants which have also been associated with the level of expression of *SNCA* - namely rs356219 and D4S3481 - in the genetic susceptibility to PD, through case-control associations tests;
2. test the potential influence of these variants on PD scales assessing motor, cognitive and non-motor symptoms, as well as on PD age-at-onset (which represent powerful PD endophenotypes), through genetic association analyses;
3. determine whether *SNCA* affects also susceptibility to L-Dopa induced dyskinesia, by testing genetic associations of rs356219 and D4S3481 with the incident risk of LIDs in survival analyses.

Then, to identify rare variants with a potential risk/protective effect on LIDs occurrence in response to low/high L-Dopa daily dosages, we performed a variant prioritization bioinformatics pipeline in a subset of 114 PD patients which underwent Whole Exome Sequencing (WES) analyses.

Chapter 3
MATERIAL AND METHODS

PD patients cohort

472 PD patients (288 males; 196 familiar cases; mean (SD) age of 66.6 (8.8) years) were recruited at the Parkinson Centre of the specialized clinics IRCCS Neuromed, Pozzilli, Italy, between June 2015 and December 2017 (57). All the cases involved in the study (hereafter called Neuromed cohort) were diagnosed with PD by a qualified neurologist, according to published diagnostic criteria (appendix 1), which included rigidity, postural instability, resting tremor and positive response to levodopa treatment (191). Where diagnosis was uncertain, dopaminergic loss observed through neuroimaging techniques (PETscan or DaTscan) was used to confirm PD diagnosis. PD patients underwent a detailed phenotypic assessment, which included neurological examination and evaluation of non-motor domains (see below). Information about family history, demographic characteristics, anamnesis, pharmacological therapy and side effects was also collected. Mean (Standard Deviation) age and age at diagnosis were 66.6 (8.8) and 58.3 (10.0) years, respectively. Among these patients, 114 samples - including 42 familiar cases and 70 males - underwent Whole Exome Sequencing analysis (mean (SD) age and age at diagnosis 65.08 (8.83) and 55.89 (9.98), respectively).

A summary description of the whole Neuromed PD cohort and of the sequenced subset is reported in Table 3.1a, b.

The project was approved by the ethical committee of IRCCS Neuromed, Pozzilli, and written informed consent was obtained from all the participating subject.

Table 3.1: Description of a) the full Neuromed PD Cohort and b) the subset of 114 sequenced PD cases.

a)

Recruiting Center	Set	N (families)	Age (mean ± SD)	AAO (mean ± SD)	Disease duration (mean ± SD)	Sex ratio (M/F/missing)	Familiarity (FPD/SPD/ missing)	Dyskinesia status (D/ND/missing)	PD clinical subtype (rigid-bradykinetic/ tremorigenic/mixed/missing)
IRCCS Neuromed	Total	472 (458)	66.63 ± 8.82	58.28 ± 9.98	8.27 ± 6.28	288/184/0	196/273/3	176/242/54	304/72/77/19
	FPD	196 (183)	66.20 ± 8.97	57.60 ± 10.50	8.58 ± 6.87	118/67/0	-	82/97/20	127/34/35/10
	SPD	273 (273)	67.00 ± 8.66	58.82 ± 9.59	8.05 ± 5.83	161/112/0	-	94/145/34	176/48/41/8

b)

Recruiting Center	Set	N (families)	Age (mean ± SD)	AAO (mean ± SD)	Disease duration (mean ± SD)	Sex ratio (M/F/missing)	Familiarity (FPD/SPD/ missing)	Dyskinesia status (D/ND/missing)	PD clinical subtype (rigid-bradykinetic/ tremorigenic/mixed/missing)
IRCCS Neuromed	Total	114 (110)	65.08 ± 8.83	55.89 ± 9.98	9.22 ± 5.41	70/44/0	72/42/0	50/51/13	57/24/26/6
	FPD	42 (38)	63.31 ± 8.39	53.68 ± 10.57	9.75 ± 6.54	25/17/0	-	20/16/6	21/7/11/3
	SPD	72 (72)	66.13 ± 8.98	57.16 ± 9.46	8.91 ± 4.67	45/27/0	-	30/35/7	36/17/15/3

Abbreviations: AAO, Age at onset; FPD, Familial Parkinson's disease; SPD, Sporadic Parkinson's disease; D, Dyskinetic; ND, Non Dyskinetic.

Phenotypic assessment of PD cases

Phenotypic assessment of PD cases recruited has been recently described in a recent paper by our group (57). The Movement Disorder Society revised version of the Unified Parkinson's Disease Rating Scale Part III (18 items, maximum score 72; hereafter called UPDRS) (48) was used to assess clinical motor symptoms. These included language, facial expressions, tremor, rigidity, agility in movements, stability, gait and bradykinesia. Cognitive abilities were tested through an Italian validated version of the Montreal Cognitive Assessment (MoCA) (12). Cognitive domains assessed include short-term memory (5 points); visuospatial abilities via clock drawing (3 points), and a cube copy task (1 point); executive functioning via an adaptation of Trail Making Test Part B (1 point), phonemic fluency (1 point), and verbal abstraction (2 points); attention, concentration, and working memory via target detection (1 point), serial subtraction (2 points), digits forward and backward (1 point each); language via confrontation naming with low-familiarity animals (3 points), and repetition of complex sentences (2 points); and orientation to time and place (6 points). The total score was given by the sum of these domains, then divided by the maximum score obtainable (30). If one or more domains could not be tested (e.g. visuospatial tasks, due to unavailability of optical devices), the corresponding score was subtracted from the maximum total score. Non motor symptoms were assessed through an Italian validated version of Non Motor Symptoms Scale (NMS) for Parkinson Disease (14). This scale tests 9 items, including cardiovascular domain, sleep/fatigue, mood/cognition, perceptual problems/hallucinations, attention/memory, gastrointestinal, urinary, sexual function, and ability to taste or smell. For each item, both severity and frequency of symptoms is measured, so that the scale accounts for both aspects. This scale is available in (14) and in Appendix 2. Here, the sleep domain was slightly modified by adding a further question on the occurrence of vivid dreams. This question was treated as all the others, i.e. the severity of impairment was scored from 0 (no symptoms) to 3 (severe impairment), and the frequency of impairment was scored from 0 (less than once a week) to 4 (daily impairment), then the total score of the sub-item was computed as the product of severity by frequency, and added to the scores of the other sub-items. For this reason, and due to the high missing rate of sub-items in the sexual domain, we computed the NMS total score as the sum of all the answered items, divided by the maximum total score obtainable. This produced a continuous score ranging between 0 and 1 (hereafter called NMS). Age-at-onset (AAO) information was also collected at the time of recruitment, since it has been reported as an endophenotype that influence the clinical course of pathology.

Since the above mentioned traits tap into specific domains affected by PD and often represent

more powerful tools to investigate its genetic underpinnings, they are considered good PD endophenotypes, and were therefore investigated in this thesis.

Levodopa (L-Dopa) dosage calculations

During the visit, the neurologist verified if the patient manifested LID and registered the therapeutic protocol followed by patients before the control, as well as drug prescriptions for the period to come. All of these informations were recorded at each visit in a proprietary software system (Novamed©), so that they can be rescued at any time for usage in any epidemiological research project involving these patients.

For each patient, the daily L-Dopa dose was calculated by summing the total quantity contained in all drug formulations which were taken during the day. Table 3.2 reported all the drug formulations used by PD patients of the Neuromed cohort. Only L-Dopa dosages were summed to obtain the total amount of active ingredient taken during the day.

Daily Levodopa dosage was computed as follows:

$$Levodopa\ (mg/die) = \sum_{i=n}^{n} mg_i * n\ forms\ of\ drug$$

Where *forms of drug* indicate either tables or cassettes of the prescribed drugs and mg indicate the amount of L-Dopa contained in each form.

Table 3.2: Pharmaceutical forms of L-Dopa in the PD Neuromed cohort.

Commercial Name	Active Ingredients	Drug Formulation 1	Drug Formulation 2	Drug Formulation 3	Drug Formulation 4	Drug Formulation 5	Drug Formulation 6
Madopar	Levodopa	100 mg	200 mg				
	Benserazide	25 mg	50 mg				
Sinemet	Levodopa	250 mg	100 mg	200 mg	100 mg		
	Carbidopa	25 mg	25 mg	50 mg	25 mg		
Sirio	Melevodopa	250 mg	125 mg	100 mg			
	Carbidopa	25 mg	12,5 mg	25 mg			
Duodopa	Levodopa	20 mg/ml					
	Carbidopa	5 mg/ml					
Stalevo	Levodopa	50 mg	75 mg	100 mg	125 mg	150 mg	200 mg
	Carbidopa	12,5 mg	18,5 mg	25 mg	31,25 mg	37,5 mg	50 mg
	Etacapone	200 mg	200 mg	200 mg	200 mg	200 mg	200 mg

Genotyping

DNA extraction

Genomic DNA was isolated from peripheral blood lymphocytes by Blood and Cell Culture DNA Midi Kit (QIAGEN, Hilden, Germany) according to manufacturer protocol, which included the following steps:

1. Prepare blood samples using PBS, adjust volume to 10 ml.
2. Equilibrate a QIAGEN Genomic-tip 500/G with10 ml of Buffer QBT, and allow the QIAGEN Genomic-tip to empty by gravity flow.
3. Vortex the sample for 10 s at maximum speed and apply it to the equilibrated QIAGEN Genomic-tip. Allow it to enter the resin by gravity flow.
4. Wash the QIAGEN Genomic-tip with 2 x 15 ml of Buffer QC.
5. Elute the genomic DNA with 1 x 15 ml of Buffer QF.
6. Add 10.5 ml (0.7 volumes) room-temperature (15–25°C) isopropanol to the eluted DNA. Precipitate the DNA and resuspend in 0.1–2 ml of a suitable buffer (e.g., TE buffer, pH 8.0, or 10 mM Tris·Cl, pH 8.5). Precipitate the DNA by inverting the tube 10 to 20 times, and by centrifuging immediately at >5000 x *g* for at least 15 min at 4°C. Carefully remove the supernatant. Wash the centrifuged DNA pellet with 4 ml of cold 70% ethanol. Vortex briefly and centrifuge at >5000 x *g* for 10 min at 4°C. Carefully remove the supernatant without disturbing the pellet. Air-dry for 5–10 min, and resuspend the DNA in buffer.
7. Dissolve the DNA overnight on a shaker or at 55°C for 1–2 h. Resuspend the DNA pellet by rinsing the walls to recover the DNA. Pipette the DNA up and down to promote resuspension should be avoided.

Genotyping of SNCA variant rs356219

The SNP rs356219 (hg19 coordinates chr4:90637601; A/G; allelic frequencies ~ 49/51%) – lying in the 3' untranslated region (3'UTR) of the *SNCA* gene (4q22.1) and previously associated with its circulating levels of expression (75–78,86–89)– was genotyped using TaqMan® custom assays (Bio-Rad, USA), according to the manufacturer's protocol, and analysed in a Bio-Rad® CFX96™ Real Time PCR detection system. About 10–50 ng of DNA were amplified with 5 μL of 2X TaqMan Universal PCR master mix, 0.5 μL of 40X primer and TaqMan probe dye mix. Cycling

conditions were 3 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 30 s at 60 °C. Genotyping was performed on 470 PD cases for which DNA samples were available at the time of genetic analyses. Along with patients, 518 controls were genotyped for the purpose of case-control association analyses, which included:

- 122 non-consanguineous family members (mean (SD) age 62.9 (9.1) years; 44 males) with no neurological signs or symptoms of PD at the time of recruitment.
- 338 unscreened controls (pseudo-controls) belonging to the general Italian population, collected at the Institute of Genetics and Biophysics of the National Research Council in Naples for the purpose of other genetic studies (122 males; age information not available);
- 58 neurological controls selected from the Moli-sani study – a large population-based cohort study of citizens from the Molise-region (192) - which showed no signs/symptoms, nor took any specific drug for neurodegenerative disorders (mean (SD) age 77 (5.4) years; 13 men).

We performed a general quality control (QC) of genotyped samples, in PLINK v1.9 (193). The SNP analysed showed a very good call rate (>98%, 17 samples with missing genotype) and was in Hardy Weinberg Equilibrium (HWE, p=0.62), suggesting the good quality of genotyping.

Genotyping of SNCA D4S3481 variant

The *SNCA* microsatellite D4S3481 (hereafter called Rep1) was analysed in the 469 PD patients of the Neuromed cohort, as well as in 518 general population controls (see above), as described in Maraganore et al, 2006 and in the following studies. Briefly, the region was amplified through Polymerase Chain Reaction (PCR) from genomic DNA, using the following primer pairs: Fam5′-CCTGGCATATTTGATTGCAA-3′ and 5′-GACTGGCCCAAGATTAACCA-3′. PCR reactions (25 µl final volume) containing 2 mmol/L MgCl2, 0.5 mol/L of each primer, 200 mol/L dNTPs, 1 unit of Taq polymerase (Life Technologies) and approximately 20 ng of genomic DNA. Thermal cycling was performed with an initial denaturation of 180 seconds at 94°C, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at melting temperature (MT), 30 seconds at 72°C, and a terminal extension of 10 min at 72°C. PCR products were then diluted 1:10 and resolved by capillary electrophoresis on an ABI-3130XL DNA Analyzer (Applied Biosystem, Foster City, CA, USA), using GeenScan-500 ROX (Applied Biosystem) as molecular weight marker. Allelic sizes were assessed using the GeneMapper® Software Version 4.0 SNPlex™ (Applied Biosystem, Foster City, CA, USA). This method allows to determine the length of dinucleotide repeats at the

investigated locus, and typically results in number of repeats ranging between 255 and 263. Since we detected only three samples (one case and two controls) carrying the 255 allele, and five samples with the 257 allele (two cases and three controls), and these alleles are usually neglected due to their low frequency (149), we removed them before the analyses, as done elsewhere (194). Also this variant showed good genotyping call rates (>97%, 29 samples with missing genotype) and was in HWE (p=0.28; Chi=3.18).

Statistical analyses

Below, we report the statistical analyses applied to test genetic associations of the *SNCA* candidate genetic variants rs356219 and D4S3481 with PD, its related endophenotypes and incident risk of LID onset. All analyses were carried out in R (https://www.r-project.org/) (195). For further theoretical background and details on these models, see Appendix 1.

Case-control genetic association tests

To test associations of rs356219 and D4S3481 with PD risk, we built logistic regression models using the formula:

$$PD \sim sex + var,$$

where *var* represents the genotyped variant (either rs356219 or D4S3481).

This was implemented through the *glm()* function in R (195), with *family=binomial(link="logit")* option. The choice of covariates was conditioned by "age" information not being available for many of the general population controls genotyped, which would have implied a notable reduction of sample size, hence of power of the analysis, if age was included as additional covariate in the model.

Three alternative genetic models were tested for rs356219, namely an additive, a dominant and a recessive model, as detailed in Table 3.3a. For D4S3481, we selected the 259 and the 263 allele to define genotype classes, since these have been more consistently reported as having a protective/risk effect on PD susceptibility (85,90,91). Although for such a multi-allelic marker as D4S3481 it would be more appropriate to call the models tested extensively (e.g. 259 allele carriers vs all others), for simplicity and brevity we will often refer to these models as Pseudo-additive/-dominant/-recessive models (see Table 3.3b), as done elsewhere (149).

For this analysis, we set significance thresholds to $\alpha = 8.3 \times 10^{-3}$, after applying a Bonferroni correction for two independent variants and three alternative genetic models tested for each variant.

Table 3.3: Alternative genetic models used to test association of the *SNCA* variants a) rs356219 and b) D4S3481, with PD case-control status.

a)

Variant	Additive	Dominant	Recessive
rs356219	Class 1: AA carriers; Class 2: AG carriers; Class 3: GG carriers	REF/Protective (A) allele carriers vs all others (GG)	ALT/Risk (G) allele carriers vs all others (AA)

b)

Variant	Pseudo-additive	Pseudo-dominant	Pseudo-recessive
REP1 (D4S3481)	Class 1: 259_259, 259_261, 259_263 carriers; Class 2: 261_261 carriers; Class 3: 261_263, 263_263 carriers.	REF/Protective (259) allele carriers vs all others	ALT/Risk (263) allele carriers vs all others

For each variant, reference (REF) and alternative (ALT) alleles are specified, as well as their effect on PD susceptibility as reported by previous literature (85,95-97) (see Introduction section for further background).

Genetic association with continuous PD endophenotypes

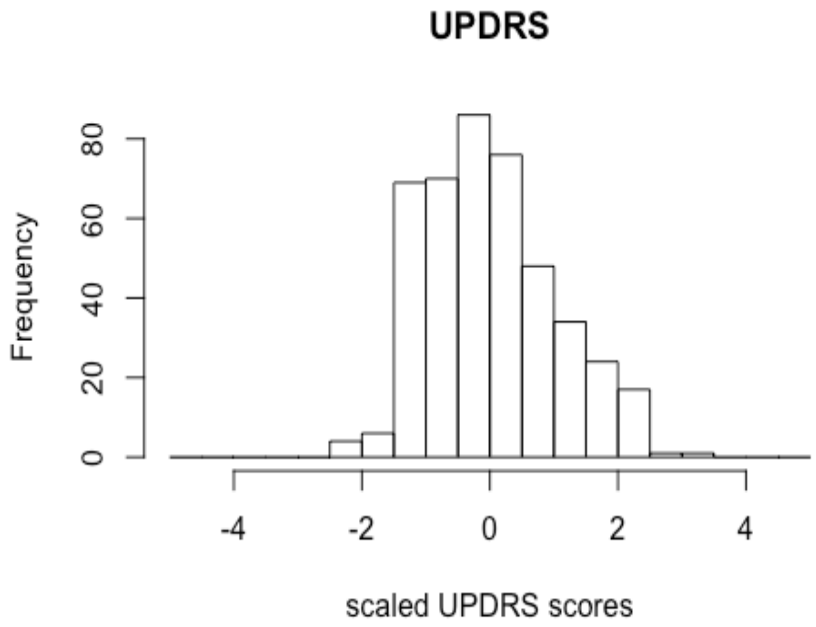
Quality Control and elaboration of continuous traits

A preliminary quality control (QC) of the continuous scales assessing neurological (UPDRS), cognitive (MoCA) and other non-motor PD symptoms (NMS), as well as PD age at onset (AAO), was carried out before association testing. More specifically, we ascertained the main basic assumptions of linear regression analyses, namely normality of distributions of the traits analysed and absence of phenotypic outliers (see Appendix 1 for further theoretical background). To this

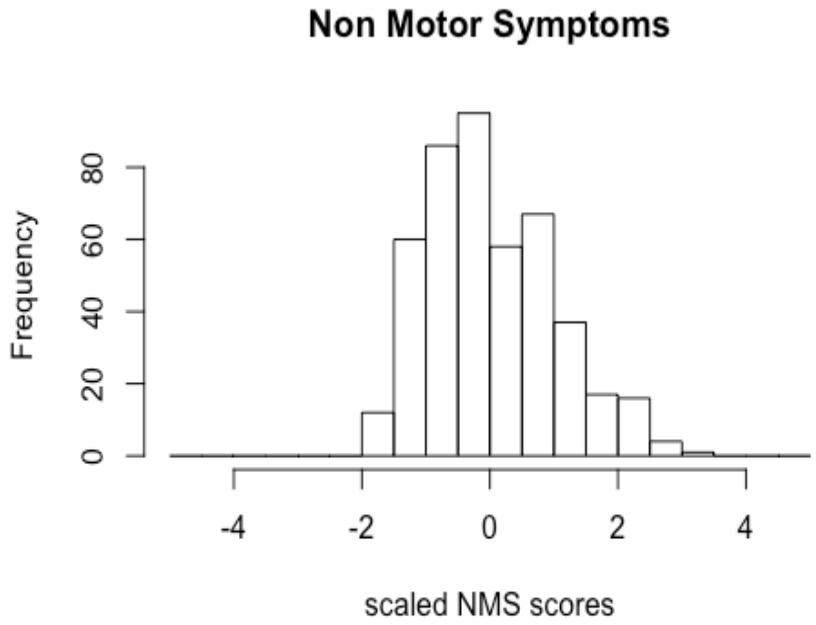
purpose, we plotted distributions of UPDRS, MOCA, NMS and AAO in the PD cohort. This revealed a substantial normality of distributions and an absence of extreme phenotypic outliers (see **Figure 3.1**), which were defined as subjects showing values at least 3 IQR (interquartile ranges) above Q3 (quartile 3) or below Q1 (quartile 1) in each distribution, respectively.

Figure 3.1: Histograms of continuous PD endophenotypes analysed, namely a) UPDRS, b) NMS, c) MOCA and d) AAO. Standardized scales are reported.

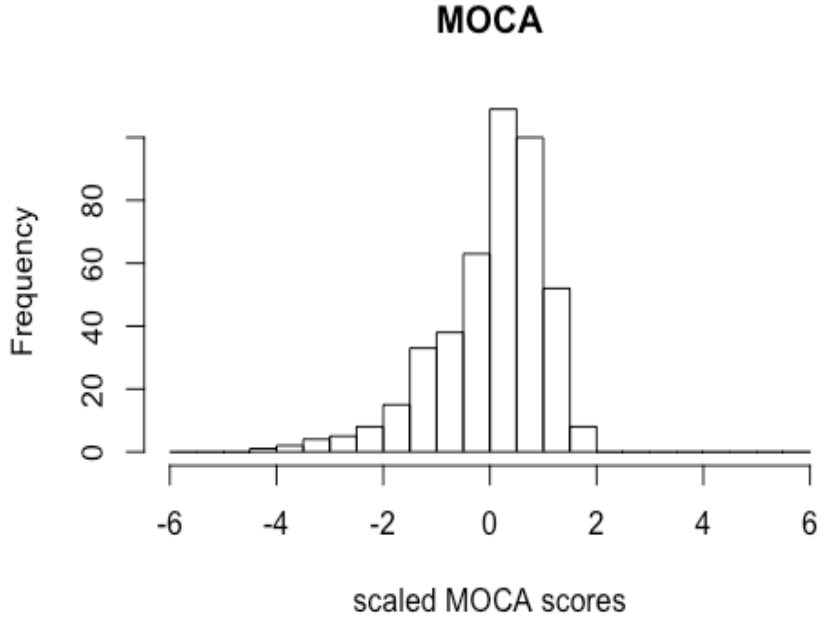
a)



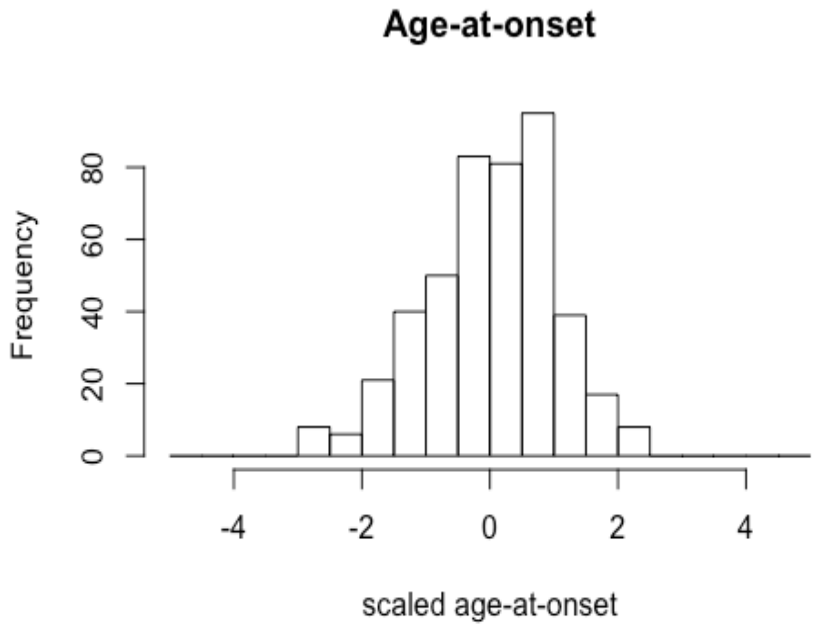
b)


Non Motor Symptoms
Frequency
0
20
40
60
80
-4
-2
0
2
4
scaled NMS scores

c)


MOCA
Frequency
0
20
40
60
80
-6
-4
-2
0
2
4
6
scaled MOCA scores

d)



Association tests with continuous PD traits

After QC and elaboration of PD endophenotypes, we tested genetic associations of the candidate variants rs356219 and D4S3481 with such traits. To do so, we built generalized linear models (*glm()* function in R, using the *family="Gaussian"* option), for each of the four endophenotypes analysed, following the formula

$$\boldsymbol{PD \sim cov + var,}$$

where *var* represents the genotyped variant (either rs356219 or D4S3481) and *cov* represents covariates used in the model (see below). For this analysis, two different statistical models were built, one adjusted for age and sex only (Model 1), and one further adjusted for PD familiarity (sporadic/familiar form), clinical subtype (tremorigenic/rigid-bradykinetic/mixed), pharmacological treatment status (ON/OFF), years of disease and daily intake of L-Dopa, in addition to age and sex (Model 2; (57)). The inclusion of these covariates was aimed at regressing out the influence of these variables on the PD symptoms and age at onset, motivated by previously reported evidence of associations between the scales analysed and sex (124), age (5,124), PD familiarity (121,196), clinical subtype (197), early onset of the disease (124) and L-Dopa treatment (124). Since Model 2 was more conservative, we took it as our main model of reference for the interpretation of results.

As in the case-control analysis, we tested three alternative genetic analyses - namely an additive, a dominant and a recessive model - both for rs356219 and for D4S3481 (see Table 3.3a, b above). Here, significance thresholds were corrected for two independent variants, three alternative genetic models and four PD endophenotypes tested ($\alpha = 0.05 / (2 \times 3 \times 4) = 2.1 \times 10^{-3}$).

Survival analyses on LID onset

In the investigation of the genetic basis of L-Dopa induced dyskinesias (LIDs), we initially focused our analysis on the investigation of the two candidate *SNCA* variants rs356219 and for D4S3481 (see below).

We tested potential influences on the incident risk of LIDs within the Neuromed PD cohort, through univariate and multivariable Cox proportional hazards (PH) regressions. In these models, the dependent variables included LID onset (Yes/No) and time-to-LID, namely the follow-up period of each PD patient (in months), starting on the date of start of L-Dopa treatment and ending when LID onset occurred. When no LIDs were reported by the patient and/or detected by the neurologist up to December 31st, 2018 (end of follow-up time), right-censoring was applied. This allowed to build a dataset structured as below (**Figure 3.2**).

Figure 3.2: Basic example of structure of a dataset used for analysis.

Subject FAM_ID	Survival time (Months)	Status (1=LID; 0=no-LID)
1_1	137	1
2_3	63	1
3_5	57	0
4_8	36	0
5_10	96	1
6_12	181	1

Note: this represents only a basic example of the mandatory data required in a database to carry out survival analyses. The dataset can be enriched with as many variables as allowed by the worksheet.

which was further enriched for other demographic, clinical and pharmacological information, such as sex, age, PD familiarity, age-at-onset and years of disease, daily L-Dopa intake, and other variables of interest for the study of LIDs (see below).

As exposure variables, we tested candidate *SNCA* genetic variants rs356219 and D4S3481 both in univariate and multivariable models, as well as other non-genetic covariates which were available in our cohort and had been previously associated with LID onset risk, in multivariable models. These covariates included sex (121,198,199), age (199), PD familiarity (121,196), PD clinical subtype (tremorigenic, bradykinetic-rigid or mixed) (197), L-Dopa intake (121,199), years of disease (121,199) and age at onset (121,199). For a brief overview of the studies implicating these covariates in LID onset, see Introduction section.

Testing basic assumptions of Cox proportional hazards (PH) models

Basic assumptions of Cox proportional hazards (PH) models were preliminarily checked for the genetic variants tested, as well as for all the covariates included in the survival models (see below). These included the proportionality of hazards (PH assumption) and the absence of outlier observations, as explained below (see Appendix 1 for theoretical background). We tested the PH assumption for each of the independent variables tested (both genetic variants and other covariates) through plotting Schoenfeld residuals of univariate cox regressions modelling LID onset as a function of each variable (200). These were computed through the *cox.zph()* function of the *survival* package (https://cran.r-project.org/web/packages/survival/index.html) (194) and through the *ggcoxzph()* function of the *survminer* package (https://cran.r-project.org/web/packages/survminer/index.html), applied to univariate cox regressions, in R. These revealed no variants or covariates with an evident and significant change in the proportionality of risks as a function of time, across the classes compared ($p>0.05$; see Table 3.4; **Figures S2.1** and **S3.1**), suggesting that the proportional hazard assumption was satisfied for all variables tested.

Similarly, we checked for the absence of outlier observations in these variables through plotting dfbeta residuals of each variant/covariate tested, computed through the *ggcoxdiagnostics()* function of the R *survminer* package (see above for URL), applied to univariate cox regressions. Again, following the rule recommended by Belsley, Kuh, and Welsch (1980), we observed no observations with $|\text{dfbetas}| > 2/\sqrt{n}$ (see **Figures S2.2** and **S3.2**), namely observations with a

significant weight in the Cox model compared to the others, which may be considered influential outliers (201).

Table 3.4: Schoenfeld residuals output for a) each covariate and b) genetic variant tested in univariate and multivariable Cox PH models.

a)

	rho	Chisq	p
Sex	0.03	0.07	0.80
Familiarity	-0.09	0.93	0.34
Phenotype (Bradykinetic vs Tremorigenic)	-0.03	0.08	0.78
Phenotype (Mixed vs Tremorigenic)	-0.06	0.42	0.52
AAO	0.05	0.32	0.57
L-Dopa dosage	0.16	32543	0.07
MOCA	0.11	10356	0.31
HY	0.12	12532	0.26
UPDRS	-0.04	0.21	0.65
Years_of_Disease	-0.14	22675	0.13
NMS	-0.08	0.96	0.33
Multivariable (GLOBAL)	NA	113258	0.42

b)

Genetic Variant	Genetic Model	rho	Chisq	p
rs356219	Additive (GG vs AG vs AA)	0.07	0.66	0.42
	Dominant (A allele carriers vs GG)	0.10	1.43	0.23
	Recessive Model (G allele carriers vs AA)	0.1	1.39	0.24
D4S3481	Pseudo-additive (263 vs 261 vs 259 allele carriers)	-0.06	0.58	0.45
	Pseudo-dominant (259 allele carriers vs all others)	-0.09	1.09	0.30
	Pseudo-recessive (263 allele carriers vs all others	-0.08	0.92	0.34

Here, Pearson product-moment correlations between the scaled Schoenfeld residuals and log(time) for each covariate and genetic model tested are reported (rho, i.e. a proxy of the slope of Schoenfeld residuals vs time curves), along with relevant Chi-squared statistics (Chisq) and p-values (p). GLOBAL gives the global test of proportionality for all the interactions of the covariates with log(time), tested at once. None of these covariates showed rhos significantly different from zero, neither in the univariate, nor in the multivariable (Global) test (α=0.05). Note: no *rho* value was computed for the GLOBAL test, as per *cox.zph()* function output.

Cox PH models

First, to investigate the relation of non-genetic covariates with the incident risk of LIDs, we carried out an exploratory survival analysis modelling LID onset as a function of non-genetic covariates, namely sex, PD familiarity, clinical subtype and age-at-onset (AAO), L-dopa dosage, UPDRS, MoCA, NMS and HY scores, and years of disease. Although only exploratory, for this analysis we set a significance threshold of $\alpha = 5.0 \times 10^{-3}$, applying a Bonferroni correction for ten different covariates tested.

Then we performed genetic Cox PH regressions, for both rs356219 and D4S3481 separately, first in crude unadjusted models (Model 1)

LID onset ~ var,

and then in conservative models adjusted for all the covariates mentioned above, which we used as reference models for interpretation of results (Model 2). As above, three alternative genetic transmission models were assumed and tested, namely an additive, a dominant and a recessive model (see Table 3.3a, b). Therefore, significance thresholds for this analysis was corrected for two genetic variants and three genetic models tested ($\alpha = 0.05 / (2\times3) = 8.3 \times 10^{-3}$).

For all of the models performed, we built Kaplan-Meier (for crude unadjusted models) and Cox curves (for adjusted models, where applicable), which showed the occurrence of LID events during follow-up in the different groups compared, for each of the independent variables tested. These plots were built through applying the *plot()* function of the *survival* package and the *ggadjusted()* function of the *survminer* package, respectively, in R.

Investigating genetic basis of LIDs at an exome-wide level

After investigating candidate *SNCA* variants, we extended the investigation on the genetic basis of LIDs at the exome-wide level, exploiting the availability of 114 samples with Whole Exome Sequencing (WES) data available within the Neuromed cohort. Since the sample size available (hence the power to detect common variants with typically small effect sizes) was relatively small, we focused on the search for rare variants which could explain strong risk/protective effects on LID onset. We did this through an innovative approach based on L-dopa dosage outlier values, which we describe below.

Whole Exome Sequencing (WES): protocol and quality control

We carried out a Whole Exome Sequencing (WES) analysis of DNA samples from 114 (42 familial and 72 sporadic) PD cases recruited within the NEUROMED cohort between June 2015 and June 2016 (see Table 3.1b). These samples underwent WES at Helmotz Zentrum, Munich, Germany. Genomic DNA was isolated from peripheral blood lymphocytes by Blood and Cell Culture DNA Midi Kit (QIAGEN, Hilden, Germany). Exonic regions were enriched using the SureSelect All Exome kit v6 (Agilent® Technologies, Santa Clara, CA, USA) based on DNA fragmentation and capture. Exomes were barcoded and sequenced using the Illumina® HiSeq2000 platform (Illumina, San Diego, CA, USA).

The alignments of the 100-bp paired-end reads to the human reference genome was performed through BWA MEM v0.7.542 (202). After removal of duplicate reads through Picard, single nucleotide variants (SNVs) and insertions/deletions (indels) were called, using HaplotypeCaller and GenotypeGVCFs in GATK v3.5-0-g36282e4 (203). Average exome coverage was 143x and at least 20x for 98.8% of the target. One sample with intraspecific contamination rate > 7% and one which was later re-defined as essential tremor were removed during QC, which was performed using vcftools v0.1.12b (204) and PLINK v1.90b3.45 (193). Variant calls with total depth (DP) < 8 and genotype quality (GQ) < 50 were set to missing, and variants with Minor Allele Count (MAC) = 0, number of alternative alleles ≠ 2 and call rate < 95% were filtered out, as well as samples with identical-by-descent sharing and sex mismatches, and samples with call rate < 90%. Similarly, samples were checked for absence of outliers in terms of genetic ancestry (through Multidimensional Scaling Analysis), genome-wide homozygosity, and of number of singleton variants per sample. 112 samples (42 FPD and 70 SPD cases) and 356,710 variants passed QC (338,278 SNPs and 18,432 indels).

Identification of rare mutations with potential risk/protective effect on LID onset

We attempted to identify rare mutations conferring a potential risk or protective effect on LID occurrence, among those PD patients which had undergone WES analyses (N=112), through two specular approaches.

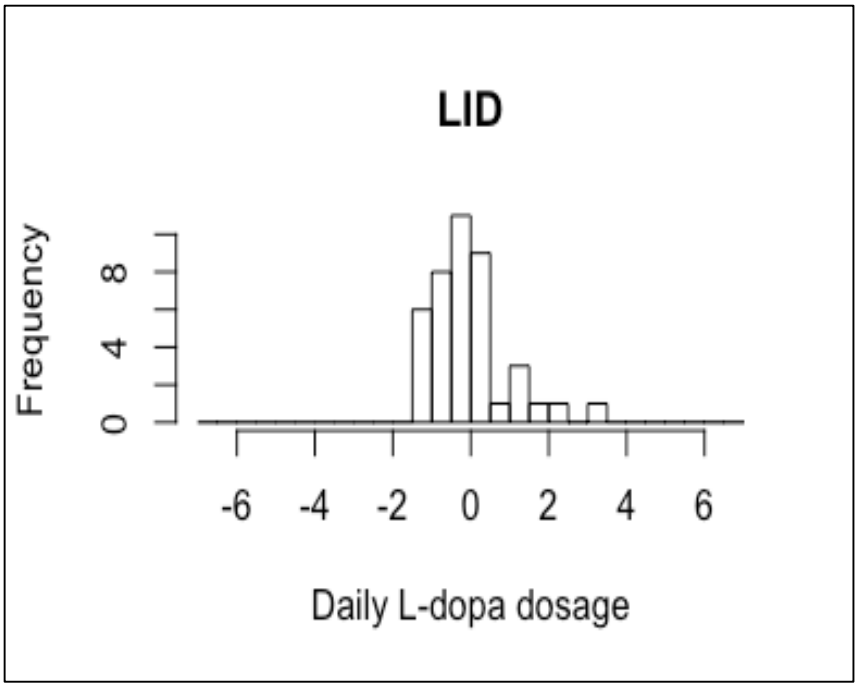
More specifically, we plotted distributions of daily L-Dopa dosage for each sequenced PD patient through the *hist()* function in R, separately for subjects with and without LID, using clinical and pharmacological information updated to December 31st, 2018 (see histograms in **Figures 3.4a, b**). Within each group (LID and non-LID), we looked for L-Dopa dosage outliers, namely those PD patients taking daily L-Dopa dosages at least 3 Standard Deviations (SDs) below the normative mean of the LID group, and patients with L-Dopa dosages at least 3 SDs above the mean of the non-LID group, respectively.

In other words, this analysis was aimed at detecting patients showing absent LID in spite of high L-Dopa intakes, and patients showing LID occurrence at low L-Dopa intakes, so to identify subjects carrying potential protective/risk mutations for LID occurrence. However, we detected no such outliers among sequenced PD patients, neither in the LID nor in the non-LID group (**Figures 3.4a, b**). Therefore, no rare variants with potential protective/risk effect on LID onset could be identified and further investigated in our sequenced sample. Similarly, when we repeated the analysis in all the PD patients of Neuromed cohort with complete clinical and pharmacological

information available (N=406), we observed no outliers L-Dopa dosages in the entire cohort study (**Figure 3.5a, b**).

Figure 3.4: L-Dopa dosage outliers detection in a) LID and b) non-LID sequenced (WES) samples.

a)



b)

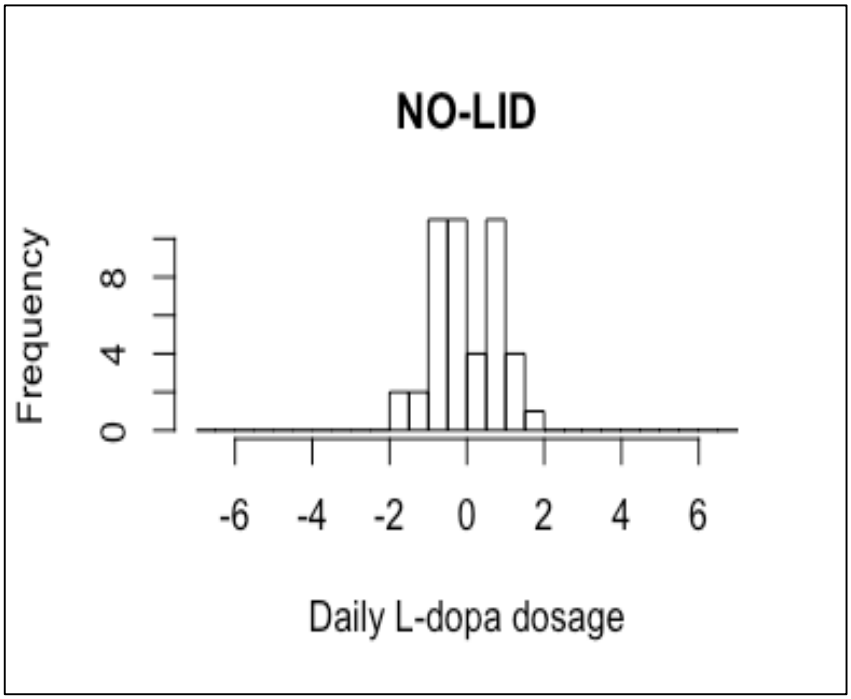
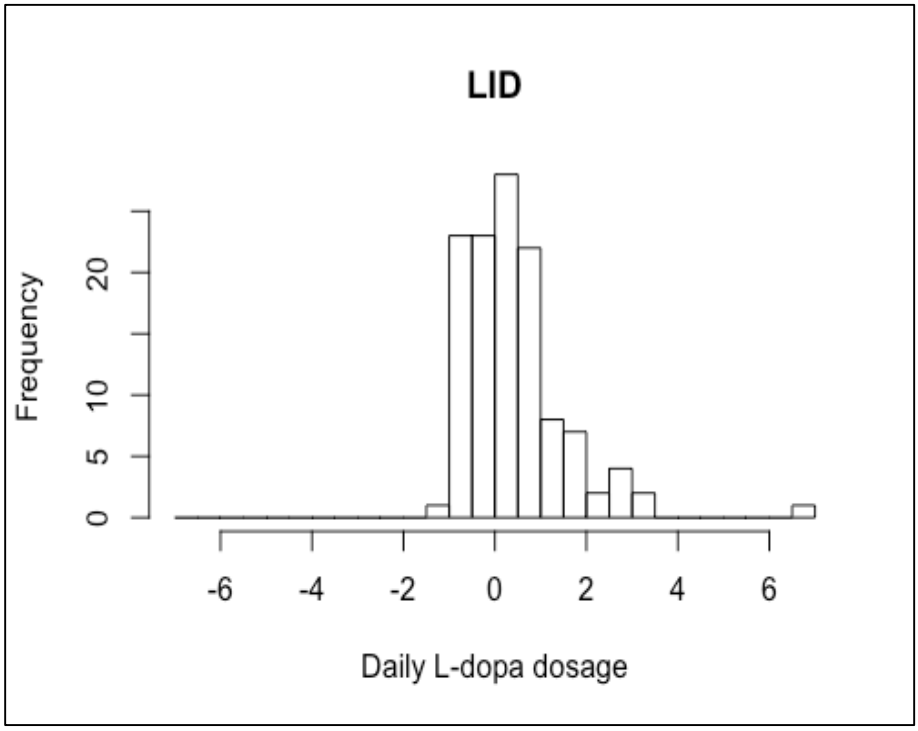
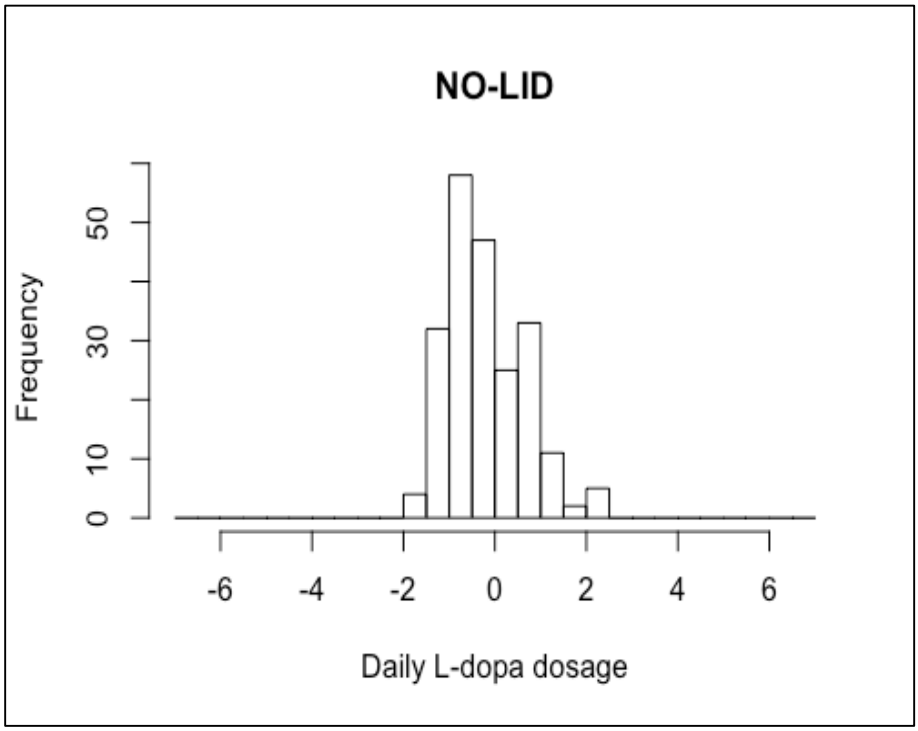


Figure 3.5: L-Dopa dosages outliers detection in a) LID and b) non-LID subsets of the whole Neuromed cohort.

a)



b)



Chapter 4

RESULTS

Genotype and allele frequencies of the two candidate variants

The two genotyped variants in the *SNCA* gene, rs356219 and D4S3481, showed the genotype frequencies reported in Table 4.1a, b (below).

Among cases, the allele frequencies of rs356219 were 59.8% for the A (reference) and 40.2% for the G (alternative) allele, 64.4% and 35.6% among controls, and 62.2% and 37.8% in the total successfully genotyped sample (N=981), respectively. For D4S3481, frequencies of 259, 261 and 263 bp alleles were 27.6%, 66.1% and 6.0% among cases, 30.5%, 64.8% and 4.2% among controls, and 29.1%,65.4% and 5.1% in the total genotyped sample (N=959). 255 and 257 bp alleles showed a cumulative allele frequency of 0.3% among cases, 0.5% among controls and 0.4% in the total overall sample, and were therefore removed before analysis (see Methods section for details).

Based on the observed genotype frequencies, the two variants were in substantial Linkage Equilibrium (pairwise $r^2 = 0.03$ and D' = 0.36 among cases, and $r^2 = 0.05$ and D' = 0.46 among controls), suggesting that the analysis of both markers was appropriate to investigate potential independent genetic effects.

Table 4.1: Genotype frequencies of the two candidate variants rs356219 and D4S3481.

a)

rs356219 genotype	Count (% frequency) Cases	Count (% frequency) Controls	Count (% frequency) Overall
A/A	168 (36.1%)	221 (42.8%)	389 (39.6%)
A/G	220 (47.3%)	223 (43.2%)	443 (45.1%)
G/G	77 (16.5%)	72 (13.9%)	149 (15.2%)

b)

D4S3481 genotypes	Count (% frequency) cases	Count (% frequency) controls	Count (% frequency) Overall
259/259	30 (6.4%)	42 (8.5%)	72 (7.5%)
259/261	183 (39.2%)	203 (41.3%)	386 (40.2%)
259/263	14 (3%)	12 (2.4%)	26 (2.7%)
261/261	198 (42.4%)	201 (40.8%)	399 (41.6%)
261/263	36 (7.7%)	29 (5.9%)	65 (6.8%)
263/263	3 (0.6%)	0 (0%)	3 (0.3%)
255/259	1 (0.2%)	1 (0.2%)	2 (0.2%)
257/261	2 (0.4%)	3 (0.6%)	5 (0.5%)
255/261	0 (0%)	1 (0.2%)	1 (0.1%)

Note: Missing genotype calls were reported for 5 cases and 12 controls for rs356219, and for 3 cases and 26 controls for D4S3481.

SNCA polymorphisms and PD risk

The results of logistic regression of PD status vs rs356219 and D4S3481 are reported in Table 4.2, for all the genetic models tested (see Table 3.3a, b in the Methods section for details). This analysis revealed a nominally significant genetic association for rs356219 in the Recessive model, where G (Alternative) allele carriers showed an increased PD risk compared to homozygotes for the reference allele (AA) (OR [CI] = 1.32 [1.01; 1.73]. This association was consistent with trends of

associations observed for the SNP in the Additive model, where AG and AA subjects showed increased risks of PD by 30% (p = 0.07) and by 41% (p = 0.08), respectively (see Table 4.2). However, this association did not survive correction for multiple testing of two independent variants and three alternative genetic models ($\alpha = 8.3 \times 10^{-3}$).

Table 4.2: Results of PD case-control genetic association tests of rs356219 and D4S3481.

Genetic Variant	Genetic Model	Contrast[a]	OR [CI]	z-score	p
rs356219	Additive	AG vs AA	1.30 [0.98; 1.72]	1.79	0.07
		GG vs AA	1.41 [0.96; 2.09]	1.73	0.08
	Dominant	A allele carriers vs GG	1.23 [0.86; 1.76]	1.12	0.26
	Recessive	G allele carriers vs AA	1.32 [1.01; 1.73]	2.06	0.04
D4S3481	Pseudo-additive	261 vs 259 allele carriers	1.08 [0.82; 1.42]	0.64	0.51
		263 vs 259 allele carriers	1.56 [0.92; 2.65]	1.69	0.09
	Pseudo-dominant	259 allele carriers vs all others	1.14 [0.88; 1.48]	0.98	0.32
	Pseudo-recessive	263 allele carriers vs all others	0.70 [0.45; 1.09]	-1.58	0.11

Here, we report Odds Ratios (OR) and 95% Confidence Interval (CI), for each genotype class compared to the reference class (see Table 3.3a, b for details), along with relevant association z-score and p-value (p). Nominally significant associations ($p < 0.05$) are highlighted in bold. No associations survived Bonferroni correction ($\alpha = 8.3 \times 10^{-3}$). [a] Note: full details on the genetic models built and on the genotype classes contrasted are reported in Table 3.3a, b (see Methods section).

SNCA polymorphisms and continuous PD endophenotypes

Linear regression analyses modelling the relation of continuous PD endophenotypes - namely UPDRS, MoCA, NMS score and AAO – revealed no significant associations surviving Bonferroni correction ($\alpha = 2.1 \times 10^{-3}$), neither in a basic model adjusted for age and sex (Model 1, Table 4.3a-d), nor in a more conservative model further adjusted for PD familiarity, clinical subtype, pharmacological treatment status, years of disease and daily intake of L-Dopa (Model 2, Table 4.4a-d), which we used as our reference model for the interpretation of results since it was more

conservative. Again, we observed two nominally significant associations in Model 1, for D4S3481 in the Pseudo-recessive model (263 allele carriers vs all others), with MoCA score (β (SE) = -0.05 (0.02); p = 0.03) and with AAO (β (SE) = -2.01 (1.00); p = 0.045). However, these did not survive correction for multiple testing of two independent variants, three alternative genetic models and four PD endophenotypes tested ($\alpha = 2.1 \times 10^{-3}$), and were not confirmed in the fully adjusted models (see Table 4.3 b, d). In Model 2, we observed another nominally significant association with AAO for D4S3481-261 bp allele vs 259 bp allele carriers (β (SE) = -2.02 (1.00); p = 0.045) in the Pseudo-additive model. Again, this did not survive Bonferroni correction. Moreover, we did not observe any evidence of association for the additional risk genotype class in the same model (i.e. 263 vs 259 bp allele carriers; see Table 4.4d).

Table 4.3: Genetic associations of rs356219 and D4S3481 with continuous PD endophenotypes, including a) UPDRS, b) MoCA, c) NMS and d) AAO, in Model 1 (adjusted for sex and age).

a) UPDRS

Genetic Variant	Genetic Model	Contrast[a]	Beta	SE	t-stat	p
rs356219	Additive	AG vs AA	-0.22	1.16	-0.19	0.85
		GG vs AA	0.13	1.50	0.09	0.93
	Dominant	A allele carriers vs GG	0.26	1.34	0.19	0.85
	Recessive	G allele carriers vs AA	-0.12	1.09	-0.11	0.91
D4S3481	Pseudo-additive	261 vs 259 allele carriers	0.87	1.10	0.80	0.43
		263 vs 259 allele carriers	0.17	1.88	0.09	0.93
	Pseudo-dominant	259 allele carriers vs all others	0.75	1.04	0.72	0.47
	Pseudo-recessive	263 allele carriers vs all others	0.35	1.58	0.22	0.83

b) MoCA

Genetic Variant	Genetic Model	Contrast[a]	Beta	SE	t-stat	p
rs356219	Additive	AG vs AA	0.02	0.01	1.30	0.19
		GG vs AA	0.03	0.02	1.78	0.08
	Dominant	A allele carriers vs GG	0.02	0.02	1.35	0.18
	Recessive	G allele carriers vs AA	0.02	0.01	1.69	0.09
D4S3481	Pseudo-additive	261 vs 259 allele carriers	-0.004	0.01	-0.27	0.79
		263 vs 259 allele carriers	0.04	0.03	1.63	0.10
	Pseudo-dominant	259 allele carriers vs all others	0.004	0.01	0.27	0.78
	Pseudo-recessive	**263 allele carriers vs all others**	**-0.05**	**0.02**	**-2.20**	**0.03**

c) NMS

Genetic Variant	Genetic Model	Contrast[a]	Beta	SE	t-stat	p
rs356219	Additive	AG vs AA	-2.07	3.66	-0.57	0.57
		GG vs AA	1.30	4.84	0.27	0.79
	Dominant	A allele carriers vs GG	2.46	4.38	0.56	0.57
	Recessive	G allele carriers vs AA	-1.16	3.43	-0.34	0.74
D4S3481	Pseudo-additive	261 vs 259 allele carriers	5.35	5.22	1.02	0.31
		263 vs 259 allele carriers	-1.65	3.32	-0.50	0.62
	Pseudo-dominant	259 allele carriers vs all others	-2.29	6.19	-0.37	0.71
	Pseudo-recessive	263 allele carriers vs all others	-1.52	3.49	-0.44	0.66

d) AAO

Genetic Variant	Genetic Model	Contrast[a]	Beta	SE	t-stat	p
rs356219	Additive	AG vs AA	0.15	1.07	0.14	0.89
		GG vs AA	0.96	1.41	0.68	0.50
	Dominant	A allele carriers vs GG	0.88	1.28	0.68	0.49
	Recessive	G allele carriers vs AA	0.37	1.00	0.37	0.72
D4S3481	Pseudo-additive	261 vs 259 allele carriers	-2.59	1.54	-1.69	0.09
		263 vs 259 allele carriers	-1.46	0.96	-1.52	0.13
	Pseudo-dominant	259 allele carriers vs all others	1.44	1.82	0.79	0.43
	Pseudo-recessive	**263 allele carriers vs all others**	**-2.01**	**1.00**	**-2.00**	**0.05[b]**

Here, we report Beta values and Standard Errors (SE), for each genotype class compared to the reference class (see Table 3.3a, b for details), along with relevant association t-statistics (t-stat) and p-value (p). Nominally significant associations ($p < 0.05$) are highlighted in bold. No associations survived Bonferroni correction ($\alpha = 2.1 \times 10^{-3}$). [a] Note: full details on the genetic models built and on the genotype classes contrasted are reported in Table 3.3a, b (see Methods section). [b] Here, the actual p-value (0.045) was rounded to 0.05 but still labelled as nominally significant. Abbreviations: UPDRS, Movement Disorder Society revised version of the Unified Parkinson's Disease Rating Scale-Part III; MoCA, Montreal Cognitive Assessment; NMS, modified version of the Non-Motor Symptoms Scale for Parkinson Disease; AAO, PD age at onset.

Table 4.4: Genetic associations of rs356219 and D4S3481 with continuous PD endophenotypes, including a) UPDRS, b) MoCA, c) NMS and d) AAO, in Model 2 (adjusted for sex, age, PD familiarity and clinical subtype, pharmacological treatment status, years of disease and daily intake of L-Dopa).

a) UPDRS

Genetic Variant	Genetic Model	Contrast[a]	Beta	SE	t-stat	p
rs356219	Additive	AG vs AA	-1.00	1.23	-0.81	0.42
		GG vs AA	0.63	1.59	0.39	0.69
	Dominant	A allele carriers vs GG	1.17	1.44	0.81	0.42
	Recessive	G allele carriers vs AA	-0.53	1.14	-0.47	0.64
D4S3481	Pseudo-additive	261 vs 259 allele carriers	0.62	1.16	0.53	0.60
		263 vs 259 allele carriers	3.12	1.96	1.59	0.11
	Pseudo-dominant	259 allele carriers vs all others	1.09	1.10	0.99	0.32
	Pseudo-recessive	263 allele carriers vs all others	-2.43	1.62	-1.50	0.13

b) MoCA

Genetic Variant	Genetic Model	Contrast[a]	Beta	SE	t-stat	p
rs356219	Additive	AG vs AA	0.02	0.02	1.12	0.27
		GG vs AA	0.03	0.02	1.44	0.15
	Dominant	A allele carriers vs GG	0.02	0.02	1.09	0.28
	Recessive	G allele carriers vs AA	0.02	0.02	1.43	0.15
D4S3481	Pseudo-additive	261 vs 259 allele carriers	0.005	0.02	-0.32	0.75
		263 vs 259 allele carriers	0.02	0.03	0.58	0.57
	Pseudo-dominant	259 allele carriers vs all others	0.001	0.02	-0.09	0.93
	Pseudo-recessive	263 allele carriers vs all others	0.03	0.02	-1.24	0.22

c) NMS

Genetic Variant	Genetic Model	Contrast[a]	Beta	SE	t-stat	p
rs356219	Additive	AG vs AA	-6.22	4.07	-1.53	0.13
		GG vs AA	-2.55	5.41	-0.47	0.64
	Dominant	A allele carriers vs GG	0.80	4.95	0.16	0.87
	Recessive	G allele carriers vs AA	-5.22	3.80	-1.37	0.17
D4S3481	Pseudo-additive	261 vs 259 allele carriers	-4.01	3.90	-1.03	0.30
		263 vs 259 allele carriers	5.57	6.75	0.83	0.41
	Pseudo-dominant	259 allele carriers vs all others	-2.28	3.70	-0.62	0.54
	Pseudo-recessive	263 allele carriers vs all others	-1.54	5.59	-0.28	0.78

d) AAO

Genetic Variant	Genetic Model	Contrast[a]	Beta	SE	t-stat	p
rs356219	Additive	AG vs AA	0.18	1.07	0.17	0.87
		GG vs AA	0.96	1.41	0.68	0.50
	Dominant	A allele carriers vs GG	0.86	1.28	0.67	0.50
	Recessive	G allele carriers vs AA	0.39	1.00	0.39	0.70
D4S3481	**Pseudo-additive**	**261 vs 259 allele carriers**	**-2.02**	**1.00**	**-2.01**	**0.05[b]**
		263 vs 259 allele carriers	1.42	1.82	0.78	0.44
	Pseudo-dominant	259 allele carriers vs all others	-1.47	0.96	-1.53	0.13
	Pseudo-recessive	263 allele carriers vs all others	-2.60	1.54	-1.69	0.09

Here, we report Beta values and Standard Errors (SE), for each genotype class compared to the reference class (see Table 3.3a, b for details), along with relevant association t-statistics (t-stat) and p-value (p). Nominally significant associations ($p < 0.05$) are highlighted in bold. No associations survived Bonferroni correction ($\alpha = 2.1 \times 10^{-3}$). [a] Note: full details on the genetic models built and on the genotype classes contrasted are reported in Table 3.3a, b (see Methods section). [b] Here, the actual p-value (0.045) was rounded to 0.05 but still labelled as nominally significant. Abbreviations: UPDRS, Movement Disorder Society revised version of the Unified Parkinson's Disease Rating Scale-Part III; MoCA, Montreal Cognitive Assessment; NMS, modified version of the Non-Motor Symptoms Scale for Parkinson Disease; AAO, PD age at onset.

Survival analyses on LID onset

Exploratory Cox PH Models using non-genetic exposures

An exploratory multivariable Cox PH regression, aimed at investigating the relation of non-genetic covariates with the incident risk of LIDs, was performed in 300 PD cases for which all phenotypic, clinical and pharmacological information was available (case-complete approach), with a total of 102 LID events. These subjects were followed for a total of 17,434 person-months (median follow-up time 49 months). 160 observations were deleted due to missing values. Among the number of variables previously implicated in LID-onset, our multivariable Cox PH regression revealed the associations with incident risk of LIDs reported in Table 4.5 (below). In particular, among categorical variables, sex showed a significant association, with women being at higher risk of LIDs compared to men (HR [CI] = 1.75 [1.16; 2.63], p-value = 0.007). Among the continuous variables, we observed a protective effect of PD age-at-onset (0.96 [0.94; 0.99] per year increase in AAO, p = 0.006) and years of disease (0.92 [0.87; 0.97] per year increase in YOD, p = 0.002), on LID onset. Similarly, increasing MoCA score was associated with a lower LID risk (0.12 [0.02; 0.54] per 1% increase in MoCA, p = 0.006). However, none of the above mentioned associations survived a correction for multiple testing of ten different covariates ($\alpha = 5 \times 10^{-3}$), except for years of disease (YOD). Cox curves for the covariates tested in a multivariable setting (see **Table 4.5** below) are reported in **Figures 4.1a-j.**

Table 4.5: Results of the exploratory multivariable Cox PH regression modelling the relation between incident LIDs and all the non-genetic covariates previously associated with LIDs.

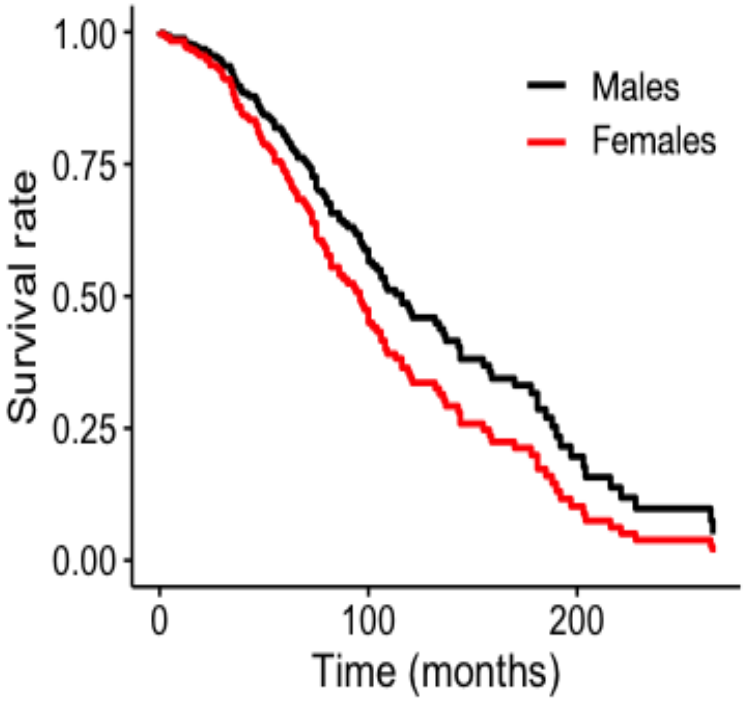
Variable	HR [CI]	z	p
Sex (F vs M)	**1.75 [1.16; 2.63]**	**2.69**	**0.007**
Familiarity (FPD vs SPD)	1.14 [0.74; 1.76]	0.61	0.543
Phenotype (Bradykinetic vs Tremorigenic)	0.72 [0.37; 1.39]	-0.99	0.322
Phenotype (Mixed vs Tremorigenic)	0.70 [0.40; 1.24]	-1.22	0.221
AAO	**0.96 [0.94; 0.99]**	**-2.77**	**0.006**
L-Dopa Dosage	1.00 [1.00; 1.00][a]	-0.37	0.71
MOCA	**0.12 [0.03; 0.54]**	**-2.76**	**0.006**
HY	1.39 [0.96; 1.99]	1.77	0.078
UPDRS	0.24 [0.04; 1.35]	-1.62	0.105
YOD	**0.92 [0.87; 0.97]**	**-3.14**	**0.002**
NMS	0.81 [0.04; 15.6]	-0.14	0.889

Here, Hazard Ratio (HR), relevant 95% Confidence Interval (CI), z-score and p-value are reported for each of the covariates tested. Variables showing nominally significant associations ($p < 0.05$) are highlighted in bold. Among these, only years of disease (YOD) survived correction for multiple testing ($p < 5 \times 10^{-3}$). [a] Note: HR and CI were rounded to two decimal places (original HR for L-dopa dosage was 0.999 [0.998-1.001].

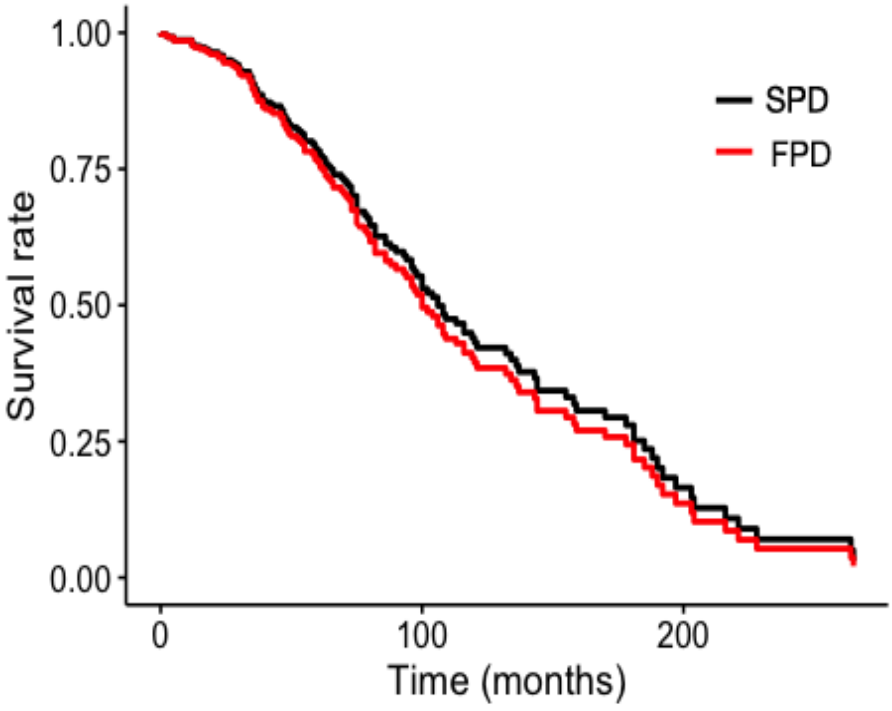
Abbreviations: AAO, PD age at onset; MoCA, Montreal Cognitive Assessment; HY, Hoehn & Yahr score; UPDRS, Movement Disorder Society revised version of the Unified Parkinson's Disease Rating Scale-Part III; NMS, modified version of the Non-Motor Symptoms Scale for Parkinson Disease; YOD, years of disease.

Figure 4.1: Cox curves of multivariable Cox PH regressions modelling incident LID risk vs a) Sex, b) PD familiarity, c) clinical subtype and d) Age-at-onset (AAO), e) L-Dopa dosage, f) MoCA, g) HY, h) UPDRS and i) NMS score, and j) Years of disease (YOD).

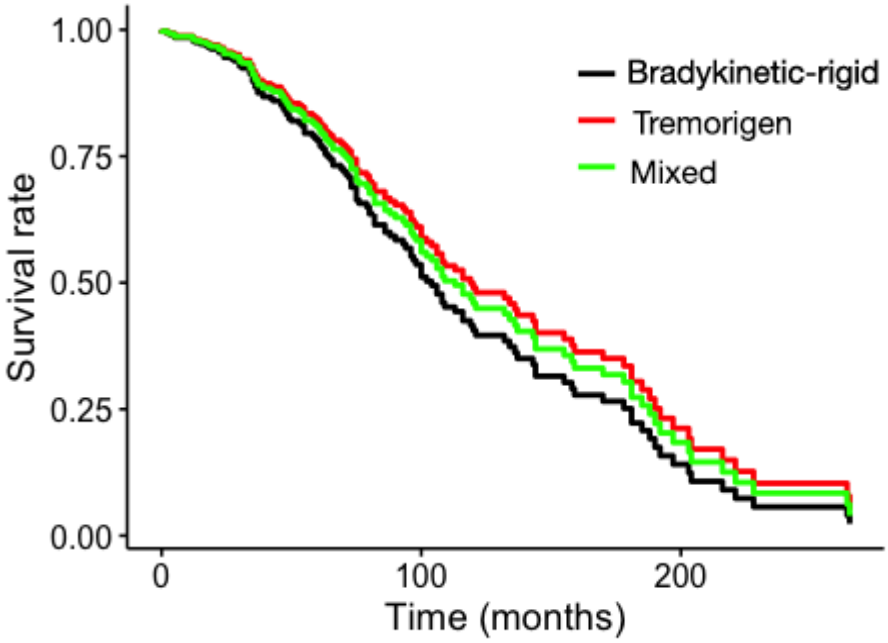
a) Sex



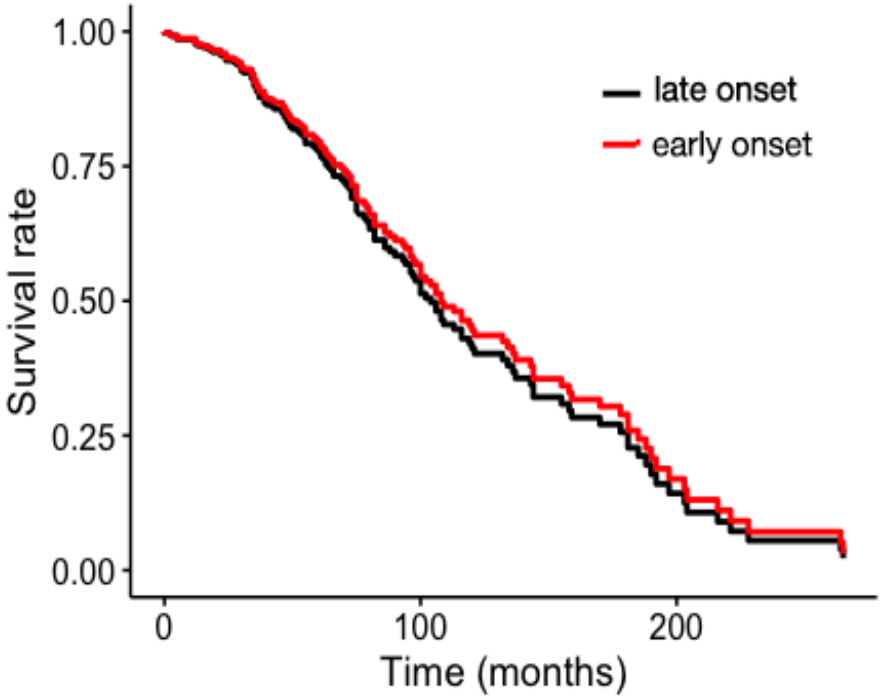
b) Familiarity



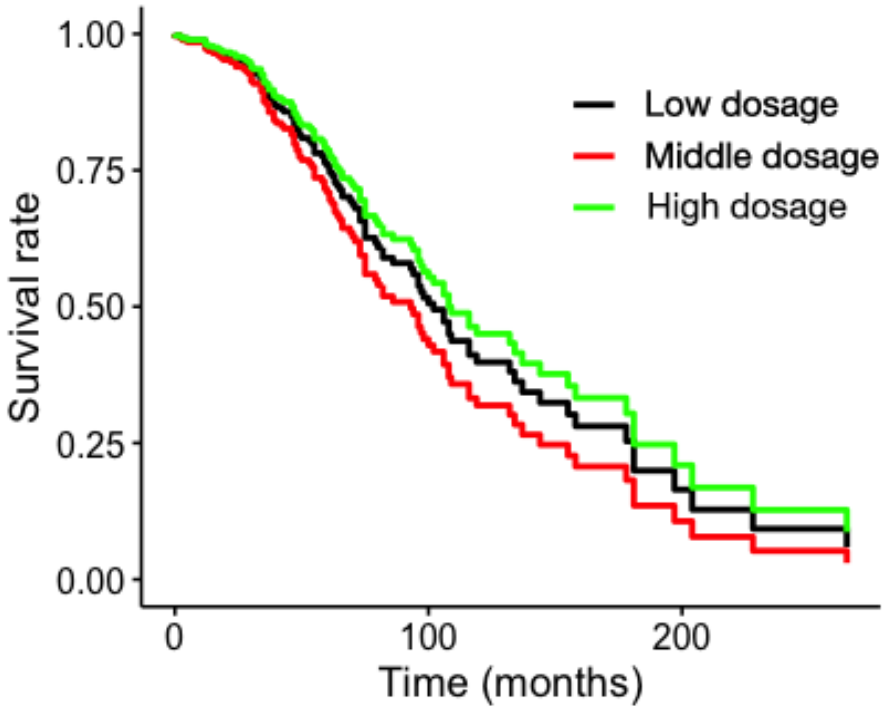
c) Clinical Subtype



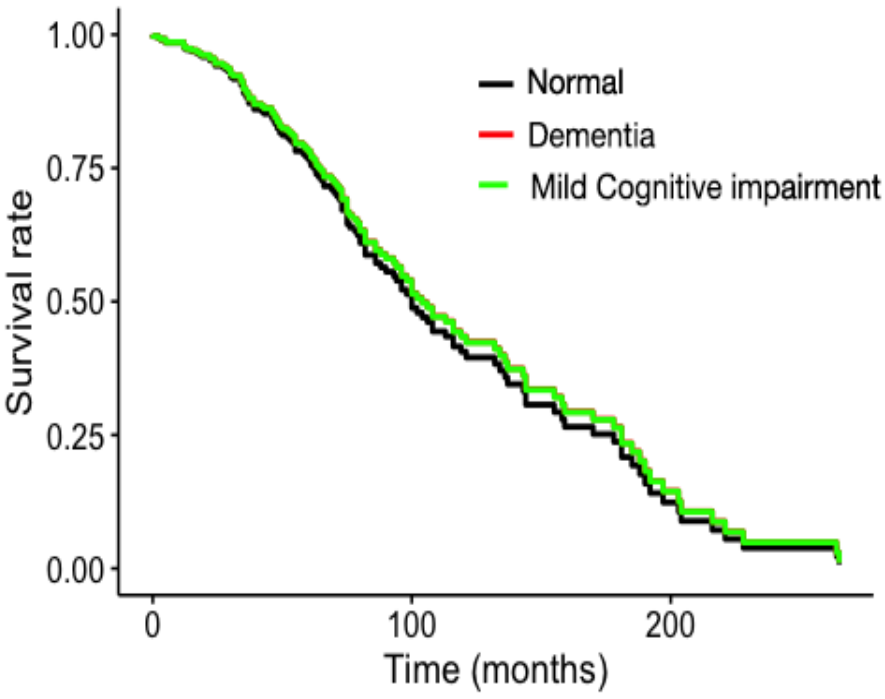
d) AAO



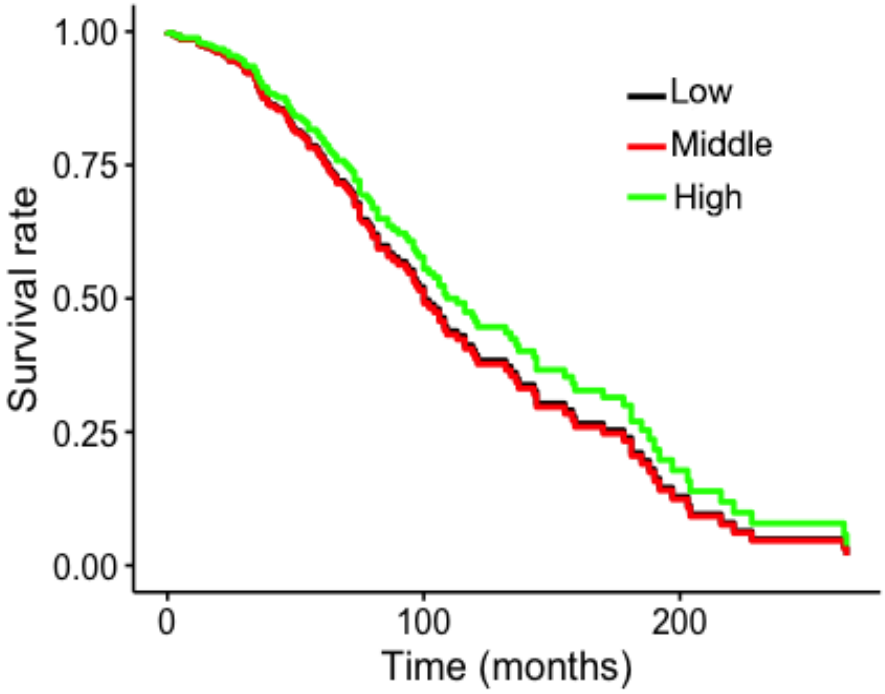
e) L-Dopa Dosage



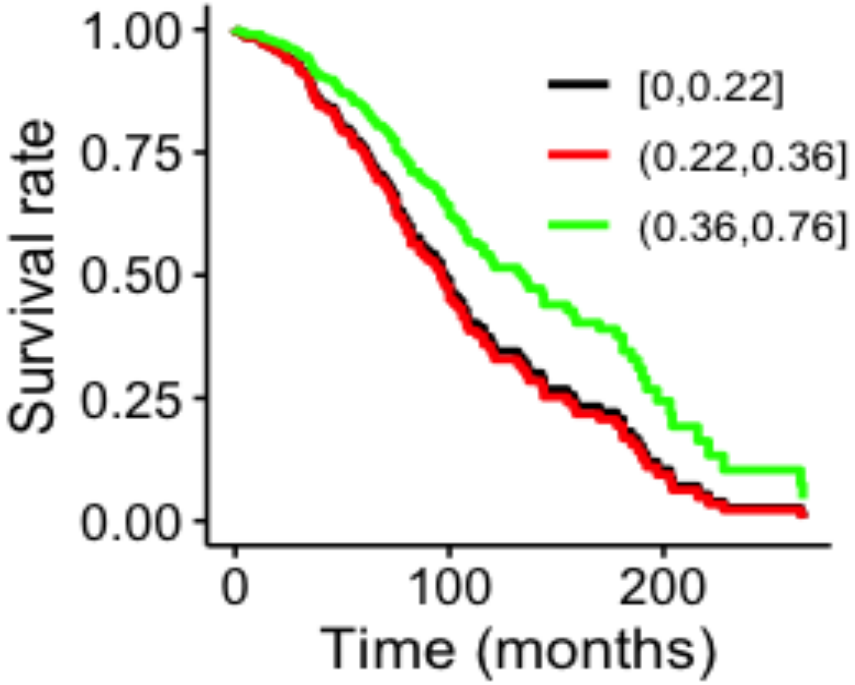
f) MoCA



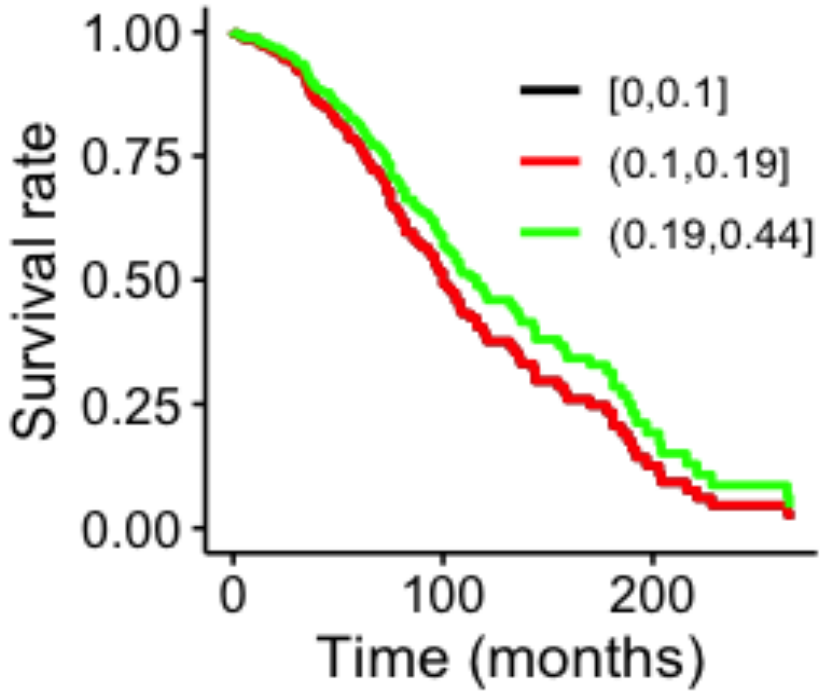
g) HY


1.00
0.75
0.50
0.25
0.00
Survival rate
0
100
200
Time (months)
Low
Middle
High

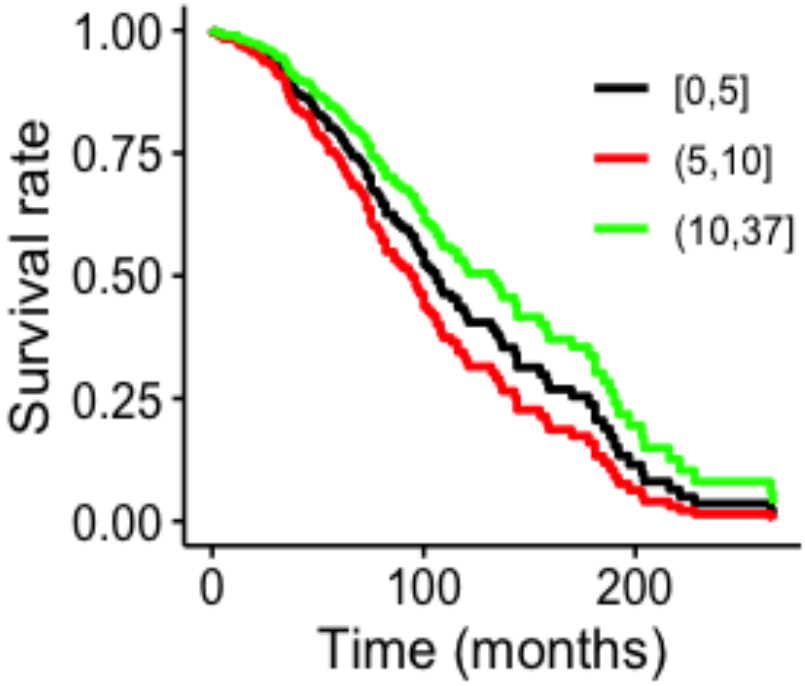
h) UPDRS


1.00
0.75
0.50
0.25
0.00
Survival rate
0
100
200
Time (months)
[0,0.22]
(0.22,0.36]
(0.36,0.76]

i) NMS


1.00
0.75
0.50
0.25
0.00
Survival rate
0
100
200
Time (months)
[0,0.1]
(0.1,0.19]
(0.19,0.44]

j) YOD


1.00
0.75
0.50
0.25
0.00
Survival rate
0
100
200
Time (months)
[0,5]
(5,10]
(10,37]

Note: to allow the *ggadjusted()* function to build the Cox curves, continuous variables were converted into categories. In particular, raw MoCA scores were classified into normal cognitive performance (MoCA ≥ 26), Mild Cognitive Impairment (MCI, 19 ≤ Moca ≤ 25) and Dementia (MoCA ≤ 18) (https://www.mocatest.org/faq/); AAO classes were defined based on common definitions of late (AAO > 50 years) and early PD onset (AAO < 50 years) (https://www.malacards.org/card/parkinson_disease_late_onset); L-Dopa intake was classified into a high (≥ 400mg/day), a moderate (401mg/day ≤ L-Dopa ≤ 601mg/day) and a low dosage class (≥ 600mg/day), as described in (205,206); Hoehn & Yahr staging was classified into initial (HY ≤ 1), medium (HY = 2) and advanced stage of the disease (HY ≥ 3), as in (207), . Where no previous classification had been proposed (i.e. for UPDRS and NMS scores, and years of disease), continuous variables were ranked into tertiles and the resulting classes were compared (see relevant Cox curves for details on intervals).

Abbreviations: AAO =, PD age at onset; MoCA, Montreal Cognitive Assessment; HY, Hoehn & Yahr score; UPDRS, Movement Disorder Society revised version of the Unified Parkinson's Disease Rating Scale-Part III; MoCA, Montreal Cognitive Assessment; NMS, modified version of the Non-Motor Symptoms Scale for Parkinson Disease; YOD, = years of disease.

SNCA polymorphisms and incident LID risk

The results of Cox PH regressions modelling LID onset as a function of the candidate genetic variants tested in the *SNCA* gene are presented below.

For rs356219, we applied these models to 296 PD cases for which all phenotypic, clinical, pharmacological and genetic information was available (164 observations deleted), with a total of 101 LID events. Total follow-up time was 17,226 person-moths (median 50 months). For D4S3481 (REP1), the regression was performed on 298 PD cases (101 LID events, 162 observations removed through case-complete approach), which were followed-up for 17,367 person months (median 50 months).

In the unadjusted models testing the relation between incident LID risk and each of the genetic variants (Model 1; Table 4.6), we observed a nominally significant association between D4S3481-263 bp allele carriers and LID onset (HR [CI] = 0.56 [0.32; 0.98], p = 0.04). However, this did not survive correction for multiple testing of two independent variants and three alternative genetic models tested for each variant ($\alpha = 8.3 \times 10^{-3}$).

No other significant association was observed for any of the alternative genetic models tested for D4S3481, nor for any of the genetic models tested for rs356219 (Table 4.7, unadjusted models). Similarly, when we analysed rs356219 and D4S3481 in conservative models fully adjusted for all the covariates previously associated with LID onset (Model 2, i.e. our model of reference for interpreting the results), we observed no significant association with incident risk of LID, neither

for rs356219 nor for D4S3481 (see Table 4.7). Cox curves of fully adjusted Cox regressions under different genetic models for the two genetic variants tested are reported in **Figure 4.2a-c** for rs356219 and in **Figure 4.3a-c** for D4S3481, while Kaplan-Meier curves of the unadjusted models are reported in **Figures S5 a-c** and **S6 a-c**, respectively (see Supplementary Results in Appendix S3).

Table 4.6: Results of univariate unadjusted Cox PH regressions modelling incident LID risk vs rs356219 and D4S3481 (Model 1).

Genetic Variant	Genetic Model	Contrast[a]	HR [CI]	z	p
rs356219	Additive	AG vs AA	0.98 [0.68; 1.41]	-0.12	0.90
		GG vs AA	1.03 [0.64; 1.66]	0.13	0.90
	Dominant	A allele carriers vs GG	1.04 [0.68; 1.60]	0.20	0.84
	Recessive	G allele carriers vs AA	0.99 [0.70; 1.40]	-0.04	0.97
D4S3481	Pseudo-additive	261 vs 259 allele carriers	0.75 [0.52; 1.08]	-1.56	0.12
		263 vs 259 allele carriers	1.33 [0.68; 2.62]	0.84	0.40
	Pseudo-dominant	259 allele carriers vs all others	0.80 [0.57; 1.14]	-1.22	0.22
	Pseudo-recessive	**263 allele carriers vs all others**	**0.56 [0.32; 0.98]**	**-2.05**	**0.04**

Hazard Ratio (HR), relevant 95% Confidence Interval (CI), z-score (z) and p-value (p) are reported for the genetic variants and each genetic model tested. Variables showing nominally significant associations ($p < 0.05$) are highlighted in bold. None of these genetic models survived correction for multiple testing ($\alpha = 8.3 \times 10^{-3}$). [a] Note: full details on the genetic models built and on the genotype classes contrasted are reported in Table 3.3a, b (see Methods section).

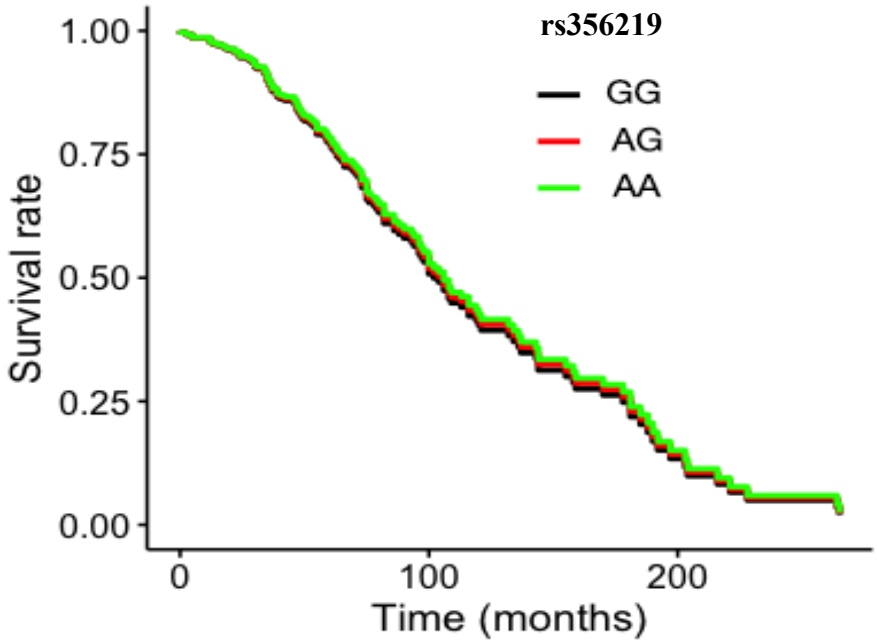
Table 4.7: Results of multivariable Cox PH regressions modelling incident LID risk vs rs356219 and D4S3481 (Model 2).

Genetic Variant	Genetic Model	Contrast[a]	HR [CI]	z	p
rs356219	Additive	AG vs AA	0.85 [0.53; 1.38]	-0.64	0.52
		GG vs AA	1.07 [0.58; 1.96]	0.22	0.83
	Dominant	A allele carriers vs GG	1.15 [0.66; 2.02]	0.51	0.61
	Recessive	G allele carriers vs AA	0.91 [0.59; 1.42]	-0.40	0.69
D4S3481	Pseudo-additive	261 vs 259 allele carriers	0.76 [0.49; 1.19]	-1.20	0.23
		263 vs 259 allele carriers	0.76 [0.29; 1.95]	-0.58	0.56
	Pseudo-dominant	259 allele carriers vs all others	0.76 [0.50; 1.16]	-1.26	0.21
	Pseudo-recessive	263 allele carriers vs all others	0.78 [0.39; 1.56]	-0.70	0.48

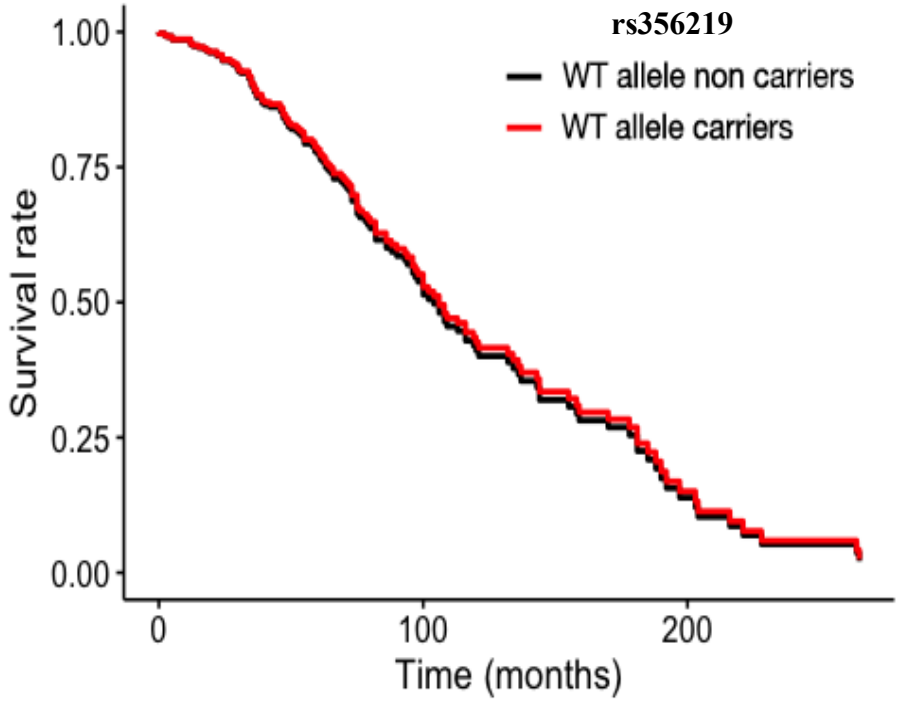
Hazard Ratio (HR), relevant 95% Confidence Interval (CI), z-score (z) and p-value (p) are reported for the genetic variants and each genetic model tested. Variables showing nominally significant associations ($p < 0.05$) are highlighted in bold. None of these genetic models survived correction for multiple testing ($\alpha = 8.3 \times 10^{-3}$). [a] Note: full details on the genetic models built and on the genotype classes contrasted are reported in Table 3.3a, b (see Methods section).

Figure 4.2: Cox curves of fully adjusted Cox PH regressions modelling incident LID risk vs rs356219 in a) Additive b) Dominant and c) Recessive model (see Table 3.3a for details).

a) Additive Model



b) Dominant Model



c) Recessive Model

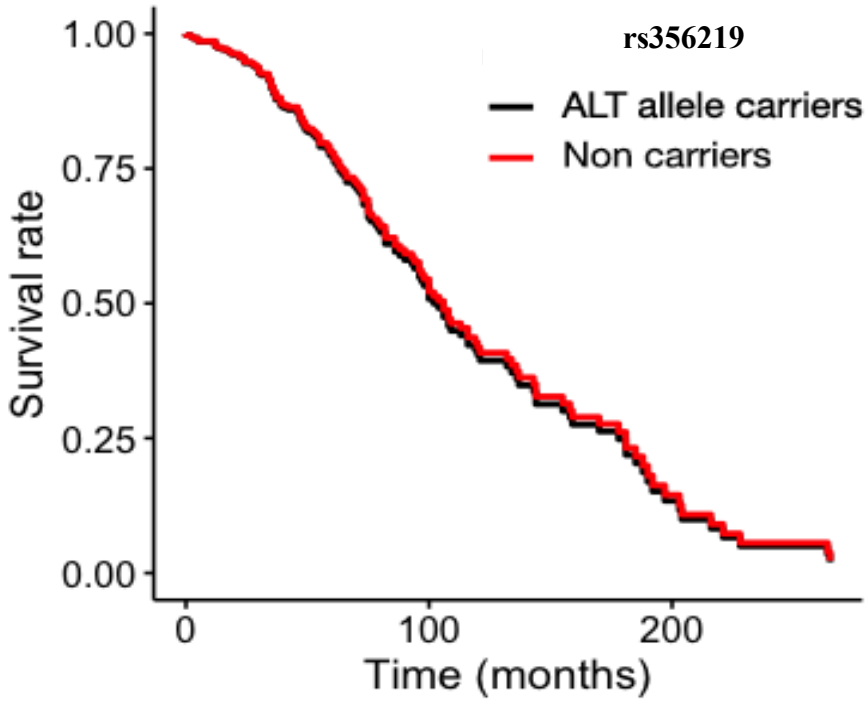
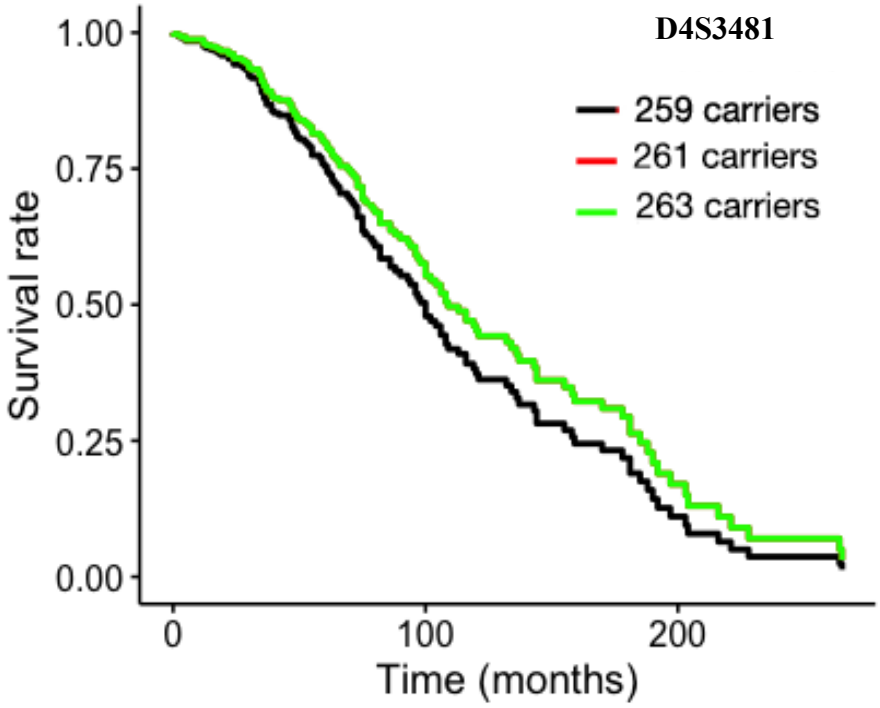
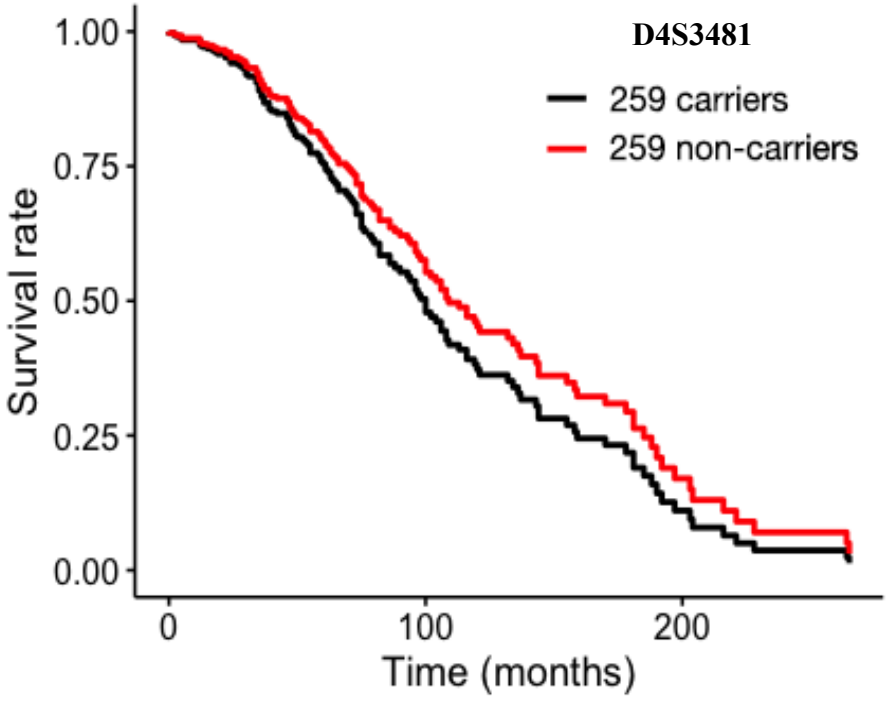


Figure 4.3: Cox curves of fully adjusted Cox PH regressions modelling incident LID risk vs D4S3481 in a) Pseudo-additive b) Pseudo-dominant and c) Pseudo-recessive model (see Table 3.3b for details).

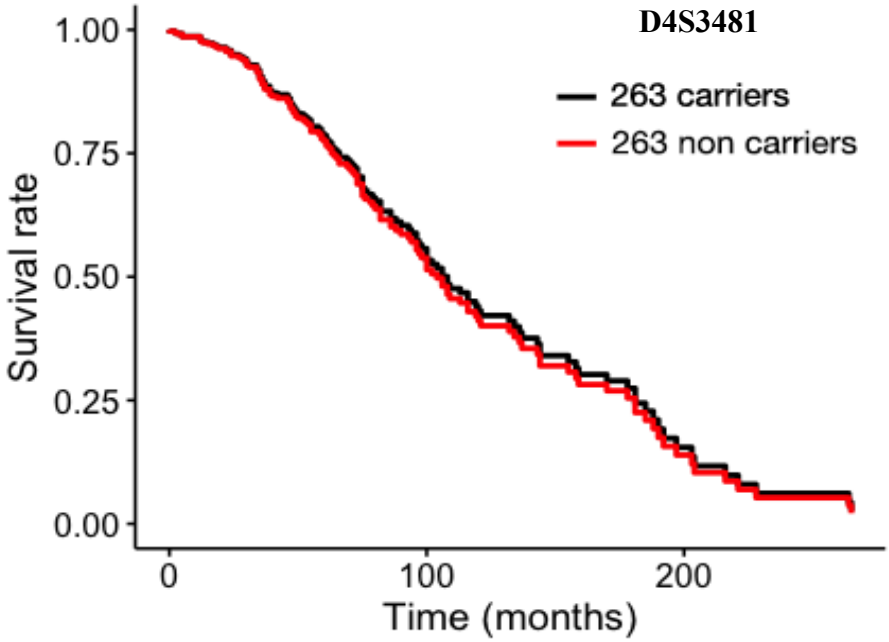
a) Pseudo-Additive Model



b) Pseudo-Dominant Model



c) Pseudo-Recessive Model



Identification of rare mutations with potential effects on LID onset

Our exome-wide approach aimed at the identification of rare genetic variants potentially affecting LID onset did not reveal any participant with absent LID in spite of high L-Dopa intakes, nor patients showing LID occurrence at low L-Dopa intakes (see **Figure 3.4a, b** in the Methods section). Therefore, no rare variant could be detected with potential risk/protective effect on LID onset.

Chapter 5

DISCUSSION

In this dissertation, we present a comprehensive genetic analysis of one of the largest cohorts of Parkinson Disease (PD) patients available in Italy (N=470). This entailed a focused analysis of two known PD susceptibility variants in the *SNCA* gene which have been associated with increased levels of expression of the gene both in the plasma and in brain tissues, namely the microsatellite D4S3481 (commonly known as REP1), and the SNP rs356219 (74-77, 85-88). Variants were initially tested for association with PD risk and related continuous endophenotypes, which included motor (UPDRS), cognitive (MoCA) and other nonmotor symptoms (NMS), as well as PD age-at-onset (AAO). Then, these variants were investigated for potential genetic influence on the incident risk of L-Dopa induced dyskinesias (LIDs), in addition to an exploratory analysis of all non-genetic risk factors which have been previously associated with LID onset. In addition to survival analyses on candidate *SNCA* genetic variants, we looked for rare mutations potentially conferring risk/protection against LID events at the exome-wide level, in a subset of 112 PD patients which had undergone Whole Exome Sequencing (WES) analyses (57).

Case-control analyses and associations with continuous PD endophenotypes

Case-control association analyses revealed no statistically significant associations, neither for rs356219 nor for D4S3481, although they both showed trends of associations ($p < 0.1$) of the putative risk alleles (G for rs356219 and 263 bp for D4S3481), in line with previous meta-analyses (86,90,95-97). The lack of evidence of association in our study may be due to different reasons, including the low power of the analysis due to the use of mostly unscreened controls and of relatively small sample sizes (465 cases vs 516 controls for rs356219, and 464 cases vs 487 controls for D4S3481). However, it is worth to note that other previous studies reported no significant associations of these markers with PD risk (90,208). Of note, both the polymorphisms tested in the present dissertation have been already analysed in an independent genetic study of 904 patients and 891 controls from the Italian population (86). Trotta and colleagues observed nominally significant associations for both markers only in crude association models, with directions of effect substantially consistent with those detected here. However, these associations disappeared after adjustment for sex, smoke and coffee consumption (86), in line with our sex-adjusted model.

Similarly, association analyses of rs356219 and D4S3481 with continuous PD endophenotypes only revealed nominally significant associations of the D4S3481-263 bp allele with cognitive performance (MoCA score) and PD age-at-onset (AAO) in a basic model adjusted for age and sex, which were not robustly supported in a conservative model further adjusted for PD familiarity,

clinical subtype, pharmacological treatment status, years of disease and daily intake of L-Dopa (Model 2). Model 2 revealed another nominally significant association between D4S3481 and AAO, where 261 bp allele carriers showed a lower AAO, compared to 259 bp allele carriers. Interestingly, 261 bp allele was the most associated D4S3481 allele with increased PD risk in a previous independent study of Italian PD case-control study mentioned above (86). However, it is worth to underline that the associations that we observed would not survive correction for multiple testing, and that 263 bp allele carriers did not show any association with AAO in our fully adjusted model. More in general, the lack of statistically significant and robust associations detected with continuous PD endophenotypes may be explained by different reasons. First, these analyses were carried out only within cases, which entailed a further reduction in sample size (N_{max}~450), hence in power. Second, previous studies testing associations with continuous PD traits have reported contrasting results. This especially applies to D4S3481, where the 263bp allele has been associated with faster disease progression, both for motor (92,208) and for non-motor symptoms (149), while other studies have reported inverse associations with motor and cognitive outcomes or no association at all (208). Although rs356219 has been less investigated with reference to PD endophenotypes, rs356219-G allele was associated with an increased risk for cognitive impairment (96), and with an earlier age at onset of the disease (101).

Analysis of L-Dopa induced dyskinesia (LID) risk

Analysis of incident risk of LIDs revealed interesting insights into their aetiology.

First, an exploratory analysis of non-genetic factors previously implicated in LID onset showed a significant association with years of disease (YOD). Patients with longer duration of disease - which is only partly dependent on age at onset - are more likely to develop LIDs (3,121,127,137). However, it is interesting to notice that in our case a longer duration of disease was associated with a protective effect, which is in contrast with positive associations previously reported (3,135,136). This may be explained by the multivariable setting of our exploratory analysis, where also PD age-at-onset (AAO) showed a significant association in the expected direction (see below for further discussion).

Second, analysis of covariates reported also significant associations of incident LID risk with sex, with women being at higher risk of LIDs compared to men, in line with previous studies (124,129). Beyond confirming sex as an important risk factor for LIDs, we observed a protective effect of AAO - the later the onset, the lower LID risk - and of cognitive performance (MoCA score) - the higher the performance, the lower LID risk. These findings support previous observations

reporting negative associations of LID risk with AAO (123–125,209) and dementia as a predictor of later LID occurrence in PD patients (210), although scales of cognitive performance have never been tested with incident LID risk. Importantly, these associations were observed in a multivariable setting and were all independent on each other, and did survive a conservative correction for multiple testing, except for YOD. Of note, we observed no evidence for an association between L-Dopa daily intake and incident LID risk, in spite of previous literature reporting it as one of the most important risk factors for LIDs (121,124,196). This may be due to the fact that different studies analysed differently exposure to L-Dopa intake. Indeed, some reported association with L-Dopa dosage at the beginning of the pharmacological treatment (ref), while others took the latest prescription as dosage of reference (199), and other works analysed Levodopa Equivalent Daily Dosage (LEDD, see *Strengths and Limitations* paragraph below for further discussion). Indeed, the prevalence of LID increases with disease and treatment duration, and usually, it takes approximately 3–5 years after administering L-Dopa for developing dyskinesias (211). Therefore, duration of treatment may also explain part of variance in LID occurrence and may represent an unaccounted factor in our analysis, since we were not able to trace the beginning of L-Dopa treatment for many of our patients.

When we examined candidate *SNCA* genetic variants, we only observed a protective effect of 263 bp allele of D4S3481 against incident LID risk, compared to carriers of all other alleles in crude models. However, this association was only nominally significant and was not confirmed by models fully adjusted for all the covariates previously associated with LID risk. More in general, we observed no significant associations with incident risk of LIDs in fully adjusted models, neither for rs356219 nor for D4S3481. At present, it is difficult to say whether this is due to the total lack of influence of these two variants - or possibly of the *SNCA* gene as a whole - on the occurrence of LIDs, since these variants and the *SNCA* gene have been under-investigated with this regard. Indeed, we are not aware of any study testing association of rs356219 with LIDs, neither with prevalent nor with incident risk. While only a recent study tested the influence of D4S3481 on the incident risk of L-Dopa motor complications in an independent Italian PD cohort (426 patients), reporting no significant effects for the 263 bp allele (149). Overall, further genetic studies on these and other *SNCA* variants are warranted to clarify the relation of this gene with LID onset and risk, which has been fairly neglected so far. More in general, if and how the inherited predisposition to PD affects the development of LIDs represent currently a largely unanswered and under-investigated issues, both in candidate gene studies and in genome or exome-wide studies with no a priori hypotheses (see *Genetics of LIDs* subsection in the Introduction section). For this reason, in this dissertation we also attempted to identify rare mutations conferring a potential risk or

protective effect on LID occurrence, among those PD patients which had undergone WES analyses in our PD cohort. However, this analysis did not reveal any PD patient with absent LID in spite of high L-Dopa intakes, nor patients showing LID occurrence at low L-Dopa intakes. Therefore, we could not proceed in the lookup for private mutations conferring protective/risk effect on LID onset. Again, this may be due to different factors, e.g. the unavailability of LEDD or the lack of adjustment of L-Dopa intake for body weight, which was not available. Future analyses will possibly attempt to include adjustment for body weight to better reflect bio-availability of L-Dopa within each patient (see below), after collecting anthropometric information in a recall of the cohort.

Strengths and Limitations

Our study presents different points of strengths, but also different limitations.

One of the main strengths of this dissertation is that we report a comprehensive analysis of D4S3481 and rs356219 with reference to PD risk, testing associations under different genetic models with PD case-control status and with relevant endophenotypes, which are known improve the power to detect genetic associations with complex disorders (102). Some of these continuous traits assessing motor, cognitive and other non-motor symptoms have been already analysed through an exome-wide association scan in a subset of our cohort and, although we did not observe any exome-wide significant association, polygenic scores associated with increasing subcortical volumes revealed interesting associations with motor symptoms (57), supporting their use in genetic analyses.

Moreover, we analyzed the relation between *SNCA* variants and incident LID risk, which so far has been mostly neglected. Indeed, although previous evidence suggested a potential implication of α-synuclein in motor complications connected to L-Dopa treatment both in human (189) and in animal studies (190) (see *Genetics of LIDs* paragraph in the Introduction section for details), the association between *SNCA* variants and LIDs has been underinvestigated so far, with only one study analysing the effect of D4S3481 on incident LIDs and reporting negative findings (149).

Third, the wealth of clinical, pharmacological and neurological information available in our cohort - which was rescued through passive follow-up - allowed us to robustly adjust association tests with continuous traits and with incident LID risk. Such a complete information is not commonly available in large scale studies, which usually result for meta-analysing different cohorts with different designs and phenotypic assessment, often representing a hindrance to power.

Finally, the longitudinal design of our PD cohort study allows us to potentially extend the follow-up to a very long time range, which will entail a further increase in power of survival analyses as the number of LID events increases.

In spite of these strengths, our study also presents several limitations.

First, the lack of availability of a high number of properly assessed neurological controls may have limited power of the case-control analysis. Indeed, we had only 58 neurological controls, which were made available in a second phase to increase the number of genotyped controls in the analysis. Sadly, this is a limitation often found in genetic studies of age-related neurodegenerative disorders such as PD (e.g. (47,57)), since such controls need to be free of any neurological sign or symptom of the disorder at a quite advanced age (usually above 70 years). Power may have been limited also by the sample size of our analysis, which was still considerable, compared to other studies (149,212,213), both for case-control association test (about 470 cases vs 500 controls) and for the analysis of continuous PD-related endophenotypes within cases (with a sample size ranging between 412 and 432). This applies also to survival analyses on incident LID risk, especially because some of the genotype classes compared (e.g. D4S3481-263 bp allele carriers) showed quite small numbers (N=53).

Further specific limitations of the survival analysis on incident LID risk may have partly affected our power, such as the lack of weight information available for the PD patients. Indeed, body weight is known to influence the levels of L-Dopa bio-availability in the organism (121,124,133), therefore this covariate may help to partial out the bias introduced by simply considering L-Dopa dosages, irrespective of the body mass of participants. The retrospective design of our study and the passive follow-up did not allow us to rescue this information, but we are planning to collect anthropometric measures in future active follow-up recruitments. Nonetheless, we believe this potential bias in our Cox models was limited at least in part by sex adjustment, since women usually report lower body mass than men. Similarly, computing a Levodopa equivalent daily dose (LEDD) for each participant may have helped to have a slightly more precise and comparable information to sum dopamine coming from different sources (e.g. carbidopa). Although different approaches have been suggested to compute LEDD, no agreement has been reached on a gold standard procedure (137,214) and different studies report different L-Dopa dosage exposures (see above). E.g., some studies reported significant associations of incident LID risk with initial L-Dopa dosage and with the duration of pharmacological treatment, which may also represent important sources of information affecting LID occurrence in our cohort. To avoid over-adjustment of already conservative Cox models and since initial doses and duration of L-Dopa

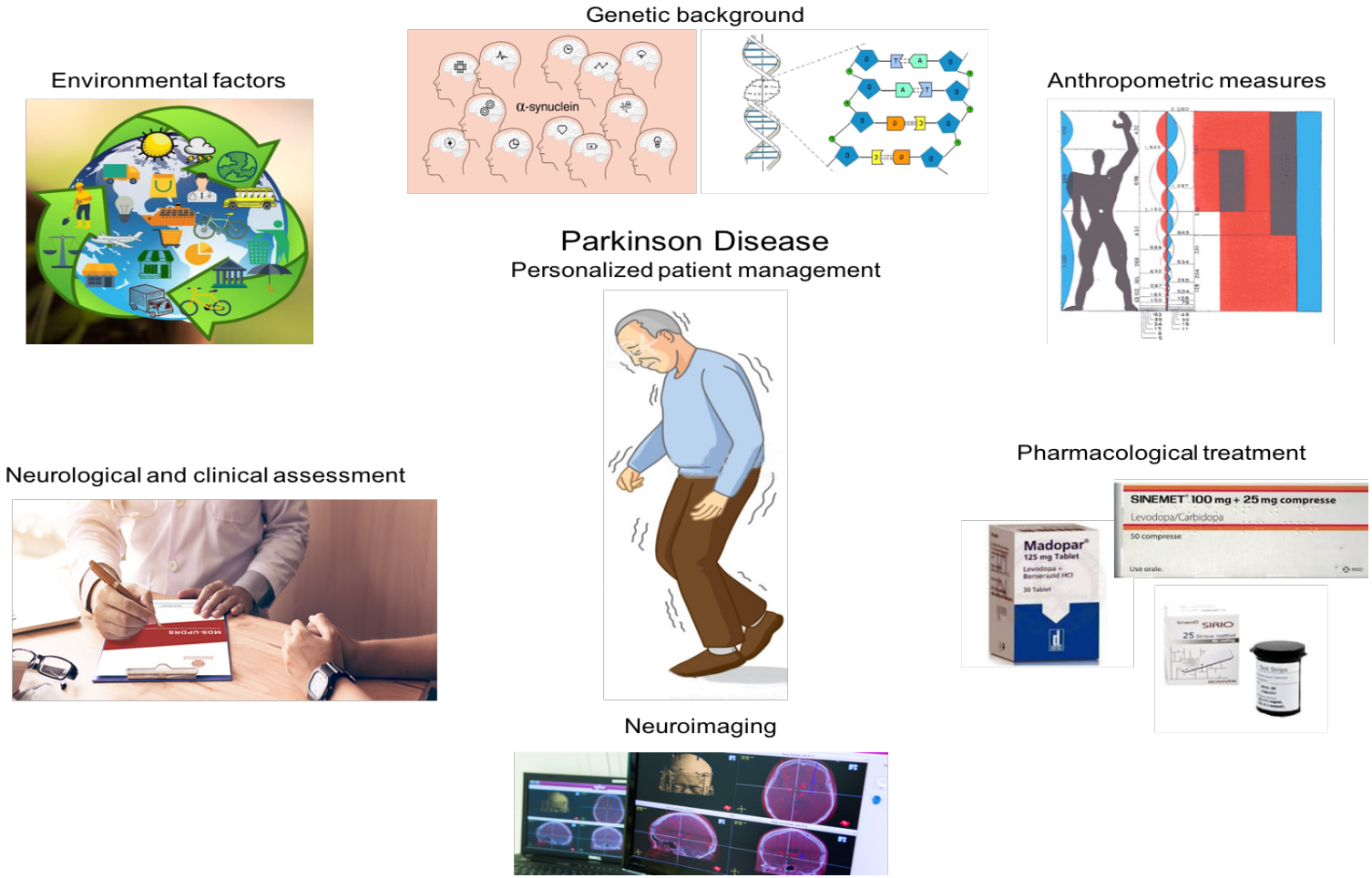
treatment were not available for all the patients, we decided to use here the most recent L-Dopa daily intake prescribed by the neurologist as the most immediate proxy of L-Dopa dosage and of the increased LID risk due this exposure. Future studies in this cohort will allow us to elaborate more refined models which may potentially take into account LEDD, possibly at different time points as the follow-up time becomes longer.

Finally, the assessment of dyskinesia made by qualified neurologists only reported the absence/presence of motor complications, hence missing precious information on the time spent with or without LIDs in the different stages of the disease, as well as on the severity of motor complications. While assessing LIDs through a dedicated scale (e.g. Rush Dyskinesia Rating Scale" (RDRS)) (215) would have helped to have continuous LID-related traits available for association analyses and a more precise outcome assessment, the nature of the phenomenon and the longitudinal retrospective design of our study make it unlikely that this information would have notably improved our power to detect significant influences on incident LID risk.

Conclusions and future perspectives

Although most of the findings of the present dissertation are only nominally significant and warrant further analyses in larger and/or independent datasets, we believe this work provides a substantial contribution to the investigation of the genetic underpinnings of PD, related endophenotypes and motor complications of L-Dopa treatment, presenting previously unreported analyses like association tests of known PD susceptibility variants within *SNCA* with cognitive and non-motor symptoms, and survival analyses to test their genetic influence on incident LID risk. Our aim is working towards improving aspects potentially limiting power of the analysis, such as the relatively low sample size, the scarcity of neurological controls and the relatively short follow-up time, to assess the robustness of the suggestive findings reported here and further improve our comprehension of the relation between *SNCA* and PD risk, symptoms and pharmacological treatment. This will represent an important translational milestone in developing future personalized strategies for the diagnosis, treatment and management of PD patients in the future (**Figure 4.1**)

Figure 4.1: Personalized PD patient management



Here, a brief representation of the future perspectives in term of personalized diagnosis, treatment and management of PD patients is reported.

Appendices 1: Theoretical and practical priming to the methods used

TaqMan probe-based assays

The TaqMan® probe principle relies on the 5′–3′ exonuclease activity of Taq polymerase to cleave a dual-labelled probe during hybridization to the complementary target sequence and fluorophore-based detection. [2] TaqMan probes consist of a fluorophore covalently attached to the 5'-end of the oligonucleotide probe, and of a quencher at the 3'-end [4], and are specifically designed to anneal within a DNA region amplified by a specific set of (custom) primers. As the Taq polymerase extends the primer and synthesizes the nascent strand, the 5' to 3' exonuclease activity of the Taq polymerase degrades the probe that has annealed to the template. Then degradation of the probe releases the fluorophore and interrupts the close proximity to the quencher, thus relieving the quenching effect and allowing the fluorophore to emit fluorescent signals. Hence, fluorescence detected in the quantitative PCR thermal cycler is directly proportional to the fluorophore released and the amount of DNA template present in the PCR. Based on the relative intensity of the two fluorescent signals, best-guess genotypes can be called for a given Single Nucleotide Polymorphism (SNP) through a proprietary software (CFX96 Touch™ Real-Time PCR Detection Systems, in our case).

Regression Models

One of the main purposes of statistical analysis of medical data is often to describe relationships between two or more variables. This is often done through associations testing, which provides a measure of the extent of statistical dependence between a dependent variable (or outcome, which in our case is represented by PD status, continuous PD endophenotypes or the occurrence of dyskinesias) and one or more independent variables (or exposure/s, e.g. the genetic variants analysed in the present dissertation). Associations can be tested through different approaches, depending on the nature of the outcome and of the exposure, and on the kind of relation we want to investigate *(*see *Box 1)*.

Regression analysis is a type of statistical evaluation that enables:

● Description: Relationships among the dependent variables and the independent variables can be statistically described by means of regression analysis.

● Estimation: The values of the dependent variables can be estimated from the observed values of the independent variables.

● Prognostication: Risk factors that influence the outcome can be identified, and individual prognoses can be determined.

Regression usually employs a model that describes the relationship between the dependent variable and the independent variable/s in a simplified mathematical form. There may be biological reasons to expect *a priori* that a certain type of mathematical function will best describe such a relationship, or simple assumptions have to be made (e.g., that blood pressure rises linearly with age). The most frequently used types of regression analyses include

● Linear regression,

● Logistic regression,

● Cox regression,

and are described in detail in the Table S1 below.

Table S1. Main types of regression models.

Regression Model	Application	Dependent variable	Independent variables
Linear Regression	Modelling a linear relationship	Continuous (e.g. cognitive performance measured through a psychometric scale)	Continuous and/or categorical
Logistic Regression	Prediction of the probability of belonging to a given group (binary outcome)	Dichotomous (e.g. PD case/control)	
Cox Regression	Modelling of survival data	Occurrence of an event of interest (e.g. dyskinesia) and time-to-event	

Logistic regressions for case-control association testing

In case-control analysis, we use logistic regression to compare either allele or genotypic frequencies between cases and controls. This analysis allows to compute Odds Ratios (OR), which

represent the ratio between Odds of an event (e.g. PD affection status, in our case) compared to the Odds of the absence of the event.

$$Odds = p/(1-p)$$

To calculate the Odds, we need to apply the following formula:

$$Logit(Odds) = a + \beta x + e$$

Where *x* is the independent variable, *a* is the intercept of the regression, *β* is the slope of the logistic regression, and *e* is the residual error term, which is not directly observed in data. To obtain the OR of the association between the independent variable and the logit function, we need to exponentiate the Beta (*β*) value resulting from the logistic regression.

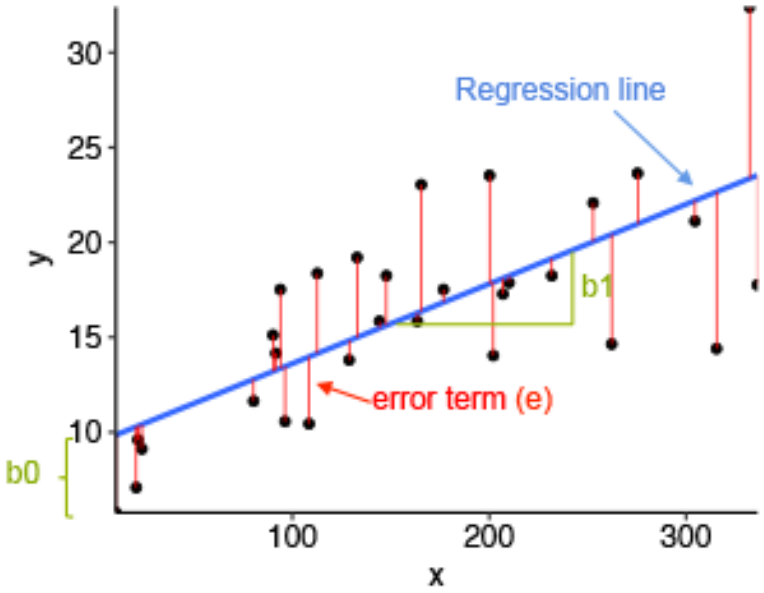
Linear regressions for testing associations with continuous variables

Similarly, in the linear regression we compute the slope of the regression line between the independent variable/s (x) and the dependent continuous variable (y). Assuming a single independent variable (x), this is accomplished through the formula:

$$y = a + \beta x + e$$

Where *a* is the intercept of the regression, β is the slope of the regression line, and *e* is the residual error term, which is not directly observed in data. In the linear regression, β represent the effect size and indicates the direction of association, between x and y, which is considered as significant if $\beta \neq 0$.

Figure S1.1. Example of a linear regression plot.



The figure above illustrates a linear regression model: the best-fit regression line is in blue, the intercept (b0) and the slope (b1) are shown in green and the error terms (e) are represented by vertical red lines. Image courtesy of Simple Linear Regression in R (www.sthda.com).

Cox regressions for survival analysis

Survival analysis

Survival analysis is the analysis of time-to-event data, which describe the length of time from a time origin to an endpoint of interest. Survival analysis methods are used to analyse data collected prospectively in time, such as data from a prospective cohort study or data collected for a clinical trial, where patients are followed-up over a given period of time (216). Within survival analysis, the dependent variable is composed of two attributes: one is a categorical variable which records if the event of interest occurred or not (in our case, the occurrence of dyskinesia), while the other one is the time to the event. An endpoint happens either when the event verifies or when the follow-up time has ended, in which case censoring is applied. Observations are defined as censored when the information about their survival time is incomplete. The most common encountered form of censoring is *right censoring*. If a patient does not experience the event of interest for the duration of the study, or when someone drops out of the study before the end of the observation time and without experiencing the event, he is defined as *right censored* and the survival time for this subject is considered to be at least as long as the duration of the study/observation time. Another type of censoring is *left censoring*, which takes place when an individual is known to have had the event before a specific time, but that could be any time before the censoring time. It is also possible

to have interval censoring where an individual is only known to have had the event between two time points, but the exact time of event is not observed. Censoring is an important issue in survival analysis, representing a particular way to treat missing data, and is usually required in order to avoid bias in the analyses.

Unlike ordinary regression models, survival methods correctly incorporate information from both censored and uncensored observations in estimating important model parameters. Then, it is possible to estimate two functions that are dependent on time: the survival and the hazard function. The survival and the hazard function represent key concepts in survival analysis to describe the distribution of time to event (*T*). The survival function *S(t)* gives, for every time point (*t*) since the start of follow-up, the probability of experiencing the event after that time. This can be described as

$$S(t) = P(T > t) = 1 - F(t), t > 0$$

Where $F(t)$ represents the Repartition function

$$F(t) = \Pr(T \leq t) = 1 - S(t),$$

namely the probability that the event occurs within time *t*.

The Hazard function *h(t)* represents the instant risk that the event will occur at a given time point (*t*), provided that an individual has not experienced the event up to that specific time.

It can be described as:

$$h(t) = \lim_{dt \to 0} \frac{\Pr\,(t \leq T < t + dt | \mathrm{T} \geq \mathrm{t})}{dt}{}' t \geq 0$$

where T represents the time to event and the interval *[t; t + dt]* represents an infinitesimal variation in follow-up time.

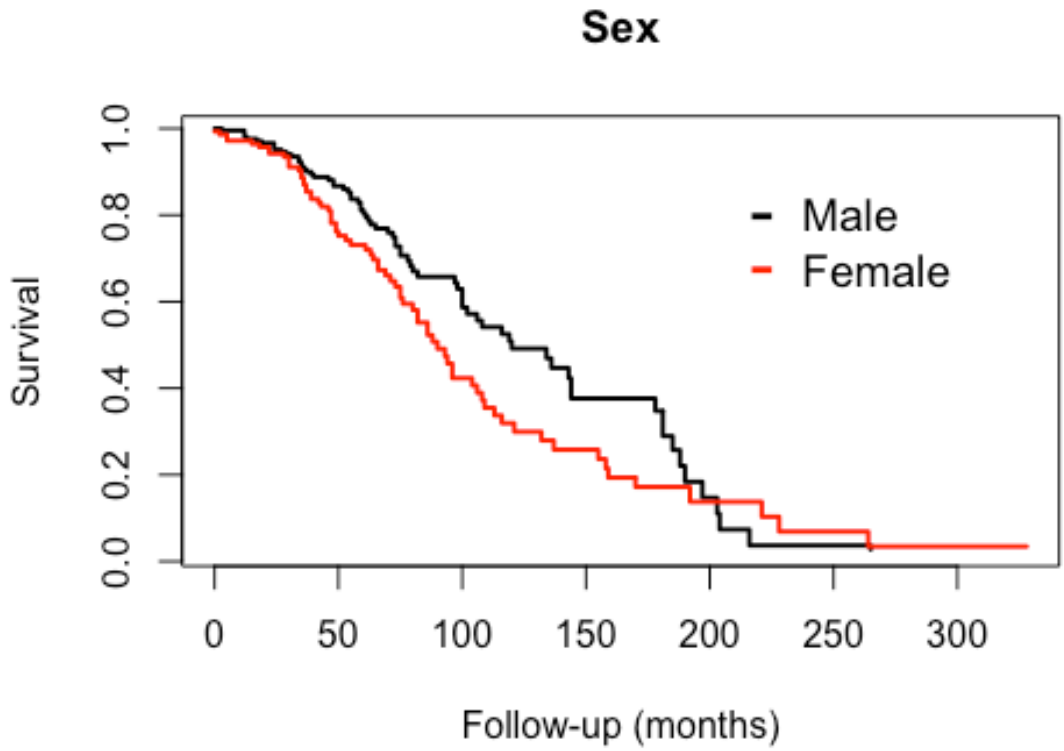
The relationship between the Hazard (*h(t)*) and the Survival function (*S(t)*) can be described as below:

$$h(t) = \frac{f(t)}{S(t)} = \frac{-dlogS(t)}{dt}$$

where $f(t)$ is the density function (dF(t)/dt) and represents the probability that the event of interest takes place in the time interval *(t, t+dt)*.

In survival analysis, the Kaplan Meier (K-M) method is widely used to estimate and graph survival probabilities as a function of follow-up time, allowing to compare two or more groups of participants which usually represent the different classes of a categorical variable (e.g. low, medium and high L-Dopa daily dosage with reference to dyskinesia onset, in our case). In K-M curves, the different groups compared are usually represented through different colours, and occurrence of the event of interest along time is represented by a step down in the curve (see example plot below, **Figure S1.2**). When two or more events occur at the same time, a deeper step is reported.

Figure S1.2: Kaplan-Meier example curve.



An example of Kaplan-Meier (K-M) curve is reported. In this case we report incident dyskinesia events in males versus females.

In other words, K-M curves are graphical representations of Cox PH regressions, and can be applied both to univariate models, where the occurrence of the event and the time-to-event are modelled as a function of a single exposure (e.g. dyskinesia ~ L-Dopa dosage), and to

multivariable models, where two or more exposures (or covariates) are present (e.g. dyskinesia ~ L-Dopa dosage). In the latter case, K-M curves are often called Cox curves.

K-M survival curves can provide an idea about the difference between survival functions among two or more groups. However, they cannot inform us whether this observed difference is statistically significant. To test the overall differences between estimated survival curves of two or more groups of subjects, such as males versus females, or treated versus untreated (control) groups, several tests are available. Among these methods, the most commonly used is probably the log-rank test. This non-parametric method is useful when the risk of an event is always greater for one group than another in order to detect a difference between groups. The log rank test is a form of Chi-square test distribution with one degree of freedom (Singh, and Mukhopadhyay, 2011) that calculates a statistic test used for testing the null hypothesis that there is no difference in survival between two groups. Essentially, the log rank test compares the observed number of events in each group to what would be expected if the null hypothesis were true.

The LOG RANK TEST formula is:

$$X^2(\log rank) = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

where O_1/ O_2 and E_1/E_2 represent the observed and the expected numbers of events within the groups of 1 and 2, respectively.

Cox regression

The most widely applicable and broadly implemented method in the survival analysis is the Cox proportional hazards (PH) regression (Cox, 1972). It allows to test for differences in survival times of two or more groups of interest, while adjusting for covariates of interest.

The Cox regression is without a doubt the most popular model for survival data analysis and is implemented in a large number of statistical software packages, including R (e.g. in the *survival* and *survminer* packages).

The basic Cox model can be described through the formula

$$h(t|Z) = h_0(t)\exp(\beta * Z),$$

where $h_0(t)$ is the baseline hazard function (e.g. the risk of LID onset at baseline), Z is a covariate vector $Z = (x_1, x_2, \ldots, x_p)$, and $\beta = (\beta_1, \beta_2, ..., \beta_p)$ is a vector of covariate coefficients. The formula above can be expressed as

$$h(t) = h_0(t) * exp(\beta_1 * x_1 + \beta_2 * x_2 + \ldots + \beta_p * x_p) \rightarrow$$

$$log(h(t)) = log(h_0(t)) + \beta_1 * x_1 + \beta_2 * x_2 + \ldots + \beta_p * x_p$$

Now, let's assume we compare two individuals, e.g. PD patients with high and low L-Dopa daily intake, which we represent through the dummy binary variable x_1 (1 for high and 0 for low L-dopa dosages). Under the assumption that these groups do not differ for any other exposure, we could apply to both the formula above

$$log(h(t))_{High\ L\text{-}Dopa} = log(h_0(t)) + \beta_1 * x_{1=0}$$

$$log(h(t)_{Low\ L\text{-}Dopa} = log(h_0(t)) + \beta_1 * x_{1=1}$$

Therefore, if we want to compare the risk of LID onset between the two patients due to L-Dopa dosage, we can compute it as Hazard Ratio (HR), namely

$$\beta_1 = log(h(t))_{High\ L\text{-}Dopa} - log(h(t))_{Low\ L\text{-}Dopa} = log(\frac{h(t))HighL-Dopa}{h(t))LowL-D\square pa}) \rightarrow$$

$$\mathbf{HR = exp(\beta_1) = \frac{h(t))HighL-Dopa}{h(t))LowL-Dopa}}$$

In other words, HR represents a measure of the increase/decrease in the risk of experiencing the event of interest (e.g. LID onset) associated with a given exposure (e.g. taking high vs low daily L-Dopa dosages), and can be simply estimated by exponentiating the slope associated with the exposure variable in the Cox regression (β_1). This model can be potentially extended to a number of covariates to test simultaneously in the same (multivariable) Cox regression. For any given variable, HR > 1 suggests increased risk associated with the exposure, HR < 1 suggests decreased risk (i.e. protective effect) and HR = 1 indicates no risk nor protective effect of the exposure on the occurrence of the event.

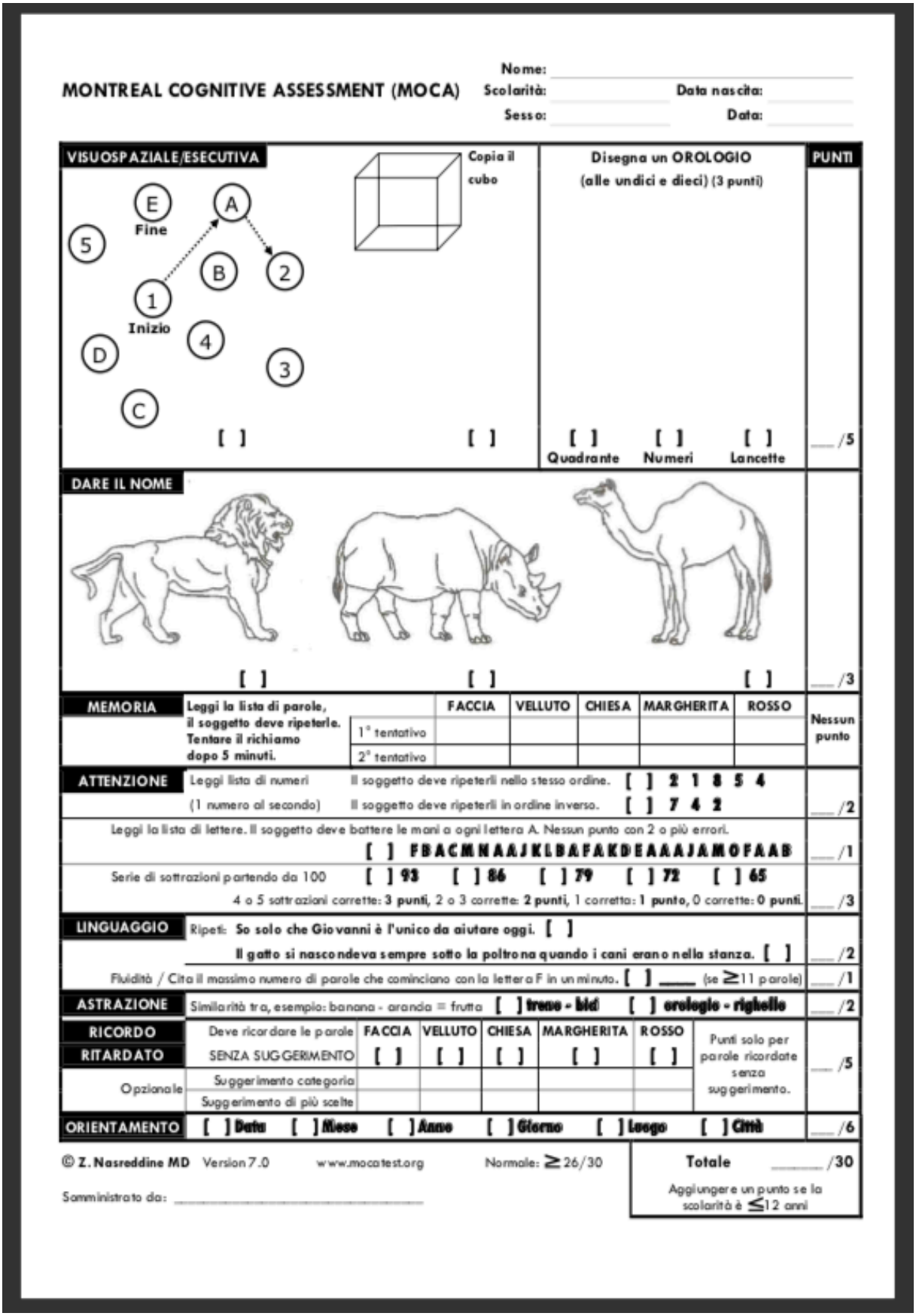
Cox PH regressions are based on three main assumptions which should be tested before performing the model:

- Proportionality of hazards: the effect of the exposure on the occurrence of the event is constant over time. This is usually assessed through analysing the relation between *Schoenfeld residuals* and time (*t*), both through scatter plots and through targeted statistical tests (217).
- Linearity of effects: continuous variables show a linear relation with the logarithm of the hazard functions (*log(h(t))*). This is usually assessed through plotting *Martingale residuals* vs the independent variable of interest (only applicable to continuous variables) (217)
- Absence of outlier observations: subjects which experience the event of interest too soon or too late should be removed before the analysis, since they may have a high weight in the regression model and introduce a bias. This is usually checked through plotting *deviance* or, alternatively, *dfbeta residuals*, and ensuring these are not higher than specific thresholds (see Methods section for details) (218).

Although it would be interesting to go into details of these assumptions, given the focus of the present dissertation we refer to theoretical works (216,217,218) and to analyses carried out in the Methods section for further details and practical examples.

Appendices 2: Supplementary Methods

Evaluation tests administered to PD patients - *Montreal Cognitive Assessment*



MONTREAL COGNITIVE ASSESSMENT (MOCA)

Nome:
Scolarità: Data nascita:
Sesso: Data:

VISUOSPAZIALE/ESECUTIVA	Copia il cubo	Disegna un OROLOGIO (alle undici e dieci) (3 punti)		PUNTI
	[ ]	[ ] Quadrante	[ ] Numeri [ ] Lancette	___ /5
DARE IL NOME	[ ] [ ] [ ]			___ /3

MEMORIA	Leggi la lista di parole, il soggetto deve ripeterle. Tentare il richiamo dopo 5 minuti.		FACCIA	VELLUTO	CHIESA	MARGHERITA	ROSSO
		1° tentativo					
		2° tentativo					

Nessun punto

ATTENZIONE	Leggi lista di numeri (1 numero al secondo)	Il soggetto deve ripeterli nello stesso ordine. [ ] 2 1 8 5 4 Il soggetto deve ripeterli in ordine inverso. [ ] 7 4 2
	Leggi la lista di lettere. Il soggetto deve battere le mani a ogni lettera A. Nessun punto con 2 o più errori.	[ ] F B A C M N A A J K L B A F A K D E A A A J A M O F A A B
	Serie di sottrazioni partendo da 100	[ ] 93 [ ] 86 [ ] 79 [ ] 72 [ ] 65 4 o 5 sottrazioni corrette: 3 punti, 2 o 3 corrette: 2 punti, 1 corretta: 1 punto, 0 corrette: 0 punti.
LINGUAGGIO	Ripeti:	So solo che Giovanni è l'unico da aiutare oggi. [ ] Il gatto si nascondeva sempre sotto la poltrona quando i cani erano nella stanza. [ ]
	Fluidità / Cita il massimo numero di parole che cominciano con la lettera F in un minuto.	[ ] ____ (se ≥11 parole)
ASTRAZIONE	Similarità tra, esempio: banana - arancia = frutta	[ ] treno - bici [ ] orologio - righello

RICORDO RITARDATO	Deve ricordare le parole SENZA SUGGERIMENTO	FACCIA [ ]	VELLUTO [ ]	CHIESA [ ]	MARGHERITA [ ]	ROSSO [ ]	Punti solo per parole ricordate senza suggerimento.
Opzionale	Suggerimento categoria						
	Suggerimento di più scelte						

ORIENTAMENTO	[ ] Data	[ ] Mese	[ ] Anno	[ ] Giorno	[ ] Luogo	[ ] Città

© Z. Nasreddine MD Version 7.0 www.mocatest.org Normale: ≥ 26/30

Somministrato da: _____

Totale ____ /30

Aggiungere un punto se la scolarità è ≤12 anni

Non-Motor Symptom assessment scale for Parkinson's Disease

Patient ID No: _____ **Initials:** _____ **Age:** _____

Symptoms assessed over the last month. Each symptom scored with respect to:

Severity: 0 = None, 1 = Mild: symptoms present but causes little distress or disturbance to patient; 2 = Moderate: some distress or disturbance to patient; 3 = Severe: major source of distress or disturbance to patient.

Frequency: 1 = Rarely (<1/wk); 2 = Often (1/wk); 3 = Frequent (several times per week); 4 = Very Frequent (daily or all the time)

Domains will be weighed differentially. Yes/ No answers are not included in final frequency x severity calculation.
(Bracketed text in questions within the scale is included as an explanatory aid).

Domain 1: Cardiovascular including falls	**Severity**	**Frequency**	**Frequency x Severity**
1. Does the patient experience light-headedness, dizziness, weakness on standing from sitting or lying position?	☐	☐	☐
2. Does the patient fall because of fainting or blacking out?	☐	☐	☐
SCORE:			☐

Domain 2: Sleep/fatigue			
3. Does the patient doze off or fall asleep unintentionally during daytime activities? (For example, during conversation, during mealtimes, or while watching television or reding).	☐	☐	☐
4. Does fatigue (tiredness) or lack of energy (not slowness) limit the patient's daytime activities?	☐	☐	☐
5. Does the patient have difficulties falling or staying asleep?	☐	☐	☐
6. Does the patient experience an urge to move the legs or restlessness in legs that improves with movement when he/she is sitting or lying down inactive?	☐	☐	☐
SCORE:			☐

Domain 3: Mood /Cognition			
7. Has the patient lost interest in his/her surroundings?	☐	☐	☐
8. Has the patient lost interest in doing things or lack motivation to start new activities?	☐	☐	☐
9. Does the patient feel nervous, worried or frightened for no apparent reason?	☐	☐	☐
10. Does the patient seem sad or depressed or has he/she reported such feelings?	☐	☐	☐
11. Does the patient have flat moods without the normal "highs" and " lows"?	☐	☐	☐
12. Does the patient have difficulty in experiencing pleasure from their usual activities or report that they lack pleasure?	☐	☐	☐
SCORE:			☐

Domain 4: Perceptual problems/hallucinations			
13. Does the patient indicate that he/she sees things that are not there?	☐	☐	☐
14. Does the patient have beliefs that you know are not true? (For example, about being harmed, being robbed or being unfaithful)	☐	☐	☐
15. Does the patient experience double vision? (2 separate real objects and not blurred vision)	☐	☐	☐
SCORE:			☐

	Severity	Frequency	Frequency x Severity
Domain 5: Attention/ Memory			
16. Does the patient have problems sustaining concentration during activities? (For example, reading or having a conversation)	☐	☐	☐
17. Does the patient forget things that he/she has been told a short time ago or events that happened in the last few days?	☐	☐	☐
18. Does the patient forget to do things? (For example, take tablets or turn off domestic appliances?)	☐	☐	☐
SCORE:			☐
Domain 6: Gastrointestinal tract			
19. Does the patient dribble saliva during the day?	☐	☐	☐
20. Does the patient having difficulty swallowing?	☐	☐	☐
21. Does the patient suffer from constipation? (Bowel action less than three times weekly)	☐	☐	☐
SCORE:			☐
Domain 7: Urinary			
22. Does the patient have difficulty holding urine? (Urgency)	☐	☐	☐
23. Does the patient have to void within 2 hours of last voiding? (Frequency)	☐	☐	☐
24. Does the patient have to get up regularly at night to pass urine? (Nocturia)	☐	☐	☐
SCORE:			☐
Domain 8: Sexual function			
25. Does the patient have altered interest in sex? (Very much increased or decreased, please underline)	☐	☐	☐
26. Does the patient have problems having sex?	☐	☐	☐
SCORE:			☐
Domain 9: Miscellaneous			
27. Does the patient suffer from pain not explained by other known conditions? (Is it related to intake of drugs and is it relieved by antiparkinson drugs?)	☐	☐	☐
28. Does the patient report a change in ability to taste or smell?	☐	☐	☐
29. Does the patient report a recent change in weight (not related to dieting)?	☐	☐	☐
30. Does the patient experience excessive sweating? (not related to hot weather)	☐	☐	☐
SCORE:			☐
TOTAL SCORE:			☐

Developed by the International Parkinson's Disease Non- Motor Group.

Evaluation tests administered to PD patients - *Unified Parkinson's Disease Rating Scale*

Part III: Motor Examination

Overview: This portion of the scale assesses the motor signs of PD. In administering Part III of the MDS-UPDRS the examiner should comply with the following guidelines:

At the top of the form, mark whether the patient is on medication for treating the symptoms of Parkinson's disease and, if on levodopa, the time since the last dose.

Also, if the patient is receiving medication for treating the symptoms of Parkinson's disease, mark the patient's clinical state using the following definitions:

ON is the typical functional state when patients are receiving medication and have a good response.
OFF is the typical functional state when patients have a poor response in spite of taking medications.

The investigator should "rate what you see." Admittedly, concurrent medical problems such as stroke, paralysis, arthritis, contracture, and orthopedic problems such as hip or knee replacement and scoliosis may interfere with individual items in the motor examination. In situations where it is absolutely impossible to test (e.g., amputations, plegia, limb in a cast), use the notation "**UR**" for Unable to Rate. Otherwise, rate the performance of each task as the patient performs in the context of co-morbidities.

All items must have an integer rating (no half points, no missing ratings).

Specific instructions are provided for the testing of each item. These should be followed in all instances. The investigator demonstrates while describing tasks the patient is to perform and rates function immediately thereafter. For Global Spontaneous Movement and Rest Tremor items (3.14 and 3.17), these items have been placed purposefully at the end of the scale because clinical information pertinent to the score will be obtained throughout the entire examination.

At the end of the rating, indicate if dyskinesia (chorea or dystonia) was present at the time of the examination, and if so, whether these movements interfered with the motor examination.

3a Is the patient on medication for treating the symptoms of Parkinson's disease? ☐ No ☐ Yes

3b If the patient is receiving medication for treating the symptoms of Parkinson's disease, mark the patient's clinical state using the following definitions:

☐ ON: On is the typical functional state when patients are receiving medication and have a good response.

☐ OFF: Off is the typical functional state when patients have a poor response in spite of taking medications.

3c Is the patient on levodopa ? ☐ No ☐ Yes

3.C1 If yes, minutes since last levodopa dose: _____

	SCORE
3.1 SPEECH Instructions to examiner: Listen to the patient's free-flowing speech and engage in conversation if necessary. Suggested topics: ask about the patient's work, hobbies, exercise, or how he got to the doctor's office. Evaluate volume, modulation (prosody) and clarity, including slurring, palilalia (repetition of syllables), and tachyphemia (rapid speech, running syllables together). 0: Normal: No speech problems. 1: Slight: Loss of modulation, diction, or volume, but still all words easy to understand. 2: Mild: Loss of modulation, diction, or volume, with a few words unclear, but the overall sentences easy to follow. 3: Moderate: Speech is difficult to understand to the point that some, but not most, sentences are poorly understood. 4: Severe: Most speech is difficult to understand or unintelligible.	☐
3.2 FACIAL EXPRESSION Instructions to examiner: Observe the patient sitting at rest for 10 seconds, without talking and also while talking. Observe eye-blink frequency, masked facies or loss of facial expression, spontaneous smiling, and parting of lips. 0: Normal: Normal facial expression. 1: Slight: Minimal masked facies manifested only by decreased frequency of blinking. 2: Mild: In addition to decreased eye-blink frequency, masked facies present in the lower face as well, namely fewer movements around the mouth, such as less spontaneous smiling, but lips not parted. 3: Moderate: Masked facies with lips parted some of the time when the mouth is at rest. 4: Severe: Masked facies with lips parted most of the time when the mouth is at rest.	☐

3.3 RIGIDITY

SCORE

Instructions to examiner: Rigidity is judged on slow passive movement of major joints with the patient in a relaxed position and the examiner manipulating the limbs and neck. First, test without an activation maneuver. Test and rate neck and each limb separately. For arms, test the wrist and elbow joints simultaneously. For legs, test the hip and knee joints simultaneously. If no rigidity is detected, use an activation maneuver such as tapping fingers, fist opening/closing, or heel tapping in a limb not being tested. Explain to the patient to go as limp as possible as you test for rigidity.

0: Normal: No rigidity.

1: Slight: Rigidity only detected with activation maneuver.

2: Mild: Rigidity detected without the activation maneuver, but full range of motion is easily achieved.

3: Moderate: Rigidity detected without the activation maneuver; full range of motion is achieved with effort.

4: Severe: Rigidity detected without the activation maneuver and full range of motion not achieved.

☐ Neck

☐ RUE

☐ LUE

☐ RLE

☐ LLE

3.4 FINGER TAPPING

Instructions to examiner: Each hand is tested separately. Demonstrate the task, but do not continue to perform the task while the patient is being tested. Instruct the patient to tap the index finger on the thumb 10 times as quickly AND as big as possible. Rate each side separately, evaluating speed, amplitude, hesitations, halts, and decrementing amplitude.

0: Normal: No problems.

1: Slight: Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the tapping movement; b) slight slowing; c) the amplitude decrements near the end of the 10 taps.

2: Mild: Any of the following: a) 3 to 5 interruptions during tapping; b) mild slowing; c) the amplitude decrements midway in the 10-tap sequence.

3: Moderate: Any of the following: a) more than 5 interruptions during tapping or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) the amplitude decrements starting after the 1st tap.

4: Severe: Cannot or can only barely perform the task because of slowing, interruptions, or decrements.

☐ R

☐ L

	SCORE
3.5 HAND MOVEMENTS Instructions to examiner: Test each hand separately. Demonstrate the task, but do not continue to perform the task while the patient is being tested. Instruct the patient to make a tight fist with the arm bent at the elbow so that the palm faces the examiner. Have the patient open the hand 10 times as fully AND as quickly as possible. If the patient fails to make a tight fist or to open the hand fully, remind him/her to do so. Rate each side separately, evaluating speed, amplitude, hesitations, halts, and decrementing amplitude. 0: Normal: No problems. 1: Slight: Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the movement; b) slight slowing; c) the amplitude decrements near the end of the task. 2: Mild: Any of the following: a) 3 to 5 interruptions during the movements; b) mild slowing; c) the amplitude decrements midway in the task. 3: Moderate: Any of the following: a) more than 5 interruptions during the movement or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) the amplitude decrements starting after the 1st open-and-close sequence. 4: Severe: Cannot or can only barely perform the task because of slowing, interruptions, or decrements.	☐ R ☐ L
3.6 PRONATION-SUPINATION MOVEMENTS OF HANDS Instructions to examiner: Test each hand separately. Demonstrate the task, but do not continue to perform the task while the patient is being tested. Instruct the patient to extend the arm out in front of his/her body with the palms down, and then to turn the palm up and down alternately 10 times as fast and as fully as possible. Rate each side separately, evaluating speed, amplitude, hesitations, halts, and decrementing amplitude. 0: Normal: No problems. 1: Slight: Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the movement; b) slight slowing; c) the amplitude decrements near the end of the sequence. 2: Mild: Any of the following: a) 3 to 5 interruptions during the movements; b) mild slowing; c) the amplitude decrements midway in the sequence. 3: Moderate: Any of the following: a) more than 5 interruptions during the movement or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) the amplitude decrements starting after the 1st supination-pronation sequence. 4: Severe: Cannot or can only barely perform the task because of slowing, interruptions, or decrements.	☐ R ☐ L

	SCORE
3.7 TOE TAPPING Instructions to examiner: Have the patient sit in a straight-backed chair with arms, both feet on the floor. Test each foot separately. Demonstrate the task, but do not continue to perform the task while the patient is being tested. Instruct the patient to place the heel on the ground in a comfortable position and then tap the toes 10 times as big and as fast as possible. Rate each side separately, evaluating speed, amplitude, hesitations, halts, and decrementing amplitude. 0: Normal: No problems. 1: Slight: Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the tapping movement; b) slight slowing; c) amplitude decrements near the end of the ten taps. 2: Mild: Any of the following: a) 3 to 5 interruptions during the tapping movements; b) mild slowing; c) amplitude decrements midway in the task. 3: Moderate: Any of the following: a) more than 5 interruptions during the tapping movements or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) amplitude decrements after the 1st tap. 4: Severe: Cannot or can only barely perform the task because of slowing, interruptions or decrements.	☐ R ☐ L
3.8 LEG AGILITY Instructions to examiner: Have the patient sit in a straight-backed chair with arms. The patient should have both feet comfortably on the floor. Test each leg separately. Demonstrate the task, but do not continue to perform the task while the patient is being tested. Instruct the patient to place the foot on the ground in a comfortable position and then raise and stomp the foot on the ground 10 times as high and as fast as possible. Rate each side separately, evaluating speed, amplitude, hesitations, halts and decrementing amplitude. 0: Normal: No problems. 1: Slight: Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the movement; b) slight slowing; c) amplitude decrements near the end of the task. 2: Mild: Any of the following: a) 3 to 5 interruptions during the movements; b) mild slowness; c) amplitude decrements midway in the task. 3: Moderate: Any of the following: a) more than 5 interruptions during the movement or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing in speed; c) amplitude decrements after the 1st tap. 4: Severe: Cannot or can only barely perform the task because of slowing, interruptions, or decrements.	☐ R ☐ L

	SCORE
3.9 ARISING FROM CHAIR Instructions to examiner: Have the patient sit in a straight-backed chair with arms, with both feet on the floor and sitting back in the chair (if the patient is not too short). Ask the patient to cross his/her arms across the chest and then to stand up. If the patient is not successful, repeat this attempt up to a maximum of two more times. If still unsuccessful, allow the patient to move forward in the chair to arise with arms folded across the chest. Allow only one attempt in this situation. If unsuccessful, allow the patient to push off using his/her hands on the arms of the chair. Allow a maximum of three trials of pushing off. If still not successful, assist the patient to arise. After the patient stands up, observe the posture for item 3.13. 0: Normal: No problems. Able to arise quickly without hesitation. 1: Slight: Arising is slower than normal; or may need more than one attempt; or may need to move forward in the chair to arise. No need to use the arms of the chair. 2: Mild: Pushes self up from the arms of the chair without difficulty. 3: Moderate: Needs to push off, but tends to fall back; or may have to try more than one time using the arms of the chair, but can get up without help. 4: Severe: Unable to arise without help.	☐
3.10 GAIT Instructions to examiner: Testing gait is best performed by having the patient walking away from and towards the examiner so that both right and left sides of the body can be easily observed simultaneously. The patient should walk at least 10 meters (30 feet), then turn around and return to the examiner. This item measures multiple behaviors: stride amplitude, stride speed, height of foot lift, heel strike during walking, turning, and arm swing, but not freezing. Assess also for "freezing of gait" (next item 3.11) while patient is walking. Observe posture for item 3.13. 0: Normal: No problems. 1: Slight: Independent walking with minor gait impairment. 2: Mild: Independent walking but with substantial gait impairment. 3: Moderate: Requires an assistance device for safe walking (walking stick, walker) but not a person. 4: Severe: Cannot walk at all or only with another person's assistance.	☐

	SCORE
3.11 FREEZING OF GAIT Instructions to examiner: While assessing gait, also assess for the presence of any gait freezing episodes. Observe for start hesitation and stuttering movements especially when turning and reaching the end of the task. To the extent that safety permits, patients may NOT use sensory tricks during the assessment. 0: Normal: No freezing. 1: Slight: Freezes on starting, turning, or walking through doorway with a single halt during any of these events, but then continues smoothly without freezing during straight walking. 2: Mild: Freezes on starting, turning, or walking through doorway with more than one halt during any of these activities, but continues smoothly without freezing during straight walking. 3: Moderate: Freezes once during straight walking. 4: Severe: Freezes multiple times during straight walking.	☐
3.12 POSTURAL STABILITY Instructions to examiner: The test examines the response to sudden body displacement produced by a quick, forceful pull on the shoulders while the patient is standing erect with eyes open and feet comfortably apart and parallel to each other. Test retropulsion. Stand behind the patient and instruct the patient on what is about to happen. Explain that s/he is allowed to take a step backwards to avoid falling. There should be a solid wall behind the examiner, at least 1-2 meters away to allow for the observation of the number of retropulsive steps. The first pull is an instructional demonstration and is purposely milder and not rated. The second time the shoulders are pulled briskly and forcefully towards the examiner with enough force to displace the center of gravity so that patient MUST take a step backwards. The examiner needs to be ready to catch the patient, but must stand sufficiently back so as to allow enough room for the patient to take several steps to recover independently. Do not allow the patient to flex the body abnormally forward in anticipation of the pull. Observe for the number of steps backwards or falling. Up to and including two steps for recovery is considered normal, so abnormal ratings begin with three steps. If the patient fails to understand the test, the examiner can repeat the test so that the rating is based on an assessment that the examiner feels reflects the patient's limitations rather than misunderstanding or lack of preparedness. Observe standing posture for item 3.13. 0: Normal: No problems. Recovers with one or two steps. 1: Slight: 3-5 steps, but subject recovers unaided. 2: Mild: More than 5 steps, but subject recovers unaided. 3: Moderate: Stands safely, but with absence of postural response; falls if not caught by examiner. 4: Severe: Very unstable, tends to lose balance spontaneously or with just a gentle pull on the shoulders.	☐

	SCORE
3.13 POSTURE Instructions to examiner: Posture is assessed with the patient standing erect after arising from a chair, during walking, and while being tested for postural reflexes. If you notice poor posture, tell the patient to stand up straight and see if the posture improves (see option 2 below). Rate the worst posture seen in these three observation points. Observe for flexion and side-to-side leaning. 0: Normal: No problems. 1: Slight: Not quite erect, but posture could be normal for older person. 2: Mild: Definite flexion, scoliosis or leaning to one side, but patient can correct posture to normal posture when asked to do so. 3: Moderate: Stooped posture, scoliosis or leaning to one side that cannot be corrected volitionally to a normal posture by the patient. 4: Severe: Flexion, scoliosis or leaning with extreme abnormality of posture.	☐
3.14 GLOBAL SPONTANEITY OF MOVEMENT (BODY BRADYKINESIA) Instructions to examiner: This global rating combines all observations on slowness, hesitancy, and small amplitude and poverty of movement in general, including a reduction of gesturing and of crossing the legs. This assessment is based on the examiner's global impression after observing for spontaneous gestures while sitting, and the nature of arising and walking. 0: Normal: No problems. 1: Slight: Slight global slowness and poverty of spontaneous movements. 2: Mild: Mild global slowness and poverty of spontaneous movements. 3: Moderate: Moderate global slowness and poverty of spontaneous movements. 4: Severe: Severe global slowness and poverty of spontaneous movements.	☐
3.15 POSTURAL TREMOR OF THE HANDS Instructions to examiner: All tremor, including re-emergent rest tremor, that is present in this posture is to be included in this rating. Rate each hand separately. Rate the highest amplitude seen. Instruct the patient to stretch the arms out in front of the body with palms down. The wrist should be straight and the fingers comfortably separated so that they do not touch each other. Observe this posture for 10 seconds. 0: Normal: No tremor. 1: Slight: Tremor is present but less than 1 cm in amplitude. 2: Mild: Tremor is at least 1 but less than 3 cm in amplitude. 3: Moderate: Tremor is at least 3 but less than 10 cm in amplitude. 4: Severe: Tremor is at least 10 cm in amplitude.	☐ R ☐ L

	SCORE
3.16 KINETIC TREMOR OF THE HANDS Instructions to examiner: This is tested by the finger-to-nose maneuver. With the arm starting from the outstretched position, have the patient perform at least three finger-to-nose maneuvers with each hand reaching as far as possible to touch the examiner's finger. The finger-to-nose maneuver should be performed slowly enough not to hide any tremor that could occur with very fast arm movements. Repeat with the other hand, rating each hand separately. The tremor can be present throughout the movement or as the tremor reaches either target (nose or finger). Rate the highest amplitude seen. 0: Normal: No tremor. 1: Slight: Tremor is present but less than 1 cm in amplitude. 2: Mild: Tremor is at least 1 but less than 3 cm in amplitude. 3: Moderate: Tremor is at least 3 but less than 10 cm in amplitude. 4: Severe: Tremor is at least 10 cm in amplitude.	☐ R ☐ L
3.17 REST TREMOR AMPLITUDE Instructions to examiner: This and the next item have been placed purposefully at the end of the examination to allow the rater to gather observations on rest tremor that may appear at any time during the exam, including when quietly sitting, during walking, and during activities when some body parts are moving but others are at rest. Score the maximum amplitude that is seen at any time as the final score. Rate only the amplitude and not the persistence or the intermittency of the tremor. As part of this rating, the patient should sit quietly in a chair with the hands placed on the arms of the chair (not in the lap) and the feet comfortably supported on the floor for 10 seconds with no other directives. Rest tremor is assessed separately for all four limbs and also for the lip/jaw. Rate only the maximum amplitude that is seen at any time as the final rating. Extremity ratings 0: Normal: No tremor. 1: Slight: < 1 cm in maximal amplitude. 2: Mild: ≥ 1 cm but < 3 cm in maximal amplitude. 3: Moderate: ≥ 3 cm but < 10 cm in maximal amplitude. 4: Severe: ≥ 10 cm in maximal amplitude. Lip/Jaw ratings 0: Normal: No tremor. 1: Slight: < 1 cm in maximal amplitude. 2: Mild: ≥ 1 cm but < 2 cm in maximal amplitude. 3: Moderate: ≥ 2 cm but < 3 cm in maximal amplitude. 4: Severe: ≥ 3 cm in maximal amplitude.	☐ RUE ☐ LUE ☐ RLE ☐ LLE ☐ Lip/Jaw

3.18 CONSTANCY OF REST TREMOR

SCORE

Instructions to examiner: This item receives one rating for all rest tremor and focuses on the constancy of rest tremor during the examination period when different body parts are variously at rest. It is rated purposefully at the end of the examination so that several minutes of information can be coalesced into the rating.

0: Normal: No tremor.

1: Slight: Tremor at rest is present ≤ 25% of the entire examination period.

2: Mild: Tremor at rest is present 26-50% of the entire examination period.

3: Moderate: Tremor at rest is present 51-75% of the entire examination period.

4: Severe: Tremor at rest is present > 75% of the entire examination period.

☐

DYSKINESIA IMPACT ON PART III RATINGS

A. Were dyskinesias (chorea or dystonia) present during examination? ☐ No ☐ Yes

B. If yes, did these movements interfere with your ratings? ☐ No ☐ Yes

HOEHN AND YAHR STAGE

0: Asymptomatic.

1: Unilateral involvement only.

2: Bilateral involvement without impairment of balance.

3: Mild to moderate involvement; some postural instability but physically independent; needs assistance to recover from pull test.

4: Severe disability; still able to walk or stand unassisted.

5: Wheelchair bound or bedridden unless aided.

☐

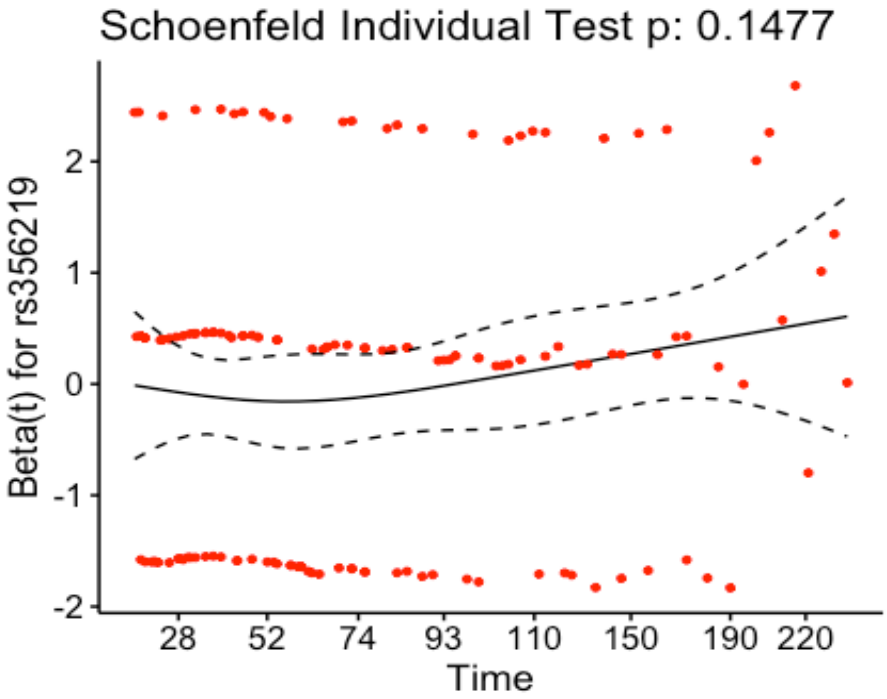
Appendices 3: Supplementary Results

Check for basic assumptions of Cox PH models

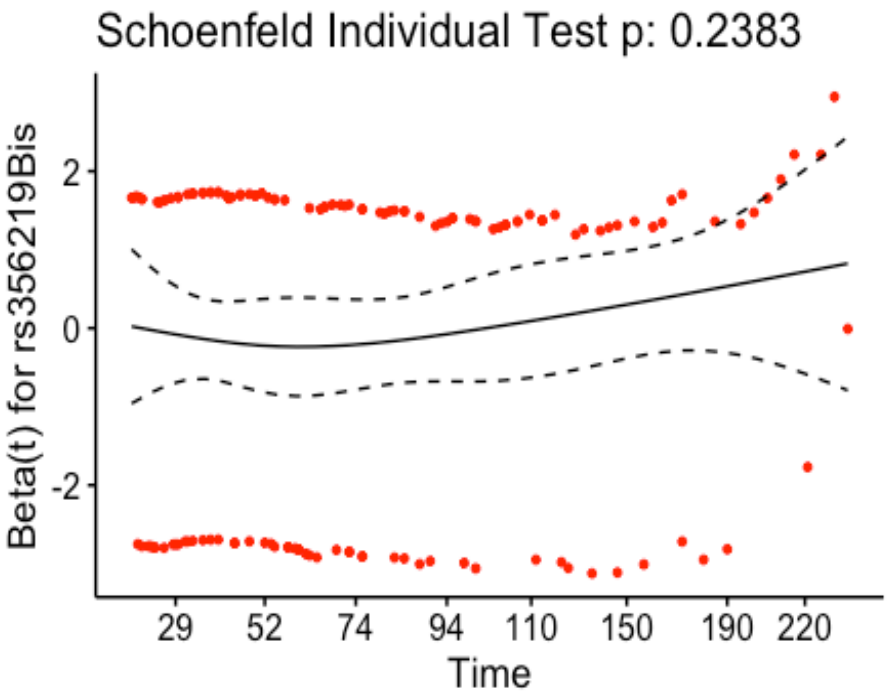
Schoenfeld residuals (proportionality of hazards)

Figure S2.1: Schoenfeld residuals computed for univariate Cox regressions of LID risk vs rs356219 in a) Additive, b) Dominant, and c) Recessive model; and vs D4S3481 in d) Pseudo-additive, e) Pseudo-dominant and f) Pseudo-recessive model (see Table 3.3a, b for details).

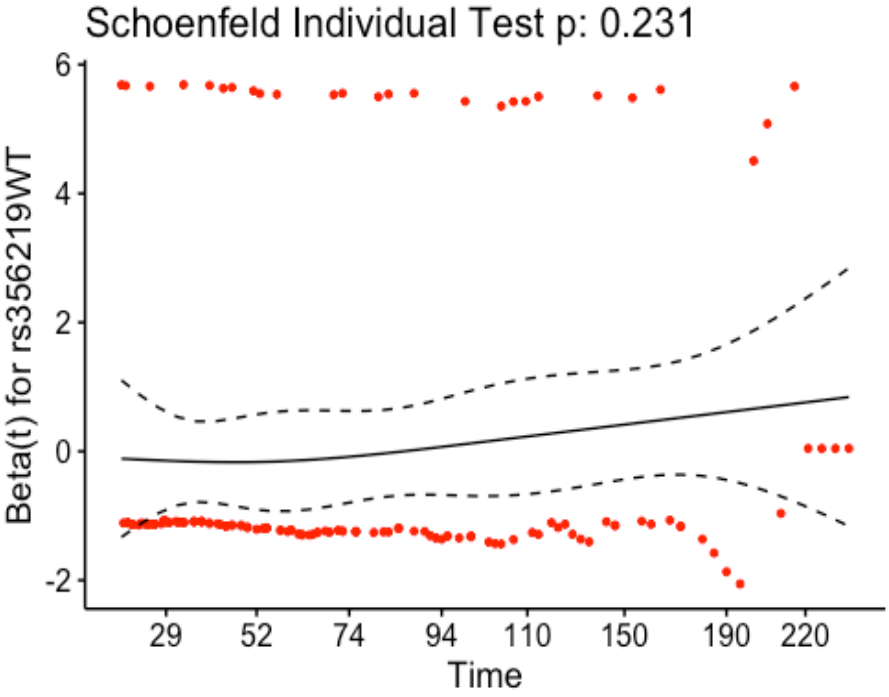
a)



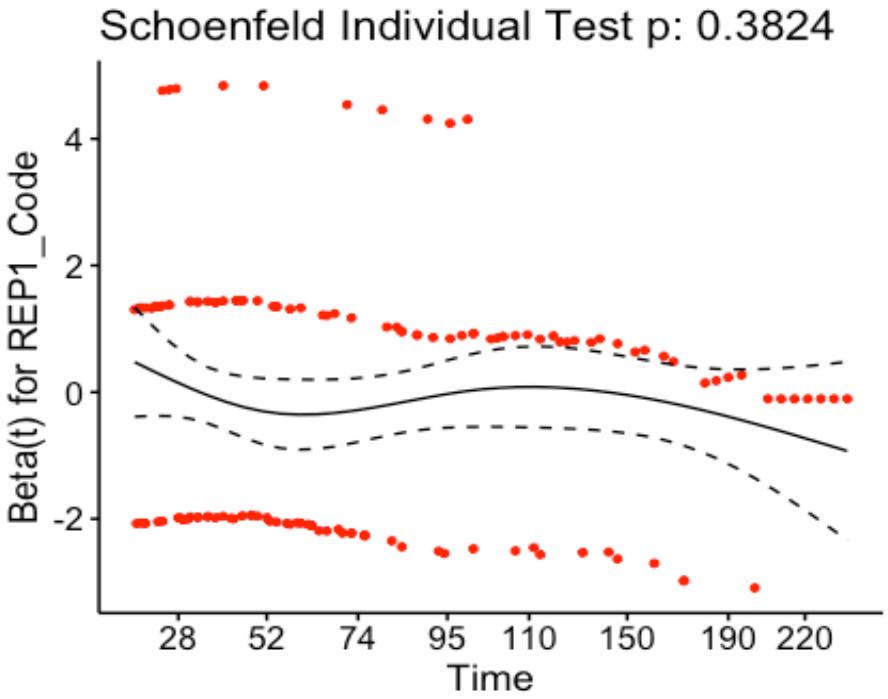
b)


Schoenfeld Individual Test p: 0.2383
Beta(t) for rs356219Bis
2
0
-2
29
52
74
94
110
150
190
220
Time

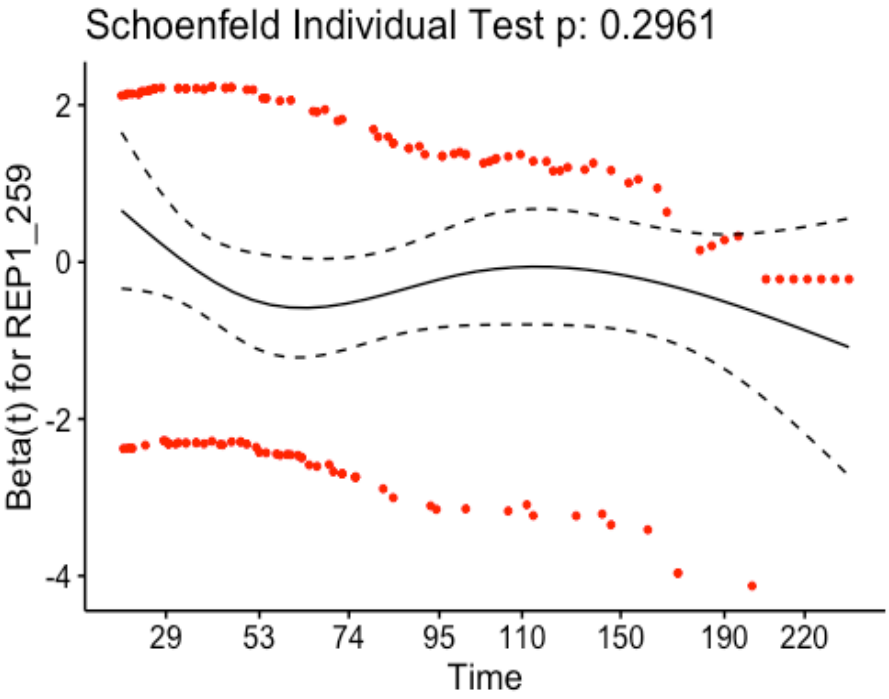
c)


Schoenfeld Individual Test p: 0.231
Beta(t) for rs356219WT
6
4
2
0
-2
29
52
74
94
110
150
190
220
Time

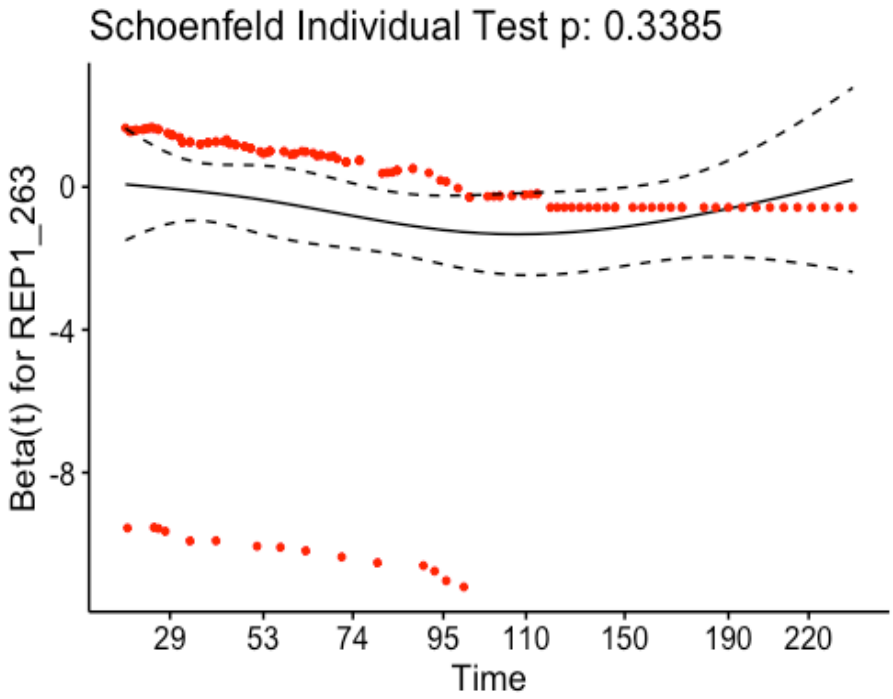
d)


Schoenfeld Individual Test p: 0.3824
Beta(t) for REP1_Code
4
2
0
-2
28
52
74
95
110
150
190
220
Time

e)


Schoenfeld Individual Test p: 0.2961
Beta(t) for REP1_259
2
0
-2
-4
29
53
74
95
110
150
190
220
Time

f)

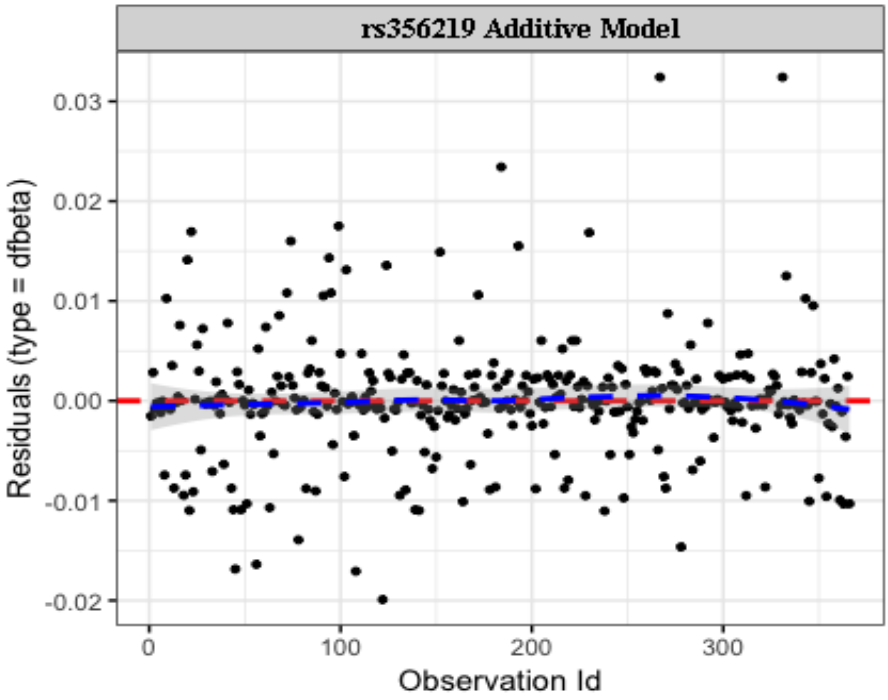


Here, Schoenfeld residuals are plotted versus time for each genetic variant and model tested (see **Table 3.3a, b** above for details).

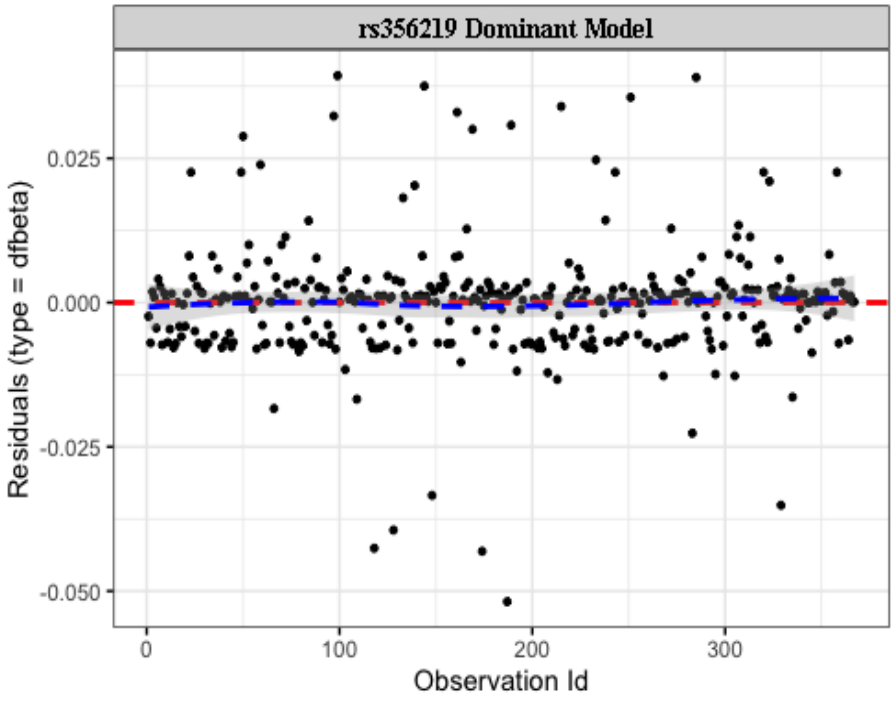
Dfbeta residuals (outlier observations)

Figure S2.2: Dfbeta residuals plots for outliers detection in rs356219 a) Additive, b) Dominant and c) Recessive model, and for D4S3481 d) Pseudo-additive, e) Pseudo-dominant and f) Pseudo-recessive model (see Table 3.3a, b for details).

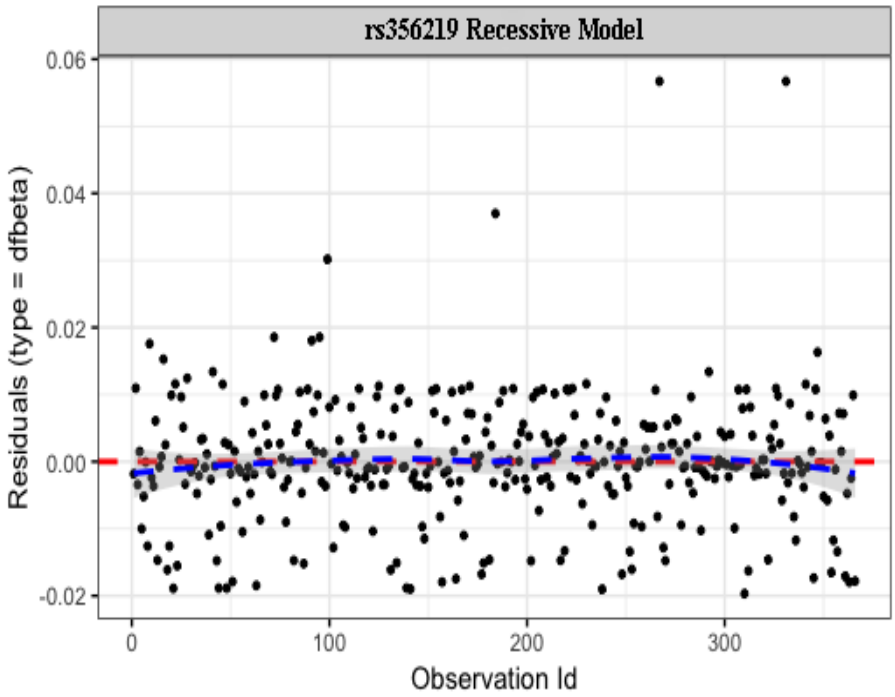
a)



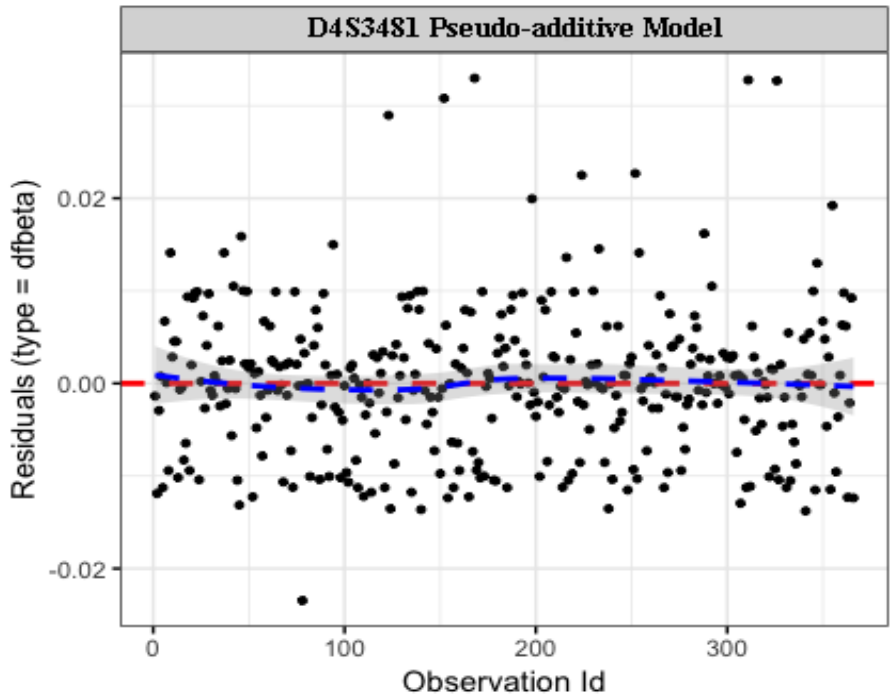
b)



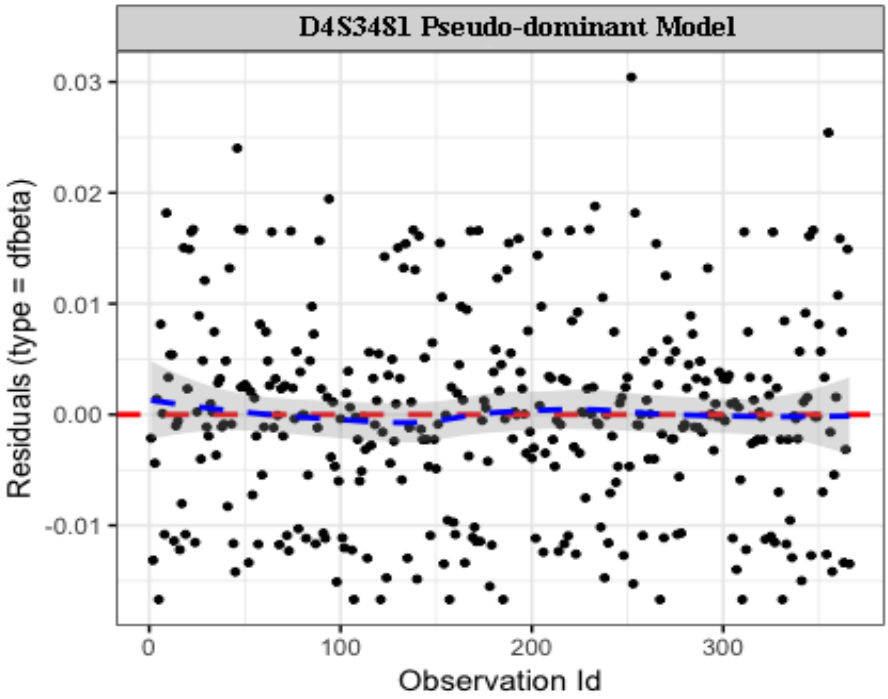
c)


rs356219 Recessive Model
Residuals (type = dfbeta)
0.06
0.04
0.02
0.00
-0.02
0
100
200
300
Observation Id

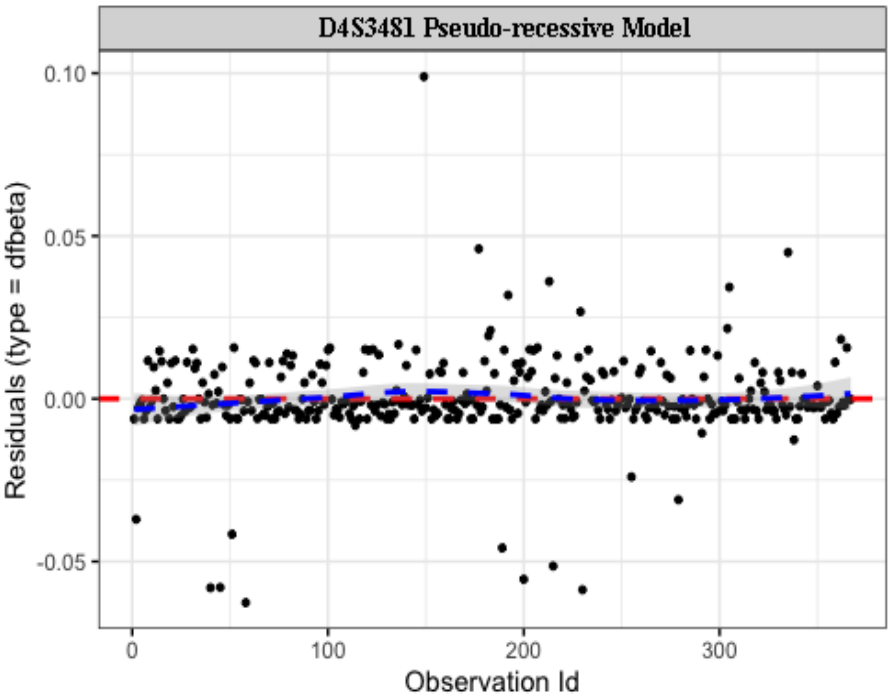
d)


D4S3481 Pseudo-additive Model
Residuals (type = dfbeta)
0.02
0.00
-0.02
0
100
200
300
Observation Id

e)



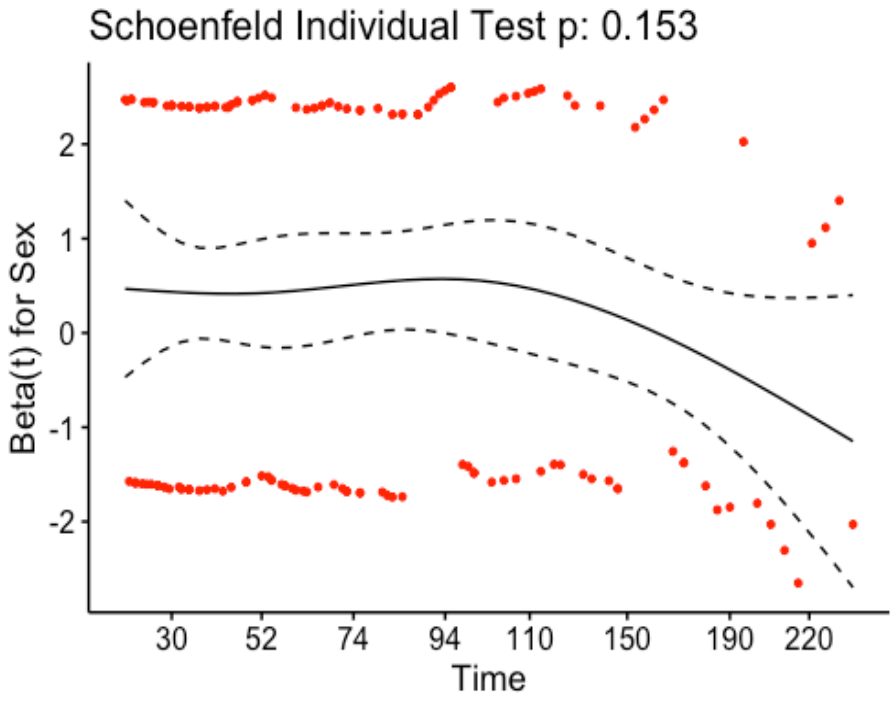
f)



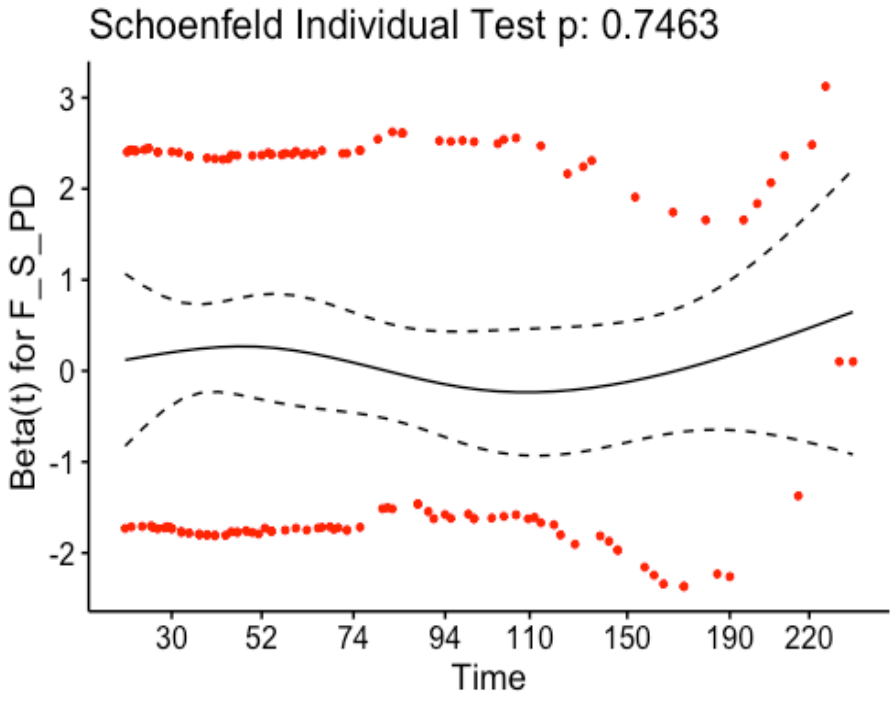
Here, Dfbeta residuals for each observation available in the PD cohort are plotted, for each genetic variant and model tested (see **Table 3.3a, b** above for details).

Figure S3.1: Schoenfeld residuals computed for univariate Cox regressions of LID risk vs a) Sex, b) PD familiarity, c) clinical subtype and d) Age-at-onset (AAO), e) L-Dopa dosage, f) MoCA, g) HY, h) UPDRS and i) NMS score, and j) years of disease (YOD).

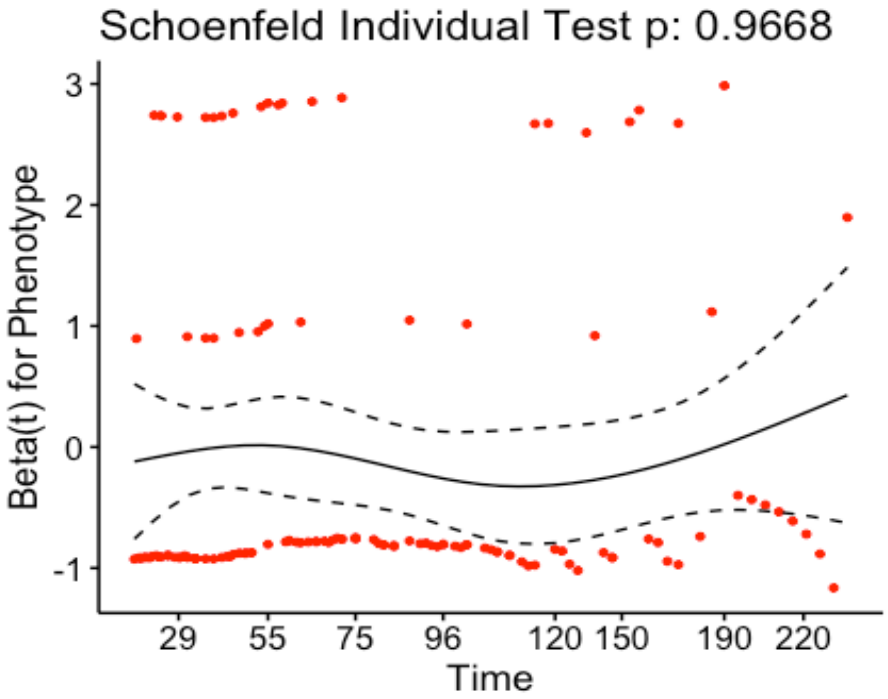
a) Sex



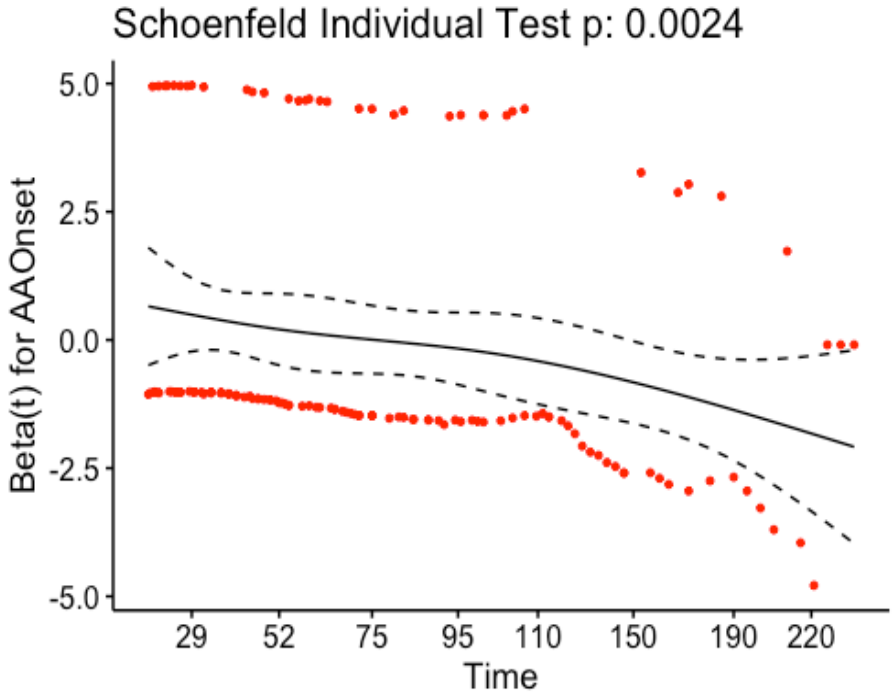
b) Familiarity



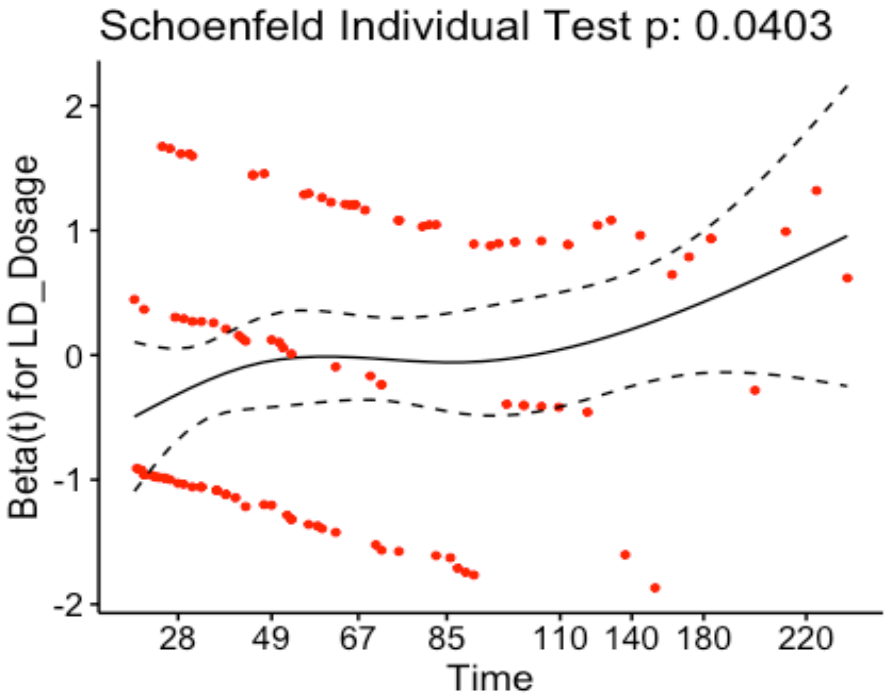
c) Clinical subtype



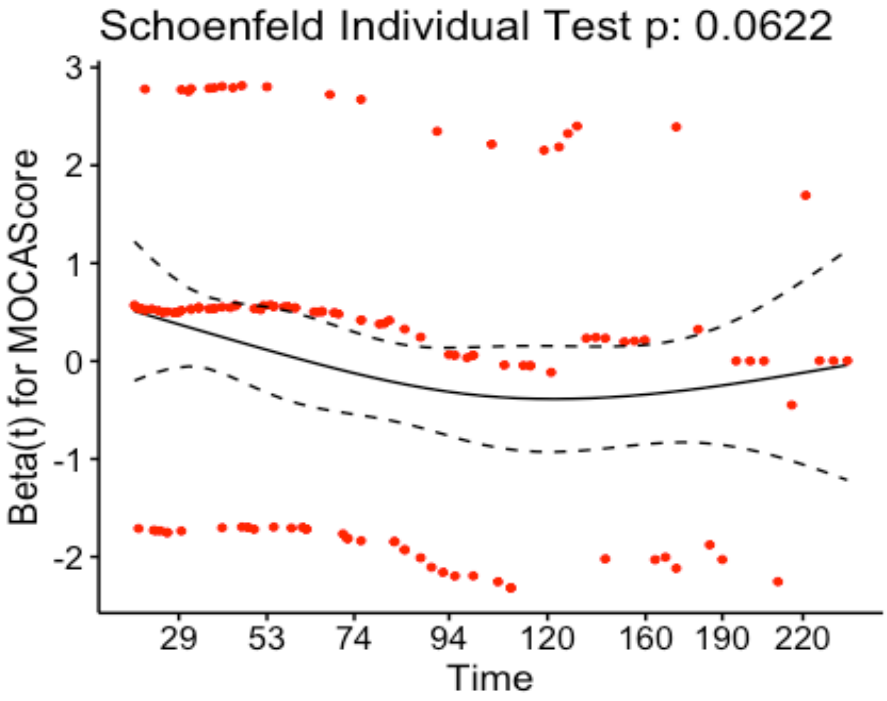
d) AAO



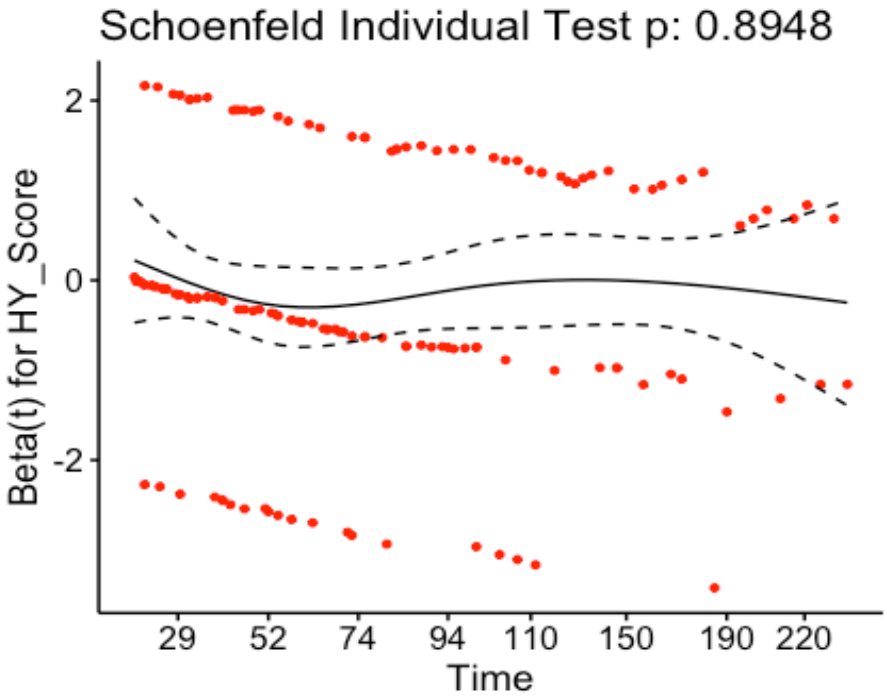
e) L-Dopa dosage



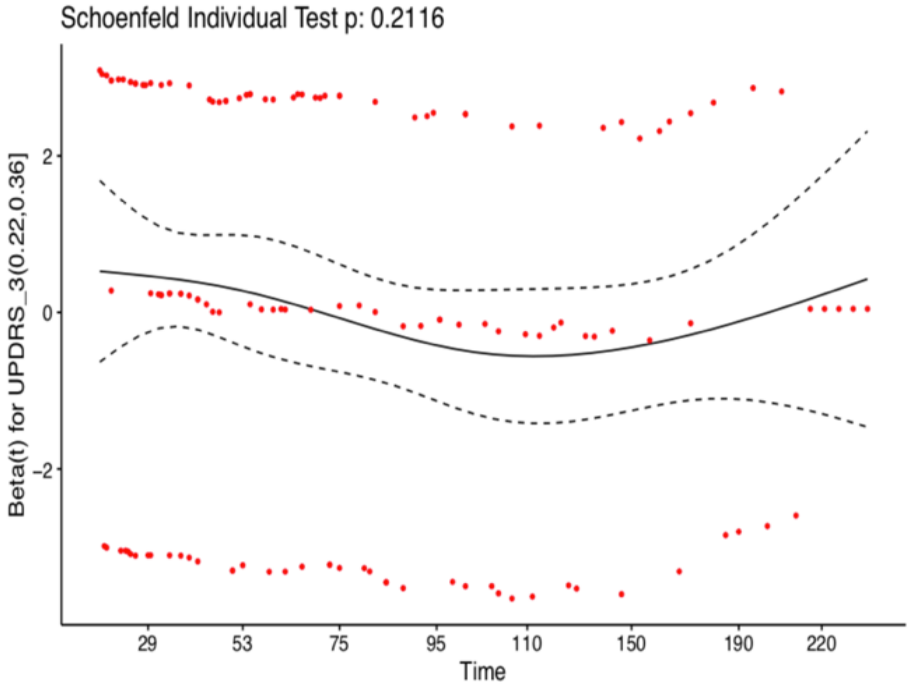
f) MoCA



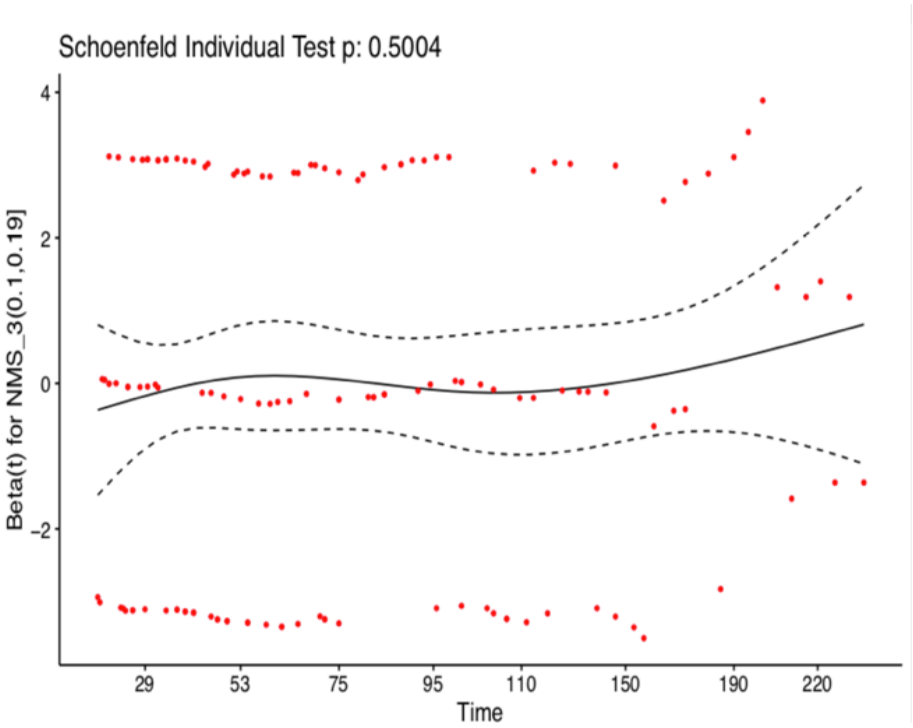
g) HY


Schoenfeld Individual Test p: 0.8948
Beta(t) for HY_Score
2
0
-2
29
52
74
94
110
150
190
220
Time

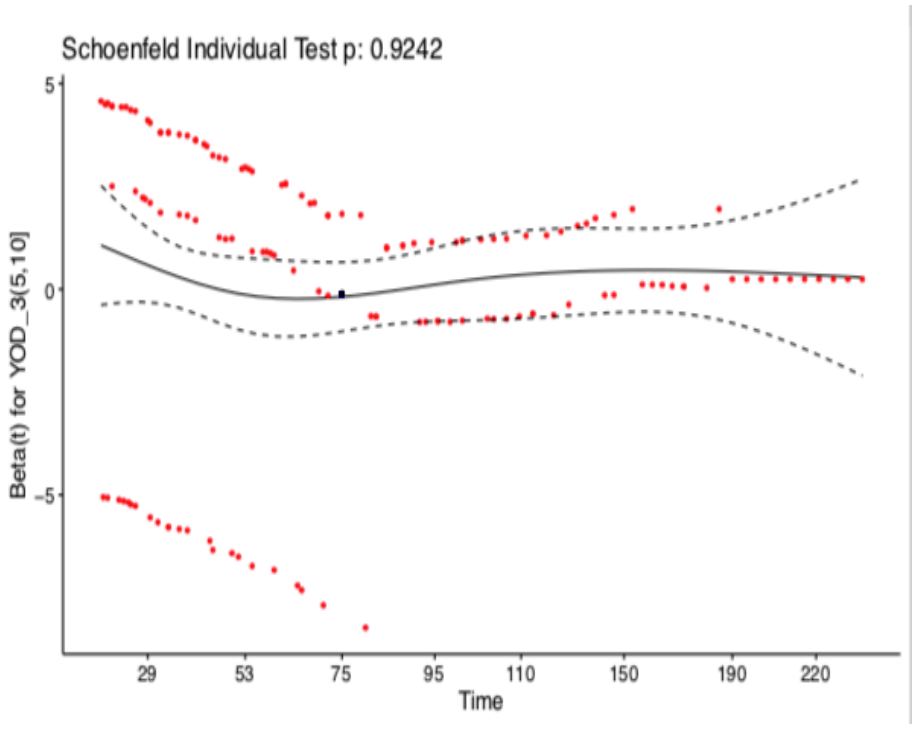
h) UPDRS


Schoenfeld Individual Test p: 0.2116
Beta(t) for UPDRS_3(0.22,0.36]
2
0
-2
29
53
75
95
110
150
190
220
Time

i) NMS



j) YOD

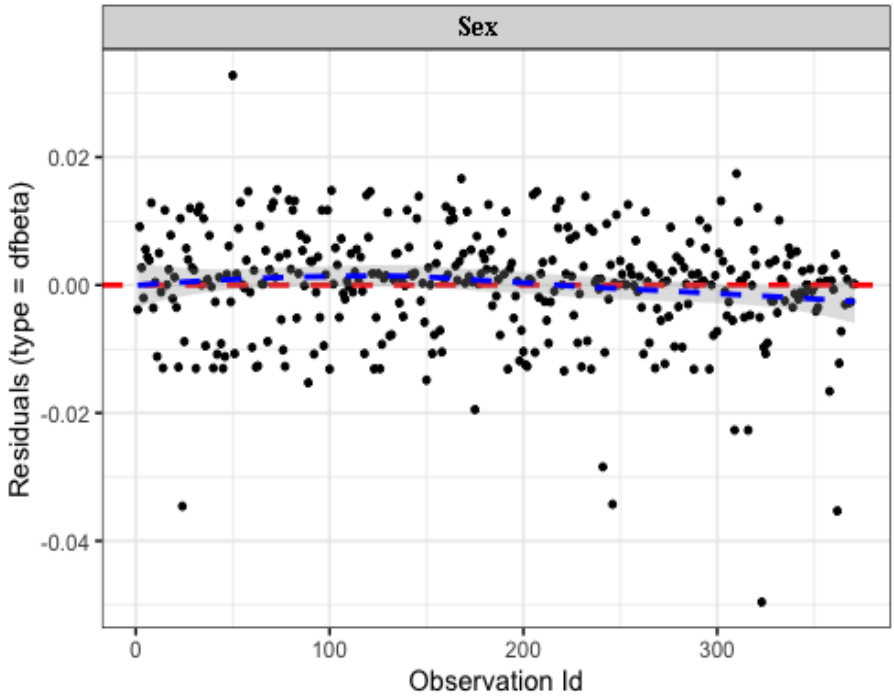


Here, Schoenfeld residuals are plotted versus time for each covariate tested.

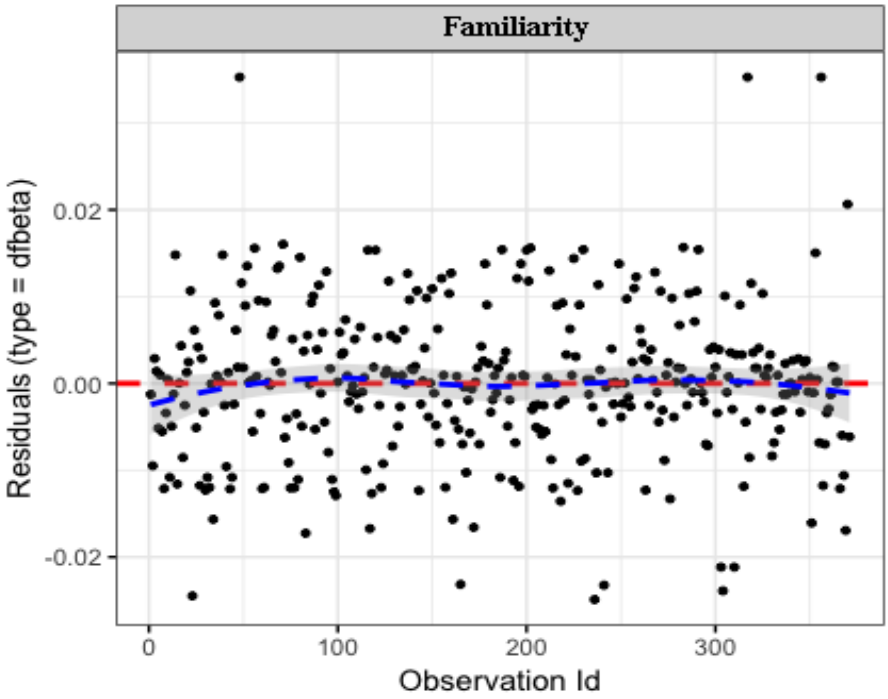
Abbreviations: AAO, PD age at onset; MoCA, Montreal Cognitive Assessment; HY, Hoehn & Yahr score; UPDRS, Movement Disorder Society revised version of the Unified Parkinson's Disease Rating Scale-Part III; MoCA, Montreal Cognitive Assessment; NMS, modified version of the Non-Motor Symptoms Scale for Parkinson Disease; YOD, years of disease.

Figure S3.2: Dfbeta residuals plots for outliers detection in a) Sex, b) PD familiarity, c) clinical subtype and d) Age-at-onset (AAO), e) L-Dopa dosage, f) MoCA, g) HY, h) UPDRS and i) NMS score, and l) years of disease (YOD).

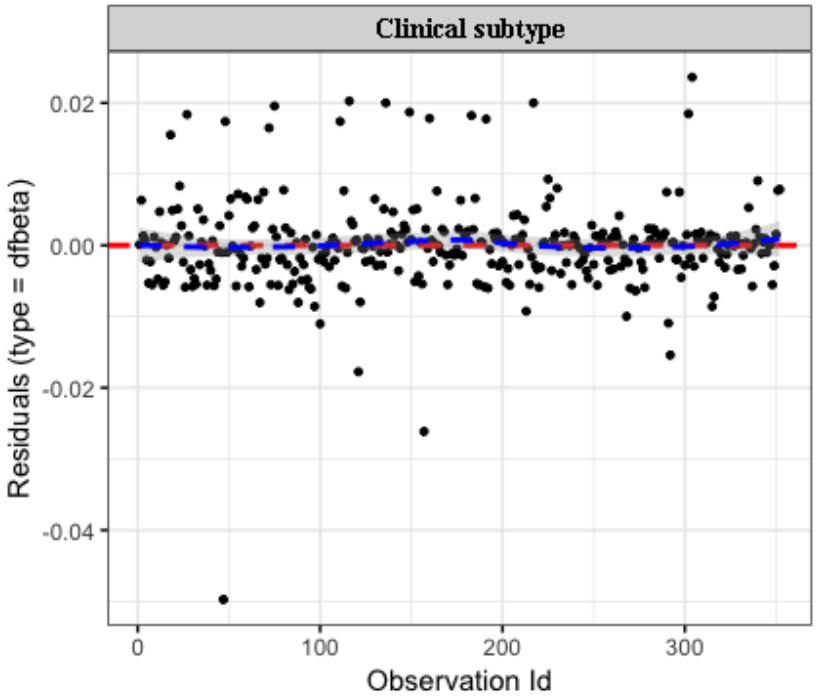
a) Sex



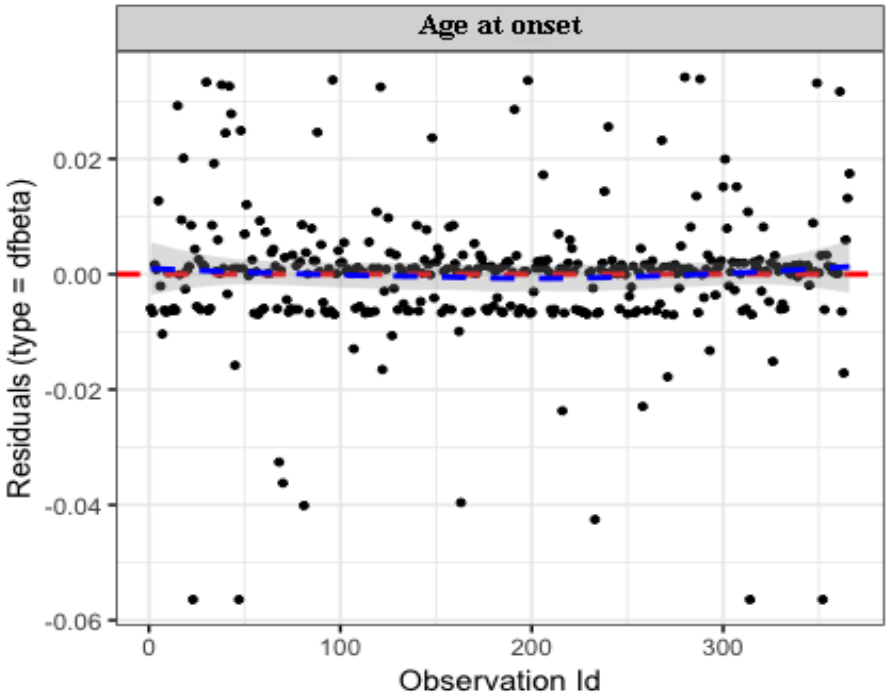
b) Familiarity



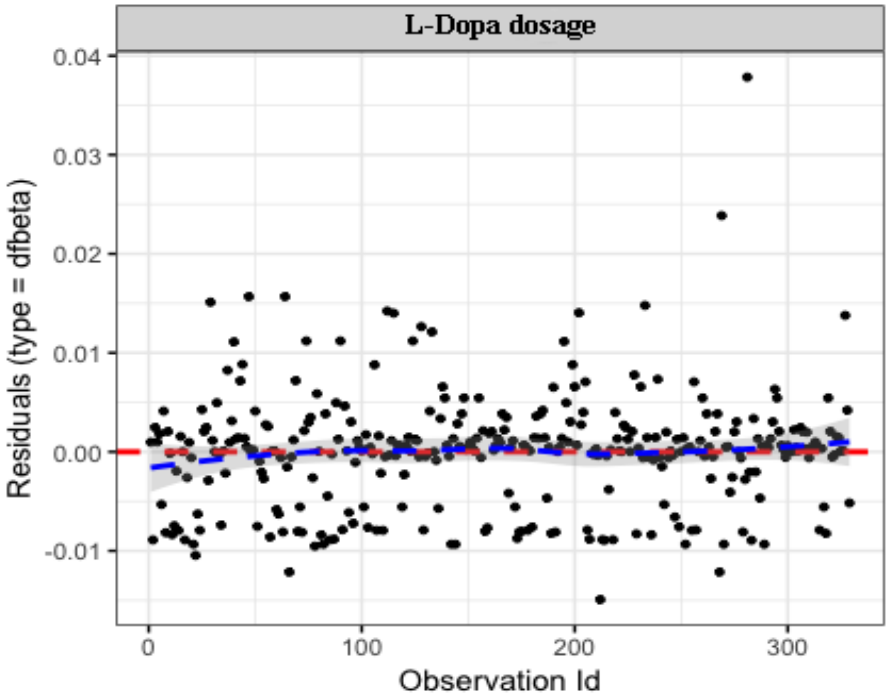
c) Clinical subtype



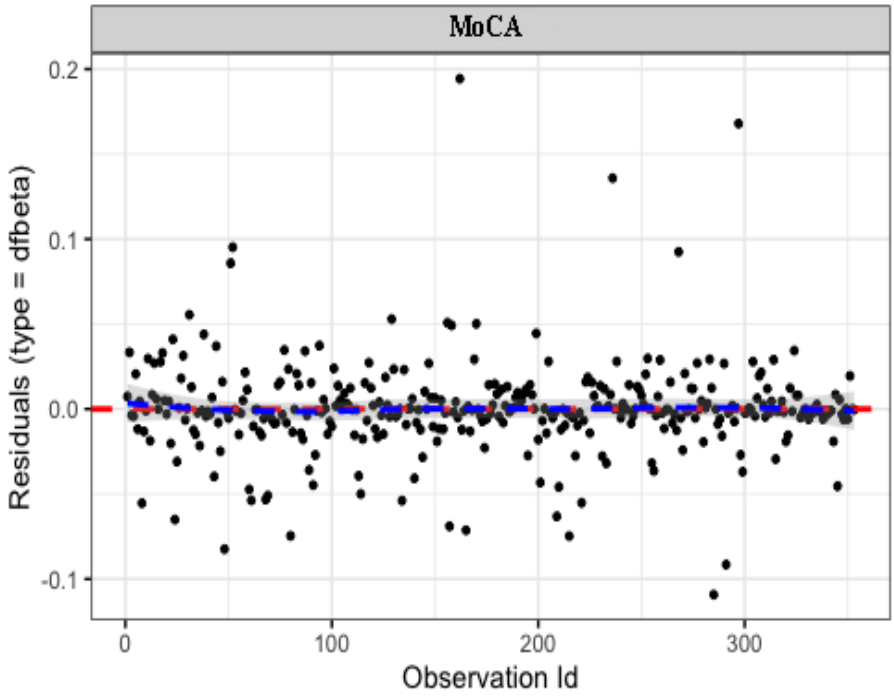
d) AAO



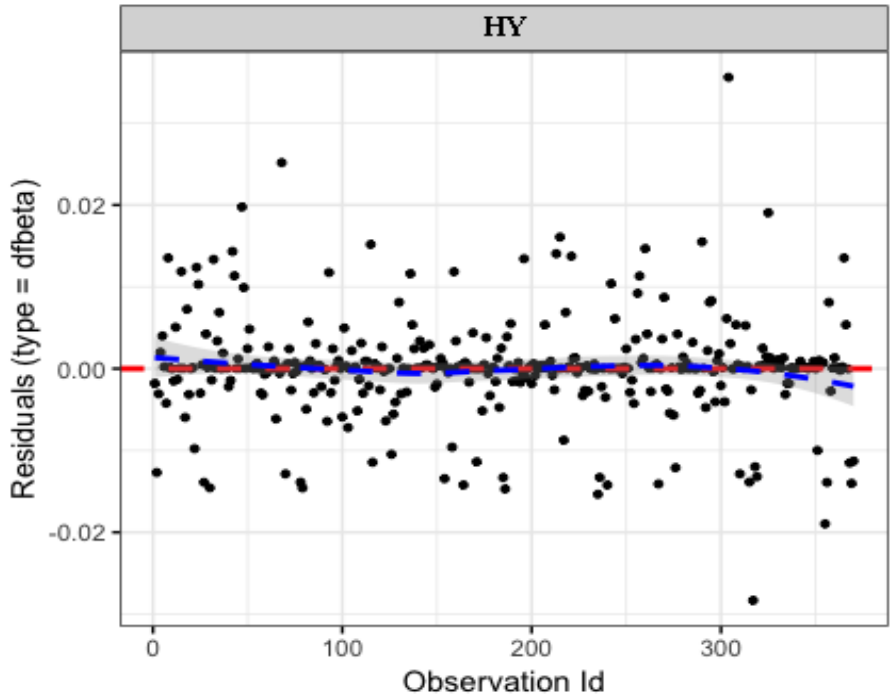
e) L-Dopa dosage



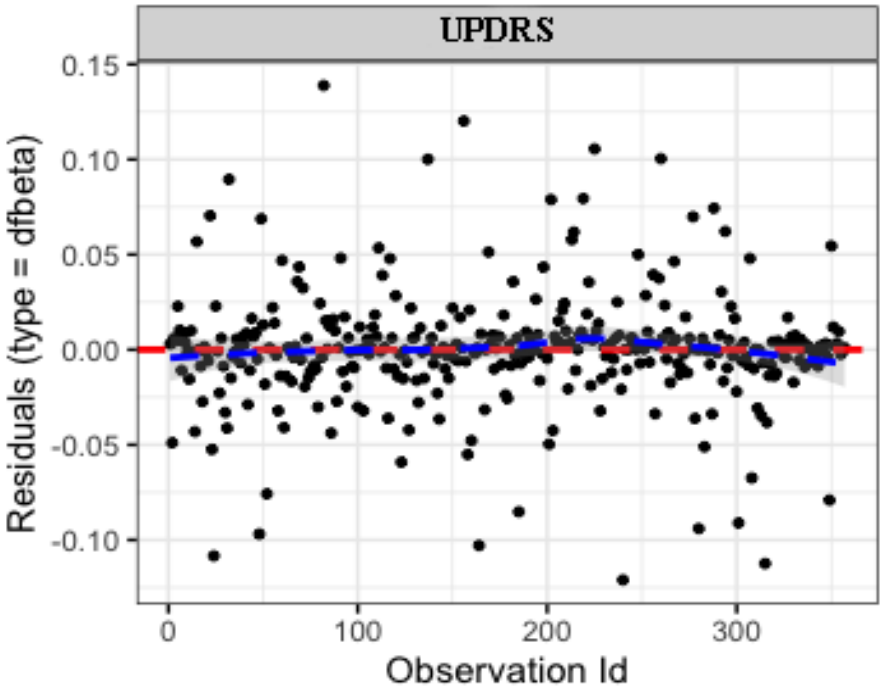
f) MoCA


MoCA
Residuals (type = dfbeta)
0.2
0.1
0.0
-0.1
0
100
200
300
Observation Id

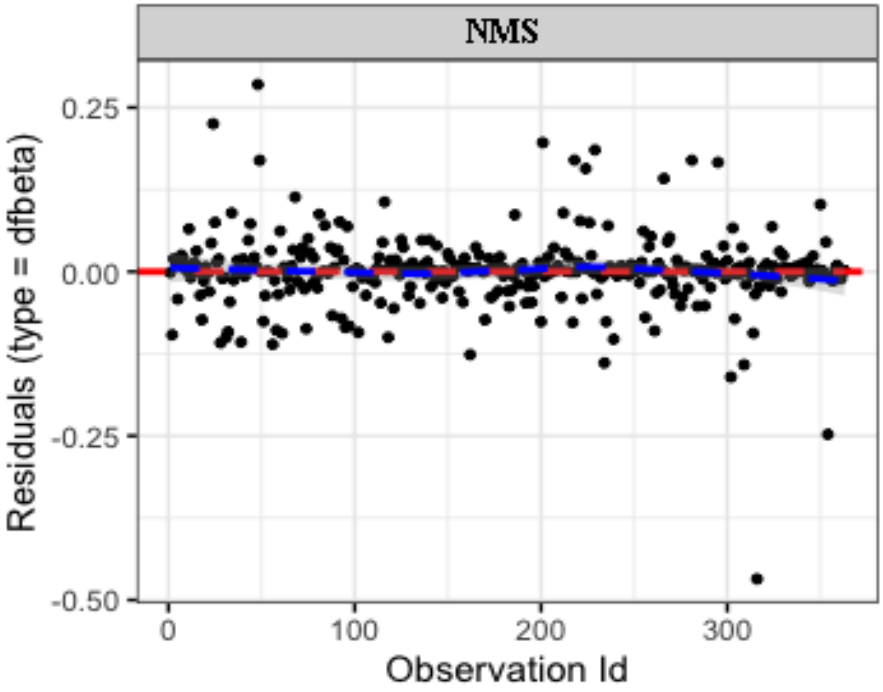
g) HY


HY
Residuals (type = dfbeta)
0.02
0.00
-0.02
0
100
200
300
Observation Id

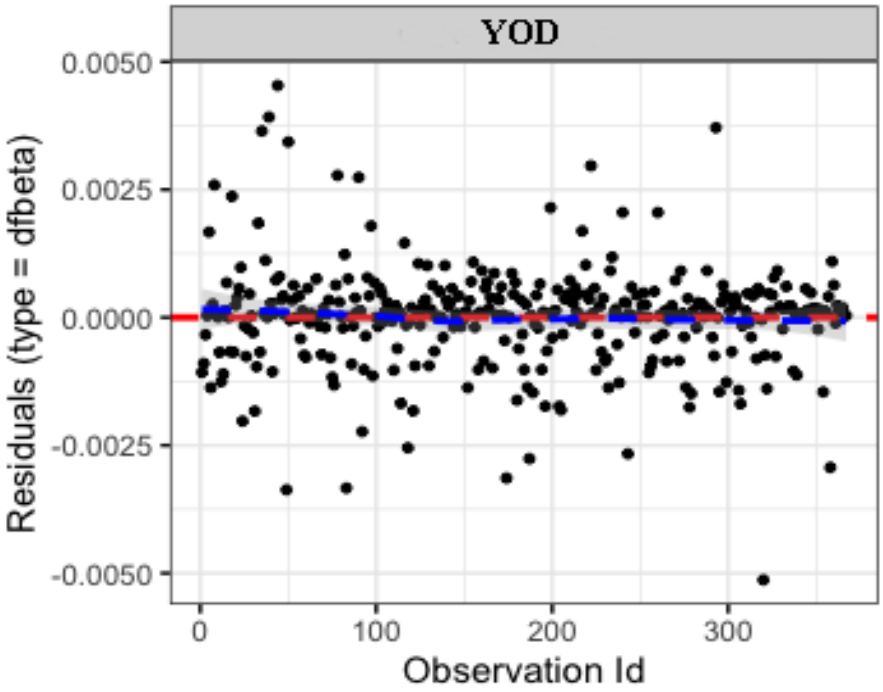
h) UPDRS


UPDRS
0.15
0.10
0.05
0.00
-0.05
-0.10
Residuals (type = dfbeta)
0
100
200
300
Observation Id

i) NMS


NMS
0.25
0.00
-0.25
-0.50
Residuals (type = dfbeta)
0
100
200
300
Observation Id

j) YOD

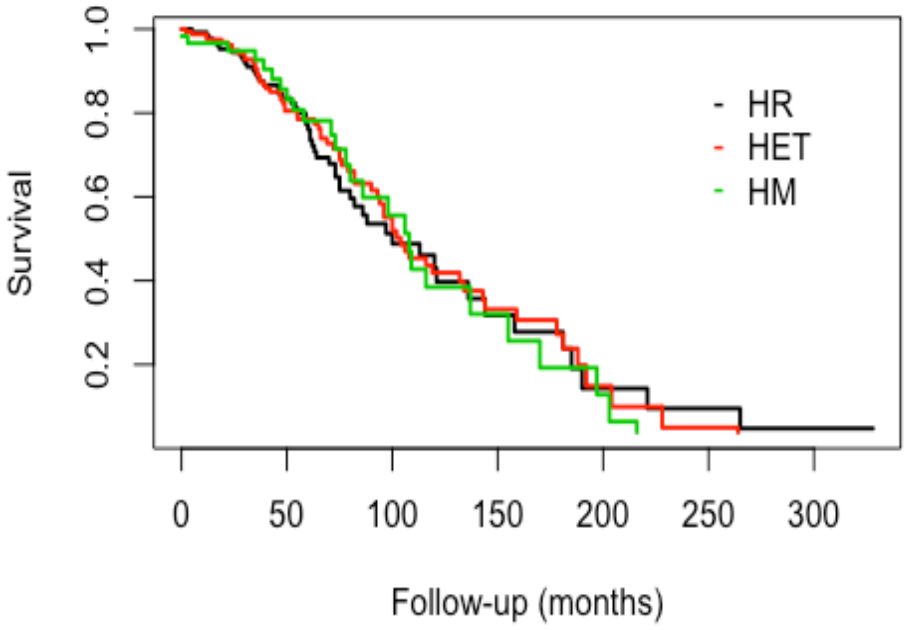


Dfbeta residuals for each observation available in the PD cohort are plotted, for each of the covariates tested.

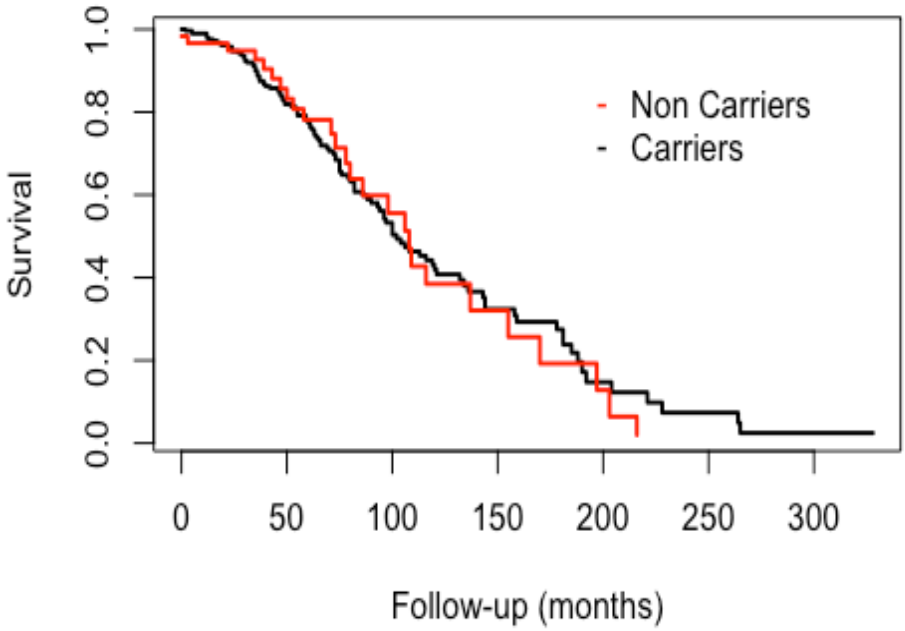
Abbreviations: AAO, PD age at onset; MoCA, Montreal Cognitive Assessment; HY, Hoehn & Yahr score; UPDRS, Movement Disorder Society revised version of the Unified Parkinson's Disease Rating Scale-Part III; MoCA, Montreal Cognitive Assessment; NMS, modified version of the Non-Motor Symptoms Scale for Parkinson Disease; YOD, years of disease.

Figure S3.3: Cox curves of crude (unadjusted) Cox PH regressions modelling incident LID risk vs rs356219 in a) Additive b) Dominant and c) Recessive model.

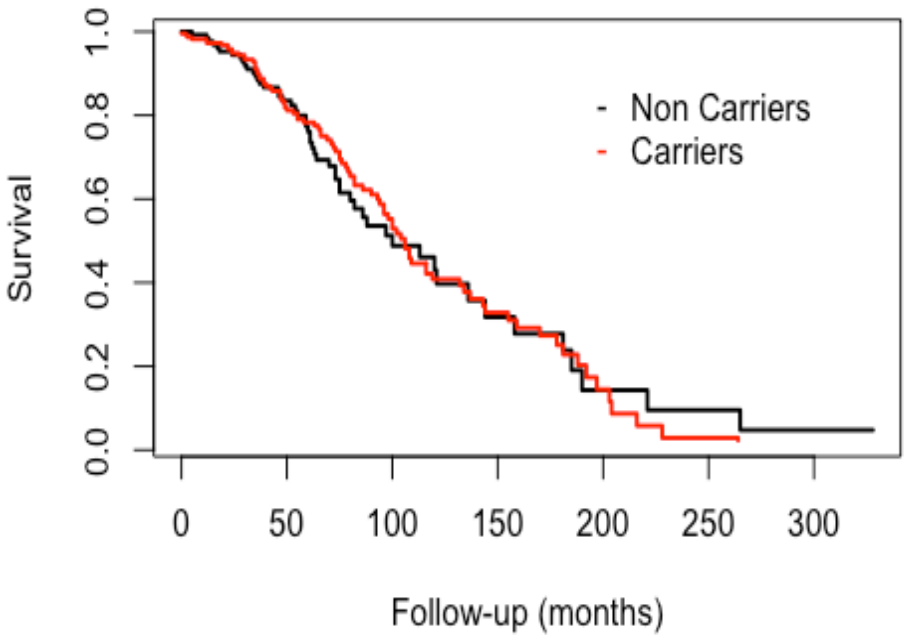
a) Additive model



b) Dominant model



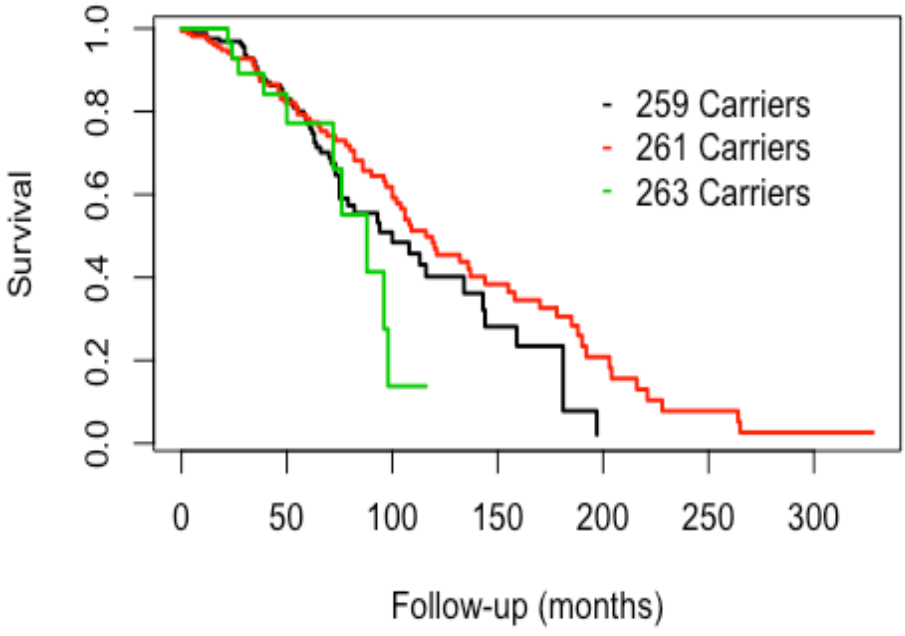
c) Recessive model



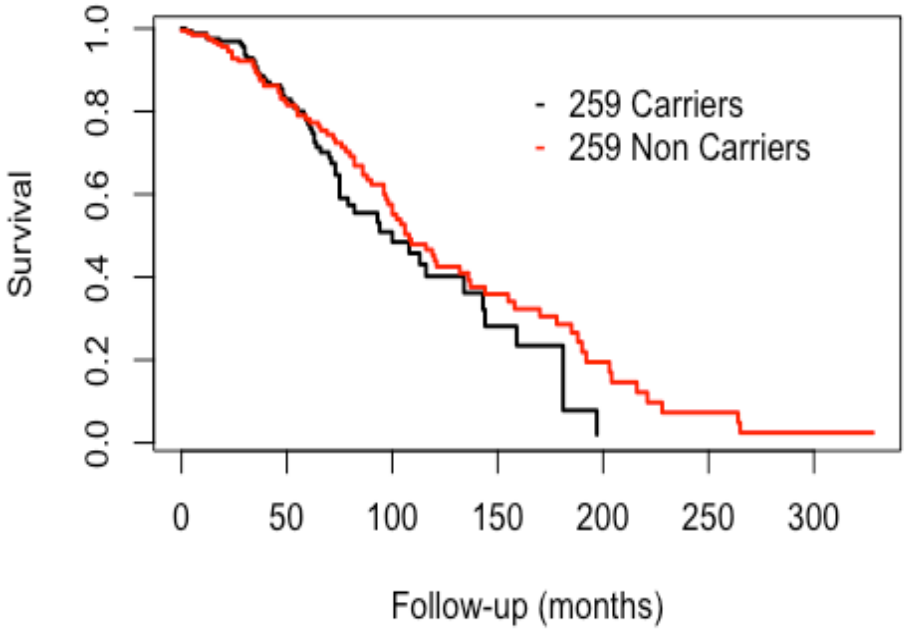
See **Table 3.3a** for further details on specification of genetic models tested.

Figure S7: Cox curves of crude (unadjusted) Cox PH regressions modelling incident LID risk vs D4S3481 in a) Pseudo-additive b) Pseudo-dominant and c) Pseudo-recessive model.

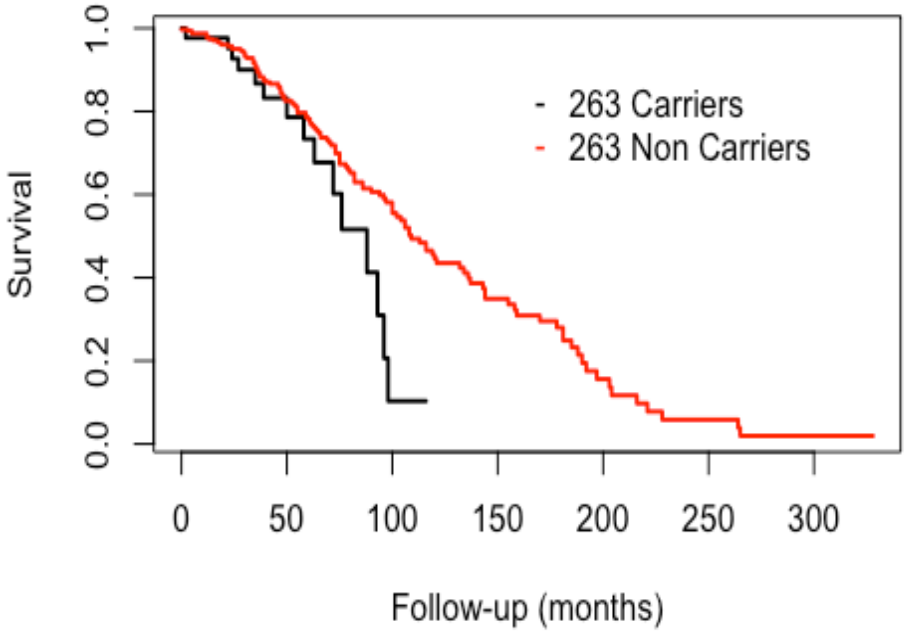
a) Pseudo-additive model



b) Pseudo-dominant model



c) Pseudo-recessive model



See **Table 3.3b** for further details on specification of genetic models tested.

Bibliography

1. Rieck M, Schumacher-Schuh AF, Altmann V, Francisconi CL, Fagundes PT, Monte TL, Callegari-Jacques SM, Rieder CR, Hutz MH. DRD2 haplotype is associated with dyskinesia induced by levodopa therapy in Parkinson's disease patients. *Pharmacogenomics* (2012) **13**:1701–1710. doi:10.2217/pgs.12.149

2. Parkinson J. An essay on the shaking palsy. 1817. *J Neuropsychiatry Clin Neurosci* (2002) **14**: doi:10.1176/jnp.14.2.223

3. Tran TN, Vo TNN, Frei K, Truong DD. Levodopa-induced dyskinesia: clinical features, incidence, and risk factors. *J Neural Transm* (2018) **125**:1109–1117. doi:10.1007/s00702-018-1900-6

4. Capriotti T, Terzakis K. Parkinson Disease. *Home Healthc now* (2016) **34**:300–307. doi:10.1097/NHH.0000000000000398

5. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. *Lancet Neurol* (2006) **5**:525–535. doi:10.1016/S1474-4422(06)70471-9

6. Olanow CW, Stern MB, Sethi K. The scientific and clinical basis for the treatment of Parkinson disease (2009). *Neurology* (2009) **72**: doi:10.1212/WNL.0b013e3181a1d44c

7. Poewe W, Gauthier S, Aarsland D, Leverenz JB, Barone P, Weintraub D, Tolosa E, Dubois B. Diagnosis and management of Parkinson's disease dementia. *Int J Clin Pract* (2008) **62**:1581–1587. doi:10.1111/j.1742-1241.2008.01869.x

8. Ferrer I. Neuropathology and neurochemistry of nonmotor symptoms in Parkinson's disease. *Parkinsons Dis* (2011) doi:10.4061/2011/708404

9. Titova N, Padmakumar C, Lewis SJG, Chaudhuri KR. Parkinson's: a syndrome rather than a disease? *J Neural Transm* (2017) **124**:907–914. doi:10.1007/s00702-016-1667-6

10. Mittur A, Gupta S, Modi NB. Pharmacokinetics of Rytary®, An Extended-Release Capsule Formulation of Carbidopa–Levodopa. *Clin Pharmacokinet* (2017) **56**:999–1014. doi:10.1007/s40262-017-0511-y

11. Hamza TH, Payami H. The heritability of risk and age at onset of Parkinson's disease after accounting for known genetic risk factors. *J Hum Genet* (2010) **55**:241–243. doi:10.1038/jhg.2010.13

12. Farlow JL, Robak LA, Hetrick K, Bowling K, Boerwinkle E, Coban-Akdemir ZH, Gambin

T, Gibbs RA, Gu S, Jain P, et al. Whole-exome sequencing in familial Parkinson disease. *JAMA Neurol* (2016) **73**:68–75. doi:10.1001/jamaneurol.2015.3266

13. Bonifati V. Genetics of Parkinson's disease - state of the art, 2013. *Park Relat Disord* (2014) **20**:S23–S28. doi:10.1016/S1353-8020(13)70009-9

14. Shulskaya M V., Alieva AK, Vlasov IN, Zyrin V V., Fedotova EY, Abramycheva NY, Usenko TS, Yakimovsky AF, Emelyanov AK, Pchelina SN, et al. Whole-exome sequencing in searching for new variants associated with the development of Parkinson's disease. *Front Aging Neurosci* (2018) **10**:1–8. doi:10.3389/fnagi.2018.00136

15. Schulte C, Gasser T. Genetic basis of Parkinson's disease: Inheritance, penetrance, and expression. *Appl Clin Genet* (2011) **4**:67–80. doi:10.2147/TACG.S11639

16. Domingo A, Klein C. Genetics of Parkinson disease. *Handb Clin Neurol* (2018) **147**:211–227. doi:10.1016/B978-0-444-63233-3.00014-2

17. Berwick DC, Heaton GR, Azeggagh S, Harvey K. LRRK2 Biology from structure to dysfunction: research progresses, but the themes remain the same. *Mol Neurodegener* (2019) **14**:49. doi:10.1186/s13024-019-0344-2

18. Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, Yokochi M, Mizuno Y, Shimizu N. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* (1998) **392**:605–608. doi:10.1038/33416

19. Periquet M, Lücking CB, Vaughan JR, Bonifati V, ürr A, De Michele G, Horstink MW, Farrer M, Illarioshkin SN, Pollak P, et al. Origin of the mutations in the parkin gene in Europe: Exon rearrangements are independent recurrent events, whereas point mutations may result from founder effects. *Am J Hum Genet* (2001) **68**:617–626. doi:10.1086/318791

20. Shimura H, Hattori N, Kubo SI, Mizuno Y, Asakawa S, Minoshima S, Shimizu N, Iwai K, Chiba T, Tanaka K, et al. Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat Genet* (2000) **25**:302–305. doi:10.1038/77060

21. Lücking CB, Dürr A, Bonifati V, Vaughan J, De Michele G, Gasser T, Harhangi BS, Meco G, Denèfle P, Wood NW, et al. Association between Early-Onset Parkinson's Disease and Mutations in the *Parkin* Gene. *N Engl J Med* (2000) **342**:1560–1567. doi:10.1056/NEJM200005253422103

22. ATP13A2 gene - Genetics Home Reference - NIH. Available at: https://ghr.nlm.nih.gov/gene/ATP13A2 [Accessed January 7, 2020]

23. Grünewald A, Arns B, Seibler P, Rakovic A, Münchau A, Ramirez A, Sue CM, Klein C. ATP13A2 mutations impair mitochondrial function in fibroblasts from patients with Kufor-Rakeb syndrome. *Neurobiol Aging* (2012) **33**:1843.e1-1843.e7. doi:10.1016/j.neurobiolaging.2011.12.035

24. Estrada-Cuzcano A, Martin S, Chamova T, Synofzik M, Timmann D, Holemans T, Andreeva A, Reichbauer J, De Rycke R, Chang DI, et al. Loss-of-function mutations in the ATP13A2/PARK9 gene cause complicated hereditary spastic paraplegia (SPG78). *Brain* (2017) **140**:287–305. doi:10.1093/brain/aww307

25. Bento CF, Ashkenazi A, Jimenez-Sanchez M, Rubinsztein DC. The Parkinson's disease-associated genes ATP13A2 and SYT11 regulate autophagy via a common pathway. *Nat Commun* (2016) **7**: doi:10.1038/ncomms11803

26. Valente EM, Salvi S, Ialongo T, Marongiu R, Elia AE, Caputo V, Romito L, Albanese A, Dallapiccola B, Bentivoglio AR. PINK1 mutations are associated with sporadic early-onset parkinsonism. *Ann Neurol* (2004) **56**:336–341. doi:10.1002/ana.20256

27. Bonifati V, Rohé CF, Breedveld GJ, Fabrizio E, De Mari M, Tassorelli C, Tavella A, Marconi R, Nicholl DJ, Chien HF, et al. Early-onset parkinsonism associated with PINK1 mutations: Frequency, genotypes, and phenotypes. *Neurology* (2005) **65**:87–95. doi:10.1212/01.wnl.0000167546.39375.82

28. Tan E-K, Yew K, Chua E, Puvan K, Shen H, Lee E, Puong K-Y, Zhao Y, Pavanni R, Wong M-C, et al. PINK1 mutations in sporadic early-onset Parkinson's disease. *Mov Disord* (2006) **21**:789–793. doi:10.1002/mds.20810

29. Li Y, Tomiyama H, Sato K, Hatano Y, Yoshino H, Atsumi M, Kitaguchi M, Sasaki S, Kawaguchi S, Miyajima H, et al. Clinicogenetic study of PINK1 mutations in autosomal recessive early-onset parkinsonism. *Neurology* (2005) **64**:1955–7. doi:10.1212/01.WNL.0000164009.36740.4E

30. Shendelman S, Jonason A, Martinat C, Leete T, Abeliovich A. DJ-1 Is a Redox-Dependent Molecular Chaperone That Inhibits α-Synuclein Aggregate Formation. *PLoS Biol* (2004) **2**:e362. doi:10.1371/journal.pbio.0020362

31. Zhou W, Zhu M, Wilson MA, Petsko GA, Fink AL. The oxidation state of DJ-1 regulates its chaperone activity toward α-synuclein. *J Mol Biol* (2006) **356**:1036–1048. doi:10.1016/j.jmb.2005.12.030

32. Wider C, Wszolek ZK. Clinical genetics of Parkinson's disease and related disorders. *Park Relat Disord* (2007) **13**: doi:10.1016/S1353-8020(08)70007-5

33. Fujibayashi A, Taguchi T, Misaki R, Ohtani M, Dohmae N, Takio K, Yamada M, Gu J, Yamakami M, Fukuda M, et al. Human RME-8 is involved in membrane trafficking through early endosomes. *Cell Struct Funct* (2008) **33**:35–50. doi:10.1247/csf.07045

34. Vilariño-Güell C, Rajput A, Milnerwood AJ, Shah B, Szu-Tu C, Trinh J, Yu I, Encarnacion M, Munsie LN, Tapia L, et al. DNAJC13 mutations in Parkinson disease. *Hum Mol Genet* (2014) **23**:1794–1801. doi:10.1093/hmg/ddt570

35. Sidransky E, Lopez G. The link between the GBA gene and parkinsonism. *Lancet Neurol* (2012) **11**:986–998. doi:10.1016/S1474-4422(12)70190-4

36. Verstraeten A, Theuns J, Van Broeckhoven C. Progress in unraveling the genetic etiology of Parkinson disease in a genomic era. *Trends Genet* (2015) **31**:140–149. doi:10.1016/J.TIG.2015.01.004

37. Kim C, Alcalay R. Genetic Forms of Parkinson's Disease. *Semin Neurol* (2017) **37**:135–146. doi:10.1055/s-0037-1601567

38. Köroğlu Ç, Baysal L, Cetinkaya M, Karasoy H, Tolun A. DNAJC6 is responsible for juvenile parkinsonism with phenotypic variability. *Parkinsonism Relat Disord* (2013) **19**:320–324. doi:10.1016/j.parkreldis.2012.11.006

39. Mizuno Y, Hattori N, Kubo SI, Sato S, Nishioka K, Hatano T, Tomiyama H, Funayama M, MacHida Y, Mochizuki H. Review. Progress in the pathogenesis and genetics of Parkinson's disease. *Philos Trans R Soc B Biol Sci* (2008) **363**:2215–2227. doi:10.1098/rstb.2008.2273

40. Valente EM. Hereditary Early-Onset Parkinson's Disease Caused by Mutations in PINK1. *Science (80-)* (2004) **304**:1158–1160. doi:10.1126/science.1096284

41. Bonifati V. Mutations in the DJ-1 Gene Associated with Autosomal Recessive Early-Onset Parkinsonism. *Science (80-)* (2003) **299**:256–259. doi:10.1126/science.1077209

42. Chang D, Nalls MA, Hallgrimsdottir IB, Hunkapiller J, van der Brug M, Cai F, Kerchner GA, Ayalon G, Bingol B, Sheng M, et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat Genet* (2017) **49**:1511–1516. doi:10.1038/ng.3955

43. Quadri M, Yang X, Cossu G, Olgiati S, Saddi VM, Breedveld GJ, Ouyang L, Hu J, Xu N,

Graafland J, et al. An exome study of Parkinson's disease in Sardinia, a Mediterranean genetic isolate. *Neurogenetics* (2015) **16**:55–64. doi:10.1007/s10048-014-0425-x

44. Ruiz-Martínez J, Azcona LJ, Bergareche A, Martí-Massó JF, Paisán-Ruiz C. Whole-exome sequencing associates novel CSMD1 gene mutations with familial Parkinson disease. *Neurol Genet* (2017) **3**:1–6. doi:10.1212/NXG.0000000000000177

45. Sandor C, Honti F, Haerty W, Szewczyk-Krolikowski K, Tomlinson P, Evetts S, Millin S, Keane T, McCarthy SA, Durbin R, et al. Whole-exome sequencing of 228 patients with sporadic Parkinson's disease. *Sci Rep* (2017) **7**: doi:10.1038/srep41188

46. Siitonen A, Nalls MA, Hernández D, Gibbs JR, Ding J, Ylikotila P, Edsall C, Singleton A, Majamaa K. Genetics of early-onset Parkinson's disease in Finland: exome sequencing and genome-wide association study. *Neurobiol Aging* (2017) **53**:195.e7-195.e10. doi:10.1016/j.neurobiolaging.2017.01.019

47. Jansen IE, Ye H, Heetveld S, Lechler MC, Michels H, Seinstra RI, Lubbe SJ, Drouet V, Lesage S, Majounie E, et al. Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing. *Genome Biol* (2017) **18**: doi:10.1186/s13059-017-1147-9

48. Kalinderi K, Bostantjopoulou S, Fidani L. The genetic background of Parkinson's disease: current progress and future prospects. *Acta Neurol Scand* (2016) **134**:314–326. doi:10.1111/ane.12563

49. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, Tan M, Kia DA, Noyce AJ, Xue A, et al. Expanding Parkinson's disease genetics: novel risk loci, genomic context, causal insights and heritable risk. *bioRxiv* (2019)388165. doi:10.1101/388165

50. Blauwendraat C, Heilbron K, Vallerga CL, Bandres-Ciga S, von Coelln R, Pihlstrøm L, Simón-Sánchez J, Schulte C, Sharma M, Krohn L, et al. Parkinson's disease age at onset genome-wide association study: Defining heritability, genetic loci, and α-synuclein mechanisms. *Mov Disord* (2019) **34**:866–875. doi:10.1002/mds.27659

51. Tan M, Hubbard L, Lawton M, Kanavou S, Wood N, Hardy J, Williams N, Grosset D, Morris H. Genome-wide association studies of motor and cognitive progression in Parkinson ' s disease. (2018) **33**:2018–2019. doi:10.1001/jamaneurol.2017.4206.

52. Chung SJ, Choi N, Kim J, Kim K, Kim M-J, Kim YJ, Ryu H-S, Park KW. Genomic variants

associated with cognitive impairment in Parkinson's disease: Ethnicity-specific genome-wide association study (P2.8-034). *Neurology* (2019) **92**:

53. Ibanez L, Farias FHG, Dube U, Mihindukulasuriya KA, Harari O. Polygenic Risk Scores in Neurodegenerative Diseases : a Review. (2019)22–29.

54. Ibanez L, Dube U, Saef B, Budde J, Black K, Medvedeva A, Del-Aguila JL, Davis AA, Perlmutter JS, Harari O, et al. Parkinson disease polygenic risk score is associated with Parkinson disease status and age at onset but not with alpha-synuclein cerebrospinal fluid levels. *BMC Neurol* (2017) **17**:198. doi:10.1186/s12883-017-0978-z

55. Escott-Price V, Nalls MA, Morris HR, Lubbe S, Brice A, Gasser T, Heutink P, Wood NW, Hardy J, Singleton AB, et al. Polygenic risk of Parkinson disease is correlated with disease age at onset. *Ann Neurol* (2015) **77**:582–591. doi:10.1002/ana.24335

56. Paul KC, Schulz J, Bronstein JM, Lill CM, Ritz BR. Association of Polygenic Risk Score With Cognitive Decline and Motor Progression in Parkinson Disease. *JAMA Neurol* (2018) **75**:360. doi:10.1001/jamaneurol.2017.4206

57. Gialluisi A, Reccia MG, Tirozzi N, Nutile T, Lombardi A, De Sanctis C, Heutink P, Varanese S, Pietracupa S, Modugno N, et al. Whole Exome Sequencing study of Parkinson Disease and related endophenotypes in the Italian Population. *Front Neurol* (2019) **10**:1362. doi:10.3389/FNEUR.2019.01362

58. Mata IF, Leverenz JB, Weintraub D, Trojanowski JQ, Chen-Plotkin A, Van Deerlin VM, Ritz B, Rausch R, Factor SA, Wood-Siverio C, et al. *GBA* Variants are associated with a distinct pattern of cognitive deficits in Parkinson's disease. *Mov Disord* (2016) **31**:95–102. doi:10.1002/mds.26359

59. Satake W, Nakabayashi Y, Mizuta I, Hirota Y, Ito C, Kubo M, Kawaguchi T, Tsunoda T, Watanabe M, Takeda A, et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat Genet* (2009) **41**:1303–1307. doi:10.1038/ng.485

60. Lill CM. Genetics of Parkinson's disease. *Mol Cell Probes* (2016) **30**:386–396. doi:10.1016/j.mcp.2016.11.001

61. Cooper CA, Jain N, Gallagher MD, Weintraub D, Xie SX, Berlyand Y, Espay AJ, Quinn J, Edwards KL, Montine T, et al. Common variant rs356182 near SNCA defines a Parkinson's disease endophenotype. *Ann Clin Transl Neurol* (2017) **4**:15–25. doi:10.1002/acn3.371

62. Shu L, Liang D, Pan H, Xu Q, Guo J, Sun Q, Tang B, Yan X. Genetic Impact on Clinical Features in Parkinson's Disease: A Study on SNCA -rs11931074. *Parkinsons Dis* (2018) **2018**: doi:10.1155/2018/2754541

63. Ibanez L, Dube U, Davis AA, Fernandez M V., Budde J, Cooper B, Diez-Fairen M, Ortega-Cubero S, Pastor P, Perlmutter JS, et al. Pleiotropic effects of variants in dementia genes in Parkinson disease. *Front Neurosci* (2018) **12**: doi:10.3389/fnins.2018.00230

64. Polymeropoulos MH, Higgins JJ, Golbe LI, Johnson WG, Ide SE, Di Iorio G, Sanges G, Stenroos ES, Pho LT, Schaffer AA, et al. Mapping of a gene for Parkinson's disease to chromosome 4q21-q23. *Science (80-)* (1996) **274**:1197–1199. doi:10.1126/science.274.5290.1197

65. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R, et al. Mutation in the α-synuclein gene identified in families with Parkinson's disease. *Science (80-)* (1997) **276**:2045–2047. doi:10.1126/science.276.5321.2045

66. Konno T, Siuda J, Wszolek ZK. Genetics of Parkinson's disease: a review of SNCA and LRRK2. *Wiad Lek* (2016) **69**:328–332.

67. George JM. The synucleins. *Genome Biol* (2002) **3**: doi:10.1186/gb-2001-3-1-reviews3002

68. Rodriguez JA, Ivanova MI, Sawaya MR, Cascio D, Reyes FE, Shi D, Sangwan S, Guenther EL, Johnson LM, Zhang M, et al. Structure of the toxic core of α-synuclein from invisible crystals. *Nature* (2015) **525**:486–490. doi:10.1038/nature15368

69. Spillantini MG, Schmidt ML, Lee VMY, Trojanowski JQ, Jakes R, Goedert M. α-synuclein in Lewy bodies [8]. *Nature* (1997) **388**:839–840. doi:10.1038/42166

70. Spillantini MG, Crowther RA, Jakes R, Hasegawa M, Goedert M. α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc Natl Acad Sci U S A* (1998) **95**:6469–6473. doi:10.1073/pnas.95.11.6469

71. Culvenor JG, McLean CA, Cutt S, Campbell BCV, Maher F, Jäkälä P, Hartmann T, Beyreuther K, Masters CL, Li QX. Non-Aβ component of Alzheimer's disease amyloid (NAC) revisited: NAC and α-Synuclein are not associated with Aβ amyloid. *Am J Pathol* (1999) **155**:1173–1181. doi:10.1016/S0002-9440(10)65220-0

72. Burré J, Sharma M, Südhof TC. Systematic mutagenesis of α-synuclein reveals distinct sequence requirements for physiological and pathological activities. *J Neurosci* (2012)

32:15227–15242. doi:10.1523/JNEUROSCI.3545-12.2012

73. Butler B, Saha K, Rana T, Becker JP, Sambo D, Davari P, Goodwin JS, Khoshbouei H. Dopamine transporter activity is modulated by α-synuclein. *J Biol Chem* (2015) **290**:29542–29554. doi:10.1074/jbc.M115.691592

74. Burré J, Sharma M, Tsetsenis T, Buchman V, Etherton MR, Südhof TC. α-Synuclein promotes SNARE-complex assembly in vivo and in vitro. *Science (80-)* (2010) **329**:1663–1667. doi:10.1126/science.1195227

75. Campêlo CLDC, Silva RH. Genetic Variants in SNCA and the Risk of Sporadic Parkinson's Disease and Clinical Outcomes: A Review. *Parkinsons Dis* (2017) **2017**: doi:10.1155/2017/4318416

76. McCarthy JJ, Linnertz C, Saucier L, Burke JR, Hulette CM, Welsh-Bohmer KA, Chiba-Falek O. The effect of SNCA 3′ region on the levels of SNCA-112 splicing variant. *Neurogenetics* (2011) **12**:59–64. doi:10.1007/s10048-010-0263-4

77. Cardo LF, Coto E, de Mena L, Ribacoba R, Mata IF, Menéndez M, Moris G, Alvarez V. Alpha-synuclein transcript isoforms in three different brain regions from Parkinson's disease and healthy subjects in relation to the SNCA rs356165/rs11931074 polymorphisms. *Neurosci Lett* (2014) **562**:45–49. doi:10.1016/j.neulet.2014.01.009

78. Fuchs J, Tichopad A, Golub Y, Munz M, Schweitzer KJ, Wolf B, Berg D, Mueller JC, Gasser T. Genetic variability in the SNCA gene influences α-synuclein levels in the blood and brain. *FASEB J* (2008) **22**:1327–1334. doi:10.1096/fj.07-9348com

79. Alkanli N, Ay A. "The Relationship between Alpha-Synuclein (SNCA) Gene Polymorphisms and Development Risk of Parkinson's Disease," in *Synucleins - Biochemistry and Role in Diseases [Working Title]* (IntechOpen). doi:10.5772/intechopen.82808

80. Rahimi M, Akbari M, Jamshidi J, Tafakhori A, Emamalizadeh B, Darvish H. Genetic analysis of SNCA gene polymorphisms in Parkinson's disease in an Iranian population. *Basal Ganglia* (2017) **10**:4–7. doi:10.1016/j.baga.2017.08.001

81. Wu-Chou YH, Chen YT, Yeh TH, Chang HC, Weng YH, Lai SC, Huang CL, Chen RS, Huang YZ, Chen CC, et al. Genetic variants of SNCA and LRRK2 genes are associated with sporadic PD susceptibility: A replication study in a Taiwanese cohort. *Park Relat Disord* (2013) **19**:251–255. doi:10.1016/j.parkreldis.2012.10.019

82. Westerlund M, Belin AC, Anvret A, Håkansson A, Nissbrandt H, Lind C, Sydow O, Olson L, Galter D. Cerebellar α-synuclein levels are decreased in Parkinson's disease and do not correlate with SNCA polymorphisms associated with disease in a Swedish material. *FASEB J* (2008) **22**:3509–3514. doi:10.1096/fj.08-110148

83. Myhre R, Toft M, Kachergus J, Hulihan MM, Aasly JO, Klungland H, Farrer MJ. Multiple *alpha-synuclein* gene polymorphisms are associated with Parkinson's disease in a Norwegian population. *Acta Neurol Scand* (2008) **118**:320–327. doi:10.1111/j.1600-0404.2008.01019.x

84. Shu L, Zhang Y, Sun Q, Pan H, Guo J, Tang B. SNCA REP1 and Parkinson's disease. *Neurosci Lett* (2018) **682**:79–84. doi:10.1016/j.neulet.2018.05.043

85. Corrado L, De Marchi FD, Tunesi S, Oggioni GD, Carecchio M, Magistrelli L, Tesei S, Riboldazzi G, Fonzo A Di, Locci C, et al. The length of SNCA Rep1 microsatellite may influence cognitive evolution in Parkinson's disease. *Front Neurol* (2018) **9**:1–7. doi:10.3389/fneur.2018.00213

86. Trotta L, Guella I, Soldà G, Sironi F, Tesei S, Canesi M, Pezzoli G, Goldwurm S, Duga S, Asselta R. SNCA and MAPT genes: Independent and joint effects in Parkinson disease in the Italian population. *Park Relat Disord* (2012) **18**:257–262. doi:10.1016/j.parkreldis.2011.10.014

87. Chiba-Falek O. Effect of allelic variation at the NACP-Rep1 repeat upstream of the alpha-synuclein gene (SNCA) on transcription in a cell culture luciferase reporter system. *Hum Mol Genet* (2001) **10**:3101–3109. doi:10.1093/hmg/10.26.3101

88. Cronin KD, Ge D, Manninger P, Linnertz C, Rossoshek A, Orrison BM, Bernard DJ, El-Agnaf OMA, Schlossmacher MG, Nussbaum RL, et al. Expansion of the Parkinson disease-associated SNCA-Rep1 allele upregulates human α-synuclein in transgenic mouse brain. *Hum Mol Genet* (2009) **18**:3274–3285. doi:10.1093/hmg/ddp265

89. Chiba-Falek O, Kowalak JA, Smulson ME, Nussbaum RL. Regulation of α-synuclein expression by poly (ADP Ribose) polymerase-1 (PARP-1) binding to the NACP-Rep1 polymorphic site upstream of the SNCA gene. *Am J Hum Genet* (2005) **76**:478–492. doi:10.1086/428655

90. Ibáñez P, Bonnet AM, Débarges B, Lohmann E, Tison F, Pollak P, Agid Y, Dürr A, Brice PA. Causal relation between α-synuclein gene duplication and familial Parkinson's disease. *Lancet* (2004) **364**:1169–1171. doi:10.1016/S0140-6736(04)17104-3

91. Ritz B, Rhodes SL, Bordelon Y, Bronstein J. α-Synuclein Genetic Variants Predict Faster Motor Symptom Progression in Idiopathic Parkinson Disease. *PLoS One* (2012) **7**:e36199. doi:10.1371/journal.pone.0036199

92. Markopoulou K, Biernacka JM, Armasu SM, Anderson KJ, Ahlskog JE, Chase BA, Chung SJ, Cunningham JM, Farrer M, Frigerio R, et al. Does α-synuclein have a dual and opposing effect in preclinical vs. clinical Parkinson's disease? *Park Relat Disord* (2014) **20**:584–589. doi:10.1016/j.parkreldis.2014.02.021

93. Luo N, Li Y, Niu M, Zhou L, Yao M, Zhu L, Ye G, Kang W, Liu J. Variants in the SNCA locus are associated with the progression of Parkinson's disease. *Front Aging Neurosci* (2019) **11**:1–7. doi:10.3389/fnagi.2019.00110

94. Pihlstrøm L, Blauwendraat C, Cappelletti C, Berge-Seidl V, Langmyhr M, Henriksen SP, van de Berg WDJ, Gibbs JR, Cookson MR, Singleton AB, et al. A comprehensive analysis of SNCA-related genetic risk in sporadic parkinson disease. *Ann Neurol* (2018) **84**:117–129. doi:10.1002/ana.25274

95. Mata IF, Shi M, Agarwal P, Chung KA, Edwards KL, Factor SA, Galasko DR, Ginghina C, Griffith A, Higgins DS, et al. SNCA variant associated with Parkinson disease and plasma α-synuclein level. *Arch Neurol* (2010) **67**:1350–1356. doi:10.1001/archneurol.2010.279

96. Campêlo CLC, Cagni FC, Figueredo D de S, Oliveira LG, Silva-Neto AB, Macêdo PT, Santos JR, Izídio GS, Ribeiro AM, de Andrade TG, et al. Variants in SNCA Gene are associated with Parkinson's disease risk and cognitive symptoms in a Brazilian sample. *Front Aging Neurosci* (2017) **9**: doi:10.3389/fnagi.2017.00198

97. Westerlund M, Ran C, Borgkvist A, Sterky FH, Lindqvist E, Lundströmer K, Pernold K, Brené S, Kallunki P, Fisone G, et al. Lrrk2 and α-synuclein are co-regulated in rodent striatum. *Mol Cell Neurosci* (2008) **39**:586–591. doi:10.1016/j.mcn.2008.08.001

98. Pankratz N, Wilk JB, Latourelle JC, DeStefano AL, Halter C, Pugh EW, Doheny KF, Gusella JF, Nichols WC, Foroud T, et al. Genomewide association study for susceptibility genes contributing to familial Parkinson disease. *Hum Genet* (2009) **124**:593–605. doi:10.1007/s00439-008-0582-9

99. Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, Destefano AL, Kara E, Bras J, Sharma M, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* (2014) **46**:989–993.

doi:10.1038/ng.3043

100. Cheng L, Wang L, Li NN, Yu WJ, Sun XY, Li JY, Zhou D, Peng R. SNCA rs356182 variant increases risk of sporadic Parkinson's disease in ethnic Chinese. *J Neurol Sci* (2016) **368**:231–234. doi:10.1016/j.jns.2016.07.032

101. Brockmann K, Schulte C, Hauser A-K, Lichtner P, Huber H, Maetzler W, Berg D, Gasser T. SNCA: Major genetic modifier of age at onset of Parkinson's disease. *Mov Disord* (2013) **28**:1217–1221. doi:10.1002/mds.25469

102. Gottesman II, Gould TD. The endophenotype concept in psychiatry: Etymology and strategic intentions. *Am J Psychiatry* (2003) **160**:636–645. doi:10.1176/appi.ajp.160.4.636

103. Cooper C, Berlyand Y, Xie S, Weintraub D, Trojanowski J, Zabetian C, Chen-Plotkin A. Genetic Variation Near the SNCA Gene Associates with Parkinson's Disease Motor Endophenotype and Progression (I1.002). *Neurology* (2016) **86**:

104. Faurbye A, Rasch P-J, Petersen PB, & GB, Pakkenberg H. NEUROLOGICAL SYMPTOMS IN PHARMAGOTHERAPY OF PSYCHOSES. *Acta Psychiatr Scand* (1964) **40**:10–27. doi:10.1111/j.1600-0447.1964.tb05731.x

105. Muenter MD, Sharpless NS, Tyce GM, Darley FL. Patterns of dystonia ("I-D-I" and "D-I-D-") in response to l-dopa therapy for Parkinson's disease. *Mayo Clin Proc* (1977) **52**:163–174.

106. Marsden CD, Parkes JD, Quinn N. "Fluctuations of disability in Parkinson's disease – clinical aspects," in *Movement Disorders* (Elsevier), 96–122. doi:10.1016/b978-0-407-02295-9.50012-6

107. Prashanth LK, Fox S, Meissner WG. "L-dopa-induced dyskinesia-clinical presentation, genetics, and treatment," in *International Review of Neurobiology* (Academic Press Inc.), 31–54. doi:10.1016/B978-0-12-381328-2.00002-X

108. Jankovic J, Nour F. Respiratory dyskinesia in Parkinson's disease. *Neurology* (1986) **36**:303–304. doi:10.1212/wnl.36.2.303-b

109. Rice JE, Antic R, Thompson PD. Disordered respiration as a levodopa-induced dyskinesia in Parkinson's disease. *Mov Disord* (2002) **17**:524–527. doi:10.1002/mds.10072

110. Linazasoro G. New ideas on the origin of L-dopa-induced dyskinesias: Age, genes and neural plasticity. *Trends Pharmacol Sci* (2005) **26**:391–397. doi:10.1016/j.tips.2005.06.007

111. Comi C, Ferrari M, Marino F, Magistrelli L, Cantello R, Riboldazzi G, Bianchi M, Bono G, Cosentino M. Polymorphisms of Dopamine Receptor Genes and Risk of L-Dopa–Induced Dyskinesia in Parkinson's Disease. *Int J Mol Sci* (2017) **18**:242. doi:10.3390/ijms18020242

112. Olanow CW, Obeso JA, Stocchi F. Drug insight: Continuous dopaminergic stimulation in the treatment of Parkinson's disease. *Nat Clin Pract Neurol* (2006) **2**:382–392. doi:10.1038/ncpneuro0222

113. Olanow CW, Kieburtz K, Odin P, Espay AJ, Standaert DG, Fernandez HH, Vanagunas A, Othman AA, Widnell KL, Robieson WZ, et al. Continuous intrajejunal infusion of levodopa-carbidopa intestinal gel for patients with advanced Parkinson's disease: A randomised, controlled, double-blind, double-dummy study. *Lancet Neurol* (2014) **13**:141–149. doi:10.1016/S1474-4422(13)70293-X

114. Hauser RA, Hsu A, Kell S, Espay AJ, Sethi K, Stacy M, Ondo W, O'Connell M, Gupta S. Extended-release carbidopa-levodopa (IPX066) compared with immediate-release carbidopa-levodopa in patients with Parkinson's disease and motor fluctuations: A phase 3 randomised, double-blind trial. *Lancet Neurol* (2013) **12**:346–356. doi:10.1016/S1474-4422(13)70025-5

115. Antonini A, Fung VSC, Boyd JT, Slevin JT, Hall C, Chatamra K, Eaton S, Benesh JA. Effect of levodopa-carbidopa intestinal gel on dyskinesia in advanced Parkinson's disease patients. *Mov Disord* (2016) **31**:530–537. doi:10.1002/mds.26528

116. Meloni B. Pathophysiology and Neuroprotective Strategies in Hypoxic-Ischemic Brain Injury and Stroke. *Brain Sci* (2017) **7**:110. doi:10.3390/brainsci7080110

117. Rascol O, Brooks DJ, Korczyn AD, De Deyn PP, Clarke CE, Lang AE. A Five-Year Study of the Incidence of Dyskinesia in Patients with Early Parkinson's Disease Who Were Treated with Ropinirole or Levodopa. *N Engl J Med* (2000) **342**:1484–1491. doi:10.1056/NEJM200005183422004

118. Holloway RG, Shoulson I, Fahn S, Kieburtz K, Lang A, Marek K, McDermott M, Seibyl J, Weiner W, Musch B, et al. Pramipexole vs levodopa as initial treatment for Parkinson disease: a 4-year randomized controlled trial. *Arch Neurol* (2004) **61**:1044–53. doi:10.1001/archneur.61.7.1044

119. F B, A B, C C, E D, O G, JF MM, JL M, Bracco F, Battaglia A, Chouza C, et al. "The long-acting dopamine receptor agonist cabergoline in early Parkinson's disease: final results of

a 5-year, double-blind, levodopa-controlled study," in *CNS drugs*, 733–746. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15330687 [Accessed January 8, 2020]

120. Chondrogiorgi M, Tatsioni A, Reichmann H, Konitsiotis S. Dopamine agonist monotherapy in Parkinson's disease and potential risk factors for dyskinesia: A meta-analysis of levodopa-controlled trials. *Eur J Neurol* (2014) **21**:433–440. doi:10.1111/ene.12318

121. Sharma JC, Bachmann CG, Linazasoro G. Classifying risk factors for dyskinesia in Parkinson's disease. *Park Relat Disord* (2010) **16**:490–497. doi:10.1016/j.parkreldis.2010.06.003

122. Wickremaratchi MM, Ben-Shlomo Y, Morris HR. The effect of onset age on the clinical features of Parkinson's disease. *Eur J Neurol* (2009) **16**:450–456. doi:10.1111/j.1468-1331.2008.02514.x

123. Ku S, Glass GA. Age of Parkinson's disease onset as a predictor for the development of dyskinesia. *Mov Disord* (2010) **25**:1177–1182. doi:10.1002/mds.23068

124. Warren Olanow C, Kieburtz K, Rascol O, Poewe W, Schapira AH, Emre M, Nissinen H, Leinonen M, Stocchi F. Factors predictive of the development of Levodopa-induced dyskinesia and wearing-off in Parkinson's disease. *Mov Disord* (2013) **28**:1064–1071. doi:10.1002/mds.25364

125. Kostic V, Przedborski S, Flaster E, Sternic N. Early development of levodopa-induced dyskinesias and response fluctuations in young-onset parkinson's disease. *Neurology* (1991) **41**:205. doi:10.1212/wnl.41.2_part_1.202

126. Kumar N, Van Gerpen JA, Bower JH, Ahlskog JE. Levodopa-dyskinesia incidence by age of Parkinson's disease onset. *Mov Disord* (2005) **20**:342–344. doi:10.1002/mds.20360

127. Ahlskog JE, Muenter MD. Frequency of levodopa-related dyskinesias and motor fluctuations as estimated from the cumulative literature. *Mov Disord* (2001) **16**:448–458. doi:10.1002/mds.1090

128. Schrag A, Quinn N. Dyskinesias and motor fluctuations in Parkinson's disease. *Brain* (2000) **123**:2297–2305. doi:10.1093/brain/123.11.2297

129. Zappia M, Annesi G, Nicoletti G, Arabia G, Annesi F, Messina D, Pugliese P, Spadafora P, Tarantino P, Carrideo S, et al. Sex differences in clinical and genetic determinants of levodopa peak-dose dyskinesias in Parkinson disease: An exploratory study. *Arch Neurol* (2005) **62**:601–605. doi:10.1001/archneur.62.4.601

130. Hassin-Baer S, Molchadski I, Cohen OS, Nitzan Z, Efrati L, Tunkel O, Kozlova E, Korczyn AD. Gender effect on time to levodopa-induced dyskinesias. *J Neurol* (2011) **258**:2048–2053. doi:10.1007/s00415-011-6067-0

131. Aubert I, Guigoni C, Håkansson K, Li Q, Dovero S, Barthe N, Bioulac BH, Gross CE, Fisone G, Bloch B, et al. Increased D $_1$ dopamine receptor signaling in levodopa-induced dyskinesia. *Ann Neurol* (2005) **57**:17–26. doi:10.1002/ana.20296

132. Sharma JC, Macnamara L, Hasoon M, Vassallo M, Ross I. Cascade of levodopa dose and weight-related dyskinesia in Parkinson's disease (LD-WD-PD cascade). *Park Relat Disord* (2006) **12**:499–505. doi:10.1016/j.parkreldis.2006.07.002

133. Arabia G, Zappia M, Bosco D, Crescibene L, Bagalà A, Bastone L, Caracciolo M, Scornaienghi M, Quattrone A. Body weight, levodopa pharmacokinetics and dyskinesia in Parkinson's disease. *Neurol Sci* (2002) **23**: doi:10.1007/s100720200066

134. Zhang Y, Meredith GE, Mendoza-Elias N, Rademacher DJ, Tseng KY, Steece-Collier K. Aberrant restoration of spines and their synapses in L-DOPA-induced dyskinesia: Involvement of corticostriatal but not thalamostriatal synapses. *J Neurosci* (2013) **33**:11655–11667. doi:10.1523/JNEUROSCI.0288-13.2013

135. Kipfer S, Stephan MA, Schüpbach WMM, Ballinari P, Kaelin-Lang A. Resting tremor in Parkinson disease: A negative predictor of levodopa-induced dyskinesia. *Arch Neurol* (2011) **68**:1037–1039. doi:10.1001/archneurol.2011.147

136. Kostic VS, Marinkovic J, Svetel M, Stefanova E, Przedborski S. The effect of stage of Parkinson's disease at the onset of levodopa therapy on development of motor complications. *Eur J Neurol* (2002) **9**:9–14. doi:10.1046/j.1468-1331.2002.00346.x

137. Zhou X, Guo J, Sun Q, Xu Q, Pan H, Yu R, Tan J, Yan X, Tang B, Fang L. Factors Associated With Dyskinesia in Parkinson's Disease in Mainland China. *Front Neurol* (2019) **10**: doi:10.3389/fneur.2019.00477

138. Schrag A, Ben-Shlomo Y, Brown R, Marsden CD, Quinn N. Young-onset Parkinson's disease revisited - Clinical features, natural history, and mortality. *Mov Disord* (1998) **13**:885–894. doi:10.1002/mds.870130605

139. Fahn S. Parkinson disease, the effect of levodopa, and the ELLDOPA trial. Earlier vs Later L-DOPA. *Arch Neurol* (1999) **56**:529–535.

140. Stoessl AJ. Central pharmacokinetics of levodopa: Lessons from imaging studies. *Mov*

Disord (2015) **30**:73–79. doi:10.1002/mds.26046

141. Hong JY, Oh JS, Lee I, Sunwoo MK, Ham JH, Lee JE, Sohn YH, Kim JS, Lee PH. Presynaptic dopamine depletion predicts levodopa-induced dyskinesia in de novo Parkinson disease. *Neurology* (2014) **82**:1597–1604. doi:10.1212/WNL.0000000000000385

142. Kalinderi K, Papaliagkas V, Fidani L. Pharmacogenetics and levodopa induced motor complications. *Int J Neurosci* (2019) **129**:384–392. doi:10.1080/00207454.2018.1538993

143. Wang X, Wang Z Bin, Luo C, Mao XY, Li X, Yin JY, Zhang W, Zhou HH, Liu ZQ. The prospective value of dopamine receptors on bio-behavior of tumor. *J Cancer* (2019) **10**:1622–1632. doi:10.7150/jca.27780

144. Rangel-Barajas C, Coronel I, Florán B. Dopamine Receptors and Neurodegeneration. *Aging Dis* (2015) **6**:349. doi:10.14336/AD.2015.0330

145. Grandy DK, Litt M, Allen L, Bunzow JR, Marchionni M, Makam H, Reed L, Magenis RE, Civelli O. The human dopamine D2 receptor gene is located on Chromosome 11 at q22-q23 and identifies a TaqI RFLP. *Am J Hum Genet* (1989) **45**:778–785.

146. Oliveri RL, Annesi G, Zappia M, Civitelli D, Montesanti R, Branca D, Nicoletti G, Spadafora P, Pasqua AA, Cittadella R, et al. Dopamine D2 receptor gene polymorphism and the risk of levodopa-induced dyskinesias in PD. *Neurology* (1999) **53**:1425–1430. doi:10.1212/wnl.53.7.1425

147. Strong JA, Dalvi A, Revilla FJ, Sahay A, Samaha FJ, Welge JA, Gong J, Gartner M, Yue X, Yu L. Genotype and smoking history affect risk of levodopa-induced dyskinesias in Parkinson's disease. *Mov Disord* (2006) **21**:654–659. doi:10.1002/mds.20785

148. Wang J, Liu ZL, Chen B. Association study of dopamine D2, D3 receptor gene polymorphisms with motor fluctuations in PD. *Neurology* (2001) **56**:1757–1759. doi:10.1212/WNL.56.12.1757

149. Corrado L, De Marchi F, Tunesi S, Oggioni GD, Carecchio M, Magistrelli L, Tesei S, Riboldazzi G, Di Fonzo A, Locci C, et al. The Length of SNCA Rep1 Microsatellite May Influence Cognitive Evolution in Parkinson's Disease. *Front Neurol* (2018) **9**: doi:10.3389/fneur.2018.00213

150. Bézard E, Ferry S, Mach U, Stark H, Leriche L, Boraud T, Gross C, Sokoloff P. Attenuation of levodopa-induced dyskinesia by normalizing dopamine D3 receptor function. *Nat Med*

(2003) **9**:762–767. doi:10.1038/nm875

151. Solís O, Garcia-Montes JR, González-Granillo A, Xu M, Moratalla R. Dopamine D3 Receptor Modulates l-DOPA-Induced Dyskinesia by Targeting D1 Receptor-Mediated Striatal Signaling. *Cereb Cortex* (2015)bhv231. doi:10.1093/cercor/bhv231

152. Lundstrom K, Turpin MP. Proposed schizophrenia-related gene polymorphism: Expression of the Ser9Gly mutant human dopamine D3 receptor with the Semliki Forest virus system. *Biochem Biophys Res Commun* (1996) **225**:1068–1072. doi:10.1006/bbrc.1996.1296

153. Lerer B, Segman RH, Fangerau H, Daly AK, Basile VS, Cavallaro R, Aschauer HN, McCreadie RG, Ohlraun S, Ferrier N, et al. Pharmacogenetics of tardive dyskinesia: Combined analysis of 780 patients supports association with dopamine D3 receptor gene Ser9Gly polymorphism. *Neuropsychopharmacology* (2002) **27**:105–119. doi:10.1016/S0893-133X(02)00293-2

154. Steen VM, Løvlie R, MacEwan T, McCreadie RG. Dopamine D3-receptor gene variant and susceptibility to tardive dyskinesia in schizophrenic patients. *Mol Psychiatry* (1997) **2**:139–145. doi:10.1038/sj.mp.4000249

155. Lee JY, Cho J, Lee EK, Park SS, Jeon BS. Differential genetic susceptibility in diphasic and peak-dose dyskinesias in Parkinson's disease. *Mov Disord* (2011) **26**:73–79. doi:10.1002/mds.23400

156. Solís O, Moratalla R. Dopamine receptors: homomeric and heteromeric complexes in l-DOPA-induced dyskinesia. *J Neural Transm* (2018) **125**:1187–1194. doi:10.1007/s00702-018-1852-x

157. Lanza K, Meadows SM, Chambers NE, Nuss E, Deak MM, Ferré S, Bishop C. Behavioral and cellular dopamine D1 and D3 receptor-mediated synergy: Implications for L-DOPA-induced dyskinesia. *Neuropharmacology* (2018) **138**:304–314. doi:10.1016/j.neuropharm.2018.06.024

158. Torvinen M, Kozell LB, Neve KA, Agnati LF, Fuxe K. Biochemical identification of the dopamine D2 receptor domains interacting with the adenosine A2A receptor. *J Mol Neurosci* (2004) **24**:173–180. doi:10.1385/JMN:24:2:173

159. Casetta I, Vincenzi F, Bencivelli D, Corciulo C, Gentile M, Granieri E, Borea PA, Varani K. A $_{2A}$ adenosine receptors and Parkinson's disease severity. *Acta Neurol Scand* (2014) **129**:276–281. doi:10.1111/ane.12181

160. Calon F. Increased adenosine A2A receptors in the brain of Parkinson's disease patients with dyskinesias. *Brain* (2004) **127**:1075–1084. doi:10.1093/brain/awh128

161. Rieck M, Schumacher-Schuh AF, Callegari-Jacques SM, Altmann V, Schneider Medeiros M, Rieder CRM, Hutz MH. Is there a role for ADORA2A polymorphisms in levodopa-induced dyskinesia in Parkinson's disease patients? *Pharmacogenomics* (2015) **16**:573–582. doi:10.2217/pgs.15.23

162. Piccini P, Weeks RA, Brooks DJ. Alterations in opioid receptor binding in Parkinson's disease patients with levodopa-induced dyskinesias. *Ann Neurol* (1997) **42**:720–726. doi:10.1002/ana.410420508

163. Peckys D, Landwehrmeyer GB. Expression of MU, KAPPA, and delta opioid receptor messenger RNA in the human CNS: A 33P in situ hybridization study. *Neuroscience* (1999) **88**:1093–1135. doi:10.1016/S0306-4522(98)00251-6

164. Bond C, Laforge KS, Tian M, Melia D, Zhang S, Borg L, Gong J, Schluger J, Strong JA, Leal SM, et al. Single-nucleotide polymorphism in the human mu opioid receptor gene alters β-endorphin binding and activity: Possible implications for opiate addiction. *Proc Natl Acad Sci U S A* (1998) **95**:9608–9613. doi:10.1073/pnas.95.16.9608

165. Henry B, Fox SH, Crossman AR, Brotchie JM. μ- and δ-Opioid receptor antagonists reduce levodopa-induced dyskinesia in the MPTP-lesioned primate model of Parkinson's disease. *Exp Neurol* (2001) **171**:139–146. doi:10.1006/exnr.2001.7727

166. Chen L, Togasaki DM, Langston JW, Di Monte DA, Quik M. Enhanced striatal opioid receptor-mediated G-protein activation in L-dopa-treated dyskinetic monkeys. *Neuroscience* (2005) **132**:409–420. doi:10.1016/j.neuroscience.2004.10.026

167. Elahi B, Phielipp N, Chen R. N-methyl-d-aspartate antagonists in levodopa induced dyskinesia: A meta-analysis. *Can J Neurol Sci* (2012) **39**:465–472. doi:10.1017/S0317167100013974

168. Weinshilboum R. Pharmacogenetics of methylation: Relationship to drug metabolism. *Clin Biochem* (1988) **21**:201–210. doi:10.1016/S0009-9120(88)80002-X

169. Roberts JW, Cora-Locatelli G, Bravi D, Amantea MA, Mouradian MM, Chase TN. Catechol-o-methyltransferase inhibitor tolcapone prolongs levodopa/carbidopa action in parkinsonian patients. *Neurology* (1993) **43**:2685–2688. doi:10.1212/wnl.43.12.2685

170. Mannisto PT, Ulmanen I, Lundstrom K, Taskinen J, Tenhunen J, Tilgmann C, Kaakkola S.

Characteristics of catechol O-methyltransferase (COMT) and properties of selective COMT inhibitors. *Prog Drug Res* (1992) **39**:291–350. doi:10.1007/978-3-0348-7144-0_9

171. de Lau LML, Verbaan D, Marinus J, Heutink P, van Hilten JJ. Catechol-O-methyltransferase Val158Met and the risk of dyskinesias in Parkinson's disease. *Mov Disord* (2012) **27**:132–135. doi:10.1002/mds.23805

172. Sampaio TF, dos Santos EUD, de Lima GDC, dos Anjos RSG, da Silva RC, Asano AGC, Asano NMJ, Crovella S, de Souza PRE. *MAO-B* and *COMT* Genetic Variations Associated With Levodopa Treatment Response in Patients With Parkinson's Disease. *J Clin Pharmacol* (2018) **58**:920–926. doi:10.1002/jcph.1096

173. Bialecka M, Kurzawski M, Klodowska-Duda G, Opala G, Tan EK, Drozdzik M. The association of functional catechol-O-methyltransferase haplotypes with risk of Parkinsons disease, levodopa treatment response, and complications. *Pharmacogenet Genomics* (2008) **18**:815–821. doi:10.1097/FPC.0b013e328306c2f2

174. Bialecka M, Drozdzik M, Klodowska-Duda G, Honczarenko K, Gawronska-Szklarz B, Opala G, Stankiewicz J. The effect of monoamine oxidase B (MAOB) and catechol-O-methyltransferase (COMT) polymorphisms on levodopa therapy in patients with sporadic Parkinson's disease. *Acta Neurol Scand* (2004) **110**:260–266. doi:10.1111/j.1600-0404.2004.00315.x

175. Lee MS, Lyoo CH, Ulmanen I, Syvänen AC, O Rinne J. Genotypes of catechol-O-methyltransferase and response to levodopa treatment in patients with Parkinson's disease. *Neurosci Lett* (2001) **298**:131–134. doi:10.1016/S0304-3940(00)01749-3

176. Watanabe M, Harada S, Nakamura T, Ohkoshi N, Yoshizawa K, Hayashi A, Shoji S. Association between Catechol-<i>O</i>-Methyltransferase Gene Polymorphisms and Wearing-Off and Dyskinesia in Parkinson's Disease. *Neuropsychobiology* (2003) **48**:190–193. doi:10.1159/000074637

177. Cheshire P, Bertram K, Ling H, O'Sullivan SS, Halliday G, McLean C, Bras J, Foltynie T, Storey E, Williams DR. Influence of Single Nucleotide Polymorphisms in COMT, MAO-A and BDNF Genes on Dyskinesias and Levodopa Use in Parkinson's Disease. *Neurodegener Dis* (2013) doi:10.1159/000351097

178. Sossi V, de la Fuente-Fernández R, Schulzer M, Troiano AR, Ruth TJ, Stoessl AJ. Dopamine transporter relation to dopamine turnover in Parkinson's disease: a positron emission tomography study. *Ann Neurol* (2007) **62**:468–474. doi:10.1002/ana.21204

179. Kaplan N, Vituri A, Korczyn AD, Cohen OS, Inzelberg R, Yahalom G, Kozlova E, Milgrom R, Laitman Y, Friedman E, et al. Sequence variants in SLC6A3, DRD2, and BDNF genes and time to levodopa-induced dyskinesias in Parkinson's disease. *J Mol Neurosci* (2014) **53**:183–188. doi:10.1007/s12031-014-0276-9

180. Kaiser R, Hofer A, Grapengiesser A, Gasser T, Kupsch A, Roots I, Brockmöller J. L-Dopa-induced adverse effects in PD and dopamine transporter gene polymorphism. *Neurology* (2003) **60**:1750–1755. doi:10.1212/01.WNL.0000068009.32067.A1

181. Tsai SJ. Role of tissue-type plasminogen activator and plasminogen activator inhibitor-1 in psychological stress and depression. *Oncotarget* (2017) **8**:113258–113268. doi:10.18632/oncotarget.19935

182. Numakawa T, Odaka H, Adachi N. Actions of brain-derived neurotrophin factor in the neurogenesis and neuronal function, and its involvement in the pathophysiology of brain diseases. *Int J Mol Sci* (2018) **19**: doi:10.3390/ijms19113650

183. Foltynie T, Cheeran B, Williams-Gray CH, Edwards MJ, Schneider SA, Weinberger D, Rothwell JC, Barker RA, Bhatia KP. BDNF val66met influences time to onset of levodopa induced dyskinesia in Parkinson's disease. *J Neurol Neurosurg Psychiatry* (2009) **80**:141–144. doi:10.1136/jnnp.2008.154294

184. Kusters CDJ, Paul KC, Guella I, Bronstein JM, Sinsheimer JS, Farrer MJ, Ritz BR. Dopamine receptors and BDNF-haplotypes predict dyskinesia in Parkinson's disease. *Park Relat Disord* (2018) **47**:39–44. doi:10.1016/j.parkreldis.2017.11.339

185. Lesage S, Belarbi S, Troiano A, Condroyer C, Hecham N, Pollak P, Lohman E, Benhassine T, Ysmail-Dahlouk F, Dürr A, et al. Is the common LRRK2 G2019S mutation related to dyskinesias in North African Parkinson disease? *Neurology* (2008) **71**:1550–1552. doi:10.1212/01.wnl.0000338460.89796.06

186. Orr-Urtreger A, Shifrin C, Rozovski U, Rosner S, Bercovich D, Gurevich T, Yagev-More H, Bar-Shira A, Giladi N. The LRRK2 G2019S mutation in Ashkenazi Jews with Parkinson disease: Is there a gender effect? *Neurology* (2007) **69**:1595–1602. doi:10.1212/01.wnl.0000277637.33328.d8

187. LRRK2 phosphorylation level correlates with abnormal motor behaviour in an experimental model of levodopa- induced dyskinesias.

188. Stanic J, Mellone M, Cirnaru MD, Perez-Carrion M, Zianni E, Di Luca M, Gardoni F,

Piccoli G. LRRK2 phosphorylation level correlates with abnormal motor behaviour in an experimental model of levodopa-induced dyskinesias. *Mol Brain* (2016) **9**:53. doi:10.1186/s13041-016-0234-2

189. Johnson NE. Whole-exome sequencing in neurologic practice. *Neurol Genet* (2015) **1**:e37. doi:10.1212/nxg.0000000000000037

190. Gupta DK, Hang X, Liu R, Hasan A, Feng Z. Levodopa-induced motor and dopamine receptor changes in caenorhabditis elegans overexpressing human alpha-synuclein. *Neurodegener Dis* (2016) **16**:179–183. doi:10.1159/000440845

191. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, Obeso J, Marek K, Litvan I, Lang AE, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* (2015) **30**:1591–1601. doi:10.1002/mds.26424

192. The Moli-Sani Project, a randomized, prospective cohort study in the Molise region in Italy; design, rationale and objectives. *Ital J Public Health* (2007) **4**: doi:10.2427/5886

193. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* (2015) **4**:7. doi:10.1186/s13742-015-0047-8

194. Chung SJ, Biernacka JM, Armasu SM, Anderson K, Frigerio R, Aasly JO, Annesi G, Bentivoglio AR, Brighina L, Chartier-Harlin M-C, et al. Alpha-synuclein repeat variants and survival in Parkinson's disease. *Mov Disord* (2014) **29**:1053–1057. doi:10.1002/mds.25841

195. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria.: R Foundation for Statistical Computing (2015). Available at: http://www.r-project.org/

196. Eusebi P, Romoli M, Paoletti FP, Tambasco N, Calabresi P, Parnetti L. Risk factors of levodopa-induced dyskinesia in Parkinson's disease: results from the PPMI cohort. *npj Park Dis* (2018) **4**: doi:10.1038/s41531-018-0069-x

197. Zhang Y han, Tang B sha, Song C yuan, Xu Q, Lou M xin, Liu Z hua, Yu R he, Yan X xiang, Guo J feng. The relationship between the phenotype of Parkinson's disease and levodopa-induced dyskinesia. *Neurosci Lett* (2013) **556**:109–112. doi:10.1016/j.neulet.2013.10.018

198. Benbir G, Özekmekçi S, Apaydin H, Delil S, Erginöz E. A hospital-based study: Risk factors in development of motor complications in 555 Parkinson's patients on levodopa

therapy. *Clin Neurol Neurosurg* (2006) **108**:726–732. doi:10.1016/j.clineuro.2006.02.002

199. Bjornestad A, Forsaa EB, Pedersen KF, Tysnes OB, Larsen JP, Alves G. Risk and course of motor complications in a population-based incident Parkinson's disease cohort. *Park Relat Disord* (2016) **22**:48–53. doi:10.1016/j.parkreldis.2015.11.007

200. Jackson D, White IR, Seaman S, Evans H, Baisley K, Carpenter J. Relaxing the independent censoring assumption in the Cox proportional hazards model using multiple imputation. *Stat Med* (2014) **33**:4681–4694. doi:10.1002/sim.6274

201. Belsley DA, Kuh E, Welsch RE. *Regression Diagnostics*. Hoboken, NJ, USA: John Wiley & Sons, Inc. (1980). doi:10.1002/0471725153

202. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* (2010) **26**:589–595. doi:10.1093/bioinformatics/btp698

203. Depristo MA, Banks E, Poplin R, Garimella K V., Maguire JR, Hartl C, Philippakis AA, Del Angel G, Rivas MA, Hanna M, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* (2011) **43**:491–501. doi:10.1038/ng.806

204. Danecek P, Auton A, Abecasis G, Albers C, Banks E, DePristo M. The variant call format and vcftools. *Bioinformatics* (2011) **27**: doi:10.1093/bioinformatics/btr330

205. Kuoppamäki M, Korpela K, Marttila R, Kaasinen V, Hartikainen P, Lyytinen J, Kaakkola S, Hänninen J, Löyttyniemi E, Kailajärvi M, et al. Comparison of pharmacokinetic profile of levodopa throughout the day between levodopa/carbidopa/entacapone and levodopa/carbidopa when administered four or five times daily. *Eur J Clin Pharmacol* (2009) **65**:443–455. doi:10.1007/s00228-009-0622-y

206. Grandas F, Galiano ML, Tabernero C. Risk factors for levodopa-induced dyskinesias in Parkinson's disease. *J Neurol* (1999) **246**:1127–1133. doi:10.1007/s004150050530

207. Zhao YJ, Wee HL, Chan YH, Seah SH, Au WL, Lau PN, Pica EC, Li SC, Luo N, Tan LCS. Progression of Parkinson's disease as evaluated by Hoehn and Yahr stage transition times. *Mov Disord* (2010) **25**:710–716. doi:10.1002/mds.22875

208. Ritz B, Rhodes SL, Bordelon Y, Bronstein J. α-Synuclein Genetic Variants Predict Faster Motor Symptom Progression in Idiopathic Parkinson Disease. *PLoS One* (2012) **7**:e36199. doi:10.1371/journal.pone.0036199

209. Hashim H, Azmin S, Razlan H, Yahya NW, Tan HJ, Manaf MRA, Ibrahim NM. Eradication

of Helicobacter pylori infection improves levodopa action, clinical symptoms and quality of life in patients with parkinson's disease. *PLoS One* (2014) **9**: doi:10.1371/journal.pone.0112330

210. Meireles J, Massano J. Cognitive impairment and dementia in Parkinson's disease: Clinical features, diagnosis, and management. *Front Neurol* (2012) **MAY**: doi:10.3389/fneur.2012.00088

211. Pandey S, Srivanitchapoom P. Levodopa-induced dyskinesia: Clinical features, pathophysiology, and medical management. *Ann Indian Acad Neurol* (2017) **20**:190–198. doi:10.4103/aian.AIAN_239_17

212. De Marco E V., Tarantino P, Rocca FE, Provenzano G, Civitelli D, De Luca V, Annesi F, Carrideo S, Cirò Candiano IC, Romeo N, et al. Alpha-synuclein promoter haplotypes and dementia in Parkinson's disease. *Am J Med Genet Part B Neuropsychiatr Genet* (2008) **147**:403–407. doi:10.1002/ajmg.b.30611

213. Chung SJ, König IR, Lohmann K, Hinrichs F, Kim J, Ryu HS, Lee HJ, Kim K, Lee JH, Jung KW, et al. Association of SNCA variants with α-synuclein of gastric and colonic mucosa in Parkinson's disease. *Park Relat Disord* (2019) **61**:151–155. doi:10.1016/j.parkreldis.2018.10.028

214. Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic Review of Levodopa Dose Equivalency Reporting in Parkinson's Disease. doi:10.1002/mds.23429

215. Colosimo C, Martínez-Martín P, Fabbrini G, Hauser RA, Merello M, Miyasaki J, Poewe W, Sampaio C, Rascol O, Stebbins GT, et al. Task force report on scales to assess dyskinesia in Parkinson's disease: Critique and recommendations. *Mov Disord* (2010) **25**:1131–1142. doi:10.1002/mds.23072

216. Kartsonaki C. Survival analysis. *Diagnostic Histopathology* (2016) **22**:263–270. https://doi.org/10.1016/j.mpdhp.2016.06.005

217. D. Y. Lin, L. J. Wei and Z. Ying Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. Biometrika (1993) **80**:557–572 doi: 0.2307/2337177

218. Xuea Y and Schifano ED. Diagnostics for the Cox model. *CSAM* (2017) **24**:583–604. https://doi.org/10.29220/CSAM.2017.24.6.583

Acknowledgments

I would like to thank my tutor, Dr. Daniela Carnevale, for teaching me the value of independence and for having accompanied me during this journey looking at me from afar.

I deeply thank Dr. Esposito Teresa for teaching me to accept and overcome challenges, even those that seemed impossible.

I deeply thank my mentor and wise teacher Dr. Alessandro Giallusi who, first of all, taught me the value of friendship within the workplace and, secondly, the intellectual rigour and the scientific correctness. I owe this hard conquest to him! Thank you so much, I would have never achieved this goal without you!

I would like to thank University of Molise for the PhD school and IRCCS Neuromed that hosts me as a PhD student.

My sincere thanks goes to all the Department of Epidemiology and Prevention that welcomed me in the last year of PhD and in particular to Dr. Giovanni de Gaetano, Dr. Maria Benedetta Donati, Dr. Chiara Cerletti and Dr. Licia Iacoviello. I would also show gratitude to Dr Benedetta Izzi and Dr. Fabrizia Noro for all the precious advices and to Annalisa, Roberta, Sabatino, Federica and Valentina for the smiles and the encouragement during this last year.

I would like to thank my first friend and colleague Dr. Mafalda Reccia, for the long scientific chatting and for her beautiful friendship. Thank you so much Gabriele and Carmela for our thousands of adventures.

I am also grateful to all my family who supports me in every challenge I encounter in my life with love and patience. Thank you Mom, Dad, Gennaro and Anna.

I would like to thank all people who shared a piece of this journey with me. Thank you Chiara, Greta, Maria and all the others.

A deep, special thanks goes to Emmanuele Jr. who endures my outbursts for hours and with his immense love never stops believing in me and reminding me of all my skills even in the darkest of times. Thank you because in spite of all you chose to share your life with me.