Exploratory analysis of the genetics of impulse control disorders in Parkinson's disease using genetic risk scores

Johann Faouzi, Baptiste Couvy-Duchesne, Samir Bekadar, Olivier Colliot, Jean-Christophe Corvol

Exploratory analysis of the genetics of impulse control disorders in Parkinson's disease using genetic risk scores

Johann Faouzi[1,2,3,4,5]; Baptiste Couvy-Duchesne[1,2,3,4,5,6], PhD; Samir Bekadar[1,2,3,4,7]; Olivier Colliot[1,2,3,4,5], PhD; Jean-Christophe Corvol[1,2,3,4,7], MD, PhD

[1] *Paris Brain Institute, F-75013, Paris, France*
[2] *Inserm, U 1127, F-75013, Paris, France*
[3] *CNRS, UMR 7225, F-75013, Paris, France*
[4] *Sorbonne Université, F-75013, Paris, France*
[5] *Inria Paris, Aramis project-team, F-75013, Paris, France*
[6] *Institute for Molecular Bioscience, the University of Queensland, Brisbane, QLD, Australia*
[7] *Assistance Publique Hôpitaux de Paris, Department of Neurology, Centre d'Investigation Clinique Neurosciences, Hôpital Pitié-Salpêtrière, F-75013, Paris, France*

Abstract

Objective: To study the association between impulse control disorders (ICDs) in Parkinsons disease (PD) and genetic risk scores (GRS) for 40 known or putative risk factors (e.g. depression, personality traits).

Background: In absence of published genome-wide association studies (GWAS), little is known about the genetics of ICDs in PD. GRS of related phenotypes, for which large GWAS are available, may help shed light on the genetic contributors of ICDs in PD.

Methods: We searched for GWAS on European ancestry populations with summary statistics publicly available for a broad range of phenotypes, including other psychiatric disorders, personality traits, and simple phenotypes. We separately tested their predictive ability in two of the largest PD cohorts with clinical and genetic available: the Parkinsons Progression Markers Initiative database ($N = 368$, 33% female, age range = $[33 - 84]$) and the Drug Interaction With Genes in Parkinsons Disease study ($N = 373$, 40% female, age range = $[29 - 85]$).

Results: We considered 40 known or putative risk factors for ICDs in PD for which large GWAS had been published. After Bonferroni correction for multiple comparisons, no GRS or the combination of the 40 GRS were significantly associated with ICDs from the analyses in each cohort separately and from the meta-analysis.

Conclusion: Albeit unsuccessful, our approach will gain power in the coming years with increasing availability of genotypes in clinical cohorts of PD, but also from future increase in GWAS sample sizes of the phenotypes we considered. Our approach may be applied to other complex disorders, for which GWAS are not available or limited.

1 Introduction

Although the cardinal symptoms of Parkinson's disease (PD) are motor, many non-motor symptoms frequently occur during the course of the disease, including psychiatric comorbidities. Impulse control disorders (ICDs), a class of psychiatric disorders characterized by impulsivity, are common in PD, with half of PD cases expected to experience them within 5 years of the disease onset (Corvol et al., 2018). The four most common ICDs in PD are pathological gambling, compulsive eating, hypersexuality, and compulsive eating disorder. ICDs are associated with reduced quality of life, strained interpersonal relationships, increased caregiver burden, and require prompt management (Weintraub and Claassen, 2017).

Numerous factors have been associated with ICDs in PD, including socio-demographic, clinical and genetic variables (Grall-Bronnec et al., 2018). Associations from candidate gene analyses between ICDs and several genetic variants have been reported in the following genes: *ANKK1* (Hoenicka et al., 2015), *DAT1* (Cormier-Dequaire et al., 2018), *DRD1* (Zainal Abidin et al., 2015), *DRD2* (Zainal Abidin et al., 2015), *DRD3* (Lee et al., 2009), *GRIN2B* (Lee et al., 2009; Zainal Abidin et al., 2015), *HTR2A* (Lee et al., 2012), *OPRK1* (Cormier-Dequaire et al., 2018), *OPRM1* (Cormier-Dequaire et al., 2018), and *SLC22A1* (Redenek et al., 2019). Several studies also reported no consistent associations with variants from some of the same genes (Cormier-Dequaire et al., 2018), highlighting the variability and the lack of replication of the reported associations.

Variation in complex traits is caused by numerous genetic variants. Each genetic variant usually provides limited information because the relative causal risk of each variant is small. On the other hand, the combined risk of numerous low-risk variants can explain a significant proportion of the genetic variance. Genetic risk scores (GRS), obtained from genome-wide association studies (GWAS), linearly summarize the contribution of these numerous variants into a single score. Using such GRS allows for studying traits not collected in the PD cohorts or diseases that would be too rare to allow direct evaluation of the comorbidities.

ICDs in PD have been rarely studied using GRS. Only one study looked for associations between GRS and ICDs in PD (Ihle et al., 2020). In this study, authors computed a GRS of PD using 90 SNPs reaching genome-wide significance in a meta-analysis of 17 GWAS and did not find an association between this GRS and ICDs in PD. Their power was limited due to the small sample size. Furthermore, the GRS may only capture part of the genetic risk factors and would benefit from larger GWAS (Dudbridge, 2013). ICDs are not associated with PD itself, and may be associated with personality traits or psychiatric endophenotypes, which has been little studied.

Our main objective was to evaluate the predictive accuracy of a broad range of GRS in order to shed light on the genetic determinants of ICDs in PD. We were particularly interested in GRS for other psychiatric disorders, but also personality traits, including

impulsivity, some of which have been associated with ICDs in PD.

2   Materials and methods

Figure 1 summarizes our methodology. A more detailed description of the materials and methods is available in the supplementary materials.

We utilized data from two large research cohorts with clinical and genetic data available: the Parkinson's Progression Markers Initiative (PPMI) database and the Drug Interaction With Genes in Parkinson's Disease (DIGPD) study. We excluded subjects with an estimated genetic ancestry different from the European genetic ancestry. We filtered the raw genetic variants and imputed missing variants using a reference panel. We searched for GWAS on European ancestry populations with summary statistics publicly available for a broad range of phenotypes, including other psychiatric disorders, personality traits, and simple phenotypes We computed genetic risk scores for each phenotype included in our analyses and assessed their association with ICDs in PD using logistic models for the analysis in each cohort independently and random effects models for the meta-analysis.

3   Results

3.1   Participants and genetic variants

Out of the 378 PPMI subjects and 382 DIGPD subjects who matched the first three inclusion criteria, we excluded 10 subjects from PPMI and 9 subjects from DIGPD for being projected too far from the European cluster (Supplementary Figure 1). Thus, we included 368 subjects from PPMI and 373 subjects from DIGPD.

Out of the 39,235,157 genetic variants of the Haplotype Reference Consortium reference panel, 601,370 SNPs in PPMI and 6,294,320 SNPs in DIGPD matched the inclusion criteria. The high discrepancy in the numbers is due to the genotyping arrays: the NeuroX array is known to have a low coverage of the genome.

3.2   Genome-wide association studies

We identified 40 GWAS that matched the inclusion criteria. Supplementary Table 1 presents the characteristics of these studies, including the phenotype of interest, the phenotype definition, the phenotype assessment, the number of SNPs and the heritability estimated from these SNPs. Supplementary Table 2 indicates the number of common SNPs between the GWAS and PPMI and DIGPD, as well as the number of SNPs used to compute the GRS in both cohorts. The included phenotypes consisted of other psychiatric disorders (anxiety, depression, obsessive compulsive, and attention-deficit hyperactivity disorders (ADHD), anorexia nervosa), personality traits (impul-
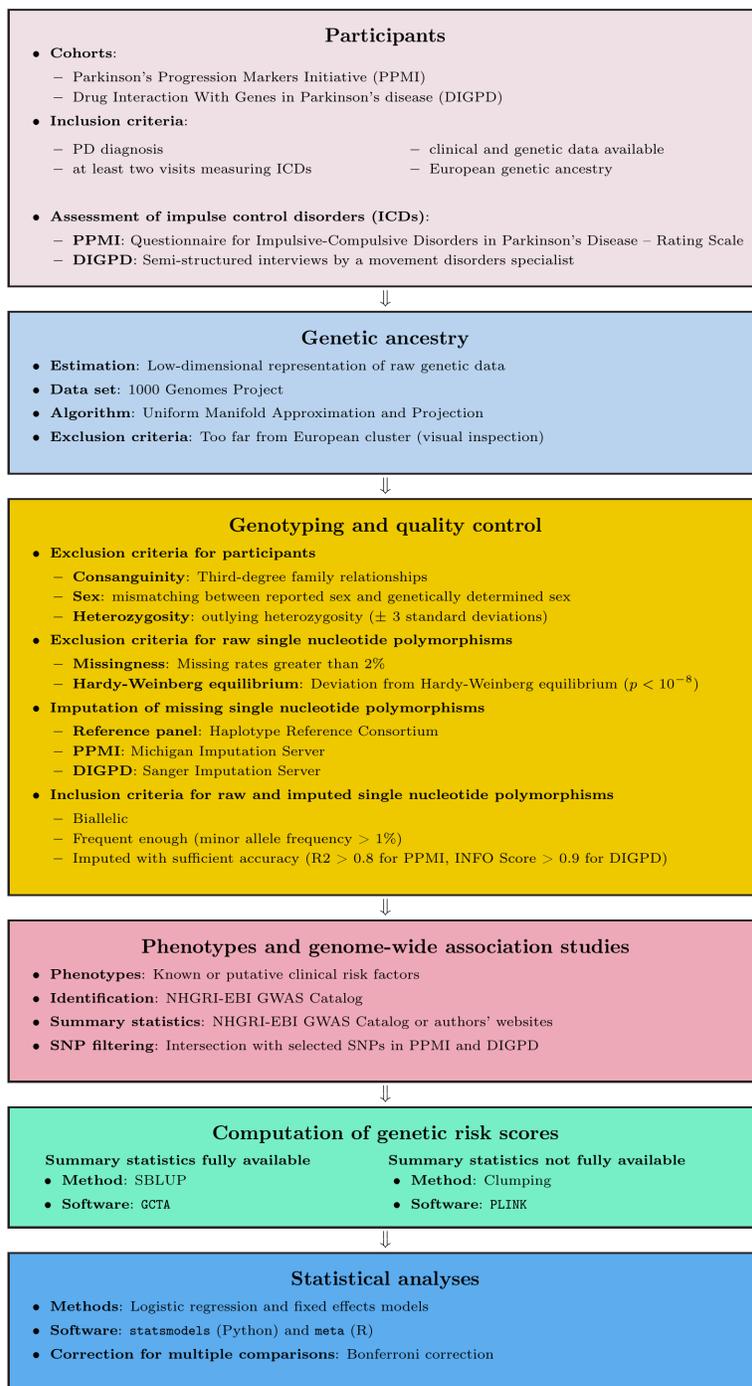
Participants

- **Cohorts**:
  - Parkinson's Progression Markers Initiative (PPMI)
  - Drug Interaction With Genes in Parkinson's disease (DIGPD)
- **Inclusion criteria**:
  - PD diagnosis
  - at least two visits measuring ICDs
  - clinical and genetic data available
  - European genetic ancestry
- **Assessment of impulse control disorders (ICDs)**:
  - **PPMI**: Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease – Rating Scale
  - **DIGPD**: Semi-structured interviews by a movement disorders specialist

⇓

Genetic ancestry

- **Estimation**: Low-dimensional representation of raw genetic data
- **Data set**: 1000 Genomes Project
- **Algorithm**: Uniform Manifold Approximation and Projection
- **Exclusion criteria**: Too far from European cluster (visual inspection)

⇓

Genotyping and quality control

- **Exclusion criteria for participants**
  - **Consanguinity**: Third-degree family relationships
  - **Sex**: mismatching between reported sex and genetically determined sex
  - **Heterozygosity**: outlying heterozygosity ($\pm$ 3 standard deviations)
- **Exclusion criteria for raw single nucleotide polymorphisms**
  - **Missingness**: Missing rates greater than 2%
  - **Hardy-Weinberg equilibrium**: Deviation from Hardy-Weinberg equilibrium ($p < 10^{-8}$)
- **Imputation of missing single nucleotide polymorphisms**
  - **Reference panel**: Haplotype Reference Consortium
  - **PPMI**: Michigan Imputation Server
  - **DIGPD**: Sanger Imputation Server
- **Inclusion criteria for raw and imputed single nucleotide polymorphisms**
  - Biallelic
  - Frequent enough (minor allele frequency $> 1\%$)
  - Imputed with sufficient accuracy (R2 > 0.8 for PPMI, INFO Score > 0.9 for DIGPD)

⇓

Phenotypes and genome-wide association studies

- **Phenotypes**: Known or putative clinical risk factors
- **Identification**: NHGRI-EBI GWAS Catalog
- **Summary statistics**: NHGRI-EBI GWAS Catalog or authors' websites
- **SNP filtering**: Intersection with selected SNPs in PPMI and DIGPD

⇓

Computation of genetic risk scores

Summary statistics fully available
- **Method**: SBLUP
- **Software**: GCTA

Summary statistics not fully available
- **Method**: Clumping
- **Software**: PLINK

⇓

Statistical analyses

- **Methods**: Logistic regression and fixed effects models
- **Software**: statsmodels (Python) and meta (R)
- **Correction for multiple comparisons**: Bonferroni correction

Figure 1: Methodology. We utilized data from two large research cohorts with clinical and genetic data available: the Parkinson's Progression Markers Initiative (PPMI) database and the Drug Interaction With Genes in Parkinson's Disease (DIGPD) study. We excluded subjects with an estimated genetic ancestry different from the European genetic ancestry. We filtered the raw genetic variants and imputed missing variants using a reference panel. We searched for GWAS on European ancestry populations with summary statistics publicly available for a broad range of phenotypes, including other psychiatric disorders, personality traits, and simple phenotypes. We computed genetic risk scores for each phenotype included in our analyses and assessed their association with ICDs in PD using logistic models for the analysis in each cohort independently and random effects models for the meta- analysis.

sivity, agreeableness, conscientiousness, extraversion, openness), risk taking behaviors (automobile speeding, alcohol consumption, smoking status, sexual activity), and simple traits (body height, body mass index, intelligence, education).

Two groups of GWAS included genetic data from 23andMe , and only the top 10,000 SNPs were made publicly available. We requested access to the whole summary statistics from 23andMe with no success.

3.3 Association analyses

Supplementary Table 3 presents the unadjusted p-values for the 40 GRS from the analyses on each cohort separately and from the meta-analysis. For the analysis in each cohort separately, among the 2 sets of 40 unadjusted p-values (correction is per-sample), only one was smaller than 0.05 (nominal significance), corresponding to the GRS of body mass index in PPMI ($p = 0.0079$). The association did not survive after Bonferroni correction. In the meta-analysis, among the 40 unadjusted p-values, only one was smaller than 0.05, corresponding to the GRS lack of premeditation ($p = 0.0107$, Supplementary Figure 2). The association did not survive after Bonferroni correction.

The combination of the 40 GRS altogether was not associated with ICDs, both from the analyses in each cohort independently ($p = 0.0969$ in PPMI, $p = 0.5166$ in DIGPD) and from the meta-analysis ($p = 0.0764$).

In order to validate our GRS calculation, we assessed the quality of the GRS of body mass index by using a linear regression model with correction for age, sex and genetic ancestry. BMI was available in both cohorts, and is well studied in genetics, leading to robust GRS-based prediction (Dudbridge, 2013). In particular, the corresponding GWAS has a very large sample size ($N = 806,834$), making the estimation of each SNP contribution more robust. In both cohorts, BMI GRS were strongly associated with the measured BMI ($p = 0.000058$ in PPMI, $p = 0.000038$ in DIGPD) and the Pearson correlation coefficients were positive and high ($r = 0.21$ in PPMI, $r = 0.19$ in DIGPD). These results gave us confidence in our methodology and in the quality of the computed GRS.

4 Discussion

To our knowledge, this is the first study investigating the association between ICDs in PD and genetic risk scores for a broad range of phenotypes, including phenotypes that have been associated with ICDs in PD (Grall-Bronnec et al., 2018). Compared to a previous study that only investigated the PD GRS computed from a small number of SNPs (Ihle et al., 2020), we explored 40 phenotypes for which we computed GRS using a large number of SNPs. However, the results were mainly negative, as we did not find a single association after correction for multiple comparisons.

The main limitation of our study is the small sample size of our clinical samples, which limits discovery of small associations. The size of the GWAS is also a limitation, as GRS are imperfect predictors of the genetic liability of traits. It is known that discouraging results in many studies were due to low number of participants, and that an increase in the sample size would allow more successful results (Dudbridge, 2013). We performed a power analysis (Supplementary Figure 3) showing that our sample sizes could only detect very strong effects (odds ratios greater than 1.7) with good power (0.80) and that smaller, more realistic effects (odds ratios between 1.1 and 1.3) would require between 1000 and 10,000 samples to be detected with good power (0.80). The genetic correlations between the traits for which GRS were calculated and ICDs are also unknown. Another limitation is the incomplete summary statistics made available for two groups of studies focusing on impulsivity and personality traits (collaborations with 23andMe, we contacted 23andMe but did not receive a response). For these traits, we had to compute GRS from a small number of SNPs. Computing the GRS using the whole summary statistics would likely increase the quality of these GRS. Another limitation is the different definitions of the phenotypes in meta- and mega-analyses. Complex phenotypes, in particular complex diseases, are not always assessed with the same instruments in different data sets. Thus, the increase in sample size is obtained at the expense of a less homogeneous definition of the phenotype.

Little is known about the genetic factors of ICDs in PD. Several studies reported associations for a few genetic variants, but they all suffer from the lack of replication, and there exists no GRS for ICDs yet. Our study could not conclude about the association between ICDs in PD and GRS for a broad range of phenotypes, but highlights the methodology to compute GRS and study their association with ICDs in PD for future studies, and shows how to investigate the genetic factors of a phenotype without performing a GWAS. Such study would deserve from being repeated when larger GWAS or clinical samples get available, which may boost power to detect significant associations.

References

Florence Cormier-Dequaire et al. Suggestive association between OPRM1 and impulse control disorders in Parkinson's disease. *Movement Disorders: Official Journal of the Movement Disorder Society*, 33(12):1878–1886, 2018. ISSN 1531-8257. doi: 10.1002/mds.27519.

Jean-Christophe Corvol et al. Longitudinal analysis of impulse control disorders in Parkinson disease. *Neurology*, 91(3):e189–e201, July 2018. ISSN 1526-632X. doi: 10.1212/WNL.0000000000005816.

Frank Dudbridge. Power and Predictive Accuracy of Polygenic Risk Scores. *PLOS Genetics*, 9(3):e1003348, March 2013. ISSN 1553-7404. doi: 10.1371/journal.pgen.1003348.

Marie Grall-Bronnec et al. Dopamine Agonists and Impulse Control Disorders: A Complex Association. *Drug Safety*, 41(1):19–75, January 2018. ISSN 1179-1942. doi: 10.1007/s40264-017-0590-6.

Janet Hoenicka et al. The addiction-related gene ANKK1 in Parkinsonian patients with impulse control disorder. *Neurotoxicity Research*, 27(3):205–208, April 2015. ISSN 1476-3524.

Jonas Ihle et al. Parkinson's disease polygenic risk score is not associated with impulse control disorders: A longitudinal study. *Parkinsonism & Related Disorders*, 75:30–33, May 2020. ISSN 1873-5126. doi: 10.1016/j.parkreldis.2020.03.017.

Jee-Young Lee, Eun Kyung Lee, Sung Sup Park, Ji-Yeon Lim, Hee Jin Kim, Ji Sun Kim, and Beom S. Jeon. Association of DRD3 and GRIN2B with impulse control and related behaviors in Parkinson's disease. *Movement Disorders: Official Journal of the Movement Disorder Society*, 24(12):1803–1810, September 2009. ISSN 1531-8257. doi: 10.1002/mds.22678.

Jee-Young Lee, Beom S. Jeon, Han-Joon Kim, and Sung-Sup Park. Genetic variant of HTR2A associates with risk of impulse control and repetitive behaviors in Parkinson's disease. *Parkinsonism & Related Disorders*, 18(1):76–78, January 2012. ISSN 1873-5126. doi: 10.1016/j.parkreldis.2011.08.009.

Sara Redenek et al. Dopaminergic Pathway Genes Influence Adverse Events Related to Dopaminergic Treatment in Parkinson's Disease. *Frontiers in Pharmacology*, 10:8, 2019. ISSN 1663-9812. doi: 10.3389/fphar.2019.00008.

Daniel Weintraub and Daniel O. Claassen. Impulse Control and Related Disorders in Parkinson's Disease. *International Review of Neurobiology*, 133:679–717, 2017. ISSN 2162-5514. doi: 10.1016/bs.irn.2017.04.006.

Shahidee Zainal Abidin et al. DRD and GRIN2B polymorphisms and their association with the development of impulse control behaviour among Malaysian Parkinson's disease patients. *BMC neurology*, 15:59, April 2015. ISSN 1471-2377. doi: 10.1186/s12883-015-0316-2.