



OPEN

Clustering of Parkinson subtypes reveals strong influence of DRD2 polymorphism and gender

Esther Annegret Pelzer^{1,6}✉, Sophie Stürmer^{5,6}, Delia-Lisa Feis¹, Corina Melzer¹, Frank Schwartz², Marcel Schrage¹, Carsten Eggers³, Marc Tittgemeyer^{1,4} & Lars Timmermann³

Most classification approaches for idiopathic Parkinson's disease subtypes primarily focus on motor and non-motor symptoms. Besides these characteristics, other features, including gender or genetic polymorphism of dopamine receptors are potential factors influencing the disease's phenotype. By utilizing a kmeans-clustering algorithm we were able to identify three subgroups mainly characterized by gender, DRD2 Taq1A (rs1800497) polymorphism—associated with changes in dopamine signaling in the brain—and disease progression. A subsequent regression analysis of these subgroups further suggests an influence of their characteristics on the daily levodopa dosage, an indicator for medication response. These findings could promote further enhancements in individualized therapies for idiopathic Parkinson's disease.

Parkinson's disease (PD) is next to Alzheimer's disease one of the most common neurodegenerative disorders¹. It manifests with heterogeneous courses of clinical symptoms² in the motor and non-motor domain. To better understand the underlying pathophysiology of the disease and its characteristic, PD patients are mostly differentiated based on their motor symptoms into akinetic-rigidic, mixed-type and tremor-dominant subtype³. Such characterizations aim to optimally fit therapeutic regimes to patients and therefore reduce undesired side effects.

So far, the classification of subtypes and their subsequent therapeutic decisions are mostly driven by empirical impressions of the physician. These are predominantly determined by motor scores such as the UPDRS⁴ or the Hoehn and Yahr Scale⁵. In the last decades further knowledge about Parkinson's disease was gained, particularly with respect to the diagnostic importance of non-motor symptoms⁶.

A common technique for subdividing data into unknown groups is cluster analysis. This unsupervised learning approach aims to separate data items (e.g. patients) into groups of similar characteristics. Such cluster analyses (CA)⁷ have been extensively used to better identify PD subtypes in the motor and non-motor domain independently from the clinical expertise of the doctor (see Refs.^{8–12}). These methods comprise the detection of new subtypes based on pre-defined clinical characteristics such as age at onset, motor phenotype, non-motor symptoms or disease progression. However, all studies highly vary in the final number of characterized clusters, sample size, inclusion criteria as well as included variables, which limits comparability.

Previous studies suggest, that beside motor and non-motor phenomena, other features could play a crucial role in the development of parkinsonian symptoms. In particular, the following measures emerge: (1) Gender differences: PD occurs 1.5 times more often in men than in women¹³. Additionally, men show an earlier disease onset and have more diverse clinical profiles than women¹⁴. (2) Gene polymorphism: Likewise, DRD2 Taq1A (rs1800497) and DRD3 Ser9Gly (rs6208) polymorphism might be of great importance when delineating PD subtypes¹⁵. An association of receptor polymorphisms—as disease causing variants—has been reported for Parkinson's disease^{16,17}. An influence of these polymorphisms on medication response and occurrence of side effects in individual patients is part of ongoing research^{18–22}.

In order to proof our hypothesis of an influence of dopamine receptor polymorphisms and gender on the daily needed levodopa dosage (LED), we wanted to explore their capability of revealing new subtypes of Parkinson disease. Additionally, we were also interested in the predictive power of these subtypes for the levodopa dosage.

¹Max Planck Institute for Metabolism Research, Translational Neurocircuitry Group, Gleulerstrasse 50, 50931 Cologne, Germany. ²Department of Neurology, University Hospital Cologne, Cologne, Germany. ³Department of Neurology, University Hospital Marburg, Marburg, Germany. ⁴Excellence Cluster on Cellular Stress Responses in Aging Associated Diseases (CECAD), University of Cologne, Cologne, Germany. ⁵Department of Psychiatry, University Hospital Cologne, Cologne, Germany. ⁶These authors contributed equally: Esther Annegret Pelzer and Sophie Stürmer. ✉email: Esther.Pelzer@sf.mpg.de

Number of patients	91
Sex men/women (% men)	62 (68%)
Age at onset (years)	58 (\pm 10)
Symptom onset left (% left)	44 (48%)
Disease progression 'off'	5 (\pm 2)
Medication response	0.41 (\pm 0.2)
Tremor-/akineti-c-rigid-score (T-/AR-score)	0.3 (\pm 0.5)
DRD2 Taq1A (rs1800497) Risk type (A/A or A/G)	27 (30%)
DRD3 Ser9Gly (rs6208) Risk type (C/C or C/T)	36 (40%)

Table 1. Patient demographics. Variables included in cluster analysis. Mean (SD) in continuous variables; absolute numbers and percentage in dichotomous variables.

Therefore, we utilize a kmeans-clustering approach on our patient cohort and subsequently perform multiple linear regressions for each of the delineated clusters to determine their predictive power of modeling LED.

Methods

All patients with Parkinson's disease or the legally authorized representative of the patients gave written informed consent. Approval of the ethics commission of Cologne University's Faculty of Medicine (Nr.: 12-268) was obtained. The study was performed in accordance with the Declaration of Helsinki. Detailed patient information, methods and any associated references are available in the supplementary information of this paper.

Results

We performed kmeans-clustering for 91 patients with idiopathic PD (iPD) including the following variables: DRD2 Taq1A (rs1800497) and DRD3 Ser9Gly (rs6208) polymorphisms, standardized medication response, the age at onset, symptom onset side, a score relating the tremor/akineti-c-rigidity score, disease progression in the OFF state, and gender (cf. Table 1). After validating the clustering results for 2 to 10 clusters (cf. Supplementary Fig. S2), the optimal solution results in three distinct groups of patients (cf. Fig. 1a). Post-hoc analyses of all input variables for the allocated groups revealed significant differences in patient's gender ($p < 0.0001$), their individual standardized medication response ($p < 0.0001$), their DRD2 Taq1A (rs1800497) polymorphism ($p < 0.0001$), disease progression OFF ($p < 0.0001$) and symptom onset side ($p = 0.017$). All other variables showed no significant group differences (cf. Supplementary Table S4).

In summary, the three clusters could be characterized mainly by gender and DRD2 Taq1A (rs1800497) polymorphism. Whereas in one cluster only men with the A1- variant of the DRD2 Taq1A (rs1800497) polymorphism were present, another cluster included only women with mostly the A1- variant of DRD2 Taq1A (rs1800497) polymorphism. The third cluster mainly comprises the A1+ variant of male patients. Furthermore, standardized medication response of the three identified clusters significantly differed, i.e. men without DRD2 Taq1A (rs1800497) risk alleles (A1- variants) have a higher standardized medication response in comparison to the two other patient groups ($p < 0.0001$; cf. Fig. 1b). Additionally, symptom onset side significantly differed between subgroups. A summary of the cluster characteristics is presented in Table 2.

Multiple linear regression model. The consecutive regression analysis revealed an optimal model for cluster 1 [including (i) symptom onset side; (ii) tremor-/akineti-c-rigidity score, (iii) DRD2 Taq1A (rs1800497) polymorphism, (iv) standardized medication response, (v) interaction of disease progression OFF and standardized medication response] to explain adjusted levodopa equivalent dosage (aLED) with a significant regression equation ($F(5,20) = 11.843$, $p < 0.0001$) and an adjusted $R^2 = 68\%$. Regression analysis of cluster 2 yielded in an optimal model [including (i) disease progression OFF, (ii) standardized medication response, (iii) interaction between disease progression OFF and standardized medication response] to predict aLED ($F(3,30) = 3.854$, $p = 0.0019$) with an adjusted R^2 of 21%. For cluster 3 no significant regression model of aLED could be determined. However, the model explaining the highest amount of variation is based on a single variable (standardized medication response) with adjusted $R^2 = 8.4\%$ ($F(1,29) = 3.763$, $p = 0.062$).

Discussion

We demonstrate the feasibility of a clustering approach resulting in three distinct iPD patient subgroups. Extending earlier studies, describing motor and non-motor characteristics in their clustering approaches, we also considered parameters for (1) gender and (2) DRD2 Taq1A (rs1800497) polymorphism. The results of our analyses suggest, that these additional features might also have an impact on the development of Parkinson's disease affecting the treatment response to dopamine or dopamine equivalents.

An extensive body of literature exists on the association of gender and iPD (for review see Picillo et al.²³). A clear trend of a 1.5 times higher incidence of iPD in men is well described¹⁴; here the neuroprotective influence of estrogen, the genetically determined structural and functional cerebral differences as well as environmental

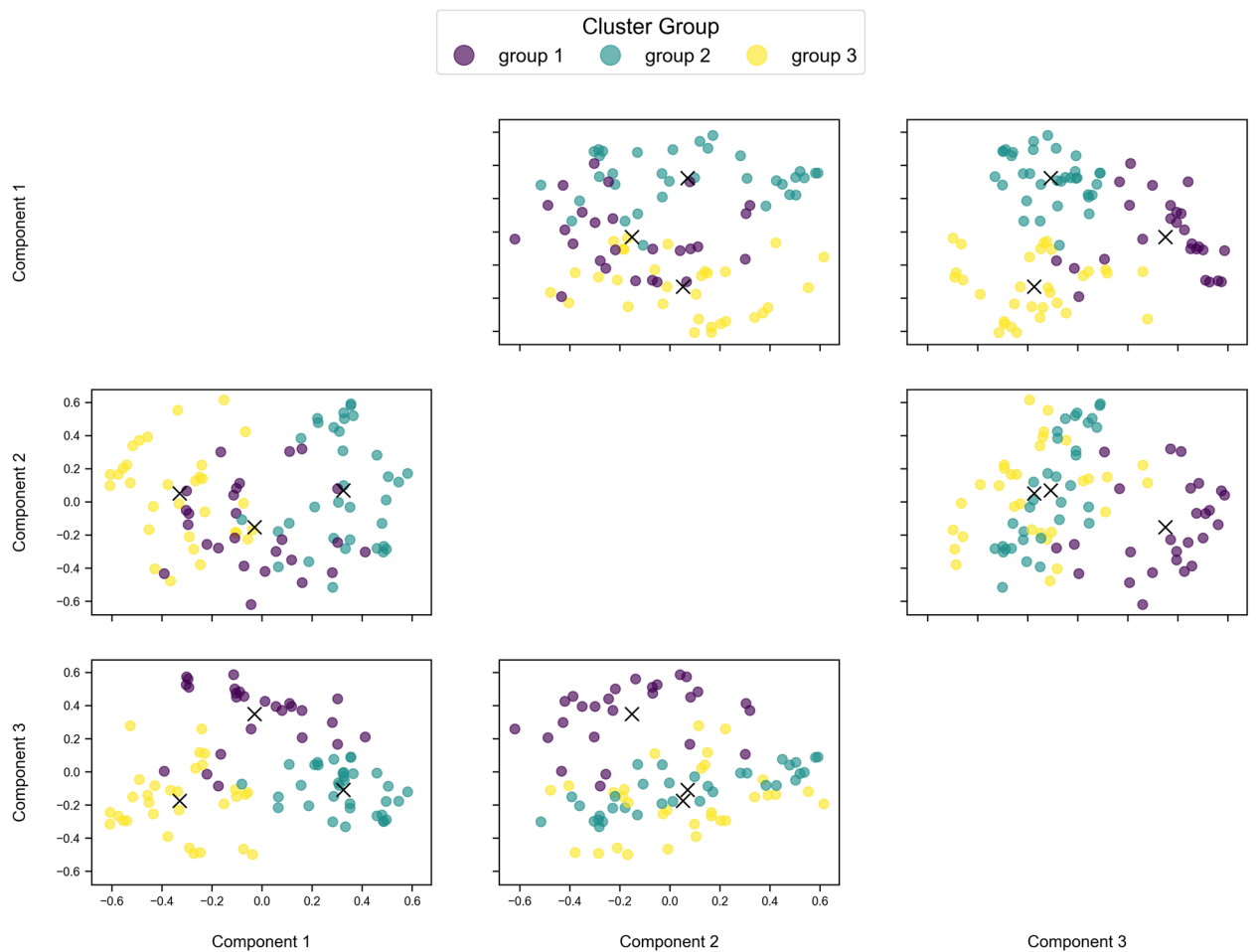
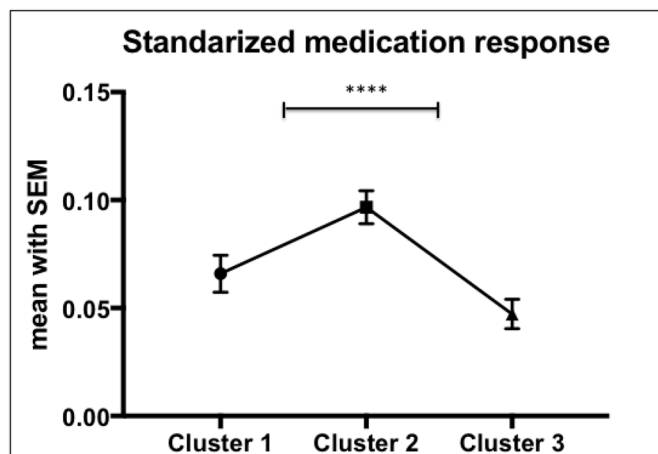
A**B**

Figure 1. Overview of key results. **(A)** Results of k-means-clustering with resulting three distinct groups (labeled in yellow, green and purple) exemplarily for the first three components (out of 8 components). **(B)** Cluster 2 showed the best medication response with a significant difference to cluster 1 and 3. Differences are marked at significance level of $p < 0.0001$ (****).

	Cluster 1	Cluster 2	Cluster 3
Gender	Women only	Men only	Mostly men
Symptom onset left	Equal	Less left	More left
Disease progression	Low	Fast	Low
Standardized medication response	Lower	High	Lower
DRD2 Taq1A (rs1800497) polymorphism	Mainly negative	All negative	Mainly positive

Table 2. Significant cluster characteristics. Overview of significant cluster characteristics.

factors are stated as putative mechanisms. Since two out of three clusters consisted only of men or women, gender seems to be a highly predictive factor in iPD and should be generally considered in the analysis of parkinsonian data.

As recently shown in Pelzer et al.²⁴, DRD2 Taq1A (rs1800497) polymorphism seems to have a neuroprotective influence on neurons in the nigro-striatal and satellite systems for patients with Parkinson's disease. Besides, the current study emphasizes an impact of DRD2 Taq1A (rs1800497) polymorphism also on the treatment response to dopaminergic drugs: (1) We were able to predict the daily levodopa intake of the individual patients in cluster 1 with very high significance. (2) Our findings, especially in cluster 2, agree with previous studies²⁵, describing a rapid-disease progression in patients with right-dominant symptom onset compared to those with left-sided symptom onset. All patients within this group had the A1- variant of the DRD2 Taq1A (rs1800497) polymorphism. (3) Although we could not identify any significant model predicting the daily levodopa dosage for patients within cluster 3, the characteristic of this cluster is of notable interest. All patients included in this group are carriers of the A1+ variant in DRD2 Taq1A (rs1800497). Current understanding from murine models report an association between expression of ANKK1, where the DRD2 Taq1A (rs1800497) is associated with, and the dopaminergic system functioning in the brain²⁶. For instance, while an increase in ANKK1 mRNA levels in the striatum was found after D1R-like stimulation (with SKF38393), the inverse reaction was observed after D2 receptor stimulation (with 7-OH-DPAT)²⁷. Hence, differences in the reaction on receptor levels are influenced by the diversity of gene polymorphisms in a complex manner. A recent paper investigating DRD2 Taq1A (rs1800497) polymorphism reported an influence of rare gene variants on the reaction to dopaminergic drugs²⁸. Accordingly, differences in the reaction to dopaminergic drugs might be explained by the diversity in the DRD2 Taq1A (rs1800497) polymorphism and the expression in the brain itself. Next to direct alterations on receptor levels, gene polymorphisms are an influencing factor on the brain structure. An interplay between basal ganglia connectivity and D2 receptor availability affects the clinical presentation and medication response of parkinsonian patients²⁴. A further sub-differentiation of DRD2 Taq1A (rs1800497) polymorphism (including the anatomical distribution, the percentage of polymorphism in the individual and the sequence variation) might be necessary to better understand the different reactions on different dopaminergic drugs (e.g. dopamine agonists or levodopa).

An altered anatomical background in basal ganglia loops for right- or left-affected parkinsonian patients has already been confirmed in a recent support vector machine analysis²⁹. In this study right-dominant expression of iPD has been connected to faster disease progression³⁰, indicating a different degeneration pattern with altered neuronal firing rates in oscillatory networks. However, the role of the DRD2 Taq1A (rs1800497) polymorphism in these different clinical situations is still unclear. Other studies investigating the association of iPD with the DRD2 Taq1A (rs1800497) genotype published contradictory findings. Whereas one study could show an increased risk of motor fluctuations³¹, another study could not state any influence of DRD2 Taq1A (rs1800497) polymorphism on dyskinesia³². An effect of DRD2 Taq1A (rs1800497) polymorphism on structural brain connectivity in patients with the wild type (A1- variant) compared to the risk type (A1+ variant) has been shown in Pelzer et al.²⁴. This correlation of structural changes in brain connectivity due to DRD2 Taq1A (rs1800497) polymorphism can further be linked to differences in the treatment response of individual genotypes. Hence, it might explain differences in the development of side effects like motor fluctuations in Parkinson's disease. Herewith these results indicate the relevance of the DRD2 Taq1A (rs1800497) polymorphism for investigating dopaminergic side effects in future studies.

In contrast to previous studies identifying subtypes of Parkinson's disease^{8,12}, we ignored non-motor scales and only focused on the implementation of parameters usually assessed in clinical routine (like age, gender, symptom onset side, motor scales, levodopa dosage per day). Beside the described encouraging results some limitations of the study have to be mentioned. One important point is the missing replication cohort in order to externally validate the study. Although the number of 91 patients seems high, the heterogeneity of the symptoms of the parkinsonian patients and the number of variables weakens the statistics and explanatory power. Hence, larger samples with a higher number of patients should be included in future analyses.

Besides increasing the sample size, the clustering process ultimately benefits from sets of parameters, suited to build computational phenotypes. Features from these parameters may be derived from neuroimaging (e.g. anatomical and diffusion magnetic resonance imaging) as well as clinical tests on non-motor symptoms; also neuropathological findings or gene polymorphisms might enhance predictability about the treatment response in Parkinson's disease. In conclusion, we consider these results as a major step to further establish predictive clinical models of iPD. The integration of more basic variables and the combination with computational algorithms could provide clinicians with reliable information to develop more individually adapted therapies to relieve parkinsonian symptoms more efficiently, while reducing potential side effects. In future approaches 'big data'

acquisitions suited for phenotyping “the parkinsonian patient” are, however, indispensable. Until then, reliably predicting the individual course of the disease in such a heterogeneous disease pattern remains preliminary. With this study, we focus attention on considering also some more basic mechanistic features influencing disease progression and therapeutic response.

Received: 22 August 2021; Accepted: 7 March 2022

Published online: 11 April 2022

References

- Erkkinen, M. G., Kim, M.-O. & Geschwind, M. D. Clinical neurology and epidemiology of the major neurodegenerative diseases. *Cold Spring Harb. Perspect. Biol.* **10**, a033118 (2018).
- Foltynie, T., Brayne, C. & Barker, R. A. The heterogeneity of idiopathic Parkinson’s disease. *J. Neurol.* **249**, 138–145 (2002).
- Spiegel, J. *et al.* Striatal FP-CIT uptake differs in the subtypes of early Parkinson’s disease. *J. Neural Transm. Gen. Sect. JNT* **114**, 331–335 (2007).
- Fahn, S. E. R. UPDRS program members. Unified Parkinsons Disease Rating Scale. *Maximil. Healthc. Inf.* **2**, 153–163 (1987).
- Goetz, C. G. *et al.* Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: Status and recommendations. *Mov. Disord.* **19**, 1020–1028 (2004).
- Pfeiffer, R. F. Non-motor symptoms in Parkinson’s disease. *Parkinsonism Relat. Disord.* **22**(Suppl 1), S119–122 (2016).
- Everitt, B. S., Landau, S. & Leese, M. *Cluster Analysis* (Taylor & Francis, 2001).
- van Rooden, S. M. *et al.* Clinical subtypes of Parkinson’s disease. *Mov. Disord.* **26**, 51–58 (2011).
- Erro, R. *et al.* The heterogeneity of early Parkinson’s disease: A cluster analysis on newly diagnosed untreated patients. *PLoS One* **8**, e70244 (2013).
- Ma, L.-Y., Chan, P., Gu, Z.-Q., Li, F.-F. & Feng, T. Heterogeneity among patients with Parkinson’s disease: Cluster analysis and genetic association. *J. Neurol. Sci.* **351**, 41–45 (2015).
- Pont-Sunyer, C. *et al.* The onset of nonmotor symptoms in Parkinson’s disease (the ONSET PD study). *Mov. Disord.* **30**, 229–237 (2015).
- Mu, J. *et al.* Parkinson’s disease subtypes identified from cluster analysis of motor and non-motor symptoms. *Front. Aging Neurosci.* **9**, 301 (2017).
- Wooten, G. F., Currie, L. J., Bovbjerg, V. E., Lee, J. K. & Patrie, J. Are men at greater risk for Parkinson’s disease than women?. *J. Neurol. Neurosurg. Psychiatry* **75**, 637–639 (2004).
- Haaxma, C. A. *et al.* Gender differences in Parkinson’s disease. *J. Neurol. Neurosurg. Psychiatry* **78**, 819–824 (2007).
- Magistrelli, L. *et al.* Polymorphisms of dopamine receptor genes and Parkinson’s disease: Clinical relevance and future perspectives. *Int. J. Mol. Sci.* **22**, 3781 (2021).
- McGuire, V. *et al.* Association of DRD2 and DRD3 polymorphisms with Parkinson’s disease in a multiethnic consortium. *J. Neurol. Sci.* **307**, 22–29 (2011).
- Yu, M., Huang, F., Wang, W. & Zhao, C. Association between the DRD2 TaqIA gene polymorphism and Parkinson disease risk: An updated meta-analysis. *Medicine* **98**, e17136 (2019).
- McDonnell, K. E., van Wouwe, N. C., Harrison, M. B., Wylie, S. A. & Claassen, D. O. TaqIA polymorphism and medication effects on inhibitory action control in Parkinson disease. *Brain Behav.* **8**, e01008 (2018).
- Paus, S. *et al.* The DRD2 TaqIA polymorphism and demand of dopaminergic medication in Parkinson’s disease. *Mov. Disord.* **23**, 599–602 (2008).
- Kwak, Y. *et al.* Task-dependent interactions between dopamine D2 receptor polymorphisms and L-DOPA in patients with Parkinson’s disease. *Behav. Brain Res.* **245**, 128–136 (2013).
- Masellis, M. *et al.* Dopamine D2 receptor gene variants and response to rasagiline in early Parkinson’s disease: A pharmacogenetic study. *Brain* **139**, 2050–2062 (2016).
- Bhattacharjee, S., Hughes, E. & Ng, K. L. Rasagiline sensitive dopamine D2 receptor gene variants: A step forward toward more personalized antiparkinsonian therapy. *Mov. Disord. Clin. Pract.* **4**, 181–182 (2016).
- Picillo, M. *et al.* The relevance of gender in Parkinson’s disease: A review. *J. Neurol.* **264**, 1583–1607 (2017).
- Pelzer, E. A. *et al.* Axonal degeneration in Parkinson’s disease—Basal ganglia circuitry and D2 receptor availability. *NeuroImage. Clin.* **23**, 101906 (2019).
- Baumann, C. R., Held, U., Valko, P. O., Wienecke, M. & Waldvogel, D. Body side and predominant motor features at the onset of Parkinson’s disease are linked to motor and nonmotor progression. *Mov. Disord.* **29**, 207–213 (2014).
- Sun, X., Luquet, S. & Small, D. M. DRD2: Bridging the genome and ingestive behavior. *Trends Cogn. Sci.* **21**, 372–384 (2017).
- Ponce, G. *et al.* The addiction-related gene Ankk1 is oppositely regulated by D1R- and D2R-like dopamine receptors. *Neurotox. Res.* **29**, 345–350 (2016).
- Pérez-Santamarina, E. *et al.* Regulatory rare variants of the dopaminergic gene ANKK1 as potential risk factors for Parkinson’s disease. *Sci. Rep.* **11**, 9879 (2021).
- Feis, D.-L., Pelzer, E. A., Timmermann, L. & Tittgemeyer, M. Classification of symptom-side predominance in idiopathic Parkinson’s disease. *NPJ Parkinsons Dis.* **1**, 15018 (2015).
- Heinrichs-Graham, E., Santamaria, P. M., Gendelman, H. E. & Wilson, T. W. The cortical signature of symptom laterality in Parkinson’s disease. *NeuroImage. Clin.* **14**, 433–440 (2017).
- Wang, J., Liu, Z. L. & Chen, B. Association study of dopamine D2, D3 receptor gene polymorphisms with motor fluctuations in PD. *Neurology* **56**, 1757–1759 (2001).
- Oliveri, R. L. *et al.* Dopamine D2 receptor gene polymorphism and the risk of levodopa-induced dyskinesias in PD. *Neurology* **53**, 1425–1430 (1999).

Author contributions

M.T., C.E. and L.T. designed the research; S.S., E.A.P. and F.S. acquired data; C.M., D.L.F. and M.S. contributed analytic tools and analyzed the data; E.A.P. and S.S. wrote the draft while M.T. and L.T. reviewed the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This study was supported from a grant of the German Research Foundation in the Clinical Research Group KFO 219. M.T. is also funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC 2030-390661388.

Competing interests

LT received payments as a consultant for Medtronic Inc, Boston Scientific, SAPIENS, St. Jude Medical, Bayer Healthcare, UCB Schwarz Pharma, Archimedes Pharma. LT received honoraria as a speaker on symposia sponsored by TEVA Pharma, Lundbeck Pharma, Bracco, Gianni PR, Medas Pharma, UCB Schwarz Pharma, Desitin Pharma, Boehringer Ingelheim, GlaxoSmithKline, Eumecom, Orion Pharma, Medtronic, Boston Scientific, Cephalon, Abbott, GE Medical, Archimedes, Bayer, TAD Pharma. EAP, SS DLF, CM, FS, MS, CE and MT had no conflicts of interest.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-09657-0>.

Correspondence and requests for materials should be addressed to E.A.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022