Title:

Genetic risk for Parkinson's disease and progression: an analysis of 13 longitudinal cohorts

Authors:

Hirotaka Iwaki M.D.[1,2], Cornelis Blauwendraat Ph.D.[1], Hampton L. Leonard M.S.[1], Ganqiang Liu Ph.D.[3,4,5], Jodi Maple-Grødem Ph.D.[6,7], Jean-Christophe Corvol M.D. Ph.D.[8], Lasse Pihlstrøm M.D. Ph.D.[9], Marlies van Nimwegen Ph.D.[10], Samantha J. Hutten Ph.D.[11], Khanh-Dung H. Nguyen Ph.D.[12], Jacqueline Rick Ph.D.[13], Shirley Eberly M.S.[14], Faraz Faghri M.S.[1,15], Peggy Auinger M.S.[16], Kirsten M. Scott MRCP, M.Phil.[17], Ruwani Wijeyekoon MRCP[17], Vivianna M. Van Deerlin M.D. Ph.D.[18], Dena G. Hernandez Ph.D.[1], Aaron G. Day-Williams Ph.D.[19,20], Alexis Brice M.D.[21,22,23], Guido Alves M.D., Ph.D.[6,24,25], Alastair J. Noyce MRCP, Ph.D.[26,27], Ole-Bjørn Tysnes M.D., Ph.D.[28,29], Jonathan R. Evans MRCP, Ph.D.[30], David P. Breen MRCP, Ph.D.[31,32,33], Karol Estrada Ph.D.[12], Claire E. Wegel MPH[34], Fabrice Danjou M.D., Ph.D.[21], David K. Simon M.D., Ph.D.[35,36], Bernard Ravina M.D.[37,38], Mathias Toft M.D., Ph.D.[9,39], Peter Heutink Ph.D.[40,41], Bastiaan R. Bloem M.D., Ph.D.[10], Daniel Weintraub M.D.[42,43], Roger A. Barker MRCP, Ph.D.[44], Caroline H. Williams-Gray MRCP, Ph.D.[44], Bart P. van de Warrenburg M.D., Ph.D.[10], Jacobus J. Van Hilten M.D., Ph.D.[45], Clemens R. Scherzer M.D.[3,4,5], Andrew B. Singleton Ph.D.[1], Mike A. Nalls Ph.D.[1,2]

1 Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA, 2 Data Tecnica International, Glen Echo, MD, USA, 3 Precision Neurology Program, Harvard Medical School, Brigham and Women's Hospital, Boston, MA, USA, 4 Neurogenomics Laboratory, Harvard Medical School, Brigham and Women's Hospital, Boston, MA, USA, 5 Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital, Boston, MA, USA, 6 The Norwegian Centre for Movement Disorders, Stavanger University Hospital, Stavanger, Norway, 7 The Centre for Organelle Research, University in Stavanger, Stavanger, Norway, 8 Assistance-Publique Hôpitaux de Paris, ICM, INSERM UMRS 1127, CNRS 7225, ICM, Department of Neurology and CIC Neurosciences, Pitié-Salpêtrière Hospital, Paris, France, 9 Department of Neurology, Oslo University Hospital, Oslo, Norway, 10 Department of Neurology, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Medical Centre, Nijmegen, The Netherlands, 11 Michael J Fox Foundation, New York, NY, USA, 12 Translational Genome Sciences, Biogen, Cambridge, MA, USA, 13 Department of Neurology University of Pennsylvania, Philadelphia, PA, USA, 14 Department of

Biostatistics and Computational Biology, University of Rochester, Rochester, NY, USA, 15 Department of Computer Science, University of Illinois Urbana-Champaign, Champaign, IL, USA, 16 Department of Neurology, Center for Health + Technology, University of Rochester, Rochester, NY, USA, 17 Department of Clinical Neurosciences, University of Cambridge, John van Geest Centre for Brain Repair, Cambridge, UK, 18 Department of Pathology and Laboratory Medicine, Center for Neurodegenerative Disease Research, Parelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA, 19 Genetics and Pharmacogenomics, Merck Research Laboratory, Boston, MA, USA, 20 Statistical Genetics, Biogen, Cambridge, MA, USA, 21 Institut du cerveau et de la moelle épinière ICM, Paris, France, 22 Sorbonne Université SU, Paris, France, 23 INSERM UMR1127, Paris, France, 24 Department of Neurology, Stavanger University Hospital, Stavanger, Norway, 25 Department of Mathematics and Natural Science, University of Stavanger, Stavanger, Norway, 26 Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, UK, 27 Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK, 28 Department of Neurology, Haukeland University Hospital, Bergen, Norway, 29 University of Bergen, Bergen, Norway, 30 Department of Neurology, Nottingham University NHS Trust, Nottingham, UK, 31 Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, Scotland, 32 Anne Rowling Regenerative Neurology Clinic, University of Edinburgh, Edinburgh, Scotland, 33 Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, Scotland, 34 Department of Medical and Molecular Genetics, Indiana University, Indianapolis, IN, USA, 35 Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA, USA, 36 Harvard Medical School, Boston, MA, USA, 37 Voyager Therapeutics, Cambridge, MA, USA, 38 Department of Neurology, University of Rochester School of Medicine, Rochester, NY, USA, 39 Institute of Clinical Medicine, University of Oslo, Oslo, Norway, 40 German Center for Neurodegenerative Diseases-Tubingen, Tuebingen, Germany, 41 HIH Tuebingen, Tubingen, Tuebingen, Germany, 42 Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, USA, 43 Department of Veterans Affairs, Philadelphia, PA, USA, 44 Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK, 45 Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands.

Corresponding author:

Mike A. Nalls Ph.D.

Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health

35 Convent Drive, Bethesda, MD 20892, USA

+1-202-468-1533

nallsm@mail.nih.gov

Statistical analysis was conducted by:

Hirotaka Iwaki M.D. Ph.D.

Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health
35 Convent Drive, Bethesda, MD 20892, USA

Character count of the title: 92
Abstract: 241 words
Manuscript: 2992 words
Number of references: 35
Number of Tables: 2
Number of Figures: 2

Funding support

The Intramural Research Program the National Institute on Aging (NIA, Z01-AG000949-02), Biogen Idec, and the Michael J Fox Foundation for Parkinson's Research

Search Terms
Parkinson's disease; Parkinson's disease with dementia; cohort study; Association studies in genetics; Prognosis

This study is supported by the Intramural Research Program the National Institute on Aging (NIA, Z01-AG000949-02), Biogen Idec, and the Michael J Fox Foundation for Parkinson's Research. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The authors had full access to the data in the study and had final responsibility for the decision to submit for publication.

Conflict of interest statements

Hirotaka Iwaki: Grants, Michael J Fox Foundation, Jodi Maple-Grødem: Grants, Norwegian Parkinson's Disease Association, Jean-Christophe Corvol: Advisory Boards, Biogen, Air Liquide, BrainEver, Theranexus, BMS, Zambon, Pfizer, Ipsen, Abbvie; Grants, MJFF, Actelion, Ipsen, Lasse Pihlstrøm: Grants, Norwegian Health Association, South-East Norway Regional Health Authority, Norwegian Parkinson Research Fund, Michael J. Fox Foundation, Khanh-Dung H. Nguyen: Stock Ownership in medically-related fields, Biotech/Pharmaceutical Industry, Kirsten M. Scott: Grants, Wellcome Trust PhD Fellowship, Vivianna M. Van Deerlin: Grants, NIH NS-053488, Aaron G. Day-Williams: Stock Ownership in medically-related fields, Biogen, Merck, Alexis Brice: Advisory Boards, FWO, ERC; Grants, JPND, ANR, Eranet Neuron, Association France Parkinson, Alastair J. Noyce: Honoraria, Britannia Pharmaceuticals; Grants, Parkinson's UK (G-1606), Jonathan R. Evans: Advisory Boards, AbbVie, Global Kinetics, Allergan; Honoraria, UCB, Allergan, AbbVie, Karol Estrada: Stock Ownership in medically-related fields, Biogen, David K. Simon: Consultancies, Lysosomal Therapeutics, Inc.; Advisory Boards, Weston Brain Institute; Honoraria, Parkinson Study Group, Harvard Medical School, Michael J Fox Foundation, Biogen; Grants, National Institutes of Health, Weston Brain Institute, Mission Therapeutics, Inc., and BioElectron Technologies, Bernard Ravina: Stock Ownership in medically-related fields, Voyager Therapeutics; Consultancies, Michael J Fox Foundation, Mathias Toft: Honoraria, Roche; Grants, Research Council of Norway, South-Eastern Norway Regional Health Authority, Michael J. Fox Foundation, Bastiaan R. Bloem: Consultancies, Abbvie, Zambon; Advisory Boards, Michael J Fox Foundation; Honoraria, Speaker fees from Abbvie, Zambon, Bial; Grants, Netherlands Organization for

Scientific Research, the Michael J Fox Foundation, UCB, Abbvie, the Stichting Parkinson Fonds, the Hersenstichting Nederland, the Parkinson's Foundation, Verily Life Sciences, the Topsector Life Sciences and Health, and the Parkinson Vereniging, Daniel Weintraub: Consultancies, Acadia, Alkahest, Anavex Life Sciences, BlackThorn Therapeutics, Bracket, Clintrex LLC, Sunovion, Theravance Biopharma, and the CHDI Foundation, Roger A. Barker: Consultancies, CDI, Oxford Biomedica; Royalties, Springer; Wiley; Grants, EU; NIHR; PUK; CPT; Rosetrees Trust; MRC; Wellcome Trust; Evelyn Trust, Caroline H. Williams-Gray: Grants, MRC Clinician Scientist fellowship, the NIHR Cambridge Biomedical Research Centre, and grants from the Michael J Fox Foundation, the Rosetrees Trust, the Evelyn Trust and Addenbrookes Charitable Trust., Bart P. van de Warrenburg: Advisory Boards, Member of medical advisory boards patient organisations; Royalties, Reed Elsevier (for chapter in Dutch Neurology textbook); Grants, Radboud university medical centre, ZonMW, Hersenstichting, Bioblast Pharma, Jacobus J. Van Hilten: Grants, The Alkemade-Keuls Foundation, Stichting Parkinson Fonds, Parkinson Vereniging, The Netherlands Organisation for Health Research and Development, Clemens R. Scherzer: Grants, NIH grants U01NS082157, U01NS095736, U01NS100603, Mike A. Nalls: Consultancies, Lysosomal Therapies Inc., Vivid Genomics Inc., Kleiner Perkins Caufield & Byers, Michael J. Fox Foundation.

Abstract

Objective To determine if any association between previously identified alleles that confer risk for Parkinson's disease and variables measuring disease progression.

Methods We evaluated the association between 31 risk variants and variables measuring disease progression. A total of 23,423 visits by 4,307 patients of European ancestry from 13 longitudinal cohorts in Europe, North America, and Australia were analyzed

Results We confirmed the importance of *GBA* on phenotypes. *GBA* variants were associated with the development of daytime sleepiness (p.N370S: HR 3.28 [1.69, 6.34]) and possible REM sleep behavior (p.T408M: OR 6.48 [2.04, 20.60]). We also replicated previously reported associations of *GBA* variants with motor/cognitive declines. The other genotype-phenotype associations include an intergenic variant near *LRRK2* and the faster development of motor symptom (Hoehn and Yahr scale 3.0 HR 1.33 [1.16, 1.52] for the C allele of rs76904798); an intronic variant in *PMVK* and the development of wearing-off effects (HR 1.66 [1.19, 2.31] for the C allele of rs114138760).Age of onset was associated with *TMEM175* variant p.M393T (-0.72 [-1.21, -0.23] in years), the C allele of rs199347 (intronic region of *GPNMB*, 0.70 [0.27, 1.14]), and G allele of rs1106180 (intronic region of *CCDC62*, 0.62 [0.21, 1.03])

Conclusions This study provides evidence that alleles associated with Parkinson's disease risk, in particular *GBA* variants, also contribute to the heterogeneity of multiple motor and non-motor aspects. Accounting for genetic variability will be a useful factor in understanding disease course and in minimizing heterogeneity in clinical trials.

Introduction

Parkinson's disease is one of the most common neurodegenerative diseases, with an estimated lifetime risk as high as 1-2%.[1] Parkinson's disease is traditionally characterized by motor features such as bradykinesia, rigidity, and tremor. However, in addition to these motor symptoms, patients with Parkinson's disease also develop non-motor symptoms, which include depression, cognitive decline, sleep abnormalities, reduced olfaction, and autonomic dysfunction.[2] Collectively, the combined spectrum of motor and non-motor symptoms more accurately reflects the multisystem nature of the disease. Patients with Parkinson's disease may present with various combinations of symptoms and show differences in the rates of progression.[3] The application of modern molecular genetic approaches over the last decade has revealed a significant number of genetic risk loci for idiopathic Parkinson's disease.[4–7] However, in comparison with case-control GWAS, analyzing how genetic factors influence clinical presentation and progression requires longitudinal cohorts with much more detailed observations. Such data are sparse and individual cohorts are often small in size and quite varied, posing a challenge both in sample size and heterogeneity.

In an attempt to address these issues, we collected data from 13 distinct longitudinal Parkinson's disease cohorts with detailed clinical data, including assessment of disease progression. We sought to determine if Parkinson's disease genetic risk factors, either in the form of known GWAS variants or an aggregate genetic risk score, are associated with changes in clinical progression and the disease features.

Materials and methods

Study design and participants

A total of 13 Parkinson's disease cohorts from North America, Europe, and Australia participated in the study. Nine were prospective observational cohorts and the rest were from randomized clinical trials. The observational cohorts were Drug Interaction with Genes in Parkinson's Disease (DIGPD), Harvard

Biomarkers Study (HBS),Oslo Parkinson's Disease study (partly including retrospective data), The Norwegian ParkWest study (ParkWest), Parkinson's Disease Biomarker Program (PDBP), Parkinsonism Incidence and Cognitive and Non-motor heterogeneity In Cambridgeshire, Parkinson's Progression Markers Initiative (PPMI), Profiling Parkinson's disease study (ProPark), and the Morris K. Udall Centers for Parkinson's Research (Udall). The four cohorts from randomized clinical trials were Deprenyl and Tocopherol Antioxidative Therapy of Parkinsonism (DATATOP), NIH Exploratory Trials in Parkinson's Disease Large Simple Study 1 (NET-PD_LS1), ParkFit study (ParkFit), and Parkinson Research Examination of CEP-1347 Trial with a subsequent prospective study (PreCEPT/PostCEPT). Information on these cohorts can be found in the Appendix e-1. Subsets of participants from the cohorts who provided DNA and were non-related participants with PD, diagnosed at the age of 18 or later, and of European ancestry, were included in the study. Participants' information and genetic samples were obtained under appropriate written consent and with local institutional and ethical approvals.

Genotyping SNPs and calculation of genetic risk score

Oslo samples were genotyped on the Illumina Infinium OmniExpress array, DIGPD samples were genotyped by Illumina Multi-Ethnic Genotyping Array, and all other samples were genotyped on the NeuroX array.[8] The quality control process of variant calling included GenTrain score < 0.7, minor allele frequency (MAF) > 0.05 (for sample QC but not in our analysis of rare risk factors), and Hardy-Weinberg equilibrium test statistic > 10^{-6}. Sample-specific quality control included a sample call rate of > 0.95, confirmation of sex through genotyping, homozygosity quantified by F within ± 3 SD from the population mean, European ancestry was confirmed by principal-components analysis with 1000 genomes data as the reference, and genetic relatedness of any two individuals < 0.125. Detailed information regarding NeuroX and the quality control process has been described previously.[9] In the present study, we investigated 31 SNPs previously shown to be significantly associated with Parkinson's disease.[10–12] In addition, we also calculated a genetic risk score (GRS) for each participant based on these variants. The scores were

transformed into *Z*-scores within each cohort and treated as an exposure, with effect estimates based on one standard deviation change from the population mean. The list of 31 SNPs and the GRS calculation method are provided in the Table e-1.

Furthermore, principal components (PCs) were created for each dataset from genotypes using PLINK. For the PC calculation, variants were filtered for minor allele frequency (> 0.05), genotype missingness (< 0.05) and Hardy–Weinberg equilibrium (P-value $\geq 10^{-5}$). The remaining variants were pruned (using a 50 kb window, with a 5 SNP shift per window and r^2 threshold of 0.5) and PCs were calculated using the pruned variants.

Measurements

The following clinical measurements and binomial outcomes were recorded longitudinally. (Table e-2) Total and sub-scores of the Unified Parkinson's Disease Rating Scale (UPDRS) or the Movement Disorder Society revised UPDRS version (MDS-UPDRS; modified Hoehn and Yahr scales (HY); modified Schwab and England Activities of Daily Living Scale (SEADL); and scores for the Mini-Mental State Examination (MMSE), SCOPA-cognition, and Montreal Cognitive Assessment (MoCA). Each was treated as a continuous outcome. For the UPDRS and MDS-UPDRS scores specifically, we took *Z*-scores of the total and sub-scores (except for part 4 at baseline) to compare the original and revised UPDRS versions. The conversion was applied to the scores for all subsequent visits. For UPDRS part 4, most participants had very low scores or 0 at baseline, so we normalized across all observations within each cohort. We also analyzed binomial outcomes. If we had access to the raw data, we used common cut-off values which had been tested and reported specificity of 85% or more in patients' population. The binomial outcomes include: existence of family history (1st degree relative. 1st and 2nd degree relatives in HBS, PreCEPT, ProPark, and Udall), hyposmia (University of Pennsylvania Smell Identification Test; UPSIT < 21,[13] or answering "yes" to question 2 in the Non-motor Symptoms (NMS) questionnaire), cognitive impairment (SCOPA-cognition < 23, MMSE < 27 or MoCA < 24,[14,15] or diagnosed with DSM-IV criteria for dementia), wearing off (UPDRS/MDS-UPDRS part 4 off time > 0 or physician's

diagnosis), dyskinesia (UPDRS/MDS-UPDRS part 4 dyskinesia time > 0 or physician's diagnosis), depression (Beck Depression Inventory (BDI) > 14 (PICNICS used 9 instead of 14), Hamilton Depression Rating Scale (HDRS) > 9, Geriatric Depression Scale (GRS) > 5,[16] or physician's diagnosis), constipation (MDS-UPDRS part 1 item 11 > 0, or answering "yes" to question 5 in the NMS questionnaire), excessive daytime sleepiness (Epworth Sleepiness scale (ESS > 9),[17] insomnia (MDS-UPDRS part 1, item 7 > 0), REM sleep behavior disorder (RBD) (answered "yes" to question 1 on the Mayo Sleep Questionnaire (MSQ),[18] or REM sleep behavior disorder screening questionnaire (RBDSQ > 5),[19] restless legs syndrome (RLS) (answered "yes" to MSQ question 3,[20] or RLS diagnosis positive by RBDSQ), and the progression to HY $\geq$ 3 (HY3, representing moderate to severe disease) and death. The individual definitions of these binomial outcomes are summarized in Table e-2. Age, sex, years of education, age at motor symptom onset, and whether or not the patient was treated with levodopa or dopamine agonists at each visit were also recorded for adjustments.

Statistical analysis

Cohort-level analysis

We analyzed the association between exposures and outcomes using appropriate additive models. Covariates of interest were not available for all cohorts; therefore, the model specifications were slightly different between cohorts (detailed in Table e-3). Briefly, the associations between a SNP/GRS and age at onset were analyzed by linear regression modeling adjusting for population stratification (PC1 and PC2). The association between family history of Parkinson's disease and SNP/GRS was analyzed with a logistic regression model adjusting for PC1/2. For continuous variables, linear regression modeling adjusting for sex, education, PC1/2, age at onset, years from diagnosis, family history, and treatment status were applied. For those who had multiple observations, random intercept was added to adjust for repeated measurements of the same individual. For binomial outcomes, the logistic regression at baseline observation was applied using the same covariates as the continuous models. Those that were negative at

baseline were further analyzed by a Cox regression with the same covariates but with treatment status as a time-varying covariate. Observations with missing variables were excluded from the analyses.

Meta-analysis

We applied inverse weighting (precision method) for each combination of outcome-predictor association and combined the estimates from the 13 different cohorts in a fixed-effect model. Multiple test correction for SNPs was controlled with an overall false discovery rate (FDR) of 0.05 per outcome being considered significant. Similarly, multiple testing of outcomes for GRS was corrected with an FDR of 0.05, but across all traits. In addition, as a test of homogeneity, I^2 indices and forest plots were used for quantitative assessment. As a sensitivity analysis, we conducted up to 13 iterations of the meta-analyses for the 12 cohorts excluding each cohort per iteration. This analysis provides information regarding heterogeneity of the cohorts and how one specific cohort exclusion impacts the results. The range of estimates and maximum *P*-values for the iterations were included. Finally, we conducted the 13-cohort meta-analysis in a random effects model with REML estimation using the same multiple testing correction.

All of the above analyses were conducted with PLINK version 1.9, and R version 3.4.4 (64-bit). Statistical tests were all two-sided.

Data availability

Qualified investigators can request raw data through the organizations' homepages (PDBP: https://pdbp.ninds.nih.gov/, PPMI: https://www.ppmi-info.org/) or collaboration.

Results

A total of 23,423 visits by 4,307 patients with a median follow-up period of 2.97 years (quartile range of [1.63, 4.94] years) were eligible for the analysis. The baseline characteristics of the cohorts are shown in Table 1. The mean ages at onset varied from 54 to 69 years old; the average disease durations at cohort entry ranged from less than one to 10 years, and the mean observation periods were between 1.2 and 6.8 years. All DATATOP, ParkWest, PPMI, and PreCEPT participants were dopaminergic therapy-naive at baseline; patients in the other cohorts were not. In the primary analysis of 13 cohorts, 17 associations were identified as significant after FDR correction (Table 2, and more information in Table e-4). Overwhelmingly, 10 were associated with *GBA* variants. In particular, *GBA* p.E365K (rs2230288) was associated with 2.37- [1.53, 3.66] (95% CI) fold higher odds of having cognitive impairment at baseline ($P = 1.09\times10^{-4}$) and 2.78- [1.88, 4.11] fold higher hazard ratio of developing cognitive impairment during follow-up among those who were negative for cognitive impairment at baseline ($P = 2.97\times10^{-7}$). This SNP was also associated with a higher mean on the HY Scale at 0.10 [0.04, 0.16] ($P = 1.53\times10^{-3}$), but the test of homogeneity was rejected ($P = 0.017$, $I^2 = 48.9\%$). In addition, it was associated with the development of a REM sleep behavior disorder among those who didn't have the disorder at baseline. Other *GBA* mutations, p.N370S (rs767763715) and p.T408M (rs75548401), were both associated with a higher HR of reaching HY3 (4.59 [2.60, 8.10] for p.N370S ($P = 1.58\times10^{-7}$) and 1.93 [1.34, 2.78] for p.T408M ($P = 4.40\times10^{-4}$)). *GBA* p.N370N was also associated with a higher risk of developing wearing-off, dyskinesia, and daytime sleepiness. p.T408M was associated with a 6.48 [2.04, 20.60] times higher odds ratio of having a REM sleep behavior disorder symptom at baseline ($P = 1.53\times10^{-3}$).

Two *LRRK2* variants in our 31 SNPs of interest were significantly associated with outcomes. *LRRK2* p.G2019S (rs34637584) was associated with higher odds of having a family history of Parkinson's disease (OR 3.54 [1.72, 7.29], $P = 6.06\times10^{-4}$) and the T allele of rs76904798 (intergenic at the 5' end of *LRRK2*) was associated with a higher HR of reaching HY3 (HR 1.33 [1.16, 1.52] for the T allele, $P = 5.27\times10^{-5}$). Age at onset was inversely associated with the *Z*-value of the genetic risk score (-0.60 [-0.89, -0.31] years per +1 SD, $P = 5.33\times10^{-5}$). Moreover, it was associated with rs34311866 (*TMEM175* p.M393T), the C

allele of rs199347 (intronic region of *GPNMB*), and the G allele of rs1106180 (intronic region of *CCDC62*).

The majority (14/17) of associations showed good accord across cohorts (I2 < 50%) and the forest plots (Fig.1 and Fig.2) also illustrate this qualitatively. Furthermore, up to 13 iterations of the leave-one-out analysis assessed 15 associations of which outcomes were measured in more than two cohorts and showed a small range of betas. The max *P*-value of 13 iterations was less than 0.05 for all associations except for rs114138769 (intron of *PMVK*) and rs76763715 (*GBA* p.N370S) for wearing-off. A meta-analysis with a random effect model also detected nine associations after the same FDR correction, even though the model is more conservative than a fixed model.

Discussion

We conducted a meta-analysis with 13 longitudinal patient cohorts and identified multiple associations between genotypes and clinical phenotypic characteristics, including progression rates. Among these, *GBA* coding variants showed clear associations with the rate of cognitive decline (binomial outcome, or UPDRS part 1 score) as well as motor symptom progression (HY, HY3), consistent with previous studies.[12,21–25]

In addition, we found associations between *GBA* variants and RBD and daytime sleepiness. A previous cross-sectional study with 120 Ashkenazi-Jewish patients reported a higher frequency of RBDSQ-detected RBD symptoms in *GBA* variant carriers,[26] Our finding suggests that *GBA* is associated not only with baseline clinical presentation but also with disease progression.

This is the first study, to our knowledge, noting an association between *GBA* and daytime sleepiness. One study reported an association between sleep problems (as assessed by the Parkinson's Disease Sleep Scale (PDSS)) and *GBA*.[27] However, this scale is a combined measure of daytime sleepiness and other aspects of sleep problems.

Finally, a *GBA* variant (p.N370S) was also associated with treatment-related complications of wearing-off and dyskinesia. Two studies have reported the association of *GBA* variants with these complications, with one positive and one negative result.[28,29] The negative result may be due to insufficient power with only 19 patients with *GBA* mutations.

Overall, our study provides a distinct clinical profile of patients with *GBA* variants compared to those without. We note that with 63 carriers for p.N370S, 166 for p.T408M, and 217 for p.E365K, we have a reasonable power, but the number is yet not enough. And this may affect the results in seemingly different magnitudes of associations and the association for different traits per variants (e.g., motor complications with p.N370S and cognitive impairment with p.E365K). Another possible explanation is that even though the effects are associated with the same gene, the biological activity or molecular mechanism could be different. Such an example has already been reported for *LRRK2* p.G2019S and p.G2385R.[30]

Aside from GBA variants, the associations between close intergenic (5'_end) variant of *LRRK2*, rs76904798, and the faster development of motor symptom. This variant is 4.3 kb upstream from 5' end of *LRRK2* and reported to be associated with *LRRK2* gene expression changes in recent blood cis-eQTL study from the eQTLGen Consortium.[31] In contrast, we did not find an association between rs34637584, *LRRK2* coding mutation (p.G2019S) and motor progression. The p.G2019 variant is a rare variant (MAF 0.5% in our study) and our sample size was not adequate barring an extremely large effect size. The intronic region variant of *PMVK*, rs114138760, and the development of wearing-off was another finding. The biological effect of *PMVK* on PD has not been reported, but the variant is also located at close proximity of the *GBA-SYT11* locus, so it is possible that its association was through a similar mechanism as *GBA*. Including the results of cross-sectional analysis, the associations of age at onset with rs34311866 (*TMEM175*, p.M393T), rs199347 (intron of *GPNMB*), and rs11060180 (intron of *CCDC62*) were found. *TMEM175* has been reported to impair lysosomal and mitochondrial function and increase α-synuclein aggregatio,[32] although no functional data for this missense variant was studied. Interestingly, the variant has recently been reported in another study as being associated with the age at onset.[33] rs199347 is an

eQTL increasing the brain expression of *GPNMB*,[34] suggesting a causal link. Regarding rs1160180, no functional data is available in this locus.

We also evaluated the association between genetic risk variants and clinical outcomes by two-step meta-analysis. This analysis is exploratory, and we acknowledge that this is biased towards the null due to power issues when partitioning studies randomly. However, we believe that it is helpful to assess the rigorousness of the associations we found in the primary analysis as well as to explore potential missed associations.

A strength of the current study was its design, incorporating multiple distinct independent Parkinson's disease cohorts with longitudinal follow-ups. Although the cohorts contained patients at different disease stages, and some of the definition of outcomes were not identical, we analyzed each cohort separately and combined the results. Thus, the significant findings are consistent and applicable to the wider Parkinson's disease populations. The forest plots showed that most of the estimates agree with each other despite the relative differences in the cohort characteristics. Another strength is the size of the study. The total number of genotyped and phenotyped Parkinson's disease patients (N = 4,307) is one of the largest to date for an investigation of disease progression.

The limitations of our study were as follows. First, we only included patients of European ancestry. It is uncertain whether the associations in the current study are also applicable to people from different ethnic backgrounds and further research is needed. Second, the current analysis could not distinguish causality, only basic associations. Different approaches, such as molecular-level assessment and Mendelian randomization, are crucial. Third, interaction effects between genes and other factors are another important research target not addressed in this report due to power constraints. For example, gene-by-smoking interactions for Parkinson's disease were indicated recently,[35] and highlight the importance of correctly modeling gene-environment interactions. Finally, compared with the typical GWAS analysis (which includes tens of thousands of cases) the number of participants was small, and the outcomes of interest were not as simple or easily defined as with case-control distinctions in GWAS. Acknowledging the limitations, the list of associations provided here is valuable as a foundation for further studies and as

an example that illustrates the potential of efforts to define the genetic basis of variability in presentation and course. Accounting for this variability, even in part, has the potential to positively impact etiology-based clinical trials by reducing variability between placebo and treatment groups, and by providing better predictions of expected individual progression.

Acknowledgements

We thank all study participants and their family, investigators and members of the following Studies: Parkinson Study Group: Deprenyl and Tocopherol Antioxidative Therapy of Parkinsonism (DATATOP); Drug Interaction with Genes in Parkinson's Disease (DIGPD); Harvard Biomarkers Study (HBS); NET-PD_LS1, NIH Exploratory Trials in Parkinson's Disease Large Simple Study 1; Oslo PD study; ParkFit study; The Norwegian ParkWest study (ParkWest); Parkinson's Disease Biomarker Program (PDBP);Parkinsonism Incidence and Cognitive and Non-motor heterogeneity In Cambridgeshire (PICNICS); Parkinson's progression markers initiative (PPMI); Parkinson Study Group: Parkinson Research Examination of CEP-1347 Trial (PreCEPT) and its following study (PostCEPT); Profiling Parkinson's disease study (ProPark); and Morris K. Udall Centers for Parkinson's Research (Udall).
We also thank the following grants and financial supporters of above studies; DATATOP was supported by a Public Health Service grant (NS24778) from the National Institute of Neurological Disorders and Stroke (NINDS); by grants from the General Clinical Research Centers Program of the National Institutes of Health at Columbia University (RR00645), the University of Virginia (RR00847), the University of Pennsylvania (RR00040), the University of Iowa (RR00059), Ohio State University (RR00034), Massachusetts General Hospital (RR01066), the University of Rochester (RR00044), Brown University (RR02038), Oregon Health Sciences University (RR00334), Baylor College of Medicine (RR00350), the University of California (RR00827), Johns Hopkins University (RR00035), the University of Michigan (RR00042), and Washington University (RR00036), the Parkinson's Disease Foundation at Columbia-Presbyterian Medical Center, the National Parkinson Foundation, the Parkinson Foundation of Canada,

the United Parkinson Foundation, Chicago, the American Parkinson's Disease Association, New York, and the University of Rochester; DIGPD is supported by Assistance Publique Hôpitaux de Paris, funded by a grant from the French Ministry of Health (PHRC 2008, AOM08010) and a grant from the Agence Nationale pour la Sécurité des Médicaments (ANSM 2013); HBS is supported by the Harvard NeuroDiscovery Center, Michael J Fox Foundation, NINDS U01NS082157, U01NS100603, and the Massachusetts Alzheimer's Disease Research Center NIA P50AG005134; NET-PD_LS1 was supported by NINDS grants U01NS043128; OSLO is supported by the Research Council of Norway and South-Eastern Norway Regional Health Authority; ParkFit is supported by ZonMw (the Netherlands Organization for Health Research and Development (75020012)) and the Michael J Fox Foundation for Parkinson's research, VGZ (health insurance company), GlaxoSmithKline, and the National Parkinson Foundation; ParkWest is supported by the Research Council of Norway, the Western Norway Regional Health Authority, Stavanger University Hospital Research Funds, and the Norwegian Parkinson's Disease Association; PDBP is a consortium with NINDS initiative; PICNICS has received funding from the Cure Parkinson's Trust, the Van Geest Foundation and is supported by the National Institute of Health Research Cambridge Biomedical Research Centre; PPMI is supported by the Michael J Fox Foundation for Parkinson's research; PreCEPT and PostCEPT were funded by NINDS 5U01NS050095-05, Department of Defense Neurotoxin Exposure Treatment Parkinson's Research Program. Grant Number: W23RRYX7022N606, the Michael J Fox Foundation for Parkinson's research, Parkinson's Disease Foundation, Lundbeck Pharmaceuticals. Cephalon Inc, Lundbeck Inc, John Blume Foundation, Smart Family Foundation, RJG Foundation, Kinetics Foundation, National Parkinson Foundation, Amarin Neuroscience LTD, CHDI Foundation Inc, National Institutes of Health (NHGRI, NINDS), Columbia Parkinson's Disease Research Center; ProPARK is funded by the Alkemade-Keuls Foundation, Stichting Parkinson Fonds, Parkinson Vereniging, The Netherlands Organisation for Health Research and Development; Udall is supported by NINDS.

Appendix 1 Authors

NAME	ALL	Role	Contribution
Hirotaka Iwaki M.D. Ph.D.	Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA	Author	Literature search; study design; data analysis; data interpretation; writings
Cornelis Blauwendraat Ph.D.	Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA	Author	Literature search: data analysis; data interpretation; critical review
Hampton L. Leonard M.S.	Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA	Author	critical review
Ganqiang Liu Ph.D.	Precision Neurology Program, Harvard Medical School, Brigham and Women's Hospital, Boston, MA, USA	Author	data collection; critical review
Jodi Maple-Grødem Ph.D.	The Norwegian Centre for Movement Disorders, Stavanger University Hospital, Stavanger, Norway	Author	data collection; critical review
Jean-Christophe Corvol M.D. Ph.D.	Assistance-Publique Hôpitaux de Paris, ICM, INSERM UMRS 1127, CNRS 7225, ICM, Department of Neurology and CIC Neurosciences, Pitié-Salpêtrière Hospital, Paris, France	Author	data collection; critical review
Lasse Pihlstrøm M.D. Ph.D.	Department of Neurology, Oslo University Hospital, Oslo, Norway	Author	data collection; critical review
Marlies van Nimwegen Ph.D.	Department of Neurology, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Medical Centre, Nijmegen, The Netherlands	Author	data collection; critical review
Samantha J. Hutten Ph.D.	Michael J Fox Foundation, New York, NY, USA	Author	data collection; critical review
Khanh-Dung H. Nguyen Ph.D.	Translational Genome Sciences, Biogen, Cambridge, MA, USA	Author	data collection; critical review
Jacqueline Rick Ph.D.	Department of Neurology University of Pennsylvania, Philadelphia, PA, USA	Author	data collection; critical review
Shirley Eberly M.S.	Department of Biostatistics and Computational Biology, University of Rochester, Rochester, NY, USA	Author	data collection; critical review
Faraz Faghri M.S.	Laboratory of Neurogenetics, National Institute on Aging,	Author	data collection; critical review

	National Institutes of Health, Bethesda, MD, USA		
Peggy Auinger M.S.	Department of Neurology, Center for Health + Technology, University of Rochester, Rochester, NY, USA	Author	data collection; critical review
Kirsten M. Scott MRCP, M.Phil.	Department of Clinical Neurosciences, University of Cambridge, John van Geest Centre for Brain Repair, Cambridge, UK	Author	data collection; critical review
Ruwani Wijeyekoon, MRCP	Department of Clinical Neurosciences, University of Cambridge, John van Geest Centre for Brain Repair, Cambridge, UK	Author	data collection; critical review
Vivianna M. Van Deerlin M.D. Ph.D.	Department of Pathology and Laboratory Medicine, Center for Neurodegenerative Disease Research, Parelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA	Author	data collection; critical review
Dena G. Hernandez Ph.D.	Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA	Author	data collection; critical review
Aaron G. Day-Williams Ph.D.	Genetics and Pharmacogenomics, Merck Research Laboratory, Boston, MA, USA	Author	data collection; critical review
Alexis Brice M.D.	Institut du cerveau et de la moelle épinière ICM, Paris, France	Author	data collection; critical review
Guido Alves M.D., Ph.D.	The Norwegian Centre for Movement Disorders, Stavanger University Hospital, Stavanger, Norway	Author	data collection; critical review
Alastair J. Noyce MRCP, Ph.D.	Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, UK	Author	data collection; critical review
Ole-Bjørn Tysnes M.D., Ph.D.	Department of Neurology, Haukeland University Hospital, Bergen, Norway	Author	data collection; critical review
Jonathan R. Evans MRCP, Ph.D.	Department of Neurology, Nottingham University NHS Trust, Nottingham, UK	Author	data collection; critical review
David P. Breen MRCP, Ph.D.	Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, Scotland	Author	data collection; critical review

Karol Estrada Ph.D.	Translational Genome Sciences, Biogen, Cambridge, MA, USA	Author	data collection; critical review
Claire E. Wegel MPH	Department of Medical and Molecular Genetics, Indiana University, Indianapolis, IN, USA	Author	data collection; critical review
Fabrice Danjou M.D., Ph.D.	Institut du cerveau et de la moelle épinière ICM, Paris, France	Author	data collection; critical review
David K. Simon M.D., Ph.D.	Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA, USA	Author	data collection; critical review
Bernard Ravina M.D.	Voyager Therapeutics, Cambridge, MA, USA	Author	data collection; critical review
Mathias Toft M.D., Ph.D.	Department of Neurology, Oslo University Hospital, Oslo, Norway	Author	data collection; critical review
Peter Heutink Ph.D.	German Center for Neurodegenerative Diseases-Tubingen, Tuebingen, Germany	Author	data collection; critical review
Bastiaan R. Bloem M.D., Ph.D.	Department of Neurology, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Medical Centre, Nijmegen, The Netherlands	Author	data collection; critical review
Daniel Weintraub M.D.	Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, USA	Author	data collection; critical review
Roger A. Barker MRCP, Ph.D.	Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK	Author	data collection; critical review
Caroline H. Williams-Gray MRCP, Ph.D.	Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK	Author	data collection; critical review
Bart P. van de Warrenburg M.D., Ph.D.	Department of Neurology, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Medical Centre, Nijmegen, The Netherlands	Author	data collection; critical review
Jacobus J. Van Hilten M.D., Ph.D.	Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands	Author	data collection; critical review
Clemens R. Scherzer M.D.	Precision Neurology Program, Harvard Medical School, Brigham and Women's Hospital, Boston, MA, USA	Author	data collection; critical review
Andrew B. Singleton Ph.D.	Laboratory of Neurogenetics, National Institute on Aging,	Author	study design; critical review

	National Institutes of Health, Bethesda, MD, USA		
Mike A. Nalls Ph.D.	Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA	Author	study design; data analysis; data interpretation; critical review

Reference

1. Elbaz A, Bower JH, Maraganore DM, et al. Risk tables for parkinsonism and Parkinson's disease. J Clin Epidemiol. 2002;55:25–31.
2. Chaudhuri KR, Healy DG, Schapira AH V, National Institute for Clinical Excellence. Non-motor symptoms of Parkinson's disease: diagnosis and management. Lancet Neurol. 2006;5:235–245.
3. Lewis SJG, Foltynie T, Blackwell AD, Robbins TW, Owen AM, Barker RA. Heterogeneity of Parkinson's disease in the early clinical stages using a data driven approach. J Neurol Neurosurg Psychiatry. 2005;76:343–348.
4. Satake W, Nakabayashi Y, Mizuta I, 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.
5. Simón-Sánchez J, Schulte C, Bras JM, et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat Genet. 2009;41:1308–1312.
6. Chang D, Nalls MA, Hallgrímsdóttir IB, et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. Nat Genet. 2017;49:1511–1516.
7. International Parkinson Disease Genomics Consortium, Nalls MA, Plagnol V, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet (London, England). 2011;377:641–649.
8. Nalls MA, Bras J, Hernandez DG, et al. NeuroX, a fast and efficient genotyping platform for investigation of neurodegenerative diseases. Neurobiol Aging. 2015;36:1605.e7-12.

9. Nalls MA, Escott-Price V, Williams NM, et al. Genetic risk and age in Parkinson's disease: Continuum not stratum. Mov Disord. 2015;30:850–854.
10. Nalls MA, McLean CY, Rick J, et al. Diagnosis of Parkinson's disease on the basis of clinical and genetic classification: A population-based modelling study. Lancet Neurol. Elsevier Ltd; 2015;14:1002–1009.
11. Pankratz N, Beecham GW, DeStefano AL, et al. Meta-analysis of Parkinson's disease: identification of a novel locus, RIT2. Ann Neurol. 2012;71:370–384.
12. Davis AA, Andruska KM, Benitez BA, Racette BA, Perlmutter JS, Cruchaga C. Variants in GBA, SNCA, and MAPT influence Parkinson disease risk, age at onset, and progression. Neurobiol Aging. 2016;37:209.e1-209.e7.
13. Picillo M, Pellecchia MT, Erro R, et al. The use of University of Pennsylvania Smell Identification Test in the diagnosis of Parkinson's disease in Italy. Neurol Sci. 2014;35:379–383.
14. Hoops S, Nazem S, Siderowf AD, et al. Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. Neurology. 2009;73:1738–1745.
15. Verbaan D, van Rooden SM, van Hilten JJ, Rijsman RM. Prevalence and clinical profile of restless legs syndrome in Parkinson's disease. Mov Disord. 2010;25:2142–2147.
16. Goodarzi Z, Mrklas KJ, Roberts DJ, Jette N, Pringsheim T, Holroyd-Leduc J. Detecting depression in Parkinson disease: A systematic review and meta-analysis. Neurology. 2016;87:426–437.
17. Simuni T, Caspell-Garcia C, Coffey C, et al. Correlates of excessive daytime sleepiness in de novo Parkinson's disease: A case control study. Mov Disord. 2015;30:1371–1381.
18. Boeve BF, Molano JR, Ferman TJ, et al. Validation of the Mayo Sleep Questionnaire to screen for REM sleep behavior disorder in an aging and dementia cohort. Sleep Med. 2011;12:445–453.
19. Nomura T, Inoue Y, Kagimura T, Uemura Y, Nakashima K. Utility of the REM sleep behavior disorder screening questionnaire (RBDSQ) in Parkinson's disease patients. Sleep Med. 2011;12:711–713.

20. Boeve BF, Molano JR, Ferman TJ, et al. Validation of the Mayo Sleep Questionnaire to screen for REM sleep behavior disorder in a community-based sample. J Clin Sleep Med. 2013;9:475–480.
21. Davis MY, Johnson CO, Leverenz JB, et al. Association of GBA Mutations and the E326K Polymorphism With Motor and Cognitive Progression in Parkinson Disease. JAMA Neurol. 2016;73:1217–1224.
22. Winder-Rhodes SE, Evans JR, Ban M, et al. Glucocerebrosidase mutations influence the natural history of Parkinson's disease in a community-based incident cohort. Brain. 2013;136:392–399.
23. Brockmann K, Srulijes K, Pflederer S, et al. GBA-associated Parkinson's disease: reduced survival and more rapid progression in a prospective longitudinal study. Mov Disord. 2015;30:407–411.
24. Liu G, Boot B, Locascio JJ, et al. Specifically neuropathic Gaucher's mutations accelerate cognitive decline in Parkinson's. Ann Neurol. 2016;80:674–685.
25. Liu G, Locascio JJ, Corvol J-C, et al. Prediction of cognition in Parkinson's disease with a clinical-genetic score: a longitudinal analysis of nine cohorts. Lancet Neurol. 2017;16:620–629.
26. Gan-Or Z, Mirelman A, Postuma RB, et al. GBA mutations are associated with Rapid Eye Movement Sleep Behavior Disorder. Ann Clin Transl Neurol. 2015;2:941–945.
27. Brockmann K, Srulijes K, Hauser AK, et al. GBA-associated PD presents with nonmotor characteristics. Neurology. 2011;77:276–280.
28. Oeda T, Umemura A, Mori Y, et al. Impact of glucocerebrosidase mutations on motor and nonmotor complications in Parkinson's disease. Neurobiol Aging. 2015;36:3306–3313.
29. Jesús S, Huertas I, Bernal-Bernal I, et al. GBA Variants Influence Motor and Non-Motor Features of Parkinson's Disease. PLoS One. 2016;11:e0167749.
30. Marras C, Alcalay RN, Caspell-Garcia C, et al. Motor and nonmotor heterogeneity of LRRK2-related and idiopathic Parkinson's disease. Mov Disord. 2016;31:1192–1202.
31. Võsa U, Claringbould P, Westra H-J, et al. Unraveling the polygenic architecture of complex traits using blood eQTL meta-analysis. bioRxiv. Epub 2018.

32. Jinn S, Drolet RE, Cramer PE, et al. TMEM175 deficiency impairs lysosomal and mitochondrial function and increases α-synuclein aggregation. Proc Natl Acad Sci U S A. 2017;114:2389–2394.
33. Blauwendraat C, Heilbron K, Vallerga CL, Bandres-Ciga S, Coelln R von, Pihlstrøm L. Parkinson disease age at onset GWAS: defining heritability, genetic loci and α-synuclein mechanisms. Mov Disord. 2019;in press.
34. Murthy MN, Blauwendraat C, UKBEC, et al. Increased brain expression of GPNMB is associated with genome wide significant risk for Parkinson's disease on chromosome 7p15.3. Neurogenetics. 2017;18:121–133.
35. Lee P-C, Ahmed I, Loriot M-A, et al. Smoking and Parkinson disease: Evidence for gene-by-smoking interactions. Neurology. 2018;90:e583–e592.

Supplementary materials

Appendix e-1: The cohort information and details

Table e-1: Allele frequency of 31 SNPs and the effect size

Table e-2: Cohort specific definitions of binomial outcomes

Table e-3: Model specifications

Table e-4: Meta-analysis for 13 cohorts and the results of sensitivity analysis (The estimated betas and FDR-corrected P)

Figure legends

Fig. 1: Forest plots for GBA variants and symptoms of Parkinson's disease

DATATOP, Deprenyl and Tocopherol Antioxidative Therapy of Parkinsonism; DIGPD, Drug Interaction with Genes in Parkinson's Disease; HBS, Harvard Biomarkers Study; NET-PD_LS1, NIH Exploratory Trials in Parkinson's Disease Large Simple Study 1; Oslo, Oslo PD study; ParkFit, ParkFit study;

ParkWest, The Norwegian ParkWest study; PDBP, Parkinson's Disease Biomarker Program; PICNICS, Parkinsonism Incidence and Cognitive and Non-motor heterogeneity In Cambridgeshire; PPMI, Parkinson's progression markers initiative; PreCEPT, Parkinson Research Examination of CEP-1347 Trial and PostCEPT; ProPark, Profiling Parkinson's disease study; and Udall, Morris K. Udall Centers for Parkinson's Research.

Fig. 2: Forest plots for non-GBA risk variants/genetic risk score and symptoms or features of Parkinson's disease.

DATATOP, Deprenyl and Tocopherol Antioxidative Therapy of Parkinsonism; DIGPD, Drug Interaction with Genes in Parkinson's Disease; HBS, Harvard Biomarkers Study; NET-PD_LS1, NIH Exploratory Trials in Parkinson's Disease Large Simple Study 1; Oslo, Oslo PD study; ParkFit, ParkFit study; ParkWest, The Norwegian ParkWest study; PDBP, Parkinson's Disease Biomarker Program; PICNICS, Parkinsonism Incidence and Cognitive and Non-motor heterogeneity In Cambridgeshire; PPMI, Parkinson's progression markers initiative; PreCEPT, Parkinson Research Examination of CEP-1347 Trial and PostCEPT; ProPark, Profiling Parkinson's disease study; and Udall, Morris K. Udall Centers for Parkinson's Research.

Table 1. Summary characteristics of 13 cohorts

	DATATOP	DIGPD	HBS	NET-PD LS1	Oslo	ParkFit	ParkWest	PDBP	PICNICS	PPMI	PreCEPT/ PostCEPT	ProPark	Udall
Cohort size, n	440	311	580	406	317	335	150	422	120	357	321	296	252
Follow-up duration, years	1.22 (0.41)	2.19 (1.51)	1.53 (0.87)	4.48 (1.45)	4.64 (3.10)	1.97 (0.00)	3.04 (0.09)	2.06 (1.70)	3.04 (1.63)	4.87 (1.35)	6.79 (0.95)	4.62 (1.14)	3.77 (1.81)
Female, n (%)	146 (33.2)	121 (38.9)	201 (34.7)	148 (36.5)	107 (33.8)	110 (32.8)	57 (38.0)	174 (41.2)	43 (35.8)	121 (33.9)	106 (33.0)	105 (35.5)	73 (29.0)
Family history, n (%)	86 (20.9)	69 (22.3)	148 (25.5)	59 (14.5)	43 (14.0)	-	17 (11.3)	54 (12.8)	19 (15.8)	48 (13.5)	93 (29.2)	76 (25.9)	71 (28.4)
Age at onset, years	58.65 (9.17)	59.41 (9.80)	62.16 (10.46)	60.64 (9.45)	54.33 (10.06)	60.79 (8.65)	67.27 (9.26)	58.51 (10.28)	68.94 (9.34)	61.45 (9.55)	59.47 (9.22)	53.14 (10.60)	64.26 (8.64)
Baseline from diagnosis, years	1.14 (1.17)	2.60 (1.57)	4.09 (4.63)	1.50 (1.00)	10.13 (6.04)	5.18 (4.44)	0.13 (0.12)	5.68 (5.64)	0.23 (0.48)	0.54 (0.54)	0.80 (0.83)	6.56 (4.67)	6.21 (5.38)
Levodopa use, n (%)	0 (0.0)	198 (63.9)	415 (71.6)	207 (51.2)	-	-	0 (0.0)	255 (60.4)	36 (30.0)	0 (0.0)	0 (0.0)	202 (68.2)	215 (85.3)
Dopamine agonist use, n (%)	0 (0.0)	228 (73.3)	224 (38.6)	280 (69.3)	-	-	0 (0.0)	61 (14.5)	22 (18.3)	0 (0.0)	1 (0.3)	222 (75.0)	118 (46.8)
Modified Hoehn and Yahr Scale	1.61 (0.53)	1.75 (0.55)	2.14 (0.64)	-	2.19 (0.64)	2.08 (0.33)	1.86 (0.58)	2.04 (0.69)	1.64 (0.67)	1.55 (0.50)	1.75 (0.48)	2.51 (0.79)	2.29 (0.68)
UPDRS1	-	7.69 (4.50)	1.70 (1.59)	1.31 (1.45)	-	-	1.95 (1.76)	9.90 (6.11)	-	5.40 (3.97)	0.84 (1.19)	-	1.92 (1.99)
UPDRS2	-	7.72 (4.66)	9.21 (5.23)	7.29 (3.86)	-	-	8.19 (4.22)	11.14 (8.01)	-	5.80 (4.11)	6.11 (3.20)	-	10.74 (7.13)
UPDRS3	-	20.37 (10.23)	19.30 (9.58)	17.77 (8.32)	15.42 (10.30)	-	22.09 (9.77)	23.64 (13.08)	-	20.88 (9.00)	18.69 (7.65)	-	22.92 (11.09)
UPDRS4	-	0.66 (2.56)	2.25 (2.05)	1.34 (1.49)	-	-	0.57 (1.14)	2.20 (3.17)	-	-	-	-	2.02 (2.75)
MDS_UPDRS total	-	36.43 (16.02)	-	-	-	-	-	46.88 (24.04)	47.27 (17.97)	-	-	-	-
UPDRS total	24.68 (11.56)	-	32.33 (14.28)	27.67 (11.62)	-	32.11 (10.10)	32.79 (13.91)	-	-	-	25.39 (10.10)	-	32.64 (18.28)
MMSE	28.99 (1.35)	28.38 (1.73)	28.35 (2.17)	-	-	28.09 (1.61)	27.88 (2.27)	-	28.71 (1.43)	-	29.29 (1.07)	27.05 (2.50)	26.83 (3.50)
MoCA	-	-	-	-	-	-	-	25.44 (3.40)	-	27.17 (2.23)	-	-	24.37 (3.63)
SEADL	91.55 (6.49)	80.55 (29.02)	-	91.59 (6.06)	-	-	89.40 (7.35)	85.11 (13.10)	-	93.18 (5.91)	92.77 (5.26)	-	80.53 (17.56)
Hyposmia, n (%)	-	89 (28.9)	-	-	-	-	54 (36.0)	276 (65.4)	-	164 (45.9)	-	173 (63.8)	69 (67.0)
Cognitive impairment, n (%)	26 (5.9)	3 (1.0)	74 (13.0)	29 (7.1)	-	55 (16.4)	27 (18.0)	96 (22.7)	11 (9.2)	28 (7.8)	3 (0.9)	77 (27.0)	29 (11.5)
Motor fluctuation, n (%)	-	40 (12.9)	228 (39.9)	103 (25.4)	-	-	4 (2.7)	129 (48.1)	1 (0.8)	-	-	94 (32.4)	75 (35.4)
Dyskinesia, n (%)	4 (0.9)	13 (4.2)	207 (36.2)	5 (1.2)	-	-	2 (1.3)	196 (46.4)	0 (0.0)	-	-	81 (27.6)	44 (22.8)
Depression, n (%)	12 (2.7)	97 (31.6)	35 (10.9)	40 (9.9)	-	-	20 (13.3)	49 (11.6)	27 (22.5)	113 (31.7)	73 (22.7)	48 (16.3)	63 (25.0)
Restless legs syndrome, n (%)	-	44 (14.5)	37 (10.9)	-	-	-	-	91 (23.3)	-	23 (6.4)	-	-	-

Constipation, n (%)	9 (2.0)	62 (20.3)	-	-	-	-	17 (11.3)	239 (56.6)	29 (24.2)	113 (31.7)	-	138 (46.6)	-
REM sleep behavior disorder, n (%)	-	-	-	-	-	-	-	197 (50.5)	-	93 (26.1)	-	-	-
Daytime Sleepiness, n (%)	5 (1.1)	138 (44.8)	-	-	-	-	25 (16.7)	165 (39.1)	25 (20.8)	55 (15.4)	-	126 (42.6)	-
Insomnia, n (%)	11 (2.5)	107 (35.1)	202 (35.1)	-	-	-	45 (30.0)	295 (69.9)	62 (51.7)	78 (21.8)	-	83 (28.0)	-
HY>=3.0, n (%)	0 (0.0)	4 (1.3)	71 (12.4)	12 (3.0)	22 (14.5)	17 (5.1)	11 (7.3)	71 (16.8)	13 (10.8)	1 (0.3)	0 (0.0)	117 (40.8)	57 (23.0)

Continuous variables were summarized in Mean (SD). MMSE, Mini Mental State Examination; Montreal Cognitive Assessment, SEADL, Schwab and England Activities of Daily Living Scale; UPDRS, Unified Parkinson Disease Rating Scale; MDS-UPDRS, Movement Disorder Society revised version of UPDRS.

Table 2. Meta-a-lysis for 13 cohorts and the results of sensitivity analysis

					Fixed effect model				Leave-one-out a-lysis		Random effect model	
Outcome	rsNo	Known gene or nearest gene	N of cohorts	Scale of the effect	Estimate [95% C.I.]	*P*	Test of Homogeneity	I² (%)	Estimate [Min, Max]	*Max P*	Estimate [95% C.I.]	*P*
Wearing-off	rs114138760	intron_PMVK	9	Multiplicative (HR)	1.66 [1.19, 2.31]	2.62E-03	0.322	12.58	1.66 [1.44, 1.81]	6.22E-02	1.65 [1.14, 2.38]	7.39E-03
Dyskinesia	rs76763715	GBA:N370S	8	Multiplicative (HR)	3.01 [1.81, 5.01]	2.17E-05	0.011	60.53	3.00 [1.98, 4.05]	2.26E-02	2.49 [1.06, 5.86]	3.73E-02
HY>=3.0	rs76763715	GBA:N370S	6	Multiplicative (HR)	4.59 [2.60, 8.10]	1.58E-07	0.654	0.00	4.59 [4.02, 5.41]	2.00E-05	4.59 [2.60, 8.10]	1.58E-07 *
Wearing-off	rs76763715	GBA:N370S	6	Multiplicative (HR)	2.03 [1.28, 3.21]	2.56E-03	0.021	62.70	2.02 [1.61, 2.65]	8.67E-02	1.92 [0.85, 4.33]	1.14E-01
Daytime sleepiness	rs76763715	GBA:N370S	6	Multiplicative (HR)	3.28 [1.69, 6.34]	4.24E-04	0.467	0.00	3.30 [2.85, 4.38]	3.75E-03	3.28 [1.69, 6.34]	4.24E-04 *
HY>=3.0	rs75548401	GBA:T408M	8	Multiplicative (HR)	1.93 [1.34, 2.78]	4.40E-04	0.208	32.43	1.93 [1.70, 2.41]	1.08E-02	1.96 [1.22, 3.14]	5.22E-03
pRBD (Baseline)	rs75548401	GBA:T408M	2	Multiplicative (OR)	6.48 [2.04, 20.60]	1.53E-03	0.118	59.06	-	-	6.25 [1.02, 38.20]	4.72E-02
HY	rs2230288	GBA:E365K	12	Continuous	0.10 [0.04, 0.16]	1.53E-03	0.017	48.90	0.10 [0.08, 0.11]	1.02E-02	0.11 [0.02, 0.21]	1.88E-02
Cognitive impairment (Baseline)	rs2230288	GBA:E365K	8	Multiplicative (OR)	2.37 [1.53, 3.66]	1.09E-04	0.794	0.00	2.37 [2.20, 2.59]	8.57E-04	2.37 [1.53, 3.66]	1.09E-04 *
Cognitive impairment	rs2230288	GBA:E365K	9	Multiplicative (HR)	2.78 [1.88, 4.11]	2.97E-07	0.555	0.00	2.78 [2.41, 2.98]	5.08E-05	2.78 [1.88, 4.11]	2.97E-07 *
pRBD	rs2230288	GBA:E365K	2	Multiplicative (HR)	2.57 [1.43, 4.63]	1.69E-03	0.665	0.00	-	-	2.57 [1.43, 4.63]	1.69E-03 *
Age at onset	rs34311866	TMEM175:M393T	13	Continuous	-0.72 [-1.21, -0.23]	3.87E-03	0.515	0.00	-0.72 [-0.83, -0.58]	2.83E-02	-0.72 [-1.21, -0.23]	
Age at onset	rs199347	intron_GPNMB	12	Continuous	0.70 [0.27, 1.14]	1.42E-03	0.824	0.00	0.70 [0.60, 0.77]	1.12E-02	0.70 [0.27, 1.14]	1.42E-03 *
HY>=3.0	rs76904798	5_LRRK2	13	Multiplicative (HR)	1.33 [1.16, 1.52]	5.27E-05	0.049	43.15	1.33 [1.26, 1.43]	1.64E-03	1.34 [1.11, 1.63]	2.80E-03 *
Family History	rs34637584	LRRK2:G2019S	8	Multiplicative (OR)	3.54 [1.72, 7.29]	6.06E-04	0.856	0.00	3.53 [2.78, 3.98]	1.66E-02	3.54 [1.72, 7.29]	6.06E-04 *
Age at onset	rs11060180	intron_CCDC62	13	Continuous	0.62 [0.21, 1.03]	3.32E-03	0.054	42.60	0.62 [0.49, 0.75]	2.74E-02	0.55 [-0.00, 1.11]	5.14E-02
Age at onset	Genetic risk score		13	Continuous	-0.60 [-0.89, -0.31]	5.33E-05	0.749	0.00	-0.60 [-0.65, -0.52]	9.02E-04	-0.60 [-0.89, -0.31]	5.33E-05 *

Possible REM sleep behavior disorder (pRBD) was only available in 2 cohorts and a leave-one-out a-lysis was not conducted for this outcome
HY, Hoehn and Yahr Scale;
* Significant after FDR adjustment in a random effect model