ONLINE METHODS

Supporting Text: Online methods and consortia members.

Participating studies

Parkinson's Progression Marker Initiative (PPMI) - The Parkinson's Progression Marker Initiative is an observational multi-center, international study sponsored by the Michael J. Fox Foundation for Parkinson's Research and partially funded by 17 industry sponsors. The overall goal of the study is to identify and validate biomarkers of PD progression (PD). The study population consists of untreated recently diagnosed PD (423), similar age and gender healthy controls (196) and subjects screened as potential PD subjects but with dopamine transporter imaging scans, measured by DaTSCAN® without evidence of dopaminergic deficit (SWEDD). Participants enrolled in PPMI undergo a series of longitudinal assessments, including standardized functional assessments, neuroimaging, and biofluid collection (including DNA from whole blood). In this report, the PPMI cohort is divided into three groups of participants: healthy controls (HC), PD subjects and SWEDD subjects. PD subjects in PPMI are required to demonstrate at least asymmetric resting tremor or asymmetric bradykinesia or some combination of bradykinesia, resting tremor and/or rigidity within two years of diagnosis [1]. They must be untreated for PD at the time of enrollment, as well as for the prior two years. PD cases must have an abnormal DaTSCAN® indicating dompamine transporter deficit. SWEDD participants have the same clinical criteria as PD cases, but DaT scanning data does not demonstrate an abnormal DaTSCAN® . HC samples are clinically defined as having no known neurologic dysfunction and a Montreal Cognitive Assessment (MoCA) > 26 [2]. The PD phenotype in this analysis excludes known SWEDD participants. The exclusion of SWEDD participants from the PD model allows us to focus our efforts on more etiologically typical PD as defined by the clinical diagnosis and DaT scanning data. All data in PPMI is available to qualified investigators via http://www.ppmi-info.org/.

Parkinson's Disease Biomarkers Program (PDBP) - Established in November 2012 by the National Institute of Neurological Disorders and Stroke (NINDS), the PDBP seeks to identify and develop potential PD biomarkers, ideally for use in clinical trials of neuroprotective agents. The PDBP includes four key components: 1) biomarker hypothesis testing and collection of clinical data and biospecimens, 2) studies to identify novel PD biomarkers, 3) biospecimen banking and distribution, and 4) data management through the Data Management Resource (DMR). The application of these goals has resulted in the establishment of a self-structured consortium consisting of 11 unique projects, 6 of which actively enroll participants. Consortium-wide protocols ensure standardization of data collection and biospecimen processing. A standard set of clinical assessments and biospecimen collection procedures are used for all participants and

---

[1] Marek, Kenneth et al. "The parkinson progression marker initiative (PPMI)." *Progress in neurobiology* 95.4 (2011): 629-635.

[2] Lindholm, Beata et al. "Prediction of Falls and/or Near Falls in People with Mild Parkinson's Disease." *PloS one* 10 (2015): e0117018-e0117018.

specified by the NINDS (see RFA NS-12-011). These clinical assessments were chosen based on the NINDS Common Data Elements (see http://www.commondataelements.ninds.nih.gov) as well as for overlap with assessments used in BioFIND and PPMI. The PDBP has enrolled over 1,000 participants to date including participants with PD based on clinical criteria, neurologically normal controls, and other individuals with parkinsonism not meeting typical PD criteria as defined by the UK Parkinson's Disease Society Brain Bank Clinical Diagnostic Criteria[3]. All sites use the DMR to record data; storage of biospecimens and quality control analysis are performed by NINDS Repository Laboratories. To access data and more information regarding PDBP, please refer to https://pdbp.ninds.nih.gov/.

Parkinson's Associated Risk Study (PARS) - The PARS study is a prospective study aiming to test a sequential biomarker strategy to identify subjects at high risk for developing motor symptoms of PD. The study enrolled approximately 10,000 individuals to examine their risk profile for PD. The study tested individual over age 60 with olfactory testing (using the University of Pennsylvania Smell Identification Test (UPSIT)) sent by mail to identify a population of hyposmic and normaosmic subjects for more intensive clinical and biomarker testing including dopamine transporter imaging. The study has demonstrated that hyposmic subjects have a 11% risk of marked dopamine transporter deficit. During four year follow-up approximately 60% (14 subjects) of subjects identified with hyposmia and dopamine transporter deficit have developed motor PD. The PARS study enrollment and follow-up is ongoing and more information is available at http://www.parsinfosource.com/. We treated PARS as a positive control since hyposmia defined by a low UPSIT score and to a lesser degree family history were criteria for recruitment of at risk individuals. We do not regard this as a true validation but more of a theoretical proof of concept for our classification model that was developed and trained on PPMI data.

23andMe - The 23andMe Parkinson's Disease cohort was described in detail previously [4]. Briefly, patients with PD were recruited through a targeted email campaign in conjunction with the Michael J. Fox Foundation, The Parkinson's Institute and Clinical Center, and numerous other PD patient groups and clinics. As part of a Michael J. Fox Foundation-funded study of PD biomarkers focusing mainly on individuals harboring at least one *LRRK2* p.G2019S allele, 20 individuals from the 23andMe Parkinson's Disease cohort without *LRRK2* p.G2019S and 20 healthy controls underwent blood draws and completed the UPSIT. Additional phenotypic information was obtained through online questionnaires and classification as cases and controls was performed as described previously. The 23andMe study protocol and consent were approved by the external Association for the Accreditation of Human Research Protection Programs, Inc. accredited Institutional Review Board, Ethical and Independent Review Services. Our consent and privacy statement preclude sharing of individual-level data without explicit consent.

---

[3] Hughes, Andrew J, Susan E Daniel, and Andrew J Lees. "Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease." *Neurology* 57.8 (2001): 1497-1499.

[4] Do, Chuong B et al. "Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease." *PLoS genetics* 7.6 (2011): e1002141.

The Longitudinal and Biomarker Study in PD (LABS-PD) - The Longitudinal and Biomarker Study in PD is an observational study designed to prospectively measure the evolution of motor and non-motor features of PD and identify promising biomarkers of progression from early to late stages of the disease. Study participants had previously been enrolled in a controlled clinical trial of a mixed lineage kinase inhibitor in early, untreated PD (PreCEPT); the average duration of illness was less than 1 year, and subjects did not require dopaminergic therapy. Many of the original trial participants were subsequently enrolled in a follow-up study (PostCEPT) and later participated in a longitudinal clinical assessment program for biomarker development in PD, with annual visits and remote follow-up. As part of the PreCEPT and PostCEPT studies, subjects underwent DAT imaging; SWEDD participants were identified as per PPMI.

Morris K. Udall Parkinson's Disease Research Center of Excellence (Penn-Udall) - The NINDS funded Penn-Udall Center was launched at the Perelman School of Medicine (Penn) at the University of Pennsylvania in 2007 (P50 NS062684 Pacific Northwest Udall Center, Zabetian PI and P50 NS053488, Trojanowski JQT-PI). The overarching goals of the Penn Udall Center are to elucidate mechanisms of disease progression and alpha–synuclein transmission through synergistic collaborations between basic and translational research.

The Clinical Core of the Penn Udall Center recruits patients with PD, PDD and DLB to participate in a longitudinal battery of neuropsychological testing and to donate plasma, whole blood for DNA, cerebrospinal fluid, structural and functional brain imaging and post-mortem brain tissue. The specific goal of the biomarker collection effort is to improve the ability to predict whether an individual patient with PD is likely to develop significant cognitive decline. To date we have enrolled over 300 patients in longitudinal neuropsychological testing, and of those over 100 have participated in complete biomarker collection, including blood for DNA.

Assignment of a cognitive diagnosis is made for each patient at baseline and at every annual or biannual visit during a consensus conference held every six months by movement disorders specialists affiliated with the Penn Udall Center. A participant is discontinued from the study if assigned a diagnosis of dementia in two consecutive years; for the purpose of these analyses, this diagnosis is carried forward if the patient is still alive at the time that a future visit would have occurred.

Factors included in the integrative model

Our integrative classifying model utilizes information from UPSIT, GRS, age, gender and family history to estimate risk. None of the factors included in our integrative classifying model are used in the clinical diagnosis of PD. Also, none of the cohorts assessed by this model were responsible for the discovery of the factors included in the classifying model. In addition, no factors in the integrative classifying model were used as part of recruitment for any cohorts except for one study (discussed below), which was used as a positive control in validating our

model. This allows us to avoid circularity, overestimation, and overfitting of the classification model. Please see Table 1 for descriptive statistics of the participating studies.

The UPSIT is a commercially available test that uses smell identification to test the function of an individual's sense of smell [5]. It is the gold standard of smell identification tests for its reliability and practicality. Smell dysfunction is known to occur in several neurodegenerative disorders, including in PD, and has been suggested as a potential biomarker [6,7]. Because of the rich clinical data available in the studies participating in this project, we were able to include UPSIT data in our analysis.

GRS were calculated by summing the risk allele counts for the 28 common risk loci identified and replicated in the most recent large-scale meta-analysis of PD genome-wide association study (GWAS) data, as well as including two additional relatively rare risk variants detected within PPMI known to be associated with PD (p.N370S in *GBA* and p.G2019S in *LRRK2*) [8,9,10,11,12,13]. In PPMI, we see expected frequencies of G2019S and N370S variants in PD cases (1.3% and 1.9% respectively) and even find one N370S in controls, with the inclusion of these variants improving the AUC of the GRS by ~1% over previous efforts only focusing on the 28 independent common risk variants [unpublished data]. Prior to summing the risk allele counts, all allele counts per variant were scaled by their log odds ratios. The effect estimates for the 28 common variants were extracted from Nalls et al., 2014, and odds ratios of rare alleles at *GBA* p.N370S (3.33) and *LRRK2* p.G2019S (9.620) were taken from the PDgene database and 23andMe [www.pdgene.org and www.23andme.com][14]. In all cohorts except 23andMe, after the scaled risk allele counts were summed and divided by the number of loci, they were transformed into Z scores using the healthy controls in PPMI as a reference. This aids in communicating effect estimates, with Z corresponding to a single standard deviation from the

[5] Doty, Richard L, Richard E Frye, and Udayan Agrawal. "Internal consistency reliability of the fractionated and whole University of Pennsylvania Smell Identification Test." *Perception & Psychophysics* 45.5 (1989): 381-384.

[6] Lazarini, Françoise et al. "Adult Neurogenesis Restores Dopaminergic Neuronal Loss in the Olfactory Bulb." *The Journal of Neuroscience* 34.43 (2014): 14430-14442.

[7] Michell, AW et al. "Biomarkers and Parkinson's disease." *Brain* 127.8 (2004): 1693-1705.

[8] Nalls, Mike A et al. "Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease." *Nature genetics* (2014).

[9] International Parkinson's Disease Genomics Consortium (IPDGC), and Wellcome Trust Case Control Consortium 2 (WTCCC2). "A Two-Stage Meta-Analysis." Greg Gibson. (2011).

[10] Do, Chuong B et al. "Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease." *PLoS genetics* 7.6 (2011): e1002141.

[11] International Parkinson Disease Genomics Consortium. "Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies." *The Lancet* 377.9766 (2011): 641-649.

[12] Sidransky, Ellen et al. "Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease." *New England Journal of Medicine* 361.17 (2009): 1651-1661.

[13] Paisán-Ruíz, Coro et al. "Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease." *Neuron* 44.4 (2004): 595-600.

[14] Lill, Christina M et al. "Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database." *PLoS genetics* 8.3 (2012): e1002548.

control mean genetic risk of PD. This method for calculating risk scores mirrors that in the software package PLINK [http://pngu.mgh.harvard.edu/~purcell/plink/profile.shtml][15]. Due to the slightly different study design and genotyping in the 23andMe cohort, imputed dosages of risk alleles were summed and divided by the number of loci and then transformed into Z scores using a reference set of 334,839 unrelated European individuals who self-reported as not having PD. Information regarding variants and effect estimates used in generating the GRS is included in the underlying data available for download by interested researchers. As a note, the GRS was used instead of single SNP estimates to improve power, as these associations were initially discovered in over 100,000 samples and effect estimates of single variants would not be accurate in cohorts of less than 1,000 samples as in this study.

Of the 40 total 23andMe samples used here, 31 were used to discover the loci comprising the GRS. However, the influence of these overlapping samples is trivial due to the immense size of the GWAS discovery efforts. Although LABS-PD was used in the replication phase of analyses for some loci, these comprise a negligible amount of the sample series contributing to the previous GWAS. Of note, the PPMI, PARS, Penn-Udall and PDBP cohorts did not have genetic data available at the time of this most recent GWAS effort used to identify and replicate the loci comprising the score and therefore were not included in that effort either.

Gender, age (at onset for cases and at last exam for controls) as well as family history were all self-reported information. To clarify, we regard family history as self-report of a first or second degree relative with a diagnosis of PD. When applicable, medical records were used to corroborate this information.

Genetic data

Genotypic data for all studies except Penn-Udall and 23andMe were generated at the National Institute on Aging's Laboratory of Neurogenetics using the NeuroX genotyping array available from Illumina Inc. Penn-Udall genotypic data was generated using the NeuroX array at the Center for Applied Genomics in Philadelphia, PA. For a detailed description of the NeuroX array, its content, and genotype calling methods, please see previously published work [16]. Quality control methods for the NeuroX genotyped samples are described in detail elsewhere [17]. In brief, all NeuroX genotyped samples met the following inclusion criteria: per variant and per sample missingness < 5%; concordance between self-reported gender and genetically ascertained gender; no first or second degree relatives within each dataset from self-report and

[15] Purcell, Shaun et al. "PLINK: a tool set for whole-genome association and population-based linkage analyses." *The American Journal of Human Genetics* 81.3 (2007): 559-575.
[16] Nalls, Mike A et al. "NeuroX, a fast and efficient genotyping platform for investigation of neurodegenerative diseases." *Neurobiology of aging* (2014).
[17] Nalls, Mike A et al. "Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease." *Nature genetics* (2014).

genetically ascertained relationships based on common polymorphisms; and European ancestry from self-report and genetic confirmation when compared to known reference populations [18].

Genotypic data for the 23andMe cohort were generated by National Genetics Institute (NGI), a Clinical Laboratory Improvements Amendments (CLIA) certified clinical laboratory and subsidiary of Laboratory Corporation of America. Samples were genotyped on either the Illumina HumanHap550+ BeadChip platform (n=10) or the Illumina HumanOmniExpress+ BeadChip platform (n=30), as described previously [19]. Every sample that did not reach a 98.5% call rate for SNPs on the standard platforms was reanalyzed. Individuals whose analyses repeatedly failed were contacted by 23andMe customer service to provide additional samples.

Model generation

We selected known non-invasive risk factors for PD to train our classification framework in the PPMI cohort. In this model, we selected three time-independent factors: family history of PD from self-report in 1st or 2nd degree relatives, female gender, and a GRS; we also selected two time-dependent factors: age (onset in cases and most recent exam in controls) and total UPSIT score. The UPSIT is scored by summing all correctly-identified odors out of the 40 items within the UPSIT test booklets. The odors are culturally adjusted for familiarity. All five factors were entered into a logistic regression model to generate estimates of case probability used to classify samples in PPMI as cases or controls at baseline. In PPMI we regard baseline as the first clinic visit for controls at enrollment, or the initial exam at recruitment that is roughly concurrent with that participant's PD diagnosis. Receiver operator characteristic (ROC) curves were used to quantify accuracy of the classification within the cohort. This probabilistic model was then applied to all other studies after training on PPMI. As a note, this model was trained on PPMI excluding known SWEDD samples. All parameters in the integrative classifying model contribute to the overall information content of the model based on Akaike information criteria (AIC) and surviving backwards and forwards stepwise modeling in PPMI[20]. We term this more complex model the "integrative model" in this report. Additional classifying models outside of the integrative model described above were created to estimate the accuracy of using only UPSIT, only the GRS, and only demographic factors (i.e. family history, age and female gender). Only AIC was used for model pruning and due to power concerns in small sample sizes, interactions were not incorporated into model generation. Standardized beta-coefficients were generated within the integrative model to compare the overall effect sizes of factors in the integrative model using PPMI data.

Internal validation within PPMI

[18] 1000 Genomes Project Consortium. "An integrated map of genetic variation from 1,092 human genomes." *Nature* 491.7422 (2012): 56-65.

[19] Hinds, David A et al. "A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci." *Nature genetics* 45.8 (2013): 907-911.

[20] Akaike, H. "A new look at the statistical model identification." 1974.
<http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1100705>

After initial specification of the predictive model in PPMI, we used resampling to validate the model within the cohort *in silico*. We resampled the PPMI PD and control samples over 10,000 iterations, generating parameters (beta-coefficients) on a randomly assigned training subset and fitting the five predictors of the integrative model to a randomly assigned validation subset to calculate AUCs. In this analysis, cases and controls were equally split at random between training and test datasets per iteration. In addition, we also modified this workflow to run a backwards stepwise pruning of the integrative model on the training subset, then fit this model (which uses 2-5 parameters based on the subset in question) to the validation subset and calculated AUCs for an additional 10,000 iterations. In this phase of analysis, subsets were approximately equal in size and partitioned for each iteration of the resampling using a random number generator to assign a sample to either the training or validation subset. For each iteration, the AUC was calculated based on training parameters derived from a randomly-generated subset and fitted to the corresponding validation subset, with no sample overlap with the training set.

Calibration of the model within PPMI

To evaluate the calibration of the model in PPMI, we used the Hosmer-Lemeshow test[21]. The Hosmer-Lemeshow test itself has a weakness of erratic results across small sampling groups so we used a variety of sampling scenarios within PPMI. We first evaluated calibration by partitioning the data into 5, 10, 25, 50, and 100 groups and then running the calibration test. Next, we repeated tests for all possible values between 5-100 groups and evaluated the distribution of the test statistics.

Additional external validation

For all studies with available data, we fit the parameter estimates trained on the entire PPMI dataset to evaluate the applicability of the integrative model's classification algorithm. We also compared the accuracy of the integrative model in the SWEDD subset of PPMI, which used controls shared with the training set, but was not used to do any additional training of the algorithm. Because the AUC could not be calculated for case-only studies using ROC (i.e. the LABS-PD and Penn-Udall studies), we quantified accuracy by the proportion of PD case classifications using a optimal prediction threshold derived from the best classification available in the training set derived from the ROC curve. Best prediction thresholds maximizing combined sensitivity and specificity were: 0.675 for the demographic model, 0.574 for the UPSIT model, 0.639 for the GRS model, and 0.655 for the integrative model.

Software note

---

[21] Lemeshow, Stanley, and David W Hosmer. "A review of goodness of fit statistics for use in the development of logistic regression models." *American journal of epidemiology* 115.1 (1982): 92-106.

Details on the generation of raw genotypes from the NeuroX array can be found elsewhere including the manual clustering of PD risk loci and the general Illumina-based pipeline. PLINK was used for management of raw genotype data. All downstream statistical analyses outside of the 23andMe dataset were carried out using R 3.0.2 on Ubuntu Linux 14.04 [22],[23]. R packages used include ggplot2, pROC, QuantPsyc, ResourceSelection, and scales [24,25,26,27,28].

CONSORTIA MEMBERS AND AFFILIATIONS

Parkinson's Disease Biomarkers Program (PDBP)

Dubois Bowman, Ph.D. Emory University (Atlanta, Georgia) Alice Chen-Plotkin, M.D. University of Pennsylvania (Philadelphia, Pennsylvania) Ted Dawson, M.D., Ph.D. Johns Hopkins University (Baltimore, Maryland) Richard Dewey, M.D. UT Southwestern Medical Center (Dallas, Texas) Dwight Charles German, Ph.D. UT Southwestern Medical Center (Dallas, Texas) Xuemei Huang, M.D., Ph.D. Pennsylvania State University Hershey Medical Center (Hershey, Pennsylvania) Vladislav Petyuk, Ph.D. Battelle Pacific Northwest Laboratories, (Richmond, Washington) Clemens Scherzer, M.D. The Brigham and Women's Hospital Inc., Harvard Medical School (Boston, Massachusetts) David Vaillancourt, Ph.D. University of Florida, (Gainesville, Florida) Andrew West, Ph.D. University of Alabama, (Birmingham, Alabama) Jing Zhang, M.D., Ph.D. University of Washington (Seattle, Washington)

Parkinson's Progression Marker Initiative (PPMI)

Steering Committee:
Kenneth Marek, MD[1] (Principal Investigator); Danna Jennings, MD[1] (Site Investigator, Olfactory Core, PI); Shirley Lasch, MBA[1] ; Caroline Tanner, MD, PhD[9] (Site Investigator);Tanya Simuni, MD[3] (Site Investigator); Christopher Coffey, PhD[4] (Statistics Core, PI); Karl Kieburtz, MD, MPH[5] (Clinical Core, PI); Renee Wilson, MA[5] ; Werner Poewe, MD[7] (Site Investigator); Brit Mollenhauer, MD[8] (Site Investigator); Douglas Galasko, MD[28]; Tatiana Foroud, PhD [16] (Genetics Coordination Core, BioRepository PI); Todd Sherer, PhD[6] ; Sohini Chowdhury[6]; Mark Frasier, PhD[6]; Catherine Kopil, PhD[6]; Vanessa Arnedo[6]

Study Cores:
Clinical Coordination Core: Cynthia Casaceli, MBA[5]
Imaging Core: John Seibyl, MD [1]; Susan Mendick, MPH[1]; Norbert Schuff, PhD[9]
Statistics Core: Christopher Coffey, PhD[4] ; Chelsea Caspell [4]; Liz Uribe[4]; Eric Foster [4]; Katherine Gloer PhD[4]; Jon Yankey MS[4]
Bioinformatics Core: Arthur Toga, PhD[10] (Principal Investigator), Karen Crawford, MLIS[10]
BioRepository: , Danielle Elise Smith[16] ; Paola Casalin[12]; Giulia Malferrari[12]
Bioanalytics Core: John Trojanowski, MD, PhD[13] (Principal Investigator); Les Shaw, PhD[13] (Co-Principal Investigator)
Genetics Core: Andrew Singleton, PhD[14] (Principal Investigator)
Genetics Coordination Core: Cheryl Halter [16]

Site Investigators:
David Russell, MD, PhD[1] Site; Stewart Factor, DO[17]; Penelope Hogarth, MD[18]; David Standaert, MD, PhD[19]; Robert Hauser, MD, MBA[20]; Joseph Jankovic, MD[21]; Matthew Stern, MD[13]; Lama Chahine, MD[13]; Shu-Ching HU, MD PhD[22]; Samuel Frank, MD[23]; Claudia Trenkwalder, MD[8],; Wolfgang Oertel MD[35]; Irene Richard, MD[24]; Klaus Seppi, MD[7]; Eva Reiter, MD[7]; Holly Shill, MD [25]; Hubert Fernandez, MD [26]; Anwar Ahmed, MD[26]; Daniela Berg, MD [27]; Isabel Wurster MD[27]; Zoltan Mari, MD[29]; David Brooks, MD[30]; Nicola Pavese, MD[30]; Paolo Barone, MD, PhD[31]; Stuart Isaacson, MD[32]; Alberto Espay, MD, MSc [33]; Dominic Rowe, MD, PhD[34]; Melanie Brandabur MD[2]; James Tetrud MD[2]; Grace Liang MD[10]Karen Marder[36]; Jean-Christophe Corvol[37]; Jose Felix Martí Masso[38]; Eduardo Tolosa[39]; Jan O. Aasly[40]; Nir Giladi[41]; Leonidas Stefanis[42];

[22] "The R Project for Statistical Computing." 20 Feb. 2015 <http://www.r-project.org/>
[23] "Ubuntu Wiki: Home." 2005. 20 Feb. 2015 <https://wiki.ubuntu.com/>
[24] "ggplot2." 2012. 20 Feb. 2015 <http://ggplot2.org/>
[25] "CRAN - Package pROC." 2010. 20 Feb. 2015 <http://cran.r-project.org/package=pROC>
[26] "CRAN - Package QuantPsyc." 2009. 20 Feb. 2015 <http://cran.r-project.org/package=QuantPsyc>
[27] "CRAN - Package ResourceSelection." 2011. 20 Feb. 2015 <http://cran.r-project.org/package=ResourceSelection>
[28] "CRAN - Package scales." 2011. 20 Feb. 2015 <http://cran.r-project.org/package=scales>

Coordinators:

Laura Leary[1]; Cheryl Riordan[1]; Linda Rees, MPH[1]; Barbara Sommerfeld, RN, MSN[17]; Cathy Wood-Siverio, MS[17]; Alicia Portillo[18]; Art Lenahan[18]; Karen Williams[3]; Stephanie Guthrie, MSN[19]; Ashlee Rawlins[19]; Sherry Harlan[20]; Christine Hunter, RN[21]; Baochan Tran[13]; Abigail Darin[13]; Carly Linder[13]; Gretchen Todd[22]; Cathi-Ann Thomas, RN, MS[23]; Raymond James, RN[23]; Cheryl Deeley, MSN[24]; Courtney Bishop BS[24]; Fabienne Sprenger, MD[7]; Diana Willeke[8]; Sanja Obradov[25]; Jennifer Mule[26]; Nancy Monahan[26]; Katharina Gauss[27]; Kathleen Comyns[9] Deborah Fontaine, BSN, MS, RN, GNP, MS [28]; Christina Gigliotti[28]; Arita McCoy[29]; Becky Dunlop[29]; Bina Shah, BSc[30]; Susan Ainscough[31]; Angela James[32]; Rebecca Silverstein[32]; Kristy Espay[33]; Madelaine Ranola[34]; Helen M. Santana[36]; Nelly Ngono[37]; Elisabet Rezola[38]; Delores Vilas Rolan[39]; Bjorg Waro[40]; Anat Mirlman[41]; Maria Stamelou[42];

ISAB (Industry Scientific Advisory Board):

Thomas Comery, PhD[43]; Spyros Papapetropoulos, MD, PhD[43]; Bernard Ravina, MD, MSCE[44]; Igor D. Grachev, MD, PhD[45]; Jordan S. Dubow, MD[46]; Michael Ahlijanian, PhD[47]; Holly Soares, PhD[47]; Suzanne Ostrowizki, MD, PhD[48]; Paulo Fontoura, MD, PhD[48]; Alison Chalker, PhD[49]; David L. Hewitt, MD[49]; Marcel van der Brug, PhD[50]; Alastair D. Reith, PhD[51]; Peggy Taylor, ScD[52]; Jan Egebjerg, PhD[53]; Mark Minton, MD[53]; Andrew Siderowf, MD, MSCE[54]; Pierandrea Muglia, PhD[55]; Robert Umek, PhD[56]; Ana Catafau, MD, PhD[57]

1 Institute for Neurodegenerative Disorders, New Haven, CT
2 The Parkinson's Institute, Sunnyvale, CA
3 Northwestern University, Chicago, IL
4 University of Iowa, Iowa City, IA
5 Clinical Trials Coordination Center, University of Rochester, Rochester, NY
6 The Michael J. Fox Foundation for Parkinson's Research, New York, NY
7 Innsbruck Medical University, Innsbruck, Austria
8 Paracelsus-Elena Klinik, Kassel, Germany
9 University of California, San Francisco, CA
10 Laboratory of Neuroimaging (LONI), University of Southern California
11 Coriell Institute for Medical Research, Camden, NJ
12 BioRep, Milan, Italy
13 University of Pennsylvania, Philadelphia, PA
14 National Institute on Aging, NIH, Bethesda, MD
16 Indiana University, Indianapolis, IN
17 Emory University of Medicine, Atlanta, GA
18 Oregon Health and Science University, Portland, OR
19 University of Alabama at Birmingham, Birmingham, AL
20 University of South Florida, Tampa, FL
21 Baylor College of Medicine, Houston, TX
22 University of Washington, Seattle, WA
23 Boston University, Boston, MA
24 University of Rochester, Rochester, NY
25 Banner Research Institute, Sun City, AZ
26 Cleveland Clinic, Cleveland, OH
27 University of Tuebingen, Tuebingen, Germany
28 University of California, San Diego, CA
29 Johns Hopkins University, Baltimore, MD
30 Imperial College of London, London, UK
31 University of Salerno, Salerno, Italy
32 Parkinson's Disease and Movement Disorders Center, Boca Raton, FL
33 University of Cincinnati, Cincinnati, OH
34 Macquarie University, Sydney Australia
35 Philipps University Marburg, Germany
36 Columbia Medical, New York, NY
37 Pitié-Salpêtrière Hospital, Paris France
38 University of Donostia-Service of Neurology Hospital, San Sebastian, Spain
39 University of Barcelona-Hospital Clinic of Barcelona, Barcelona, Spain
40 Norwegian University of Science and Technology, Trondheim, Norway
41 Tel Aviv Sourasky Medical Center, Tel Aviv, Isreal
42 Foundation for Biomedical research of the Academy of Athens, Athens, Greece
43 Pfizer, Inc., Groton, CT

44 Biogen Idec, Cambridge, MA
45 GE Healthcare, Princeton, NJ
46 AbbVie, Abbot Park, IL
47 Bristol-Myers Squibb Company
48 F.Hoffmann La-Roche, Basel, Switzerland
49 Merck & Co., North Wales, PA
50 Genentech, Inc., South San Francisco, CA
51 GlaxoSmithKline, Stevenage, United Kingdom
52 Covance, Dedham, MA
53 H. Lundbeck A/S
54 Avid Radiopharmaceuticals, Philadelphia , PA
55 UCB Pharma S.A., Brussels, Belgium
56 Meso Scale Discovery
57 Piramal Life Sciences, Berlin, Germany

Supporting Table 1: Details of the stepwise regression. Logistic regression was used as the basis for the integrative predictive model trained on the PPMI dataset. Parameter estimates are sorted in descending order of Akaike information criterion.

Parameter	Beta	SE	Z	P-value	AIC	Standardized Beta	Percentage of variance explained by parameter
UPSIT	-0.304	0.030	-10.077	< 2.00E-16	603.27	-6.039	63.149
GRS	0.514	0.134	3.847	1.20E-04	371.55	1.303	13.626
Family history	1.283	0.471	2.723	0.006	362.41	1.088	11.379
Female	0.561	0.287	1.953	0.051	357.92	0.571	5.972
Age	-0.026	0.014	-1.923	0.054	357.79	-0.562	5.875

Supporting Table 2: Variants included in the GRS calculations as they appear in the PPMI and PDBP NeuroX genotype datasets.

Variant Name	NeuroX_ID	Other Name	Data Dictionary Entry
rs114138760	NeuroX_dbSNP_rs114138760		rs114138760 C/G (FWD) G:Ancestral C:Minor
rs76763715	exm106217	*GBA p.N370S*	rs76763715 C/T (FWD) T:Ancestral C:Minor
rs71628662	NeuroX_rs71628662		rs71628662 C/T (FWD) T:Ancestral C:Minor
rs823118	NeuroX_rs823118		rs823118 C/T (FWD) C:Ancestral T:Minor
rs10797576	NeuroX_rs10797576		rs10797576 C/T (FWD) C:Ancestral T:Minor
rs6430538	NeuroX_rs6430538		rs6430538 C/T (FWD) T:Ancestral C:Minor
rs1955337	NeuroX_rs1955337		rs1955337 G/T (FWD) G:Ancestral T:Minor
rs12637471	NeuroX_rs12637471		rs12637471 A/G (FWD) G:Ancestral A:Minor
rs34884217	NeuroX_rs34884217		rs34884217 (G/T) REV T:Ancestral C:Minor

rs34311866	NeuroX_rs34311866		rs34311866 A/G (REV) A:Ancestral C:Minor
rs11724635	NeuroX_rs11724635		rs11724635 A/C (FWD) A:Ancestral A:Minor
rs6812193	exm-rs6812193		rs6812193 C/T (FWD) C:Ancestral T:Minor
rs356181	NeuroX_rs356181		rs356181 C/T (REV) T:Ancestral A:Minor
rs3910105	NeuroX_rs3910105		rs3910105 C/T (REV) T:Ancestral G:Minor
rs8192591	exm535099		rs8192591 A/G (REV) G:Ancestral T:Minor
rs115462410	NeuroX_dbSNP_rs115462410		rs9275326 (was rs115462410) C/T (FWD) C:Ancestral T:Minor
rs199347	NeuroX_rs199347		rs199347 C/T (REV) C:Ancestral G:Minor
rs591323	NeuroX_rs591323		rs591323 A/G (FWD) G:Ancestral A:Minor
rs118117788	NeuroX_dbSNP_rs118117788		rs118117788 C/T (FWD) C:Ancestral T:Minor
rs329648	NeuroX_rs329648		rs329648 C/T (FWD) T:Ancestral T:Minor
rs76904798	NeuroX_rs76904798		rs76904798 C/T (FWD) T:Ancestral T:Minor
rs34637584	exm994671	*LRRK2 p.G2019S*	rs34637584 A/G (FWD) G:Ancestral A:Minor
rs11060180	NeuroX_rs11060180		rs11060180 A/G (FWD) A:Ancestral G:Minor
rs11158026	NeuroX_rs11158026		rs11158026 C/T (FWD) T:Ancestral T:Minor
rs2414739	NeuroX_rs2414739		rs2414739 A/G (FWD) G:Ancestral G:Minor
rs14235	NeuroX_dbSNP_rs14235_replciate_1		rs14235 A/G (FWD) G:Ancestral A:Minor
rs11868035	exm-rs11868035		rs11868035 A/G (FWD) G:Ancestral A:Minor
rs17649553	NeuroX_rs17649553		rs17649553 C/T (FWD) T:Ancestral T:Minor
rs12456492	NeuroX_rs12456492		rs12456492 A/G (FWD) G:Ancestral G:Minor
rs55785911	NeuroX_rs55785911		rs55785911 A/G (FWD) G:Ancestral A:Minor