Title
Diagnosis of Parkinson's disease on the basis of clinical and genetic classification: a population-based modelling study

Journal
The Lancet Neurology, 14(10)

ISSN
1474-4422

Authors
Nalls, Mike A
McLean, Cory Y
Rick, Jacqueline
et al.

Publication Date
2015-10-01

DOI
10.1016/s1474-4422(15)00178-7

Peer reviewed

Diagnosis of Parkinson's disease on the basis of clinical and genetic classification: a population-based modelling study

*Mike A Nalls, Cory Y McLean, Jacqueline Rick, Shirley Eberly, Samantha J Hutten, Katrina Gwinn, Margaret Sutherland, Maria Martinez, Peter Heutink, Nigel M Williams, John Hardy, Thomas Gasser, Alexis Brice, T Ryan Price, Aude Nicolas, Margaux F Keller, Cliona Molony, J Raphael Gibbs, Alice Chen-Plotkin, Eunran Suh, Christopher Letson, Massimo S Fiandaca, Mark Mapstone, Howard J Federoff, Alastair J Noyce, Huw Morris, Vivianna M Van Deerlin, Daniel Weintraub, Cyrus Zabetian, Dena G Hernandez, Suzanne Lesage, Meghan Mullins, Emily Drabant Conley, Carrie A M Northover, Mark Frasier, Ken Marek, Aaron G Day-Williams, David J Stone, John P A Ioannidis, Andrew B Singleton, for the Parkinson's Disease Biomarkers Program and Parkinson's Progression Marker Initiative investigators**

Summary

Background **Accurate diagnosis and early detection of complex diseases, such as Parkinson's disease, has the potential to be of great benefit for researchers and clinical practice. We aimed to create a non-invasive, accurate classification model for the diagnosis of Parkinson's disease, which could serve as a basis for future disease prediction studies in longitudinal cohorts.**

Methods **We developed a model for disease classification using data from the Parkinson's Progression Marker Initiative (PPMI) study for 367 patients with Parkinson's disease and phenotypically typical imaging data and 165 controls without neurological disease. Olfactory function, genetic risk, family history of Parkinson's disease, age, and gender were algorithmically selected by stepwise logistic regression as significant contributors to our classifying model. We then tested the model with data from 825 patients with Parkinson's disease and 261 controls from five independent cohorts with varying recruitment strategies and designs: the Parkinson's Disease Biomarkers Program (PDBP), the Parkinson's Associated Risk Study (PARS), 23andMe, the Longitudinal and Biomarker Study in PD (LABS-PD), and the Morris K Udall Parkinson's Disease Research Center of Excellence cohort (Penn-Udall). Additionally, we used our model to investigate patients who had imaging scans without evidence of dopaminergic deficit (SWEDD).**

Findings **In the population from PPMI, our initial model correctly distinguished patients with Parkinson's disease from controls at an area under the curve (AUC) of 0·923 (95% CI 0·900–0·946) with high sensitivity (0·834, 95% CI 0·711–0·883) and specificity (0·903, 95% CI 0·824–0·946) at its optimum AUC threshold (0·655). All Hosmer-Lemeshow simulations suggested that when parsed into random subgroups, the subgroup data matched that of the overall cohort. External validation showed good classification of Parkinson's disease, with AUCs of 0·894 (95% CI 0·867–0·921) in the PDBP cohort, 0·998 (0·992–1·000) in PARS, 0·955 (no 95% CI available) in 23andMe, 0·929 (0·896–0·962) in LABS-PD, and 0·939 (0·891–0·986) in the Penn-Udall cohort. Four of 17 SWEDD participants who our model classified as having Parkinson's disease converted to Parkinson's disease within 1 year, whereas only one of 38 SWEDD participants who were not classified as having Parkinson's disease underwent conversion (test of proportions, p=0·003).**

Interpretation **Our model provides a potential new approach to distinguish participants with Parkinson's disease from controls. If the model can also identify individuals with prodromal or preclinical Parkinson's disease in prospective cohorts, it could facilitate identification of biomarkers and interventions.**

Funding **National Institute on Aging, National Institute of Neurological Disorders and Stroke, and the Michael J Fox Foundation.**

Introduction

Accurate diagnosis or prediction of risk by use of simple, non-invasive measures is a rarely realised goal for many complex diseases. For complex progressive diseases such as Parkinson's disease, preclinical diagnosis and low error rates in diagnosis are crucial in clinical trials and the study of disease-altering therapeutic approaches.

Imaging is often deemed the gold standard for identification of typical Parkinson's disease pre-mortem, however, high cost and restricted portability limit the use of this approach. We aimed to develop a portable method to identify patients with Parkinson's disease who show aetiologically typical disease presentation (confirmed by dopamine transporter [DAT] imaging data). We used a combination of factors that vary over the life of an individual, factors that are constant and do not change with time, general indicators of neurodegeneration, and Parkinson's disease-specific measures to create our classification algorithm.

Methods

Study design and participants

Figure 1 shows a summary of our workflow. Table 1 describes the cohorts we used and further details are available in the appendix: the Parkinson's Progression Marker Initiative (PPMI), the Parkinson's Disease

Lancet Neurol 2015

Published Online
August 11, 2015
http://dx.doi.org/10.1016/S1474-4422(15)00178-7

See Online/Articles
http://dx.doi.org/10.1016/S1474-4422(15)00192-1

*See appendix for a complete list of Parkinson's Disease Biomarkers Program and Parkinson's Progression Marker Initiative investigators and their affiliations

Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA (M A Nalls PhD, T R Price MS, A Nicolas PhD, J R Gibbs MS, C Letson BS, D G Hernandez MS, A B Singleton PhD); **23andMe, Mountain View, CA, USA** (C Y McLean PhD, M Mullins BA, E Drabant Conley PhD, C A M Northover PhD); **Department of Neurology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA** (J Rick PhD, A Chen-Plotkin MD, D Weintraub MD); **Deptartment of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, NY, USA** (S Eberly MS); **The Michael J Fox Foundation for Parkinson's Research, New York, NY 10018, USA** (S J Hutten PhD, M Frasier PhD); **National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA** (K Gwinn MD, M Sutherland PhD); **INSERM, UMR 1043, Centre de Physiopathologie de Toulouse-Purpan, Toulouse, France** (M Martinez PhD); **Paul Sabatier University, Toulouse, France** (M Martinez); **Genome Biology of Neurodegenerative Diseases,**

German Center for Neurodegenerative Diseases, Tübingen, Germany (P Heutink PhD); Institute of Psychological Medicine and Clinical Neurosciences, MRCCentre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK (N M Williams PhD); Reta Lila Weston Institute, University College London Institute of Neurology, London, UK (J Hardy PhD, A J Noyce MRCP, H Morris MD); Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany (T Gasser MD); Sorbonne Université, UPMC Univ Paris 06, UM 1127, ICM, Paris, France (A Brice MD, S Lesage PhD); INSERM, U 1127, and CNRS, UMR 7225, Institut du Cerveau et de la Moelle Epinière, Paris, France (A Brice, S Lesage); AP-HP, Hôpital de la Salpêtrière, Département de Génétique et Cytogénétique, Paris, France (A Brice, S Lesage); Genetics and Pharmacogenomics, Merck Research Laboratories, West Point, PA, USA (M F Keller PhD, C Molony PhD, D J Stone PhD); Merck Research Laboratories, Boston, MA, USA (M F Keller, C Molony, D J Stone); Center for Neurodegenerative Disease Research, Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA (E Suh PhD, V M Van Deerlin MD); Department of Neurology, Georgetown University Medical Center, Washington, DC, USA (M S Fiandaca MD, H J Federoff MD); Department of Neurology, University of Rochester Medical Center, Rochester, NY, USA (M Mapstone PhD); Department of Psychiatry, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA (D Weintraub); Department of Neurology, Division of Neurogenetics, VA Puget Sound Health Care System, Seattle, WA, USA (C Zabetian MD); Institute for Neurodegenerative Disorders, New Haven, CT, USA (K Marek MD); Precision Medicine, Biogen, Cambridge, MA, USA (A G Day-Williams PhD); Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford,

Research in context

Evidence before this study

We searched PubMed for articles up to Jan 1, 2015, containing possible combinations of the terms "Parkinson's disease", "neurodegeneration", "biomarker", and "risk prediction". Previous studies have tried to identify accurate biomarkers of Parkinson's disease to enable risk quantification and classification outside or before entering a clinic for traditional symptomatic diagnosis. To build our model, we used clinically derived data from several sources within the Parkinson's Progression Marker Initiative (PPMI), such as olfactory function, imaging, genetic risk estimates, and demographics.

Added value of this study

Our model uses an algorithm based on data that are cheap to collect and can be remotely administered. This algorithm is also accurate for classification of cases based on datapoints outside of both study recruitment and Parkinson's disease diagnostic criteria (area under the curve >0·89 in all independent datasets, except in one study used as a positive control). Additionally, this study has suggested that participants classed as having scans without evidence of dopaminergic deficit (SWEDD) who would later be diagnosed with Parkinson's disease might be distinguished from those who would not develop Parkinson's disease. If these findings are confirmed in independent cohorts, they might be useful to researchers in the clinical trial setting, especially when combined with imaging data. Our model was able to discriminate patients without evidence of dopaminergic deficit typical of Parkinson's disease from those patients with aetiologically typical disease.

Implications of all available evidence

The development and primary validation of this classification algorithm using publicly available data shows the usefulness of public datasets such as PPMI and the Parkinson's Disease Biomarkers Program (PDBP). Within these two large case-control cohorts, our model significantly outperforms any single classifier. As the pace of Parkinson's disease genomics advances with added precision from sequencing studies, the genetic contribution to risk prediction is expected to grow rapidly, and we hope this study can serve as a foundation. We show some success in predicting which of the patients who present with SWEDD will progress to typical Parkinson's disease with evidence for dopaminergic deficit. Easy identification of this group will probably be important in clinical trials.

Biomarkers Program (PDBP), the Parkinson's Associated Risk Study (PARS), 23andMe, the Longitudinal and Biomarker Study in PD (LABS-PD), and the Morris K Udall Parkinson's Disease Research Center of Excellence cohort (Penn-Udall). PPMI and PDBP are case-control studies that use a shared set of common data elements and publicly available data to help to identify biomarkers for Parkinson's disease. PARS is a study of incident Parkinson's disease cases, at risk participants, and population controls, which focuses on screening by smell tests and other risk factors, and it served as a planned positive control in our study. The 23andMe cohort in our analysis is part of a small study of *LRRK2* mutation carriers that enrolled cases and controls who had *LRRK2* risk variants, additional participants with idiopathic Parkinson's disease, and healthy controls not carrying *LRRK2* disease risk variants. LABS-PD is a case-only study derived from clinical trial participants and is focused on biomarker development. Penn-Udall is a longitudinal cohort of patients with Parkinson's disease, which aimed to develop biomarkers for disease progression. Each contributing study abided by the ethics guidelines set out by their institutional review boards and all participants gave informed consent for inclusion in both their initial cohorts and subsequent studies.

Procedures and statistical analysis

The appendix includes further details of the statistical methods, beginning with model development by stepwise logistic regression in the PPMI dataset, which we used to train the initial model. Before beginning our primary analyses, our collaborators identified more than 100 common risk factors or potential biomarkers for Parkinson's disease that were available in PPMI (with genetic data available either by genome-wide association studies [GWAS] or NeuroX-derived genotyping) through manually auditing these datasets. Factors that did not survive the stepwise logistic regression modelling for use in the integrative model because of low independent information contribution (assessed with the Akaike information criterion) included handedness, self-report of constipation, self-report of sleep disturbances, caffeine intake, and smoking status.

The resulting model retained five factors—testing of sense of smell, self-reported family history of Parkinson's disease (first or second degree relative with Parkinson's disease), age, sex, and a Parkinson's disease-specific genetic risk score (GRS)—to develop an integrative predictive model to discriminate patients with typical Parkinson's disease (clinical diagnosis with evidence of dopaminergic dysfunction) from neurologically normal controls.[1–6] All smell testing used in this analysis was done with the University of Pennsylvania Smell Identification Test (UPSIT). This test is a so-called scratch and sniff exam for scent identification, which scores individuals from 0 to 40, and is considered to be an objective and well validated measurement of olfactory function. The use of this test to quantify olfactory function as an indicator of neurodegeneration has been suggested previously.[7,8] The appendix includes further details about the

generation and analysis of genetic data and the creation of the GRS from 30 genetic risk factors implicated and replicated in one or more studies, in addition to information about the other model parameters that we used.[9–16] None of the factors in our model is part of the general diagnostic criteria for Parkinson's disease or recruitment into PPMI, the dataset used for training and development of the model.

We assessed model calibration by resampling and used the Hosmer-Lemeshow test[17] to assess goodness of fit. We then applied our integrative model, attempting external validation in the five independent cohorts. We also attempted to extend our model to atypical Parkinson's disease by evaluating the performance of the model in subsets of patients who were screened as potentially having Parkinson's disease but who had dopamine transporter imaging scans without evidence of dopaminergic deficit (SWEDD), suggesting possible aetiological differences.

Our study used the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health (Bethesda, MD, USA) and DNA panels, samples and clinical data from the National Institute of Neurological Disorders and Stroke Human Genetics Resource Center DNA and Cell Line Repository, human subjects protocol 2003-077.

Role of the funding source

The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. All authors could access all data and statistical programming code used in this project for the analyses and results generation, except for the small amount of data contributed by 23andMe. Data from 23andMe could be accessed only by their employees (CYM, MM, EDC, and CAMN) as part of their unique consent process, so only summary statistics and area under the curve (AUC) data could be provided for this report. MN takes final responsibility for the decision to submit the paper for publication.

Results

To accompany this report, and to help with replication and extension of our work, the code and training data for this predictive model and some validation data have been made publicly available online.

Each of five factors that we included made significant contributions to the information content of the integrative predictive model. In comparisons of the standardised beta coefficients within the regression model, the UPSIT score was responsible for 63·1% of the explained variance, followed by the genetic risk score (13·6%), family history (11·4%), sex (6·0%), and age (5·9%). The appendix (p 10) shows additional information about parameter estimates for these factors.

For discrimination of participants with Parkinson's disease from healthy controls in the PPMI cohort, the AUC of the integrative model was 0·923 (95% CI 0·900–0·946; table 2, figure 2). Sensitivity was 0·834 (95% CI 0·711–0·883) and specificity was 0·903 (95% CI 0·824–0·946) in PPMI at the best threshold for classification, which was 0·655 in the receiver operating curve (ROC). However, the low prevalence of Parkinson's disease—2% in populations older than 60 years—results in a positive predictive value (PPV) of 0·149 despite an AUC more than 0·9.[18] Although the AUC of the UPSIT-only model in PPMI was individually strong (AUC 0·901, 95% CI 0·874–0·928), the integrative model was significantly more informative based on DeLong's test for correlated ROC curves[19] (Z=3·027, p=0·002). When we used the integrative model to classify SWEDD participants and controls in PPMI, classification accuracy decreased, with an AUC of 0·707 (95% CI 0·630–0·783; table 2).

For in-silico validation of the integrative model in the PPMI dataset, we used 10 000 randomly generated subsets of data to train integrative predictive models specific to the randomly generated subsets to evaluate the distribution of the AUC through resampling. We fitted these models to matched, non-overlapping validation sets for each iteration. We noted a normal distribution of AUCs across all iterations, suggesting good model fit (figure 3). The mean AUC estimate was 0·918 (SD 0·012, range 0·830–0·959).

CA, USA (Prof J P A Ioannidis MD); **Department of Medicine, Stanford Prevention Research Center, Stanford, CA, USA** (Prof J P A Ioannidis); **Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA** (Prof J P A Ioannidis); **and Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, CA, USA** (Prof J P A Ioannidis)

Correspondence to:
Dr Andrew B Singleton, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA
singleta@mail.nih.gov

See Online for appendix

For details of the **Biowulf Linux cluster** see http://biowulf.nih.gov

For **data and code from PPMI and PDBP** see http://www.ppmi-info.org/ and https://pdbp.ninds.nih.gov/

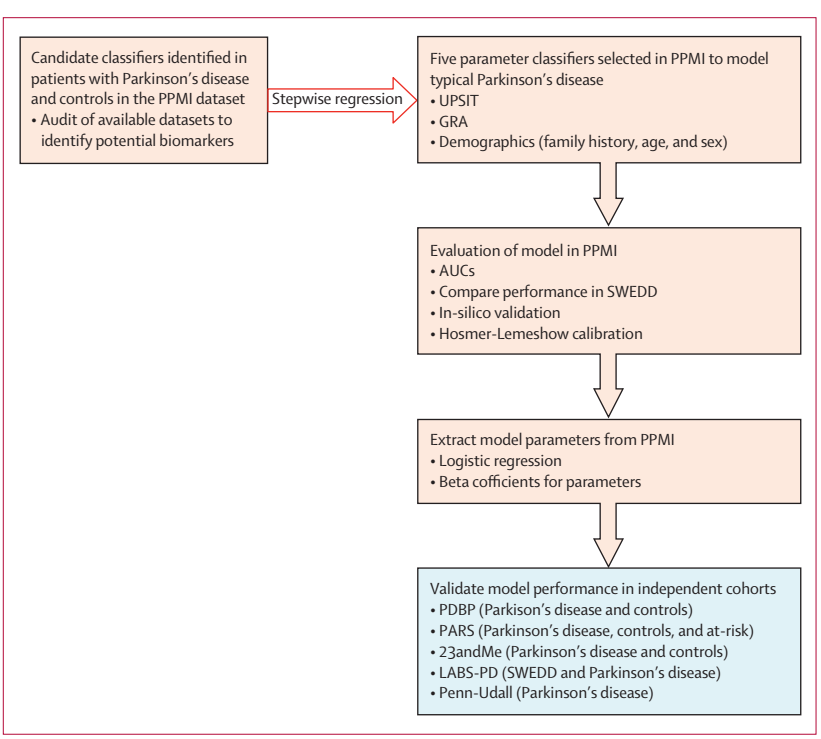

Figure 1: **Profile of model development and validation**
Pink boxes show steps of the workflow specific to the PPMI study and the blue box shows the validation phase. PPMI=Parkinson's Progression Marker Initiative. PDBP=Parkinson's Disease Biomarkers Program. PARS=Parkinson's Associated Risk Study. LABS-PD=Longitudinal and Biomarker Study in PD. Penn-Udall=Morris K Udall Parkinson's Disease Research Center of Excellence cohort. UPSIT=University of Pennsylvania Smell Identificaiton Test. GRS=genetic risk score. SWEDD=scans without evidence of dopaminergic deficit. AUC=area under the curve.

	PPMI (training dataset)			PDBP		PARS			23andMe		LABS-PD		Penn-Udall
	Case	Control	SWEDD	Case	Control	Case	Control	At risk	Case	Control	Case	SWEDD	Case
Participants	367	165	55	453	156	15	85	146	20	20	239	13	98
Female	122 (33%)	55 (33%)	19 (35%)	183 (40%)	78 (50%)	5 (33%)	46 (54%)	76 (52%)	9 (45%)	11 (55%)	84 (35%)	8 (62%)	31 (32%)
Age (years)	64·256 (9·598)	63·794 (10·588)	63·018 (10·117)	64·936 (9·046)	62·301 (10·958)	67·400 (6·909)	63·467 (10·269)	64·290 (8·338)	52·650 (9·959)	60·150 (10·148)	59·127 (9·075)	60·041 (12·927)	63·816 (8·813)
Family history of Parkinson's disease	92 (25%)	9 (5%)	17 (31%)	115 (25%)	17 (11%)	10 (67%)	37 (44%)	64 (44%)	5 (25%)	5 (25%)	64 (27%)	5 (38%)	29 (30%)
Total UPSIT score	22·196 (8·261)	34·345 (4·372)	32·181 (5·754)	19·984 (7·959)	32·590 (5·924)	17·667 (5·827)	35·118 (3·311)	22·897 (6·283)	20·550 (7·067)	33·050 (3·634)	19·422 (7·591)	28·692 (8·420)	18·316 (6·929)
GRS (Z units)	0·595 (1·201)	0 (1)	0·347 (1·065)	0·600 (1·229)	0·094 (1·030)	0·685 (0·912)	0·140 (1·071)	0·263 (1·070)	0·531 (1·471)	–0·159 (0·936)	0·625 (1·114)	0·328 (0·668)	0·398 (1·189)
Diagnostic criteria	UK Brain Bank criteria for Parkinson's disease plus DAT scan showing dopaminergic dysfunction at baseline	No clinical indication of Parkinson's disease	UK Brain Bank criteria for Parkinson's disease plus DAT scan showing normal dopaminergic function at baseline	UK Brain Bank criteria for Parkinson's disease	No clinical indication of Parkinson's disease	UK Brain Bank criteria for Parkinson's disease	No clinical indication of Parkinson's disease	No clinical indication of Parkinson's disease but some degree of hyposmia	Self-report of Parkinson's disease diagnosis	Self-report of neurologically normal status	UK Brain Bank criteria for Parkinson's disease plus DAT scan showing dopaminergic dysfunction at baseline	UK Brain Bank criteria for Parkinson's disease plus DAT scan showing normal dopaminergic function at baseline	UK Brain Bank criteria for Parkinson's disease

Data are n, n (%), or mean (SD). PPMI=Parkinson's Progression Marker Initiative. PDBP=Parkinson's Disease Biomarkers Program. PARS=Parkinson's Associated Risk Study. LABS-PD=Longitudinal and Biomarker Study in PD. Penn-Udall=Morris K Udall Parkinson's Disease Research Center of Excellence cohort. UPSIT=University of Pennsylvania Smell Identification Test. GRS=genetic risk score. SWEDD=scans without evidence of dopaminergic deficit. DAT=dopamine transporter positron emission tomography.

Table 1: **Descriptive statistics by study and status within study**

We did a subsequent resampling exercise, repeating the previous analysis but using backward stepwise pruning of the integrative model that was informed by the Akaike information criterion in the training subsets. We then applied this version of the integrative model to the additional randomly generated validation subsets. In 10 000 iterations, the UPSIT score always remained after stepwise pruning, whereas the GRS remained in 98·6% of the iterations, family history in 89·6%, sex in 49·9%, and age in 49·4%. Of the iterations, 49·1% contained four factors, 29·6% contained three factors, 19·9% contained five factors, 1·4% contained two factors, and only one iteration contained a single factor (the UPSIT score) based on resampling. Across the resampling iterations, the AUC was a mean of 0·915 (SD 0·013, range 0·826–0·960).

In our Hosmer-Lemeshow test to investigate model calibration in the PPMI dataset, we first iterated across possibilities of five, ten, 25, 50, and 100 random subsets within the dataset. Each grouping returned p values between 0·286 and 0·592 for the Hosmer-Lemeshow test, showing that no outlier subgroups were identified and that calibration was good. We also repeated this analysis for all possible numbers of groupings ranging from five to 100. All p values were greater than 0·05, showing that the integrative model does not suffer from any subset of the data disproportionately affecting the results (figure 3).

Our integrative model showed high accuracy (quantified by AUC estimates) in discrimination of patients with Parkinson's disease from healthy controls when applied to additional cross-sectional case-control studies (table 2, figure 4). In PDBP, the AUC of the UPSIT-only model was less than that of the integrative model (table 2): the integrative model was significantly more powerful than the UPSIT-only model when used to discriminate patients with Parkinson's disease from controls in PDBP (Z=2·154, p=0·0313). The AUC was slightly lower for the integrative model than for the UPSIT-only model in the 23andMe cohort, but this decrease was not statistically significant in DeLong's test (p=0·44) and might result from the increased recruitment of patients with *LRRK2* risk variants in this study subset and its small sample size. We used PARS as a positive control because recruitment of patients and at-risk participants to this study included the UPSIT score, therefore biasing estimates and introducing some circularity, and making UPSIT scores more different between patients and controls in this cohort than might otherwise have been the case.

To further validate this integrative model, we attempted to predict Parkinson's disease case status in the Penn-Udall and LABS-PD datasets. The integrative model was able to categorise 93% (222/239) of cases correctly in LABS-PD and was 94% (92/98) correct in the Penn-Udall dataset. Classification accuracy in the Penn-Udall cohort was slightly lower for the integrative model than for the UPSIT-only model, possibly because of the small sample size (table 2).

	PPMI		PDBP	PARS		23andMe	LABS-PD*		Penn-Udall*
	PD	SWEDD	PD	PD	At risk	PD	PD	SWEDD	PD
Demographic model	0·604 (0·555–0·653)	0·609 (0·514–0·705)	0·602 (0·552–0·652)	0·678 (0·521–0·834)	0·479 (0·402–0·556)	0·385	0·276 (0·219–0·333)	0·385 (0·109–0·660)	0·316 (0·224–0·409)
UPSIT model	0·901 (0·874–0·928)	0·624 (0·539–0·708)	0·881 (0·850–0·912)	0·994 (0·981–1·000)	0·976 (0·959–0·992)	0·962	0·925 (0·891–0·958)	0·538 (0·256–0·821)	0·959 (0·920–0·999)
GRS model	0·639 (0·589–0·688)	0·569 (0·481–0·656)	0·619 (0·568–0·670)	0·657 (0·529–0·786)	0·533 (0·454–0·612)	0·620	0·489 (0·426–0·553)	0·231 (0·000–0·469)	0·327 (0·252–0·442)
Integrative model	0·923 (0·900–0·946)	0·707 (0·630–0·783)	0·894 (0·867–0·921)	0·998 (0·992–1·000)	0·962 (0·941–0·984)	0·955	0·929 (0·896–0·962)	0·692 (0·431–0·953)	0·939 (0·891–0·986)

Data are AUC (95% CI), except where indicated otherwise. 95% CIs could not be reported for the 23andMe cohort because of their consent process for publishing data. The demographic model includes estimates for a logistic model containing parameters of female sex, family history, and age. The UPSIT model includes estimates for a logistic model containing only the total UPSIT parameter. The GRS model includes estimates for a logistic model only containing the GRS parameter. The integrative model includes all parameters from the previous three models into the estimate of classification accuracy between cases and controls. For PPMI, PDBP, PARS, and 23andMe, the AUC estimates were generated by comparing Parkinson's disease cases with controls. In PPMI, participants designated as SWEDD were compared with the same controls as Parkinson's disease patients. In PARS, participants designated at risk were compared with the same controls as Parkinson's disease patients. PPMI=Parkinson's Progression Marker Initiative. PDBP=Parkinson's Disease Biomarkers Program. PARS=Parkinson's Associated Risk Study. LABS-PD=Longitudinal and Biomarker Study in PD. Penn-Udall=Morris K Udall Parkinson's Disease Research Center of Excellence cohort. UPSIT=University of Pennsylvania Smell Identificaiton Test. GRS=genetic risk score. SWEDD=participants with scans without evidence of dopaminergic deficit. AUC=area under the curve. *These studies were case-only so instead of AUC the proportion of correctly predicted cases (95% CI) is reported as a measure of classification accuracy.

Table 2: **Performance of classification, by study and disease status within study**

In the PPMI dataset, application of the integrative model to SWEDD participants and the same controls as used in the previous categorisation of patients with Parkinson's disease did not generally result in the SWEDD participants being classified as having Parkinson's disease (table 2). Additionally, our model classified only 69% (9/13) of the SWEDD participants in LABS-PD as having Parkinson's disease, suggesting that many SWEDD participants are aetiologically distinct from Parkinson's disease cases with respect to all factors included in our model, not just those related to functional imaging.

Our integrative model classified the SWEDD participants in a bimodal distribution, suggesting that this group represents a heterogeneous mixture of typical cases of Parkinson's disease and people without Parkinson's disease, rather than a distinct disease entity (figure 4). In the PPMI cohort, a second round of longitudinal DAT scan to identify Parkinson's disease is underway. So far, in the available DAT scanning data at 1–2 years after enrolment, five of 55 SWEDD participants have been identified as showing evidence of dopaminergic dysfunction, and would therefore no longer be counted as SWEDD. At baseline the integrative model classified 17 SWEDD participants as having Parkinson's disease, including four of the five participants who showed evidence of dopaminergic dysfunction (probabilities of having Parkinson's disease: 0·868, 0·907, 0·995, and 0·997), and the fifth participant was close to being classified as having Parkinson's disease (probability 0·651, slightly lower than our threshold of 0·655). If we used a cutoff of 0·651, a further 12 SWEDD participants from PPMI would fall above the Parkinson's disease threshold, although none of these participants showed evidence of dopaminergic dysfunction after 1–2 years of follow-up. We used a χ^2 test of proportions to investigate the enrichment of latent dopaminergic deficit—ie, the process by which detection of dopaminergic dysfunction in follow-up imaging causes SWEDD to be reclassified as typical Parkinson's. The test showed that the classification of SWEDD participants as having Parkinson's disease was unlikely to have been due to chance; we detected significant differences in the prevalence of dopaminergic deficit during follow-up in SWEDD participants whose probabilities of having Parkinson's disease were above our classification threshold of 0·655 (4/17) compared with those below the threshold (1/38; p=0·003).

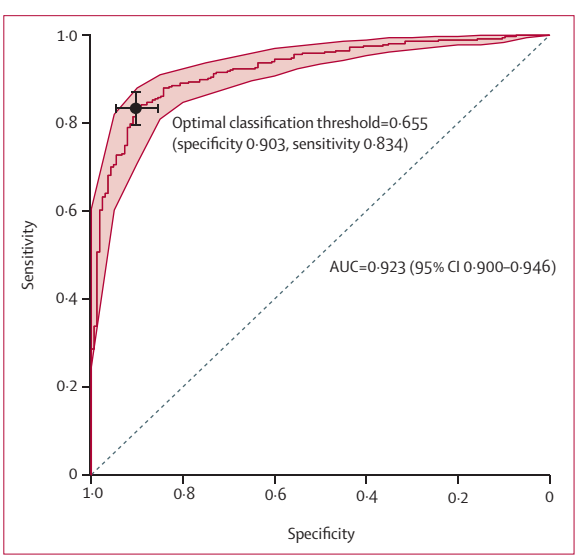

Figure 2: **Receiver operating characteristic curves**
The receiver operating characteristic curve for the integrative model as developed in PPMI cohort. Red shading shows the bootstrap estimated 95% CI with the AUC. Crosshair marks the optimum threshold for classification. AUC=area under the curve. PPMI=Parkinson's Progression Marker Initiative.

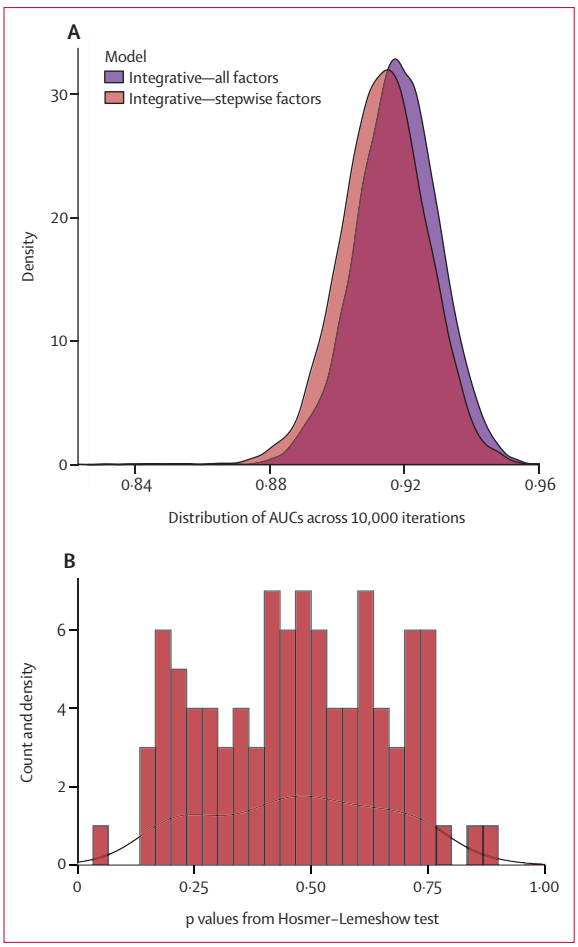

Figure 3: **Internal validation and calibration**
Densities of AUCs from in-silico validation of the integrative model within the PPMI dataset **(A)**. This plot includes all five parameters (UPSIT, GRS, family history of Parkinson's disease, age, and sex) for 10 000 iterations, in addition to another 10 000 iterations of training using backwards stepwise pruning. Distribution of Hosmer-Lemeshow p values for the integrative model in PPMI **(B)** for 5–100 randomly generated groupings to test calibration; a p value of less than 0·05 suggests that bias was detected in that subset of the data. The histogram of the p value distribution from the Hosmer-Lemeshow tests has been overlaid with a linear representation of gausian kernel smoothed density. p values ranged from 0·055 to 0·876, with a mean of 0·471 (SD 0·193) and a median of 0·480 (IQR 0·310–0·624).

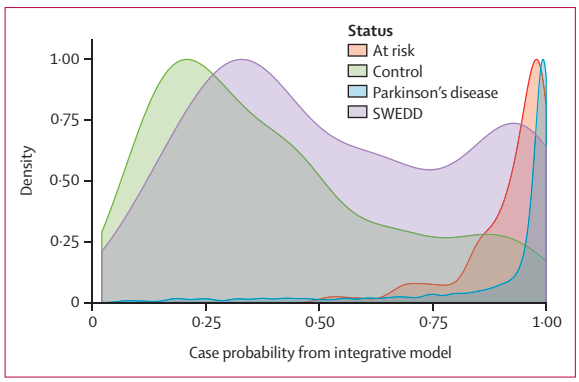

Figure 4: **Probability of Parkinson's disease**
Density plots of the predicted probability of Parkinson's disease estimated from the integrative model by participant status. Participant-level data from all studies analysed are included; these data were not available from 23andMe. Densities were smoothed using a gaussian kernel and no participants had probability estimates of zero. At risk participants were defined in the Parkinson's Associated Risk Study.

Discussion

We have designed an accurate, non-invasive method to discriminate patients with Parkinson's disease from controls. The studies that we assessed vary in their design, recruitment, and implementation; however, our results and validation suggest that the model might be useful in future. The model we developed includes hyposmia, which is often considered an indicator of neurodegeneration, in addition to genetic, clinical, and demographic data.[1] This approach makes use of the growing wealth of data from different aspects of genetic, clinical, and biomarker research.

The main strengths of this model are in its high classification accuracy (AUC about 0·9 or higher) and ease of implementation. This model could be used to refine phenotypes in large research studies by identification of SWEDD participants who overlap with Parkinson's disease in the spectrum of predicted risk and might later show evidence of dopamine dysfunction. As additional DAT scan data become available from the follow-up of SWEDD patients, these data should test the ability of the integrative model to distinguish SWEDD patients who will go on to develop a dopaminergic deficit from those who will not. Our data generally suggest that, within a study, patients incorrectly classified as having Parkinson's disease might need additional, more detailed follow-up than do correctly classified patients, as they might not have aetiologically typical Parkinson's disease and their inclusion might have a negative effect on the power of future biomarker or interventional studies One key element in the application of our integrative model to clinical studies and interventional trials could be the accurate identification of groups of patients with homogeneous disease, such as to exclude SWEDD patients, who typically represent 15% of a clinically acquired cohort of patients with Parkinson's disease. Unlike DAT scanning, our model is portable and can be administered remotely at a fraction of the cost: our model costs around US$100 per sample versus DAT scanning, which can cost thousands of dollars per patient and needs to be administered on site. Additionally, this model might be useful as part of a diagnostic path towards more accurate preclinical detection of Parkinson's disease: our model could potentially be used for disease prediction within populations, although this would require follow-up studies in prospective cohorts.

We have validated this classification model in three case-control studies (with PARS as a positive control) and two case-only studies of Parkinson's disease. We hope to improve the accuracy of this model by identifying more disease-specific biomarkers and genetic risk loci, and by resequencing known loci to generate more accurate estimates of genetic risk. 93% (N=28/30) of the genetic

risk variants that we used to create our GRS are from GWAS and are therefore probably surrogates for true functional variants because of the inherent nature of these imputation-based studies. Identification of the true functional variants within loci would improve our algorithmic classification of Parkinson's disease. Resequencing studies of genetic loci that are now underway might help to refine the genetic aspects of this model. We also hope to expand the model to increase accuracy as more data are being accumulated in our training and validation datasets, especially within PPMI and PD-BP. In this report we have shown that, in these two larger studies, our integrative model significantly outperforms its components if they are assessed independently. In 23andMe and PARS, which had targeted recruitment, hyposmia alone had such a high AUC that the addition of other factors did not significantly change the accuracy of the model.

Our model is specific to Parkinson's disease, and incorporates the classification power of the UPSIT score, a known proxy for generalised neurodegeneration, and the Parkinson's disease-specific factors of family history and GRS. We have reported how this model is focused towards identification of typical Parkinson's disease, as shown by its bimodal classification of SWEDD participants. If the SWEDD participants were typical, they would have been classified as having Parkinson's disease because the model was trained on aetiologically typical Parkinson's disease confirmed by DAT scan. We intend to expand our model to other neurodegenerative diseases, by incorporating multiple disease-specific genetic risk profiles and family histories. If sufficient genetics data become available, adapted versions of our model might be tested as a potential way to minimise misdiagnosis of conditions such as frontotemporal dementia, multiple system atrophy, and dementia with Lewy bodies.

A shortcoming of this analysis is the fact that this study included only participants with genetically ascertained European ancestry, because some genetic heterogeneity might exist across different continental ancestries with respect to risk factors for Parkinson's disease. To address this limitation, we hope to build cohorts of adequate size to investigate Parkinson's disease risk in more diverse populations. This next step should help to refine and improve our predictive models and make them more applicable worldwide. Another shortcoming is the use of age-dependent factors, especially hyposmia, which is common in old age and might affect model performance in populations older than those included in this study. The high proportion of patients with Parkinson's disease in our study who reported a family history of the disorder might be a potential source of bias not seen in some population-based studies or the general population.

Currently, this integrative model has restricted application as a general screening method for Parkinson's disease. Even among populations older than 60 years, prevalence of Parkinson's disease is low, at 2%—despite an AUC of 0·923, the integrative model would probably falsely identify six individuals as having Parkinson's disease for every real case identified.[18] At this prevalence, PPV is also low, at 0·149. Application of Bayes' theorem suggests that, if the prevalence were 10%, the model would falsely classify one individual for every true case of Parkinson's disease detected (PPV 0·489). If prevalence in a population were 20%, one false classification of Parkinson's disease would be made for every two correct classifications (PPV 0·682). These data show that the integrative model might be most useful to identify Parkinson's disease in high-risk populations—eg, in a sample of people with symptoms or other features that might suggest the onset of Parkinson's disease, even though the disease criteria are not yet met. Conversely, the PPV would be low if the model were to be used as a screening test for the general public or by a medical practitioner in a routine clinical setting.

Future research should be directed towards the development of predictive and classification models based on data from prospective studies. Such data will allow the assessment, modification, and reassessment of these predictive models with temporally-developed information, rather than simulations based on retrospective statistics, which might depend on too many assumptions.[20] We expect that models can be refined, evaluated, and tuned for varying rates of disease progression in established patients as datasets grow in size and depth of information content, then be evaluated and validated further in prospective cohort studies. We acknowledge that by basing this model on cross-sectional case-control data from PPMI, we might have caused the results to be slightly conservative, especially with respect to the predictive power of the age and sex parameters in such a well matched study. However, the strength of the imaging-confirmed diagnoses probably contributed to the model's classification accuracy, helping to avoid misdiagnosis. Through future prospective studies in which participants are well characterised, we hope to refine and extend this work to identify a viable timeframe for accurate prediagnostic screening. Another clear area of interest is the application of this model to a broader range of neurodegenerative diseases.

Contributors

MAN and ABS contributed to the concept and design of the study. MAN, CYM, JR, SE, SJH, KG, MS, MMar, PH, NMW, JH, TG, AB, TRP, MFK, CM, JRG, AC-P, ES, CL, MSF, MMap, HJF, AJN, VMVD, DW, CZ, DGH, SL, MMu, EDC, CAMN, MF, KM, AGD-W, DJS, and ABS were involved in the acquisition of data, data generation, and data cleaning. MAN, CYM, MSF, MMap, HJF, AJN, AGD-W, DJS, JPAI, and ABS did the analysis and interpretation of data. MAN, CYM, JR, SE, SJH, KG, MS, MMar, PH, NMW, JH, TG, AB, TRP, AN, MFK, CM, JRG, AC-P, ES, CL, MSF, MMap, HJF, AJN, HM, VMVD, DW, CZ, DGH, SL, MMu, CAMN, MF, KM, AGD-W, DJS, JPAI, and ABS contributed to the drafting of the article and revising it critically.

Declaration of interests

CYM, MM, CAMN, and EDC are current or former employees of, and own stock or stock options in, 23andMe. MFK, CM, and DJS are employees of Merck Pharmaceuticals and hold stocks. AGD-W is an employee of Biogen and holds stock. MSF, MMap, and HJF hold a number of provisional patents related to dementia biomarkers. SJH and MF are employees of the

Michael J Fox Foundation. HM reports grants from the Medical Research Council UK, Wellcome Trust, Parkinson's UK, and the Ipsen Fund, during the conduct of the study grants from the Motor Neuron Disease Association and the Welsh Assembly Government, lecture fees from Teva, advisory board fees from Abbvie, Teva, and Boerhinger Ingelheim, travel expenses from Teva and Medtronic, lecture fees from UCB, and webcast fees from GSK, all outside the scope of this work. AJN is funded by Parkinson's UK (ref K1201) and has received grants from GE Healthcare and Prothena. All other authors declare no competing interests.

Acknowledgments

This project relied mainly on publicly available data for model development and testing or through collaborators who were not compensated for participation in this project specifically. This work was supported in part by the Intramural Research Program of the National Institute on Aging and National Institute of Neurological Disorders and Stroke (project number Z01-AG000949-02). PPMI is supported by the Michael J Fox Foundation for Parkinson's Research and is co-funded by the Michael J Fox Foundation for Parkinson's Research, Abbvie, Avid Radiopharmaceuticals, Biogen Idec, Bristol-Myers Squibb, Covance, Eli Lilly & Co, F Hoffman-La Roche, GE Healthcare, Genentech, GlaxoSmithKline, Lundbeck, Merck, MesoScale, Piramal, Pfizer, and UCB. LABS-PD was supported by the National Institute for Neurological Disorders and Stroke, Department of Defense Neurotoxin Exposure Treatment Parkinson's Research Program, Michael J Fox Foundation, Parkinson's Disease Foundation, Cephalon/Teva, and Lundbeck. PDBP sample and clinical data collection is supported under grants by NINDS: U01NS082134, U01NS082157, U01NS082151, U01NS082137, U01NS082148, U01NS082133. The University of Pennsylvania Udall centre is supported by NINDS under grant P50NS053488 and P50 NS062684 and genotyping of the Penn-Udall cohort was funded by P50 NS062684 (the Pacific Northwest Udall Center). AJN is funded in part by Parkinson's UK (Career Development Award reference F-1201). The 23andMe work was supported in part by the Michael J Fox Foundation for Parkinson's Research (Project title "23andMe Blood Collection for LRRK2 Consortium"). A C-P is also supported by the Doris Duke Clinician Scientist Development Award, the Burroughs Wellcome Fund Career Award for Medical Scientists, and the Benaroya Fund. This work was supported in part (JH) by the Wellcome Trust/Medical Research Council (MRC) Joint Call in Neurodegeneration award (WT089698) to the UK Parkinson's Disease Consortium (UKPDC) whose members are from University College London/Institute of Neurology, the University of Sheffield, and the MRC Protein Phosphorylation Unit at the University of Dundee. HM is funded by Parkinson's UK (Grants 8047) and the Medical Research Council UK (G0700943, G1100643). The authors from 23andMe thank the customers of 23andMe who answered surveys and the other employees of 23andMe, who together made a portion of this research possible. All authors would like to thank all study participants for making this research possible.

References

1 Sobel N, Thomason ME, Stappen I, et al. An impairment in sniffing contributes to the olfactory impairment in Parkinson's disease. *Proc Natl Acad Sci USA* 2001; **98:** 4154–59.

2 Goldstein DS, Sewell L. Olfactory dysfunction in pure autonomic failure: implications for the pathogenesis of Lewy body diseases. *Parkinsonism Relat Disord* 2009; **15:** 516–20.

3 Adler CH, Gwinn KA, Newman S. Olfactory function in restless legs syndrome. *Mov Disord* 1998; **13:** 563–65.

4 Siderowf A, Jennings D, Connolly J, Doty RL, Marek K, Stern MB. Risk factors for Parkinson's disease and impaired olfaction in relatives of patients with Parkinson's disease. *Mov Disord* 2007; **22:** 2249–55.

5 Doty RL, Reyes PF, Gregor T. Presence of both odor identification and detection deficits in Alzheimer's disease. *Brain Res Bull* 1987; **18:** 597–600.

6 Li S, Okonkwo O, Albert M, Wang M-C. Variation in Variables that Predict Progression from MCI to AD Dementia over Duration of Follow-up. *Am J Alzheimer's Dis* 2013; **2:** 12–28.

7 Michell AW, Lewis SJG, Foltynie T, Barker RA. Biomarkers and Parkinson's disease. *Brain* 2004; **127:** 1693–705.

8 Lazarini F, Gabellec M-M, Moigneu C, de Chaumont F, Olivo-Marin J-C, Lledo P-M. Adult neurogenesis restores dopaminergic neuronal loss in the olfactory bulb. *J Neurosci* 2014; **34:** 14430–42.

9 Nalls MA, Pankratz N, Lill CM, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* 2014; **46:** 989–93.

10 International Parkinson's Disease Genomics Consortium (IPDGC), Wellcome Trust Case Control Consortium 2 (WTCCC2). A two-stage meta-analysis identifies several new loci for Parkinson's disease. *PLoS Genet* 2011; **7:** e1002142.

11 Do CB, Tung JY, Dorfman E, et al. Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet* 2011; **7:** e1002141.

12 Nalls MA, Plagnol V, Hernandez DG, 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. *Lancet* 2011; **377:** 641–49.

13 Sidransky E, Nalls MA, Aasly JO, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med* 2009; **361:** 1651–61.

14 Paisán-Ruíz C, Jain S, Evans EW, et al. Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron* 2004; **44:** 595–600.

15 Zimprich A, Biskup S, Leitner P, et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* 2004; **44:** 601–07.

16 Lill CM, Roehr JT, McQueen MB, et al. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: the PDGene database. *PLoS Genet* 2012; **8:** e1002548.

17 Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982; **115:** 92–106.

18 Gasser T. Genetics of Parkinson's disease. *Curr Opin Neurol* 2005; **18:** 363–69.

19 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; **44:** 837–45.

20 Postuma RB, Aarsland D, Barone P, et al. Identifying prodromal Parkinson's disease: pre-motor disorders in Parkinson's disease. *Mov Disord* 2012; **27:** 617–26.