WEB EXTRA MATERIAL

Prediction of cognition in Parkinson's disease with a clinical-genetic score: longitudinal analysis of nine cohorts

SUPPLEMENTARY MATERIAL

CONTENTS

Supplementary Methods

Cohorts and diagnostic criteria

Patients with Parkinson's disease for whom DNA samples were available were included from seven longitudinal studies conducted at several sites from North America and Europe (table 1): Harvard Biomarker Study (HBS)[1-4]; Deprenyl and Tocopherol Antioxidative Therapy of Parkinsonism (DATATOP)[5,6]; Parkinson Research Examination of CEP-1347 Trial/A Longitudinal Follow-up of the PRECEPT Study Cohort (PreCEPT/PostCEPT)[7]; Cambridgeshire Parkinson's Incidence from GP to Neurologist (CamPaIGN)[8], Parkinsonism: Incidence, Cognition and Non-motor heterogeneity in Cambridgeshire (PICNICS)[9,10], Drug Interaction with Genes in Parkinson's Disease (DIGPD)[11], PROfiling PARKinson's disease (PROPARK) study[12], Parkinson's Disease Biomarkers Program (PDBP)[13] and Parkinson's Progression Marker Initiative (PPMI)[14]. Seven cohorts enrolled patients with a diagnosis of PD established according to modified UK PD Society Brain Bank diagnostic criteria as previously reported[4,5,7,9,11,12,15,16]. Diagnostic criteria in PPMI included a positive DATscan and are described in Ref.[14]. In DATATOP, the eligibility criteria required a clinical diagnosis of early, idiopathic PD (HY stages 1 or 2) with patients not on anti-parkinsonian medications[5]. Diagnostic certainty was increased by confirming the clinical diagnosis of PD during longitudinal follow-up visits[17] in all cohorts. Patients whose longitudinal follow-up evaluations were not consistent with a diagnosis of PD were excluded. The diagnosis was further supported by dopamine transporter neuroimaging in PreCEPT and PPMI. 556 subjects were recruited into HBS, between 2005 and 2016; 437 subjects were recruited into DATATOP between 1986 and 1994; 409 subjects were recruited into DIGPD between 2009 and 2014; 114 subjects were studied in CamPaIGN between 2000 and 2014; 129 subjects were studied in PICNICS between 2007 and 2014; 327 subjects in PROPARK were visited between 2003 and 2010; and 332 subjects were recruited into PreCEPT between 2002 and 2010. 500 subjects from PDBP were visited between 2012 and 2016, data was downloaded on the June 30th, 2016; and 396 subjects in PPMI were visited between 2010 and 2016, data was downloaded on the June 30th, 2016.

GBA Sequencing Analysis

For the purpose of this analysis we refer to mutations in *GBA* as those unequivocally identified as pathogenic in Gaucher's disease (GD)[18] as well as E326K, T369M, and E388K mutations. E326K, T369M, and E388K mutations are associated with the risk of PD[3,19] and are linked to GD when occurring in conjunction with other *GBA* mutations, but it is controversial whether they are *per se* pathogenic for GD[20-22]. Participants were identified as carriers (with one or more *GBA* mutations) or non-carriers (no *GBA* mutation). Mutations were identified through targeted next-generation sequencing of the entire *GBA* coding sequence and flanking intronic regions in four data sets (table 1). For 173 PD samples in HBS, 332 PD samples in PreCEPT/PostCEPT and 437 PD samples in DATATOP, as well as 164 PD samples from PROPARK, *GBA* mutations were systematically identified through full sequencing of the exons and flanking intronic regions of *GBA* in RefSeq (NM_001005741.2). To avoid sequencing its neighboring pseudogene, the entire locus was amplified in a single long range PCR reaction using the LA PCR Kit version 2.1 (Takara Bio Inc., Otsu, Japan). Template DNA (100 ng) was added to a 50 µL reaction along with primers (final concentration 0.4 µM) with the following sequences: forward primer (5'-CGACTTTACAAACCTCCCTG-3') and reverse primer (5'-CCAGATCCTATCTGTGCTGG-3'), and cycling conditions were 94° for 1 min (1 cycle), 98° for 10 sec followed by 68°C for 15 min (30 cycles), 72°C for 10 min (1 cycle). This long range PCR assay uses primers that target sequences that uniquely flank *GBA* and produces a single 7,755 bp PCR product. PCR products were visualized on a 0·8% agarose gel with ethidium bromide to confirm successful amplification, which were then used to construct Illumina ready sequencing libraries using the NexteraXT kit (Illumina, San Diego, US) following the manufacturer's instructions. Uniquely indexed samples were pooled (up to 384 samples/pool) and run on the

Illumina MiSeq instrument to generate 150 bp paired-end reads. Sequencing reads were aligned to the human assembly genome (GRCh37/hg19) using BWA[16] (version 0.6.1). *GBA* mutations were called by the GATK [23] (version 1.6) toolkit. A genotype quality of at least 50, and coverage of at least 10x coverage was achieved for all samples. As a quality control, reproducibility of mutation detection was assessed by sequencing 57 samples across different batches in replicates, and the concordance rate was 100% across the *GBA* mutation locus for all replicates. For 383 PD sample in HBS, E326K, N370S and T369M were previously genotyped[3]. Genotypes for 339 of the 383 samples were confirmed on the Illumina NeuroX[24] chip with a genotyping concordance rate of 100%; for 114 PD samples in CamPaIGN and 129 PD samples in PICNICS, mutations and common genetic variants had been identified through full exonic sequencing of *GBA* after two-stage PCR, as part of a previous study[25]. For 409 subjects in DIGPD, exons and flanking intronic regions of *GBA* were sequenced. To avoid amplifying and sequencing the neighbouring pseudogene, *GBA* was amplified in three large fragments (a 2972 bp fragment encompassing exons 1 – 5; a 2049 bp fragment encompassing exons 5 – 7 and a 1682 bp fragment encompassing exons 8 – 11), using previously described primers and a unique 648C to 548C touch-down PCR program[26]. PCR products were sequenced with internal primers, adjacent to coding exons and exon – intron boundaries, using the Big Dye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems), as prescribed. Sequencing products were purified using the Big Dye XTerminator Purification kit (Applied Biosystems), then electrophoresed on an ABI 3730 automated sequencer and analyzed with DNA Sequencing Analysis (version 5.1) and Seqscape (version 2.6) software (Applied Biosystems). 163 samples from PROPARK, which are part of a larger PD exome cohort belonging to the International Parkinson's Disease Genomics Consortium, were sequenced using the EZ Exome Library v2.0 (Roche/Nimblegen) targeting 44.1 Mb. Sequencing reads were aligned to the human reference genome (hg19) with BWA[16]. Single nucleotide variants and small insertions/deletions (indels) were called and filtered using the GATK (version 3.x)[23]. For 91 of the 163 samples, the BAM files were available and calculation of *GBA* coverage was performed. On average, 99·5% of the exonic *GBA* regions was covered for at least 15x. For the remaining samples the overall targeted exome coverage (10x) is 90·8%. For 396 subjects from the Parkinson Progression Marker Initiative (PPMI) cohort, *GBA* genotypes were retrieved from the publicly accessible NeuroX and exome datasets (41 carriers identified by exome sequencing; for 36 of these carriers NeuroX chip data was also assayed with a genotype concordance rate of 100%). For 500 subjects from the Parkinson's Disease Biomarkers Program (PDBP) cohort, *GBA* genotypes accessible through the Data Management Resource (DMR) from the NeuroX array were used.

Statistical Analysis

Cohort differences in covariate values, and unexplained residual heterogeneity among cohorts in their effects on the dependent variable were addressed. Cochran's Q-test and the I^2 index[27,28] were used to test for the presence and degree of heterogeneity across studies (appendix table 1). In addition, differences in continuous and categorical baseline characteristics were compared at baseline between patients with either high and low (cut off 0·196) global cognitive impairment scores using Student's t-tests, and distribution-free Mann-Whitney or Fisher Exact tests, as appropriate (appendix table 2). Unless otherwise stated, in all analyses, p values less than or equal to 0·05 were considered statistically significant.

The primary analyses used the Cox proportional hazards model to estimate the PD global cognitive impairment risk. To ensure consistency across studies, MMSE scores were used to determine global cognitive impairment, with the cut-off of $\leq$ 25 taken as an indicator of significant global cognitive impairment, as recommended by the consensus recommendations of the Movement Disorders Society Task Force[29]. PD patients with global cognitive impairment at enrollment were removed from the Cox proportional hazards analyses.

The Cox proportional hazards statistic was used to estimate the influence of these risk factors on time (years from PD disease onset) to reaching the endpoint of global cognitive impairment (i.e., duration of PD onset at point of global

cognitive impairment) during longitudinal follow-up in the discovery population. Age at onset was defined as the patient's age at the time of PD diagnosis for eight studies. PROPARK defined age at onset as the patient's age at the time of first patient-reported motor symptoms. The Cox regression coefficients which were incorporated into the global cognitive impairment risk score each index the hazard rate throughout the time period (12 years) analyzed, which is assumed to be constant throughout that period (the "proportional hazard assumption"). The Cox regression analysis was performed using the coxph function in the Survival package (Version 2.38-1)[30], and the "Breslow" method was used for handling observations that have tied survival times. The available nine risk factors: gender, age at onset, years of education, *GBA* mutation carrier status, baseline MDS-UPDRS II, III, Hoehn & Yahr stage, Depression, MMSE were in the initial model, and we used the cox.zph function in the Survival package to test the validity of the proportional hazard assumption and found it to be true for all predictors expect MDS-UPDRS II score (p = 1·21e-07), which was discarded for that reason. Age at onset of PD, gender, *GBA* mutation carrier status, years of education, baseline MMSE, baseline MDS-UPDRS III and depression status were selected from backward elimination based on the lowest value of Akaike's information criterion (AIC), as predictors of global cognitive impairment in the multivariable Cox regression model. The proportional hazards assumption of the Cox regression model was tested and not violated by any of the seven predictors in multivariable Cox analysis. (Note also that age at onset of PD illness was found to be one of the predictors in the Cox model. Thus, holding age at PD onset as constant, makes the time since PD onset, i.e., duration of illness, at point of cognitive impairment, perfectly correlate with age at time of cognitive impairment. Age at onset, age at visit and duration of illness at visit are interdependent – if one is held constant, the other two perfectly correlate.

The cumulative/dynamic receiver operating characteristic (ROC) curves and area under curves (AUC), sensitivity, specificity was calculated using the timeROC package (Version 0.2)[31,32]. A bootstrap method was used to calculate the confidence intervals for AUC, and the sweep for the risk threshold averaging method was used to generate confidence bands: we discounted the confidence interval for false positive (1-specificity) and only used the confidence intervals for sensitivity. The Incident/dynamic AUC was calculated using the package RisksetROC package (Version 1.0.4)[33].

The calculation of the PD cognitive risk score can be best described as a series of steps.

(1) Firstly, the age at diagnosis of PD, years of education (first grade = one year, etc.), current MMSE, current MDS-UPDRS III scores, as well as the values for gender, *GBA* carrier status, and depression are multiplied by the coefficients from the Cox model (β_i corresponding to ith coefficient × X_i corresponding to ith covariate value). The sum of the individual "coefficient × value" products I is calculated.

$$I = \sum_i b_i \times X_i$$

(2) And the overall mean "coefficient × value" sum *G* is then calculated from discovery population and combined population in discovery and optimized model, respectively.

$$G = \sum_i b_i \times \overline{X}_i$$

Here, $\overline{X}_i$ is the mean value of ith covariate.

(3) One then calculates the exponent of the individual "coefficient × value" sum minus the overall mean "coefficient × value" sum.

$$B = e^{I-G}$$

(4) The estimated t-year risk of PD global cognitive impairment risk is formally calculated as 1 minus the survival rate at t years (where s(t) corresponds to baseline survival rate of time t), raised to the power of the *B*. For example, an estimate of the 10-year risk of global cognitive impairment is calculated as one minus the survival rate at 10 years.

$$Risk_{(t)} = 1 - [s(t)]^{B}$$

Individual contributions of the seven predictor variables to the score. In order to delineate the individual contributions of the seven predictor variables --- *GBA* mutation status, age at onset of PD, gender, years of education at enrollment, enrollment MMSE, MDS-UPDRS III scores and depression status --- to the cognitive risk score, we performed a multiple regression of the log-transformed risk score on these seven variables simultaneously, based on 1,350 subjects in the discovery cohort. This analysis provides the percent variance in the risk score linearly accounted for by the model as a whole, i.e., by the optimal ordinary least squares linear combination of the seven predictors, as well as the percent variance accounted for by each predictor uniquely, holding all other predictors constant. The former is equal to the sample multiple R squared and the latter are equal to the square of the semi-partial correlations. The log-transformed risk score was used rather than the raw score to better meet test assumptions.

Hypothetical power analysis for a clinical trial targeting PD patients with high cognitive risk scores. We used a repeated measures ANOVA design of two groups versus four time points (enrollment, one year, two years and three years in study). One hypothetical group was assigned to placebo and therefore stipulated to have a MMSE (or MoCA) score across time predicted by our mixed effects model, and the second group was assigned to treatment with an experimental drug, which has the hypothetical ability to halt a decline in the scores (all scores set equal to the predicted MMSE (or MoCA) scores at enrollment)[34]. We computed analyses assuming a two-tailed of 0·05 to detect the difference in trajectories across time for the two groups (group × time interaction), and assuming a within group/timepoint standard deviation of two and a one year test/retest correlation of 0·7 as approximate to those values found empirically, with a first-order autoregressive decay across longer periods[34]. The conservative Greenhouse-Geisser correction for degrees of freedom for correlated error was employed. For comparison, analogous computations were performed for a hypothetical clinical trial scenario including any PD patient without enrichment based on the cognitive risk score.

Supplementary Results

A clinical variables-only version of the risk score

A clinical variables-only version of the risk score can be used in settings, where *GBA* genotyping is not easily obtained, while the clinical-genetic score provides superior prediction where *GBA* status is available. Because *GBA* genotyping requires specific laboratory expertise and is not widely available to many clinicians, we explored a variation of the score comprising the six clinical features only (without *GBA*). This clinical score was informative and predicted global cognitive decline with high accuracy in both the discovery population with an AUC of 0·859 (95% CI, 0·816 - 0·898) and the validation population with an AUC of 0·827 (95% CI, 0·741 - 0·893) (appendix figure 5A,B). As for the clinical-genetic score, the clinical score (with its optimal cutoff of 0·199) was built as predictor of global cognitive impairment in the discovery population and replicated in the independent validation population. To evaluate the stability of the clinical score for predicting cognitive impairment and dementia, we repeated the random re-sampling analysis using 10,000 randomly re-sampled training and test sets and rebuilding the model *ab initio* in the training set in each iteration starting with the seven clinical variables using stepwise backwards elimination based on Akaike's information criterion. Six of the seven clinical variables were stable (with the exception of HY stage) and selected for inclusion into the model in more than 75% of iterations (appendix figure 5C). The prediction accuracy for dementia had an average AUC of 0·860 (95% CI, 0·803 – 0·910) (appendix figure 5D), sensitivity of 0·764 (95%CI 0·567-0·943), and specificity of 0·743 (95%CI 0·573-0·888) in the test sets. The negative predictive value for dementia in the test sets ranged from 0·875 (95% CI, 0·812-0·926) to 0·906 (95% C.I., 0·872-0·937) based on estimates of prevalence of dementia amongst patients with PD ranging from 31·1% (high quality studies)[35] to 24·5% (all studies)[35]. In a head-to-head comparison with the clinical-genetic score, however, the prediction accuracy of the clinical-only score significantly underperformed compared to that of the clinical-genetic score with $p < 0·0001$ (appendix figure 5E). Thus, the informative clinical-only score allows facile implementation in settings where *GBA* status is not easily obtained, while the clinical-genetic score can provide superior prediction where *GBA* status is available.

Table e1: Definition of Parkinson's disease dementia across studies.

Study (Country)	Definition of PD dementia (PD-D)
HBS (USA)	Dementia was defined using operationalized level 1 MDS dementia criteria[36]. These criteria required *1*, an MMSE < 26; *2*, cognitive deficits severe enough to impact daily living (UPDRS sub-score I item 1, Intellectual impairment score ≥ 2 indicating 'Dementia has impact on active daily living scale'); *3*, impairment in at least two cognitive domains operationalized as impairment in two of the following four tasks: ≤ 3 of 5 points in the MMSE Seven backward test (attention); abnormal clock drawing test (executive dysfunction); subscore = 0 in the MMSE Pentagons (visuo-constructive ability); and ≤ 2 of 3 points in the 3-Word Recall of the MMSE (memory performance). A Geriatric Depression Scale-15 (GDS-15) score <10 was used to indicate the absence of severe depression.
DATATOP (USA, Canada)	For DATATOP published criteria for disabling cognitive impairment were used (cognitive impairment leading to functional impairment) as in Ref.[37].
DIGPD (France)	Dementia was defined using the diagnostic criteria and checklist recommended for the diagnosis of PD-D by the Movement Disorder task force as in Ref[29] as well as interview-based assessments with the patient and caregiver.
CamPaIGN (UK)	Dementia was diagnosed on the basis of a MMSE of less than or equal to 24 and fulfillment of the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV criteria for dementia as previously reported in Ref[16].
PICNICS (UK)	Dementia was diagnosed using level 1 MDS dementia criteria[36], which were operationalized in this cohort using the Addenbrooke's Cognitive Examination-Revised, tests of semantic and phonemic fluency, and the pentagon copying test as well as interview-based assessments with the patient and caregiver.
PROPARK (Netherlands)	Diagnosis of dementia was based on a SCOPA-COG cut-off value of 22/23. In Ref[38], using the MDS criteria for dementia as the gold standard, maximum accuracy was attained with this cut-off value.
PreCEPT (USA, Canada)	PreCEPT defined PDD as a score of 4 on the MDS-UPDRS subscale 1 item 1 defined as "cognitive dysfunction [that] precludes the patient's ability to carry out normal activities and social interactions"
PDBP (USA)	Dementia was defined using operationalized level 1 MDS dementia criteria[36]. These criteria required *1*, a Montreal Cognitive Assessment (MoCA) score < 21[39]; *2*, cognitive deficits severe enough to impact daily living (MDS-UPDRS sub-score I item 1, Cognitive impairment score ≥ 2 as criteria for 'Dementia has impact on active daily living scale'); *3*, impairment in at least two cognitive domains operationalized as impairment in two of the following four tasks: ≤ 2 of 3 points in the MoCA seven backwards test (attention); 0 points in the MoCA language fluency test item (language); ≤ 4 of 5 points in the word recall of the MoCA (delayed recall); ≤ 4 of 5 on the MoCA visuospatial/executive test. A Hamilton Depression Rating Scale (HDRS-17) < 24[40] was used as indicating the absence of severe depression.
PPMI (USA, Europe)	Dementia was extracted from PPMI database Cognitive_Categorization table, where Cognitive State score (COGSTATE) =3; Cognitive decline marked as 'Yes'; Any 2 or more of the following cognitive tests are >1.5 SD below the standardized mean: 1,HVLT Total Recall ≤ 35; 2,HVLT Recognition Discrimination ≤ 35; 3,Benton Judgment of Line Orientation ≤ 6; 4,Letter Number Sequencing ≤ 6; 5,Semantic Fluency Test ≤35; 6,Symbol Digit Modalities ≤ 35; Functional impairment marked as 'Yes'

Table e2: Test for residual heterogeneity for each covariate.

	Heterogeneity Q	p value[#]	I^2
Age at onset	19·98	0·01	59·95%
Gender, male versus female	6·57	0·58	0%
Years of education	14·10	0·08	43·27%
GBA mutation carrier	2·79	0·90	0%
MMSE score*	7·81	0·45	0%
MDS-UPDRS III*	10·93	0·21	26·82%
Depression status*	10·37	0·24	22·82%

#The Cochran's Q-test was used to test for residual heterogeneity across studies via R metafor package. I^2 index (100%× (Q-df)/Q) was used to quantify the degree of heterogeneity. *Baseline visits.

Table e3: Examples for calculating the probability of global cognitive impairment.

	Coefficient	Mean value of population	Mr. X		Ms. Y	
			Individual value	Coefficient × value	Individual value	Coefficient × value
Age at visit (years)	N/A	N/A	67	N/A	76	N/A
Age at onset (years)	0·0813	60·4	65	5·2867	74	6·0187
Gender (male:1, female:0)	0·3803	0·619	1	0·3803	0	0·0000
Years of education	-0·0863	13·7	16	-1·3814	5	-0·4317
GBA mutation carrier (yes:1, no:0)	0·4599	0·097	0	0·0000	1	0·4599
MMSE score (26-30)	-0·2819	28·6	28	-7·8919	28	-7·8919
MDS-UPDRS III score (0-132)	0·0219	26·5	16	0·3501	30	0·6565
Depression status (yes:1, no:0)	0·4287	0·207	0	0·0000	1	0·4287
Individual sum				-3·2562		-0·7598
Mean G (coefficient × value)	N/A	-3·3828	N/A	-3·3828	N/A	-3·3828
Baseline survival at 10 years S(t)	N/A	0·7989	N/A	0·7989	N/A	0·7989
Estimated 10-year risk of global cognitive impairment	N/A	N/A	N/A	22·5%	N/A	95·5%

N/A, age at the current visit is not included in the model.

For Mr. X, (1) I =0·0813×65 (age at onset) + 0·3803×1 (male) - 0·0863×16 (years of education) + 0·4599×0 (*GBA* mutation non-carrier) - 0·2819×28 (MMSE score) + 0·0219×16 (MDS-UPDRS III) + 0·4287×0 (no depression) = -3·2562.

(2) G = 0·0813×60·4 (mean of age at onset) + 0·3803×0·619 (proportion of male) - 0·0863×13·7 (mean of years of education) + 0·4599×0·097 (proportion of *GBA* mutation carrier) - 0·2819×28·6 (mean of baseline MMSE score) + 0·0219×26·5 (mean of baseline MDS-UPDRS III) + 0·4287×0·207 (proportion of depression) = -3·3828.

(3) I-G = -3·2562-(-3·3828)=0·1266; and $B = e^{I-G} = e^{0·1249}$=1·135.

(4) $Risk_{(10)} = 1\text{-}S(t)^{B} = 1\text{-}0·7989^{1·135}$ = 0·225. S(t) = 0·7989, which is the 10-year survival rate $S(10)$ derived from the optimized Cox model.

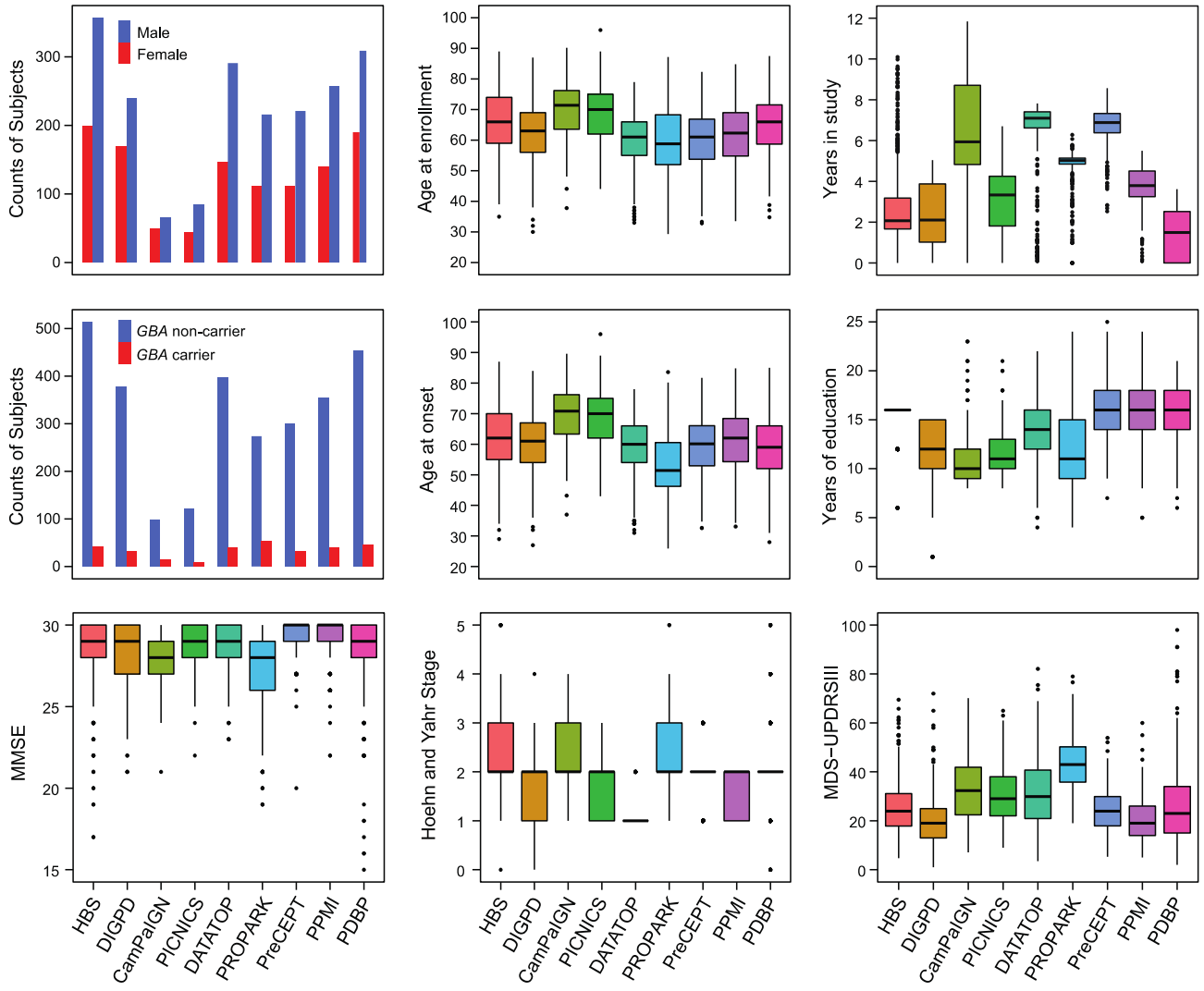
For Ms. Y, (1) I=0·0813×74 (age at onset) + 0·3803×0 (female) - 0·0863×5 (years of education) + 0·4599×1 (*GBA* mutation carrier) - 0·2819×28 (MMSE score) + 0·0219×30 (MDS-UPDRS III) + 0·4287×1 (no depression) = -0·7598.

(2) G = 0·0813×60·4 (mean of age at onset) + 0·3803×0·619 (proportion of male) - 0·0863×13·7 (mean of years of education) + 0·4599×0·097 (proportion of *GBA* mutation carrier) - 0·2819×28·6 (mean of baseline MMSE score) + 0·0219×26·5 (mean of baseline MDS-UPDRS III) + 0·4287×0·207 (proportion of depression) = -3·3828.

(3) I-G = -0·7598-(-3·3828)=2·6229; and $B = e^{I-G} = e^{2·6229}$=13·776.

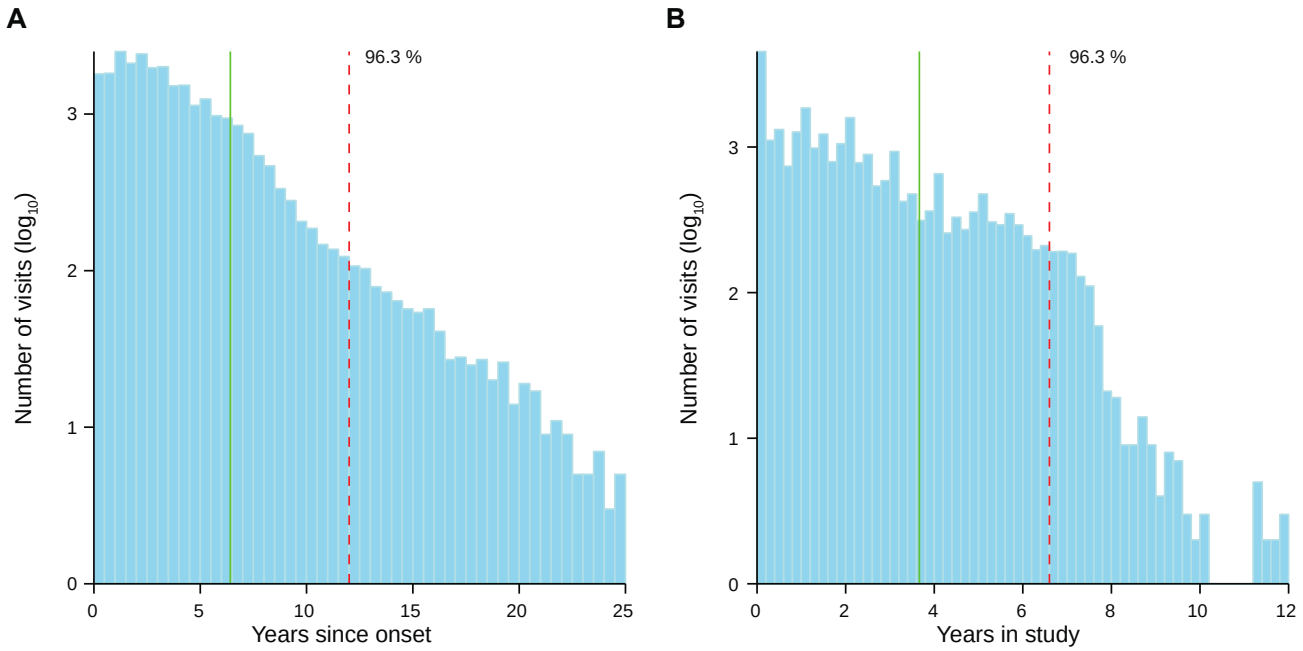
(4) $Risk_{(10)} = 1\text{-}S(t)^{B} = 1\text{-}0·7989^{13·776}$ = 0·955.

Figure e1: Baseline characteristics of patients with PD enrolled in the nine studies.



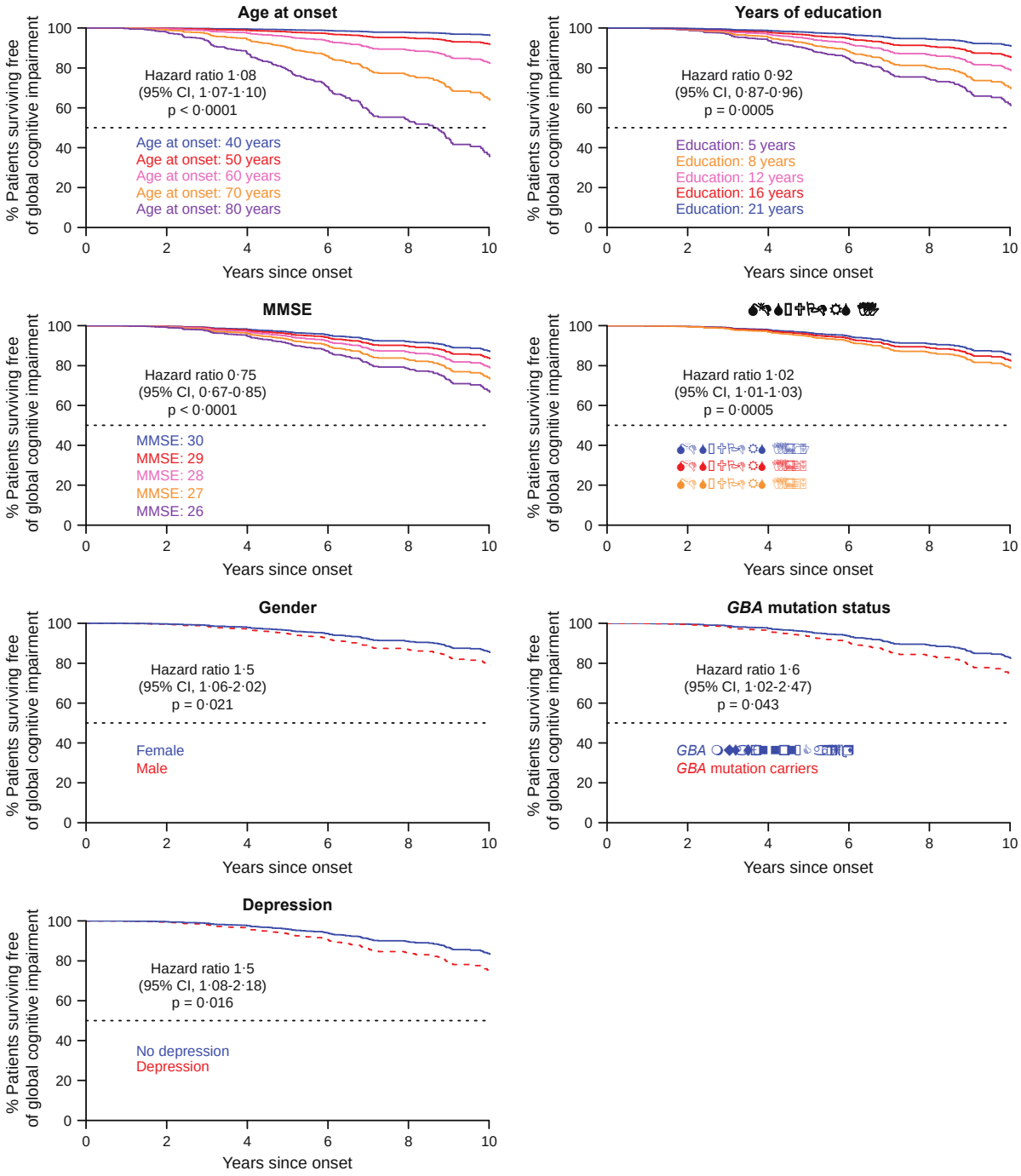
Box plots visualize first, third quartiles and median value, the ends of the whiskers represent the lowest (or highest) value still within 1.5-times the interquartile range. Data beyond the end of the whiskers are outliers and plotted as points.

Figure e2: Longitudinal follow-up visits in the nine cohorts.



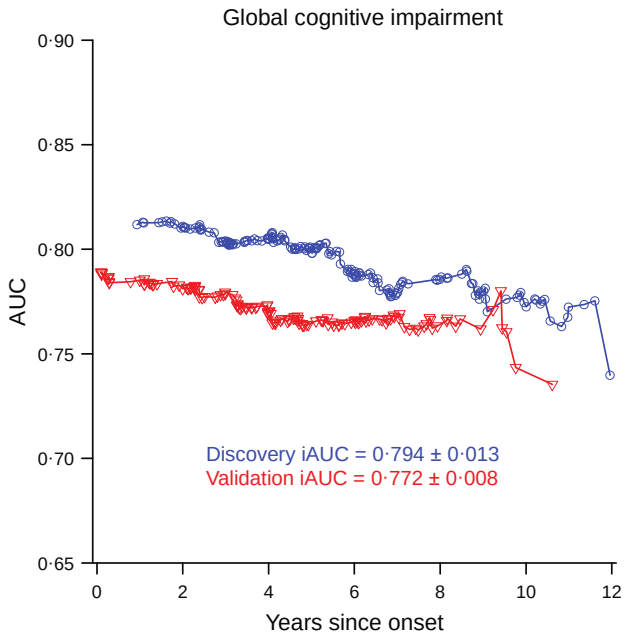
(A) 3,200 patients with PD were longitudinally assessed with 27,022 study visits. 96·3% of study visits occurred within 12 years of longitudinal follow-up from disease onset with a median follow-up time of 6·4 years (inter-quartile range, 4.6 years). 61·1% of all patients (1,957 of 3,200) were enrolled into a cohort within two years of disease onset. **(B)** Here visits are displayed showing "years in study" on the x-axis. Median years in study was 3·7 years (inter-quartile range, 3.7 years). Note, y-axes are shown in log10 scale, bars show number of visits per 0·5 year interval in (**A**) and 0·2 year interval in (**B**).

Figure e3: Covariate adjusted survival curves for seven predictors of cognitive decline.



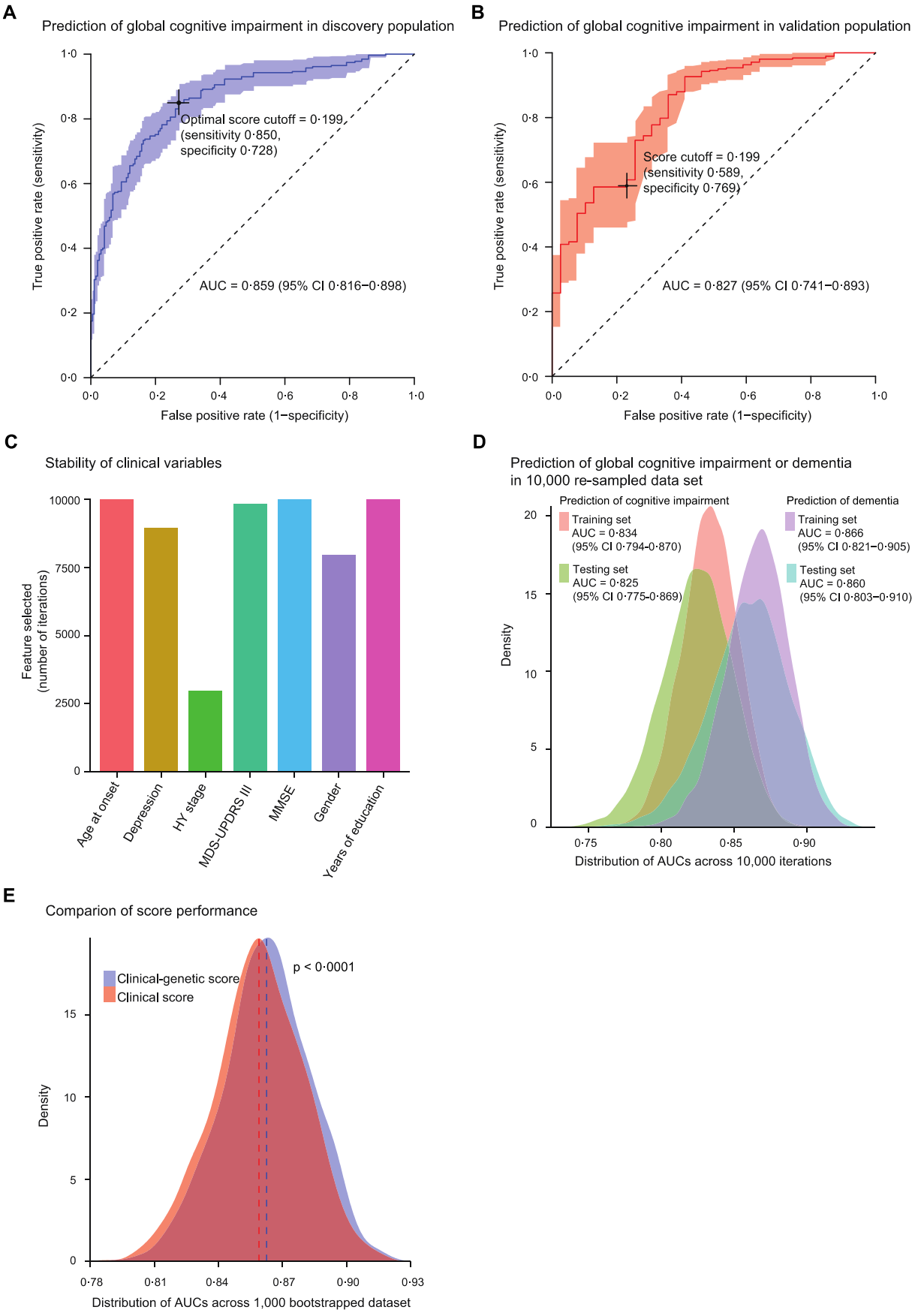
GBA carrier status, age at onset of PD, gender, depression, years of education at baseline, baseline MMSE, and baseline MDS-UPDRS III scores showed significant associations with risk of global cognitive decline in the multivariate Cox proportional hazard analysis. In each panel, predicted values of the probability of survival free of global cognitive impairment are shown for various indicated levels of the stated covariate, with other covariates held constant at their respective grand means. The hazard ratios (HR) shown in each of the panels indicate respectively: HR per year increase in onset age; per additional year of education; per point increase in baseline MMSE; per point increase in baseline MDS-UPDRS III; for male vs. female gender; for *GBA* mutation carrier vs. non-carrier status; for presence vs. absence of baseline depression.

Figure e4: Time-dependent AUCs of the PD cognitive risk predictor.



The prediction accuracy at various follow-up time points using *time-dependent*, incident/dynamic ROC curves was stable up to twelve years since onset in both the discovery and validation populations. A decline in the incident/dynamic AUC was observed around the twelve-year time point, which is likely due to the relatively small number of follow-up visits available for these time points in the combined data set. Estimates of the time-dependent, incident/dynamic AUC (t) versus time based on the PD cognitive risk score are shown under the assumption of proportional hazards estimated as in Ref.[33].

Figure e5: Performance of Clinical variables-only score model.

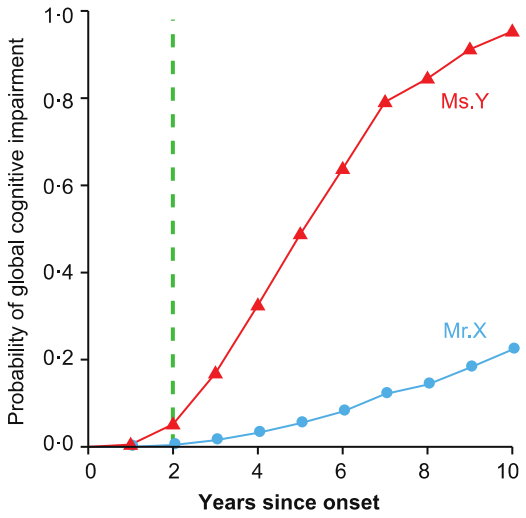


Receiver operating characteristic curves for clinical-only risk score model. The PD cognitive risk (clinical only) score predicts the ten-year cumulative risk of global cognitive decline with a substantial area under the receiver operating characteristic curve (AUC) in **(A)** discovery and **(B)** validation cohort. The shaded area shows the bootstrap estimated 95% CI with the AUC. **(C)** Stability of clinical variables score: to evaluate the clinical score as a predictor of dementia we repeated the random re-sampling analysis using 10,000 randomly re-sampled training and test sets and rebuilt the model *ab initio* in the training set in each iteration starting with the seven clinical variables using stepwise backwards elimination based on Akaike's information criterion. Six of the seven clinical variables were stable (with the exception of HY stage) and selected for inclusion in the model in more than 75% of iterations. (**D**) The clinical score was a strong predictor of dementia with an average AUC of 0·860 (95% CI, 0·803 - 0·910) in the test sets. **(E)** Comparing bootstrapped AUCs between clinical-genetic and clinical-only model: densities of AUCs across 1,000 bootstrapped dataset (discovery cohort) on both of clinical-genetic multivariable Cox regression model (7 risk factors: age onset, gender, years of education, *GBA* mutation carrier or not, presence of depression at baseline, MDS-UPDRS III score and MMSE score at baseline) and clinical-only multivariable Cox regression model (6 risk factors, excluding *GBA* mutations). p value was calculated using paired t-test.

Figure e6: A PD cognitive risk calculator with clinical-genetic feature.

	Coefficient	Mr. X	Ms. Y
Age at visit	N/A	67	76
Age at onset	0·0813	65	74
Gender (male: 1, female: 0)	0·3803	1	0
Years of education (≥ 0)	-0·0863	16	5
GBA mutation carrier (yes: 1, no: 0)	0·4599	0	1
Current MMSE score (must be 26 to 30)	-0·2819	28	28
Current MDS-UPDRS III score (0-132)	0·0219	16	30
Current depression status (yes: 1, no: 0)	0·4287	0	1

Time from diagnosis (years)	Probability of global cognitive impairment	
	Mr. X	Ms. Y
1	0·000	0·005
2	0·004	0·052
3	0·015	0·171
4	0·032	0·329
5	0·054	0·492
6	0·081	0·642
7	0·122	0·795
8	0·144	0·849
9	0·183	0·914
10	0·225	0·955



The calculator estimates a PD patient's probability of global cognitive impairment for each year following his or her diagnosis of PD (up to ten years from diagnosis). A patient's age at the current visit plus seven predictors (age at diagnosis of PD, years of education (first grade = one year, etc.), current MMSE, current MDS-UPDRS III scores, as well as gender, and *GBA* carrier status) all of which can be easily obtained in a clinical setting are entered into the calculator. The cognitive risk estimate is then computed based on the patient's current age, the seven predictors, the β coefficients (derived from the Cox model of the combined discovery and validation populations) and mean predictor variable values (from the combined populations) using the survival rate statistic. See appendix for a detailed step by step method. Values for two hypothetical patients are shown for illustrative purposes. Both patients have MMSE scores of 28 in the normal range at the current visit and both are within two years of disease duration. Mr. X, a 67 year-old man, has only one "high risk" factor out of seven, that is, male gender. By contrast, Ms. Y has an elevated risk profile with five of seven risk factors indicating high risk (older age at diagnosis of 74, less education, GBA mutation present, higher current motor score, depression). The green interrupted line indicates the year of the current visit (e.g. two years since diagnosis). Estimated probabilities of cumulative risk of global cognitive impairment are outputted in table format for each year post diagnosis and plotted as line graphs. While both patients had the same disease duration and same MMSE scores at the current visit, their ten-year risk of global cognitive decline is quite different (22·5% for Mr. X vs. 95·5% for Ms. Y).

References

1. Ding H, Sarokhan AK, Roderick SS, et al. Association of SNCA with Parkinson: replication in the Harvard NeuroDiscovery Center Biomarker Study. *Mov Disord* 2011; **26**(12): 2283-6.
2. Scherzer CR. Chipping away at diagnostics for neurodegenerative diseases. *Neurobiol Dis* 2009; **35**(2): 148-56.
3. Pankratz N, Beecham GW, DeStefano AL, et al. Meta-analysis of Parkinson's disease: identification of a novel locus, RIT2. *Ann Neurol* 2012; **71**(3): 370-84.
4. Ding H, Dhima K, Lockhart KC, et al. Unrecognized vitamin D3 deficiency is common in Parkinson disease: Harvard Biomarker Study. *Neurology* 2013; **81**(17): 1531-7.
5. DATATOP: a multicenter controlled clinical trial in early Parkinson's disease. Parkinson Study Group. *Arch Neurol* 1989; **46**(10): 1052-60.
6. Jankovic J, McDermott M, Carter J, et al. Variable expression of Parkinson's disease: a base-line analysis of the DATATOP cohort. The Parkinson Study Group. *Neurology* 1990; **40**(10): 1529-34.
7. Ravina B, Tanner C, Dieuliis D, et al. A longitudinal program for biomarker development in Parkinson's disease: a feasibility study. *Mov Disord* 2009; **24**(14): 2081-90.
8. Foltynie T, Brayne CE, Robbins TW, Barker RA. The cognitive ability of an incident cohort of Parkinson's patients in the UK. The CamPaIGN study. *Brain* 2004; **127**(Pt 3): 550-60.
9. Breen DP, Evans JR, Farrell K, Brayne C, Barker RA. Determinants of delayed diagnosis in Parkinson's disease. *J Neurol* 2013; **260**(8): 1978-81.
10. Breen DP, Vuono R, Nawarathna U, et al. Sleep and circadian rhythm regulation in early Parkinson disease. *JAMA Neurol* 2014; **71**(5): 589-95.
11. 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**(9): 989-93.
12. Verbaan D, Marinus J, Visser M, van Rooden SM, Stiggelbout AM, van Hilten JJ. Patient-reported autonomic symptoms in Parkinson disease. *Neurology* 2007; **69**(4): 333-41.
13. Rosenthal LS, Drake D, Alcalay RN, et al. The NINDS Parkinson's disease biomarkers program. *Mov Disord* 2015.
14. Parkinson Progression Marker I. The Parkinson Progression Marker Initiative (PPMI). *Prog Neurobiol* 2011; **95**(4): 629-35.
15. Locascio JJ, Eberly S, Liao Z, et al. Association between alpha-synuclein blood transcripts and early, neuroimaging-supported Parkinson's disease. *Brain* 2015; **138**(Pt 9): 2659-71.
16. Williams-Gray CH, Evans JR, Goris A, et al. The distinct cognitive syndromes of Parkinson's disease: 5 year follow-up of the CamPaIGN cohort. *Brain* 2009; **132**(Pt 11): 2958-69.
17. Hughes AJ, Daniel SE, Ben-Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain* 2002; **125**(Pt 4): 861-70.
18. Beutler E, Gelbart T, Scott CR. Hematologically important mutations: Gaucher disease. *Blood Cells Mol Dis* 2005; **35**(3): 355-64.
19. Nichols WC, Pankratz N, Marek DK, et al. Mutations in GBA are associated with familial Parkinson disease susceptibility and age at onset. *Neurology* 2009; **72**(4): 310-6.
20. Horowitz M, Pasmanik-Chor M, Ron I, Kolodny EH. The enigma of the E326K mutation in acid beta-glucocerebrosidase. *Mol Genet Metab* 2011; **104**(1-2): 35-8.
21. Liou B, Grabowski GA. Is E326K glucocerebrosidase a polymorphic or pathological variant? *Mol Genet Metab* 2012; **105**(3): 528-9.

22. Paciotti S, Persichetti E, Pagliardini S, et al. First pilot newborn screening for four lysosomal storage diseases in an Italian region: identification and analysis of a putative causative mutation in the GBA gene. *Clin Chim Acta* 2012; **413**(23-24): 1827-31.

23. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; **20**(9): 1297-303.

24. Nalls MA, Bras J, Hernandez DG, et al. NeuroX, a fast and efficient genotyping platform for investigation of neurodegenerative diseases. *Neurobiol Aging* 2015; **36**(3): 1605 e7-12.

25. 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**(Pt 2): 392-9.

26. Lesage S, Anheim M, Condroyer C, et al. Large-scale screening of the Gaucher's disease-related glucocerebrosidase gene in Europeans with Parkinson's disease. *Hum Mol Genet* 2011; **20**(1): 202-10.

27. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychological Methods* 1998; **3**: 486-504.

28. Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychol Methods* 2006; **11**(2): 193-206.

29. Dubois B, Burn D, Goetz C, et al. Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. *Mov Disord* 2007; **22**(16): 2314-24.

30. Therneau TM. A Package for Survival Analysis in S.; 2014.

31. Blanche P. timeROC: Time-dependent ROC curve and AUC for censored survival data. 2013.

32. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 2013; **32**(30): 5381-97.

33. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics* 2005; **61**(1): 92-105.

34. Liu G, Boot B, Locascio JJ, et al. Specifically neuropathic Gaucher's mutations accelerate cognitive decline in Parkinson's. *Ann Neurol* 2016; **80**(5): 674-85.

35. Aarsland D, Zaccai J, Brayne C. A systematic review of prevalence studies of dementia in Parkinson's disease. *Mov Disord* 2005; **20**(10): 1255-63.

36. Emre M, Aarsland D, Brown R, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov Disord* 2007; **22**(12): 1689-707; quiz 837.

37. Uc EY, McDermott MP, Marder KS, et al. Incidence of and risk factors for cognitive impairment in an early Parkinson disease clinical trial cohort. *Neurology* 2009; **73**(18): 1469-77.

38. Verbaan D, Jeukens-Visser M, Van Laar T, et al. SCOPA-cognition cutoff value for detection of Parkinson's disease dementia. *Mov Disord* 2011; **26**(10): 1881-6.

39. Dalrymple-Alford JC, MacAskill MR, Nakas CT, et al. The MoCA: well-suited screen for cognitive impairment in Parkinson disease. *Neurology* 2010; **75**(19): 1717-25.

40. Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton Depression Rating Scale. *J Affect Disord* 2013; **150**(2): 384-8.