ARTICLE OPEN ACCESS

The SPID-GBA study

Sex distribution, Penetrance, Incidence, and Dementia in GBA-PD

Letizia Straniero, PhD, Rosanna Asselta, PhD, Salvatore Bonvegna, MD, Valeria Rimoldi, PhD, Giada Melistaccio, MSc, Giulia Soldà, PhD, Massimo Aureli, PhD, Matteo Della Porta, MD, Ugo Lucca, MSc, Alessio Di Fonzo, MD, PhD, Anna Zecchinelli, MD, Gianni Pezzoli, MD, Roberto Cilia, MD,* and Stefano Duga, PhD*

Correspondence
Dr. Pezzoli
Gianni.Pezzoli@asst-pini-cto.it or
Prof. Duga
Stefano.duga@hunimed.eu

Neurol Genet 2020;6:e523. doi:10.1212/NXG.0000000000000523

Abstract

Objective

To provide a variant-specific estimate of incidence, penetrance, sex distribution, and association with dementia of the 4 most common Parkinson disease (PD)-associated *GBA* variants, we analyzed a large cohort of 4,923 Italian unrelated patients with primary degenerative parkinsonism (including 3,832 PD) enrolled in a single tertiary care center and 7,757 ethnically matched controls.

Methods

The p.E326K, p.T369M, p.N370S, and p.L444P variants were screened using an allele-specific multiplexed PCR approach. All statistical procedures were performed using R or Plink v1.07.

Results

Among the 4 analyzed variants, the p.L444P confirmed to be the most strongly associated with disease risk for PD, PD dementia (PDD), and dementia with Lewy bodies (DLB) (odds ratio [OR] for PD 15.63, 95% confidence interval [CI] = 8.04–30.37, $p = 4.97*10^{-16}$; OR for PDD 29.57, 95% CI = 14.07–62.13, $p = 3.86*10^{-19}$; OR for DLB 102.7, 95% CI = 31.38–336.1, $p = 1.91*10^{-14}$). However, an unexpectedly high risk for dementia was conferred by p.E326K (OR for PDD 4.80, 95% CI = 2.87–8.02, $p = 2.12*10^{-9}$; OR for DLB 12.24, 95% CI = 4.95–30.24, $p = 5.71*10^{-8}$), which, on the basis of the impact on glucocerebrosidase activity, would be expected to be mild. The 1.5–2:1 male sex bias described in sporadic PD was lost in p.T369M carriers. We also showed that PD penetrance for p.L444P could reach the 15% at age 75 years.

Conclusions

We report a large monocentric study on *GBA*-PD assessing mutation-specific data on the sex distribution, penetrance, incidence, and association with dementia of the 4 most frequent deleterious variants in *GBA*.

*These authors contributed equally.

From the Department of Biomedical Sciences (L.S., R.A., V.R., G.M., G.S., M.D.P., S.D.), Humanitas University; Humanitas Clinical and Research Center (R.A., V.R., G.S., M.D.P., S.D.), IRCCS, Rozzano; Fondazione Grigioni per il Morbo di Parkinson (S.B., A.Z., G.P.); Parkinson Institute (S.B., A.Z., G.P., R.C.), ASST "Gaetano Pini-CTO"; Department of Medical Biotechnology and Translational Medicine (M.A.), University of Milan; Laboratory of Geriatric Neuropsychiatry (U.L.), Istituto di Ricerche Farmacologiche Mario Negri IRCCS; IRCCS Foundation Ca' Granda Ospedale Maggiore Policlinico (A.D.F.), Dino Ferrari Center, Neuroscience Section, Department of Pathophysiology and Transplantation, University of Milan; and Fondazione IRCCS Istituto Neurologico "Carlo Besta" (R.C.), Milan, Italy.

Go to Neurology.org/NG for full disclosures. Funding information is provided at the end of the article.

The Article Processing Charge was funded by PRIN (Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale, Grant no. 2017228L3J, Program coordinator S.D.).

Glossary

CBD = corticobasal degeneration; **CI** = confidence interval; **DLB** = dementia with Lewy body; **FTD** = frontotemporal dementia; **GBA** = glucocerebrosidase; **MSA** = multiple system atrophy; **OR** = odds ratio; **PD** = Parkinson disease; **PDD** = PD dementia; **PKS** = parkinsonism.

Parkinson disease (PD) is a neurodegenerative disorder characterized by motor and nonmotor symptoms,[1,2] with dementia often contributing to poor life quality and reduced survival.[3] Besides aging, epidemiologic studies reported a higher risk of PD for males.[4]

Among predisposing genes, *GBA* (encoding the lysosomal beta-glucocerebrosidase) represents the most common large-effect genetic factor associated with PD in all analyzed populations.[5,6] Patients with heterozygous *GBA* mutations have peculiar characteristics, including earlier onset, more rapid progression of motor impairment, and greater risk for dementia and death.[7–9] Considering other synucleinopathies, *GBA* mutations were even more strongly associated with PD dementia (PDD) and dementia with Lewy bodies (DLB).[10,11]

Among mutations reported in *GBA*, 4 missense variants (p.E326K, p.T369M, p.N370S, and p.L444P) account for ~82% of PD deleterious alleles.[12] Data on their penetrance are disparate and sometimes conflicting, especially for variations having a lower impact on enzyme activity (p.E326K and p.T369M; figure e-1, links.lww.com/NXG/A327).[8,13–18] Reasons for such discrepancies likely stem from (1) differences in the effect size of different mutations; (2) effects of aging/sex on mutation-specific penetrance; and (3) population-specific differences in *GBA* variant frequencies (figure e-1, links.lww.com/NXG/A327).

Here, we propose an extension of our previous study,[11] both in terms of number of patients/controls and of analyzed *GBA* variants (besides p.N370S and p.L444P, we genotyped also p.E326K and p.T369M). We aimed to define the prevalence/penetrance of the most common *GBA*-PD predisposing variants in a large monocentric PD cohort (Sex, Penetrance, Incidence, and Dementia, SPID-*GBA* study), especially in the view of ongoing trials, which will hopefully indicate novel treatments targeting *GBA*-PD.[19]

Methods

Standard protocol approvals, registrations, and informed consents

This study was approved by the Ethics Committee of the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico (study ID 483, March 8, 2018) and conducted according to the Declaration of Helsinki. All participants signed an informed consent.

Study participants

Patients

We enrolled 4,923 unrelated patients who contributed to the Parkinson Institute Biobank (www.parkinson.it/biobanca) of Milan, Italy, from 2002 to 2015, regardless of their sex, family history, age at onset, or other clinical characteristics. Apart from 26 patients having frontotemporal dementia (FTD), all the others received a diagnosis of primary degenerative parkinsonism (PKS) by expert neurologists: 3,832 fulfilled the criteria for PD, 95 for DLB, 216 for multiple system atrophy (MSA), 199 for progressive supranuclear palsy (PSP), and 56 for corticobasal degeneration (CBD). Diagnosis of PD was made according to the UK Brain Bank criteria,[20] and diagnosis of DLB was made for those fulfilling the 1-year rule between the dementia onset and PKS.[21] Patients with suspect of secondary PKS were excluded. A total of 323 patients, who did not meet the validated diagnostic criteria of clinically probable for the various forms of primary (degenerative) PKS (i.e., PD vs DLB vs MSA vs PSP vs CBD vs FTD) and, at the same time, who fulfilled the diagnostic criteria of a clinically possible PKS for more than one among the etiologies listed above, were classified in the category of undefined PKS.

For each participant, the following demographic and clinical data were collected: sex, place of birth, age at onset (the age at which the patient noticed the first motor symptom), disease duration (calculated on the basis of the last examination), and dementia status (diagnosed according to the criteria of the *Diagnostic and Statistical Manual of Mental Disorders, DSM-IV*)[7]; whenever full neuropsychological testing was available, the Movement Disorder Society–recommended diagnostic criteria were applied.[22] For each patient, a detailed family history was collected, and patients were classified as familial if at least 1 relative among their 1st- or 2nd-degree relatives had a diagnosis of PD. Only 1 proband was considered for each family. Among patients with PD, those who developed dementia during the course of the disease were defined as PDD.[22] All clinical data were independently reviewed by an additional experienced neurologist (R.C. or S.B.).

Among the 3,832 patients fulfilling the criteria for PD, 30% were previously screened for several PD-related genes, such as *LRRK2* (p.G2019S, p.R1441C/G, and p.I2020L), *Parkin*, *PINK1*, *DJ1*, and *SNCA*. The 129 carriers of mutations in these genes were excluded from subsequent analyses.

Controls

Our control cohort consisted of 7,757 subjects, deriving from 3 different Italian series. In particular, a total of 1,625 controls

were recruited among partners and caregivers of patients with PD at the Parkinson Institute of Milan. They were not affected by neurodegenerative disorders and denied any family history for movement disorders in first-degree relatives.

A total of 4,016 individuals, along with information about their sex, age, ancestry, and variant calls from whole-exome sequencing data, were collected as part of the Myocardial Infarction Genetics Consortium (MIGen cohort).[23]

A total of 2,116 subjects, deriving from the Health & Anemia study,[24,25] were also recruited. Health & Anemia is a prospective population-based observational study (performed between years 2003 and 2017) of all residents in the municipality of Biella (Italy), aged 65 years or older, aimed to investigate the clinical consequences of anemia in the elderly.[24,25] Information on sex, age, and ancestry was available for all the analyzed subjects.

GBA mutation screening

DNA was extracted from peripheral blood (Maxwell-16 System; Promega, Madison, WI; or Chemagic-Star workstation; Hamilton, ON, Canada).

The mutational screening for the 4 most common *GBA* genetic defects (p.E326K, p. T369M, p.N370S, and p.L444P; legacy names) in cases and controls was performed by a multiplex allele-specific PCR approach designed to avoid the coamplification of the highly homologous *GBAP1* pseudogene.[26] All variants were confirmed by conventional Sanger sequencing. All PCRs were performed on 10 ng of genomic DNA using the GoTaq DNA Polymerase (Promega). Direct sequencing of PCR products was performed using the BigDye Terminator Cycle Sequencing Ready Reaction Kit v1.1 (Thermo Fisher Scientific; Carlsbad, CA) and analyzed on an ABI-3500 Genetic Analyzer (Thermo Fisher Scientific). All oligonucleotides used in the genotyping and validation steps were purchased from Sigma (St Louis, MO).

Statistical analysis

All statistical procedures were performed using R (r-project.org/) or Plink v1.07 (zzz.bwh.harvard.edu/plink/).

For each *GBA* variant, standard case-control analyses on allele frequency data were performed with χ^2 statistics (Fisher exact test), using the minor allele as reference; p values are presented as noncorrected for multiple testing.

The association of variables with the presence of *GBA* variants in patients was evaluated by testing for differences between carriers and noncarriers using either a χ^2 test for categorical data or the Student t test for continuous data (variables were analyzed as quantitative factors, after having controlled their nonsignificant departure from linearity). The effect of each factor was expressed as the odds ratio (OR) and 95% confidence interval (CI). Unadjusted ORs were calculated using a logistic regression model, which included only the factor of interest; adjusted ORs were obtained using a model including the variable of interest plus all variables found significant in the first stage of analysis.

Cumulative incidence was calculated and visualized combining data for different variants/phenotype/sex in a graphical representation that we named Vinicunca plot (after the amazing rainbow mountains of Peru), using the incidence R package. Penetrance was calculated according to[27] using PD prevalence data of the Italian population.[28]

This work conforms the STrengthening the REporting of Genetic Association Studies guidelines (goodreports.org/strega; see supplemental materials, links.lww.com/NXG/A327).

Data availability

Any data not published within the article will be shared by request from any qualified investigator as anonymized data.

Results

Distribution of the 4 most common PD-associated *GBA* variants in cases and controls

We screened a total of 4,923 patients and 7,757 controls for the presence of the 4 most common PD-associated *GBA* variants (p.E326K, p.T369M, p.N370S, and p.L444P, the latter including carriers of the RecNcil allele). Overall, 308 patients and 170 controls were carriers of at least 1 *GBA* variant (carrier frequency: 6.7% and 2.2%, respectively). Only patients with a single variant in *GBA* were considered for further analyses.

Among synucleinopathies, a gradient in carrier frequency is observable, going from 25.8% in DLB, to 6.7% in PD (11.2% in PDD), to 3.7% in MSA; increased carrier frequencies compared with controls were also observed for the tauopathies CBD, FTD, PSP, and for undefined PKSs (table 1).

Comparing the carrier frequency of the 4 variants in individuals affected by the different synucleinopathies, in patients with PD the 4 genetic variations have similar frequencies (range: 1.3%–2.1%), while in PDD, and even more so in DLB, the p.L444P and the p.E326K are overrepresented.

Association of *GBA* variants with PD, PDD, and DLB: the weak p.E326K is not so weak

Compared with the allele frequencies observed in controls, all variants had an allelic OR > 1.56 for PD (range: 1.56–15.63, corrected for age and sex; table 2). As expected, the p.L444P shows, by far, the strongest association with disease risk for all analyzed phenotypes (up to OR = 102.70, 95% CI = 31.38–336.1 in DLB), except for MSA, FTD, and CBD, where no p.L444P carrier was found. Among non–GD-associated variants, p.E326K has an OR for PDD (4.80, 95% CI = 2.87–8.02) and DLB (12.24, 95% CI = 4.95–30.24) similar to those of p.N370S. This is confirmed

Table 1 *GBA* mutation carriers according to clinical diagnosis

		GBA carriers, n (%)				
Diagnosis	**Subjects, n**	**Total**	**p.E326K**	**p.T369M**	**p.N370S**	**p.L444P**
PD	3,691	248 (6.7%)	61 (1.6%)	49 (1.3%)	76 (2.1%)	62 (1.7%)
PDD	807	90 (11.2%)	26 (3.2%)	11 (1.4%)	26 (3.2%)	27 (3.3%)
DLB	93	24 (25.8%)	7 (7.4%)	4 (4.2%)	6 (6.3%)	7 (7.4%)
MSA	216	8 (3.7%)	4 (1.9%)	3 (1.4%)	1 (0.5%)	—
CBD	56	3 (5.4%)	—	2 (3.6%)	1 (1.8%)	—
FTD	26	2 (7.7%)	1 (3.8%)	1 (3.8%)	—	—
PSP	199	9 (4.5%)	2 (1.0%)	3 (1.5%)	3 (1.5%)	1 (0.5%)
Undefined PKS	323	14 (4.3%)	6 (1.9%)	2 (0.6%)	4 (1.2%)	2 (0.6%)
Controls	7,755	170 (2.2%)	55 (0.7%)	61 (0.8%)	43 (0.6%)	11 (0.1%)

Abbreviations: CBD = corticobasal degeneration; DLB = dementia with Lewy body; FTD = frontotemporal dementia; MSA = multiple system atrophy; PD = Parkinson disease; PDD = Parkinson with dementia; PKS = parkinsonism; PSP = progressive supranuclear palsy.
Information on a total of 308 heterozygous patients and 170 controls is reported (collectively, or stratified according to the p.E326K, p.T369M, p.N370S, or p.L444P). Data are reported also for patients with PD classified as PDD.

by the observation that p.E326K is significantly associated also with undefined PKSs.

Genotype-phenotype correlations in PD: p.T369M is the only variant not associated with dementia but showing a sex bias

Given the relatively low number of carriers with other diagnoses (table 1), we compared the clinical phenotype between *GBA* carriers (considering all variants together or stratified according to the 4 variants) and noncarriers only in PD (table 3).

This analysis highlighted that:

1. There is no significant sex unbalance among carriers of p.E326K, p.N370S, and p.L444P compared with noncarriers, whereas p.T369M shows a significant skewing toward females (OR = 0.53, 95% CI = 0.30–0.94);
2. An earlier age at onset is significantly associated only with the 2 more severe mutations (p.N370S and p.L444P; 53.9 and 51 years, respectively, against 58.3 years in noncarriers);
3. Positive family history is significantly associated only with p.L444P (OR = 2.3, 95% CI = 1.36–3.89);
4. A higher risk of dementia was significantly associated not only with the 2 more severe mutations but also with the p.E326K variant (OR = 2.83, 95% CI = 1.69–4.72).

Mutation-, sex-, and disease-specific cumulative incidence: the sex matters

Vinicunca plots in figure 1 report the calculated cumulative incidence by variant and sex in PD. According to the stronger association of p.N370S and p.L444P with earlier age at onset, at 40 years, 70% of patients with *GBA*-PD were carriers of these variants. Conversely, at age 70 years, the contribution of the 2 non–GD-associated variants (p.E326K and p.T369M) was proportionally increased (from 30 to 43%). For 3 of the 4 analyzed variants, a plateau was visible at an age ranging from 71 to 75 years, with the p.L444P reaching the plateau first (figure 1A).

Looking at the sex distribution of patients carrying the different variants, a bimodal curve was observed for the p.T369M variant, with a first plateau reached at age 52 years and a second increase in disease prevalence observable after 57 years. Of interest, males carrying the variant reached the plateau at age 66 years, whereas the incidence in females continued to increase up to 80, an effect that contributes to the increased OR for female sex described above for this variant (figure 1B). Moreover, considering all genetic defects together, the mean age at onset in females was unexpectedly lower than in males (53.24 ± 10.9 vs 55.28 ± 10.9). This is significantly different ($p = 0.0074$) from what we observed in our PD cohort of noncarriers (males have an anticipated disease onset: 57.59 ± 10.8 vs 59.34 ± 10.5).

Looking at the different analyzed phenotypes (figure 2, A and B), a sex bias was observed in patients with PDD and DLB with *GBA* variants, with a larger fraction of male patients with dementia at all ages (percentage of males: 49, 64, and 79% in PD only, PDD, and DLB, respectively; $p = 0.0037$; figure 2B). Finally, the distribution of variant frequencies in PDD is more similar to that of DLB than that of the PD-only group (figure 2C).

Age-dependent penetrance and relative risk in *GBA*-PD

The age-dependent penetrance was estimated using population allele frequencies (from our cohort of 7,755 Italian

Table 2 Allele frequency differences of *GBA* variants between cases and controls

		Allele frequency		Uncorrected analysis		Adjusted analysis	
Diagnosis	**Variants**	**Cases**	**Controls**	**OR (95% CI)**	***p***	**OR (95% CI)**	***p***
PD	All	0.034	0.011	**3.18 (2.61–3.87)**	**1.89E-33**	**3.65 (2.94–4.45)**	**1.25E-31**
	p.E326K	0.0083	0.0035	**2.34 (1.63–3.37)**	**2.62E-06**	**2.48 (1.67–3.70)**	**7.99–06**
	p.T369M	0.0066	0.0039	**1.69 (1.16–2.47)**	**0.0057**	**1.56 (1.02–2.39)**	**0.039**
	p.N370S	0.010	0.0028	**3.74 (2.57–5.44)**	**1.38E-13**	**4.59 (3.07–6.86)**	**1.22E-13**
	p.L444P	0.0084	0.00071	**11.93 (6.28–22.67)**	**5.13E-22**	**15.63 (8.04–30.37)**	**4.97E-16**
PDD	All	0.056	0.011	**5.39 (4.15–7.00)**	**3.69E-45**	**5.89 (4.38–7.91)**	**5.70E-32**
	p.E326K	0.016	0.0035	**4.60 (2.88–7.35)**	**2.55E-12**	**4.80 (2.87–8.02)**	**2.12E-09**
	p.T369M	0.0068	0.0039	1.738 (0.91–3.31)	0.089	1.16 (0.51–2.61)	0.73
	p.N370S	0.016	0.0028	**5.89 (3.61–9.61)**	**8.33E-16**	**6.68 (3.86–11.54)**	**1.03E-11**
	p.L444P	0.017	0.00071	**23.97 (11.87–48.42)**	**9.86E-39**	**29.57 (14.07–62.13)**	**3.86E-19**
DLB	All	0.13	0.011	**13.53 (8.59–21.32)**	**3.87E-48**	**17.53 (9.99–30.76)**	**1.74E-23**
	p.E326K	0.038	0.0035	**10.99 (4.94–24.46)**	**1.74E-13**	**12.24 (4.95–30.24)**	**5.71E-08**
	p.T369M	0.022	0.0039	**5.57 (2.01–15.47)**	**0.00021**	**4.02 (1.16–13.98)**	**0.029**
	p.N370S	0.032	0.0028	**11.99 (5.04–28.52)**	**7.74E-13**	**16.3 (5.74–46.28)**	**1.57E-07**
	p.L444P	0.038	0.00071	**55.10 (21.12–143.8)**	**1.67E-49**	**102.7 (31.38–336.1)**	**1.91E-14**
MSA	All	0.019	0.011	1.72 (0.84–3.53)	0.13	1.81 (0.87–3.74)	0.11
	p.E326K	0.003	0.0035	2.63 (0.95–7.28)	0.054	2.66 (0.95–7.49)	0.064
	p.T369M	0.0069	0.0039	1.77 (0.55–5.67)	0.33	1.86 (0.58–6.04)	0.30
	p.N370S	0.0023	0.0028	0.84 (0.12–6.07)	0.86	0.88 (0.12–6.51)	0.90
	p.L444P	—	0.00071	—	—	—	—
FTD	All	0.039	0.011	3.65 (0.88–15.14)	0.056	3.85 (0.90–16.54)	0.070
	p.E326K	0.019	0.0035	5.51 (0.75–40.58)	0.059	5.58 (0.74–42.34)	0.097
	p.T369M	0.019	0.0039	4.97 (0.68–36.51)	0.080	5.12 (0.68–38.71)	0.11
	p.N370S	—	0.0028	—	—	—	—
	p.L444P	—	0.00071	—	—	—	—
CBD	All	0.027	0.011	2.51 (0.79–7.99)	0.11	1.88 (0.45–7.87)	0.39
	p.E326K	—	0.0035	—	—	—	—
	p.T369M	0.018	0.0039	**4.61 (1.11–19.07)**	**0.021**	2.59 (0.35–19.29)	0.35
	p.N370S	0.0089	0.0028	3.24 (0.44–23.74)	0.22	3.71 (0.49–28.28)	0.21
	p.L444P	—	0.00071	—	—	—	—
PSP	All	0.023	0.011	**2.11 (1.07–4.16)**	**0.027**	**2.32 (1.15–4.43)**	**0.019**
	p.E326K	0.0050	0.0035	1.45 (0.35–5.84)	0.63	1.47 (0.35–6.20)	0.60
	p.T369M	0.0075	0.0039	1.92 (0.60–6.16)	0.26	1.98 (0.60–6.51)	0.26
	p.N370S	0.0075	0.0028	2.73 (0.84–8.84)	0.080	3.25 (0.96–11.01)	0.059
	p.L444P	0.0025	0.00071	3.55 (0.46–27.56)	0.16	5.22 (0.64–42.64)	0.12
Undefined PKS	All	0.022	0.011	**2.02 (1.17–3.51)**	**0.011**	**2.12 (1.17–3.84)**	**0.014**
	p.E326K	0.0093	0.0035	**2.63 (1.13–6.14)**	**0.020**	**2.77(1.14–6.74)**	**0.024**

Continued

Table 2 Allele frequency differences of *GBA* variants between cases and controls *(continued)*

Diagnosis	Variants	Allele frequency		Uncorrected analysis		Adjusted analysis	
		Cases	Controls	OR (95% CI)	*p*	OR (95% CI)	*p*
	p.T369M	0.0031	0.0039	0.79 (0.19–3.22)	0.74	0.82 (0.19–3.44)	0.78
	p.N370S	0.0062	0.0028	2.24 (0.80–6.26)	0.11	2.07 (0.61–7.01)	0.24
	p.L444P	0.0031	0.00071	**4.38 (0.97–19.78)**	**0.036**	**7.23 (1.50–34.9)**	**0.014**

Abbreviations: CI = confidence interval; CBD = corticobasal degeneration; DLB = dementia with Lewy body; PD = Parkinson disease; MSA = multiple system atrophy; OR = odds ratio; PDD = Parkinson with dementia; PKS = parkinsonism; PSP = progressive supranuclear palsy.
p Values < 0.05 and the corresponding ORs and 95% CI are indicated in bold. In the adjusted analyses, values are corrected for age and sex.

controls), which are in line with those observed among Southern Europeans (table 1; figure e-1, links.lww.com/NXG/A327), as well as age-dependent PD prevalence estimates for the Italian population.[28] The most striking results concern p.L444P, which reaches a penetrance of 15.07% above 75 years. All the other 3 variants show a comparable penetrance, ranging from 2.15% to 4.72% for p.T369M and p.N370S, respectively (>75 years; table 4).

Age-dependent increasing penetrance is paralleled by decreasing relative risk estimates for all variants. In particular, p.L444P shows a peak in relative risk at an age comprised between 35 and 44 years (RR = 7.51, 95% CI = 5.91–9.54), whereas above 75 years, the relative risk for this mutation is null (table 4).

Discussion

Although there is consensus on the fraction of patients with PD bearing *GBA* defects (5%–10%),[6] the effect of each variant changes according to different authors, with p.E326K being reported as having the highest or the lowest OR among the 4 most frequent PD-associated *GBA* variants.[29,30] Moreover, p.L444P was reported to confer an OR ranging from 2.62[18] to 8.17.[30] This is likely due to population-specific effects[31] and technical difficulties in genotyping these variants, which is also present in the *GBAP1* pseudogene. In this frame, we tackled a large monocentric study on the penetrance, incidence, and association with dementia of the 4 most frequent *GBA* variants in PD. The main strengths of this work are the meticulous clinical characterization of patients, all followed in a single center, the ethnical match between cases and controls, and the accurate genotyping of variants, which were all analyzed by allele-specific PCR and confirmed by Sanger sequencing. Moreover, recent literature on the predisposing role of *GBA* variants on PD suggests that non–GD-causing variants (p.E326K and p.T369M) might be associated with a second genetic defect, pointing to an oligogenic model for PD predisposition,[17,18] further complicating the interpretation of the specific role of single variants. We therefore chose to focus as much as possible on single-allele carriers, excluding from the analysis all *GBA*-positive patients and controls homozygous or compound heterozygous for one of the analyzed *GBA* variants as well as all carriers of PD-causing mutations in known PD genes. Having taken these considerations into account, we could calculate an unbiased estimate of the true risk, penetrance, cumulative incidence, and sex-specific effects associated with the most frequent PD-related variants in *GBA*.

When *GBA* variants are ranked according to the predicted impact on enzyme activity,[13–15] a gradient in the conferred risk is evident not only for PD but also for PDD and DLB for p.T369M, p.N370S, and p.L444P (figure e-2, links.lww.com/NXG/A327). Conversely, p.E326K appears to have a more severe impact than expected. Indeed, in our study, p.E326K and p.L444P are the 2 variants with the highest OR for PDD (OR = 4.80, 95% CI = 2.87–8.02, $p = 2.12*10^{-9}$ and OR = 29.57, 95% CI = 14.07–62.13, $p = 3.86*10^{-19}$, respectively). This is further confirmed by an OR for p.E326K of 12.24 (95% CI = 4.95–30.24, $p = 5.71*10^{-8}$) for DLB and by the recently reported frequency of p.E326K in a large study on DLB among Caucasians (3.9%),[32] almost overlapping with the p.E326K allele frequency found in our DLB cohort (3.8%). However, in their study, Orme and colleagues[32] failed to report the frequency of p.L444P, the most strongly associated with DLB in our cohort (OR = 102.7, 95% CI = 31.38–336.1, $p = 1.91*10^{-14}$).

Reliable penetrance estimates are hugely important, particularly for individuals undergoing predictive genetic testing and in the view of the clinical trials on *GBA*-PD. Family-based methods to estimate penetrance have been used, but they have ascertainment bias. A practical alternative to estimate penetrance is to use orthogonal, population-based methods.[27] To this aim, a good estimate of allele frequency, even in the case of uncommon variants, is crucial. However, despite the vast amount of information in databases (Exome Aggregation Consortium, ExAC and Genome Aggregation Database, gnomAD), these data are scarcely available in case-control matched populations. The situation is even more complicated for *GBA*, as the presence of the *GBAP1* pseudogene complicates genotyping, thus reducing the availability of reliable information from exome data. This is why we decided to gather a large number of ethnically matched controls to have a

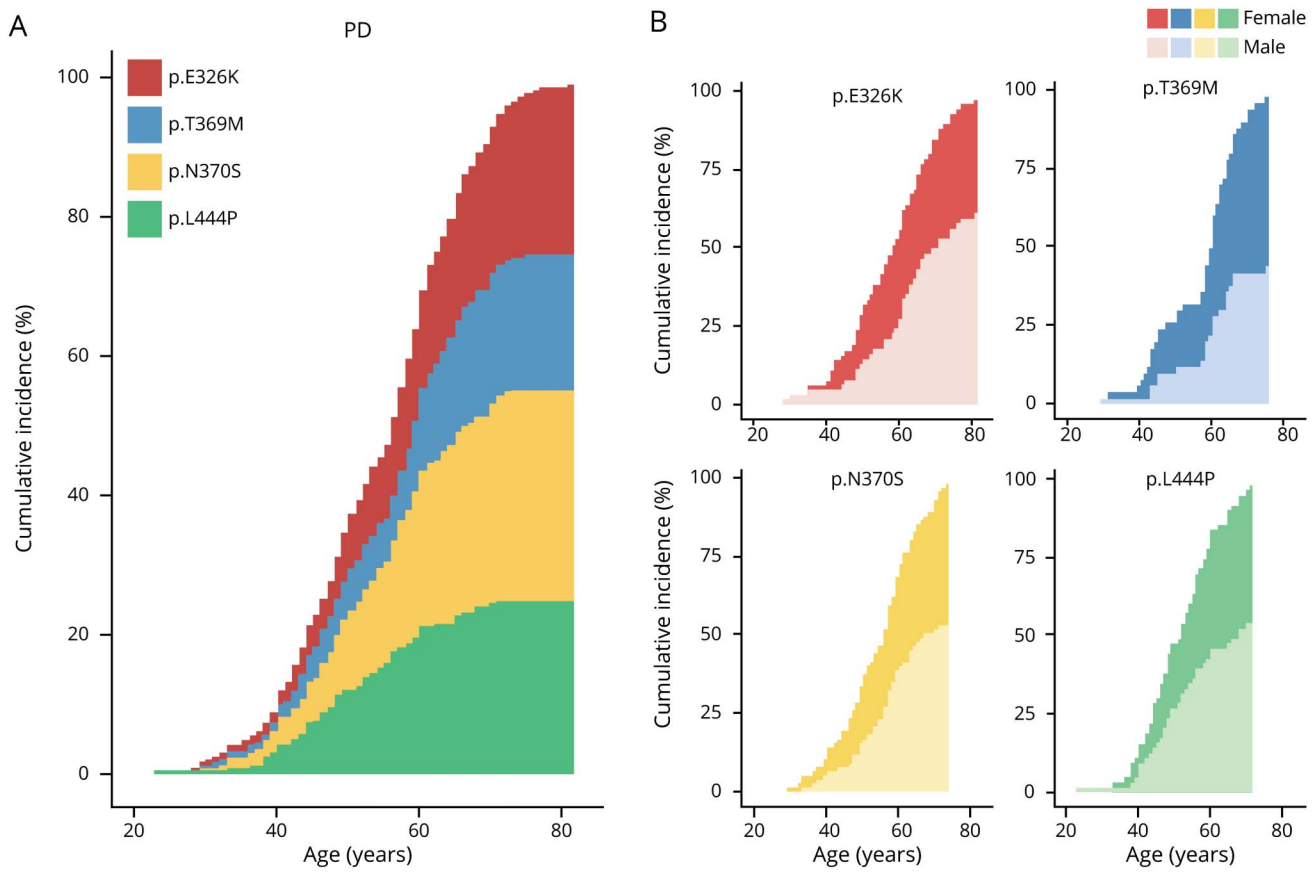
Table 3 Analysis of demographic and general clinical features of 3,691 patients with PD according to *GBA* mutation status

Feature	Noncarriers (n = 3,433)	All variants (n = 248)		
			OR (95% CI) adj (95% CI)	*p* adj *p*
Sex (Male %, [n])	60.30% (2,076)	54.44% (135)	0.79 (0.61–1.02)	0.069
Age at onset (mean ± SD)	58.24 ± 10.78	54.35 ± 10.96	**1.39 (1.24-1.56)** 1.40 (1.24–1.58)	**2.48E-08** 2.94E-08
PD family history (% [n])	19.38% (653)	28.69% (70)	**1.67 (1.25-2.24)** 1.56 (1.16–2.10)	**0.00051** 0.0030
Disease duration (mean ± SD)	13.55 ± 6.89	14.05 ± 7.29	1.01 (0.99–1.03)	0.29
Dementia (% [n])	20.82% (717)	36.29% (90)	**2.17 (1.65-2.84)** 2.25 (1.71–2.977)	**2.45-08** 9.01–09

	p.E326K (n = 61)			p.T369M (n = 49)			p.N370S (n = 76)			p.L444P (n = 62)		
Feature		**OR (95% CI)**	***p***		**OR (95% CI)**	***p***		**OR (95% CI)** **adj OR (95% CI)**	***p*** **adj *p***		**OR (95% CI)** **adj OR (95% CI)**	***p*** **adj *p***
Sex (Male %, [n])	62.30% (38)	1.09 (0.64–1.81)	0.762	44.90% (22)	**0.53 (0.30–0.94)**	**0.0305**	53.95% (41)	0.77 (0.49–1.22)	0.258	54.84% (34)	0.80 (0.48–1.31)	0.378
Age at onset (mean ± SD)	56.75 ± 11.88	1.10 (0.88–1.39)	0.396	56.22 ± 10.65	1.20 (0.93–1.56)	0.149	53.92 ± 10.48	**1.47 (1.20–1.81)** 1.49 (1.22–1.85)	**0.000185** 0.000108	51.03 ± 9.87	**1.85 (1.47–2.32)** 1.81 (1.44–2.32)	**1.74E-07** 4.66E-07
PD family history (% [n])	20.00% (12)	1.04 (0.55–1.97)	0.898	27.66% (13)	1.59 (0.84–3.04)	0.156	28.00% (21)	1.62 (0.97–2.71)	0.0637	38.71% (24)	**2.63 (1.57–4.42)** 2.30 (1.36–3.89)	**0.000248** 0.00199
Disease duration (mean ± SD)	13.86 ± 6.80	1.02 (0.98–1.059)	0.404	14.52 ± 6.97	1.02 (0.98–1.06)	0.339	13.69 ± 7.33	1.00 (0.97–1.04)	0.865	13.86 ± 6.80	1.01 (0.97–1.05)	0.736
Dementia (% [n])	43.55% (27)	**2.83 (1.69–4.72)**	**7.50E-05**	22.45% (11)	1.10 (0.56–2.16)	0.78	34.21% (26)	**1.98 (1.22–3.20)** 2.08 (1.28–3.37)	**0.00545** 0.00296	43.55% (27)	**2.93 (1.76–4.88)** 3.14 (1.87–5.25)	**3.37E-05** 1.38E-05

Abbreviations: CI = confidence interval; OR = odds ratio; PD = Parkinson disease.
Values are expressed as % (and counts) or as mean ± SD. Significant values are in bold.
If significantly different in the unadjusted analysis, variables were also analyzed in a multivariate context through multivariate logistic regression (adjusting for all covariates that resulted significantly different between carriers and noncarriers in the crude analysis), with the presence of *GBA* variants as dichotomous response variable.

Figure 1 Overall and sex-specific age-related cumulative incidences of PD according to *GBA* genotype



(A) Vinicunca plot showing the cumulative incidence of PD in *GBA*-mutated patients. Each colored area represents the individual contribution of each analyzed *GBA* variant. (B) Cumulative incidence of PD according to genotype (heterozygosity for p.E326K, p.T369M, p.N370S, and p.L444P) and sex (the number of male and female individuals is presented in table 3). In all cases, cumulative incidence is reported as percentage (Y axis), and age reported in 20-year intervals on the X axis. PD = Parkinson disease.

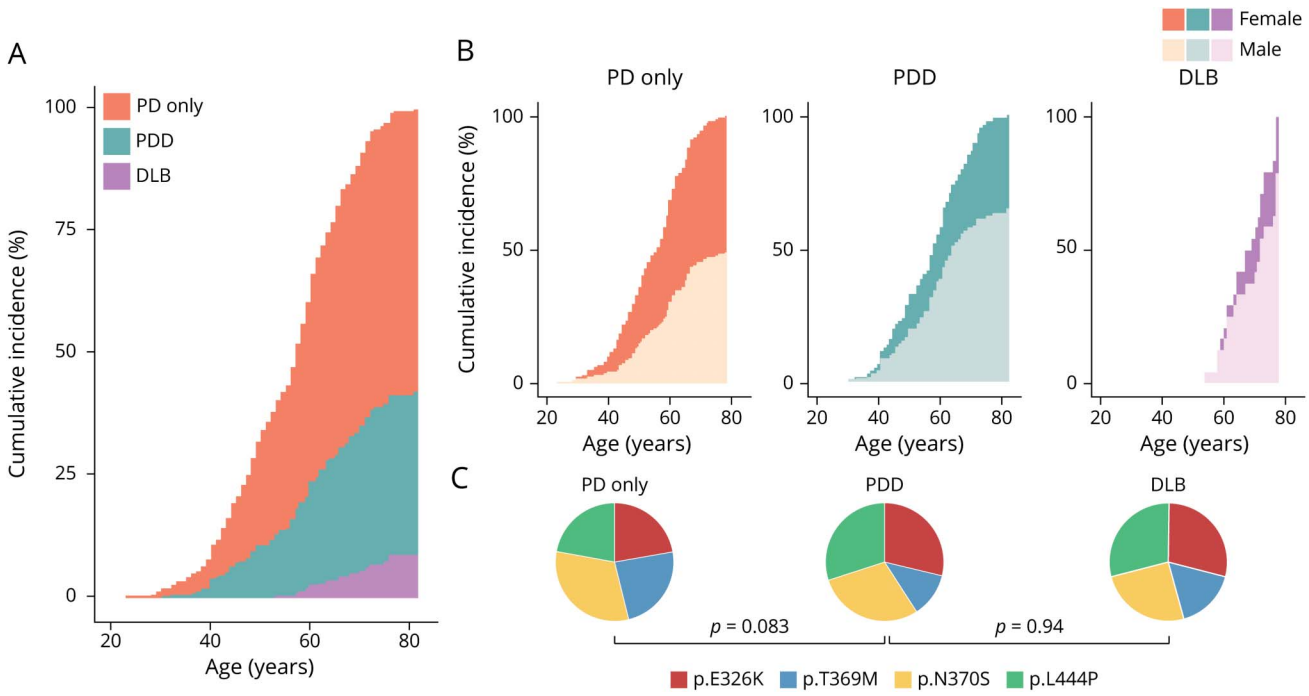
reliable estimate of the Italian population frequencies in individuals without PD for the 4 most common PD-associated *GBA* variants.

Comparing the penetrance of the 4 variants, p.L444P stands out as the mutation conferring the higher PD risk, with a lifelong (at 75 years) penetrance of 15%. This is not far from the penetrance reported for the *LRRK2* p.G2019S mutation among Ashkenazi Jews (26% to age 80 years).[33] Of interest, the risk for PD was similar among male and female p.G2019S carriers, whereas it is quite higher in males among noncarriers.[33] This is similar to what we observed in *GBA* carriers of p.N370S, p.L444P, and, even more so, of p.T369M, for which the OR of being female when affected and carrier of the variant reached statistical significance (OR = 1.89, 95% CI = 1.06–3.33). Taken together, these observations suggest that in mutation carriers, the contribution of a deleterious variant in *LRRK2* or in *GBA* is the major cause of PD in both sexes, overriding the differential environmental exposure and/or hormonal influences that have been proposed to explain the higher prevalence of PD in men.[34] This evidence is further corroborated by the finding that the mean age at onset in female carriers of a *GBA* variant was almost 2 years anticipated compared with men, at difference with what reported in sporadic PD, where females have a comparable or even later onset.[4]

We are aware of the potential limitations of our study, which cannot rule out the possibility that rare deleterious variants in *GBA*, or other mutations in PD-causing genes, might have an impact on variant penetrance. However, our study design is focused to estimate the disease penetrance, which can only be reliably calculated on frequent variants. In addition, considering the ethnic-specific differences in genetic variation frequencies, we chose to limit our analysis to Italians, who, however, show variant frequencies similar to those of Southern Europeans and Caucasians, so that our calculated age-related penetrance and cumulative incidence might be reasonably extended to an average Caucasian. Finally, patients referred to a tertiary care center are likely to be enriched in severe cases; this selection bias is counterbalanced by the high accuracy of clinical diagnoses.

The complex nature of PD etiology, its long presymptomatic phase, and its heterogeneous clinical manifestations present major challenges for formulating successful personalized

Figure 2 Age-related cumulative incidence of PD, PDD, and DLB in male and female carriers of a *GBA* variant



Vinicunca plot showing the cumulative incidence of the *GBA*-associated synucleinopathies in *GBA*-mutated patients (A) and according to diagnosis and sex (B). In all cases, cumulative incidence is reported as percentage (Y axis), and age reported in 20-year intervals on the X axis. (C) Pie charts report the distribution of the 4 analyzed *GBA* variants in patients with respect to clinical diagnosis. *p* Values were calculated using the χ^2 test. DLB = dementia with Lewy body; PD only = Parkinson disease without dementia; PDD = Parkinson with dementia.

Table 4 Overall, variant-specific penetrance and relative risk of *GBA*-PD

	Penetrance, % (95% CI)				
Age (y)	**All variants**	**p.E326K**	**p.T369M**	**p.N370S**	**p.L444P**
35	0.010 (0.0034–0.027)	0.0055 (0.0007–0.041)	0.0060 (0.00070–0.041)	0.017 (0.0025–0.10)	0.033 (0.0016–0.65)
45	0.034 (0.021–0.054)	0.021 (0.0079–0.055)	0.018 (0.0070–0.047)	0.034 (0.014–0.082)	0.20 (0.055–0.74)
55	0.089 (0.061–0.13)	0.062 (0.029–0.13)	0.038 (0.016–0.086)	0.090 (0.046–0.18)	0.48 (0.16–1.4)
65	0.39 (0.28–0.54)	0.28 (0.15–0.53)	0.23 (0.12–0.42)	0.44 (0.24–0.80)	1.40 (0.55–3.60)
75	1.51 (1.10–2.00)	1.13 (0.65–1.90)	0.84 (0.48–1.50)	1.90 (1.10–3.30)	5.42 (2.30–13.00)
≥75	3.90 (3.00–5.10)	2.96 (1.80–4.90)	2.15 (1.30–3.60)	4.72 (2.80–7.90)	15.07 (6.50–34.00)
	Relative risk, (95% CI)				
≤34	3.83 (2.22–6.61)	2.36 (0.70–7.99)	2.36 (0.70–7.99)	4.85 (2.34–10.06)	6.33 (2.76–14.53)
35–44	3.72 (2.88–4.80)	2.53 (1.39–4.60)	2.22 (1.21–4.09)	3.17 (1.95–5.17)	7.51(5.91–9.54)
45–54	1.56 (1.41–1.73)	1.54 (1.27–1.86)	1.50 (1.11–2.04)	1.42 (1.17–1.73)	1.68 (1.51–1.87)
55–64	1.33 (1.23–1.43)	1.25 (1.05–1.48)	1.23 (1.05–1.45)	1.45 (1.41–1.50)	1.28 (1.09–1.52)
65–74	1.54 (1.26–1.89)	1.47 (1.03–2.08)	1.07 (0.64–1.78)	1.93 (1.42–2.64)	2.19 (1.60–3.00)
≥75	0.94 (0.37–2.39)	1.81 (0.67–4.94)	0.65 (0.1–4.25)	0	0

Abbreviations: CI = confidence interval; PD = Parkinson disease.
For penetrance calculations, prevalence for the Italian population at different ages has been obtained from Pupillo et al.[28]

disease-modifying therapies. Hence, the identification of *GBA* mutation carriers combined with a thorough understanding of the risk conferred by the different genetic defects alone and in the context of other genetic and nongenetic risk factors is mandatory to select subjects who could benefit from *GBA*-targeted therapies, currently in clinical trials.

Acknowledgment

Luigi Onorati and Francesca Cancellieri are acknowledged for their invaluable assistance and technical support. The authors thank (in alphabetical order) Drs. Margherita Canesi, Francesca Del Sorbo, Claudio B. Mariani, Nicoletta Meucci, Giorgio Sacilotto, Silvana Tesei, and Michela Zini of the Parkinson Institute for patient referral. Finally, the authors thank all patients and their relatives for their contribution.
The DNA samples from patients were obtained from the Parkinson Institute Biobank (parkinson.it/biobanca), member of the Telethon Network of Genetic Biobanks (biobanknetwork.telethon.it/), and supported by Fondazione Grigioni per il Morbo di Parkinson, Milan, Italy.

Study funding

This work was supported by PRIN (Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale, Grant no. 2017228L3J, program coordinator S. Duga) and by the Fondazione Grigioni per il Morbo di Parkinson, Milan, Italy. L. Straniero was supported by fellowships from the Fondazione Veronesi and Fondazione Grigioni per il Morbo di Parkinson.

Disclosure

The authors report no disclosures. Go to Neurology.org/NG for full disclosures.

Publication history

Received by *Neurology: Genetics* July 6, 2020. Accepted in final form August 31, 2020.

Appendix Authors

Name	Location	Contribution
Letizia Straniero, PhD	Humanitas University, Milan, Italy	Designed and conceptualized the study; analyzed the data; and drafted the manuscript for intellectual content
Rosanna Asselta, PhD	Humanitas University, Milan, Italy	Was in charge of the statistical analysis of the data and revised the manuscript for intellectual content
Salvatore Bonvegna, MD	Fondazione IRCCS Istituto Neurologico "Carlo Besta," Milan, Italy	Interpreted the data and revised the manuscript for intellectual content
Valeria Rimoldi, PhD	Humanitas University, Milan, Italy	Performed genotyping and drew the figures
Giada Melistaccio, MSc	Humanitas University, Milan, Italy	Performed genotyping
Giulia Soldà, PhD	Humanitas University, Milan, Italy	Interpreted the data and revised the manuscript for intellectual content
Massimo Aureli, PhD	Department of Medical Biotechnology and Translational Medicine, University of Milan, Milan, Italy	Participated in the manuscript writing and critical reviewing
Matteo Della Porta, MD	Humanitas University, Milan, Italy	Major role in control cohort collection
Ugo Lucca, MSc	Laboratory of Geriatric Neuropsychiatry, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy	Major role in control cohort collection
Alessio Di Fonzo, MD, PhD	IRCCS Foundation Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy	Participated in the manuscript writing and critical reviewing
Anna Zecchinelli, MD	Parkinson Institute, ASST "Gaetano Pini-CTO", Milan, Italy	Obtained samples and clinical data from patients
Gianni Pezzoli, MD	Fondazione Grigioni per il Morbo di Parkinson, Milan, Italy	Interpreted the data and revised the manuscript and coordinated the study
Roberto Cilia, MD	Fondazione IRCCS Istituto Neurologico "Carlo Besta", Milan, Italy	Obtained samples and clinical data from patients and revised the manuscript
Stefano Duga, PhD	Humanitas University, Milan, Italy	Interpreted the data, revised the manuscript, and coordinated the entire study

References

1. Lees AJ, Hardy J, Revesz T. Parkinson's disease. Lancet 2009;373:2055–2066.
2. Mendoza-Velásquez JJ, Flores-Vázquez JF, Barrón-Velázquez E, Sosa-Ortiz AL, Illigens BW, Siepmann T. Autonomic dysfunction in α-synucleinopathies. Front Neurol 2019;10:363.
3. Hanagasi HA, Tufekcioglu Z, Emre M. Dementia in Parkinson's disease. J Neurol Sci 2017;374:26–31.
4. Georgiev D, Hamberg K, Hariz M, Forsgren L, Hariz GM. Gender differences in Parkinson's disease: a clinical perspective. Acta Neurol Scand 2017;136:570–584.
5. Sidransky E, Nalls MA, Aasly JO, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. N Engl J Med 2009;361:1651–1661.
6. Migdalska-Richards A, Schapira AH. The relationship between glucocerebrosidase mutations and Parkinson disease. J Neurochem 2016;139(suppl 1):77–90.
7. Cilia R, Tunesi S, Marotta G, et al. Survival and dementia in GBA-associated Parkinson's disease: the mutation matters. Ann Neurol 2016;80:662–673.
8. Iwaki H, Blauwendraat C, Leonard HL, et al. Genetic risk of Parkinson disease and progression:: an analysis of 13 longitudinal cohorts. Neurol Genet 2019;5:e348.
9. Stoker TB, Camacho M, Winder-Rhodes S, et al. Impact of GBA1 variants on long-term clinical progression and mortality in incident Parkinson's disease. J Neurol Neurosurg Psychiatry 2020;91:695–702.
10. Nalls MA, Duran R, Lopez G, et al. A multicenter study of glucocerebrosidase mutations in dementia with Lewy bodies. JAMA Neurol 2013;70:727–735.

11. Asselta R, Rimoldi V, Siri C, et al. Glucocerebrosidase mutations in primary parkinsonism. Parkinsonism Relat Disord 2014;20:1215–1220.
12. Malek N, Weil RS, Bresner C, et al. Features of GBA-associated Parkinson's disease at presentation in the UK tracking Parkinson's study. J Neurol Neurosurg Psychiatry 2018;89:702–709.
13. Malini E, Grossi S, Deganuto M, et al. Functional analysis of 11 novel GBA alleles. Eur J Hum Genet 2014;22:511–516.
14. Alcalay RN, Levy OA, Waters CC, et al. Glucocerebrosidase activity in Parkinson's disease with and without GBA mutations. Brain 2015;138(pt 9):2648–2658.
15. Sanchez-Martinez A, Beavan M, Gegg ME, Chau KY, Whitworth AJ, Schapira AH. Parkinson disease-linked GBA mutation effects reversed by molecular chaperones in human cell and fly models. Sci Rep 2016;6:31380.
16. Berge-Seidl V, Pihlstrøm L, Maple-Grødem J, et al. The GBA variant E326K is associated with Parkinson's disease and explains a genome-wide association signal. Neurosci Lett 2017;658:48–52.
17. Mallett V, Ross JP, Alcalay RN, et al. GBA p.T369M substitution in Parkinson disease: polymorphism or association? A meta-analysis. Neurol Genet 2016;2:e104.
18. Goldstein O, Gana-Weisz M, Cohen-Avinoam D,et al. Revisiting the non-Gaucher-GBA-E326K carrier state: is it sufficient to increase Parkinson's disease risk? Mol Genet Metab 2019;128:470–475.
19. Mullin S, Smith L, Lee K, et al. Ambroxol for the treatment of patients with Parkinson disease with and without glucocerebrosidase gene mutations: a nonrandomized, noncontrolled trial. JAMA Neurol 2020;77:427–434.
20. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 1992;55:181–184.
21. McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology 2005;65:1863–1872.
22. Barton B, Grabli D, Bernard B, et al. Clinical validation of Movement Disorder Society-recommended diagnostic criteria for Parkinson's disease with dementia. Mov Disord 2012;27:248–253.
23. Do R, Stitziel NO, Won HH, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature 2015;518:102–106.
24. Riva E, Tettamanti M, Mosconi P, et al. Association of mild anemia with hospitalization and mortality in the elderly: the Health and Anemia population-based study. Haematologica 2009;94:22–28.
25. Tettamanti M, Lucca U, Gandini F, et al. Prevalence, incidence and types of mild anemia in the elderly: the "Health and Anemia" population-based study. Haematologica 2010;95:1849–1856.
26. Straniero L, Rimoldi V, Melistaccio G, et al. A rapid and low-cost test for screening the most common Parkinson's disease-related GBA variants. Parkinsonism Relat Disord 2020;80:138–141.
27. Minikel EV, Vallabh SM, Lek M, et al. Quantifying prion disease penetrance using large population control cohorts. Sci Transl Med 2016;8:322ra9.
28. Pupillo E, Cricelli C, Mazzoleni F, et al. Epidemiology of Parkinson's disease: a population-based study in primary care in Italy. Neuroepidemiology 2016;47:38–45.
29. Blauwendraat C, Bras JM, Nalls MA, Lewis PA, Hernandez DG, Singleton AB. Coding variation in GBA explains the majority of the SYT11-GBA Parkinson's disease GWAS locus. Mov Disord 2018;33:1821–1823.
30. Ran C, Brodin L, Forsgren L, et al. Strong association between glucocerebrosidase mutations and Parkinson's disease in Sweden. Neurobiol Aging 2016;45:212.
31. Zhang Y, Shu L, Sun Q, et al. Integrated genetic analysis of racial differences of common GBA variants in Parkinson's disease: a meta-analysis. Front Mol Neurosci 2018;11:43.
32. Orme T, Hernandez D, Ross OA, et al. Analysis of neurodegenerative disease-causing genes in dementia with Lewy bodies. Acta Neuropathol Commun 2020;8:5.
33. Marder K, Wang Y, Alcalay RN, et al Age-specific penetrance of LRRK2 G2019S in the Michael J. Fox Ashkenazi Jewish LRRK2 Consortium. Neurology 2015;85:89–95.
34. Marras C, Saunders-Pullman R. The complexities of hormonal influences and risk of Parkinson's disease. Mov Disord 2014;29:845–848.

The SPID-GBA study: Sex distribution, Penetrance, Incidence, and Dementia in GBA-PD

Letizia Straniero, Rosanna Asselta, Salvatore Bonvegna, et al.

Neurol Genet 2020;6;
DOI 10.1212/NXG.0000000000000523