Genetic Associations of Parkinson's Disease Clinical, Pathological, and Data-driven Subtypes

Ahmed Negida *, Moaz Elsayed Aboelmagd, Belal Mohamed Hamed, Yousef Hawas, Aya Dziri, Yasmin Negida, Brian D. Berman, Matthew J. Barrett

Posted Date: 10 March 2026

doi: 10.20944/preprints202603.0734.v1

Keywords:

Parkinson's disease; subtypes; genetics; LRRK2; GBA; SNCA; APOE; alpha-synuclein; seed amplification assay; brain-first; body-first; diffuse malignant

 doi:10.20944/preprints202603.0734.v1

Article

Genetic Associations of Parkinson's Disease Clinical, Pathological, and Data-driven Subtypes

Ahmed Negida [1,*], Moaz Elsayed Aboelmagd [2], Belal Mohamed Hamed [3], Yousef Hawas [4], Aya Dziri [5], Yasmin Negida [1], Brian D. Berman [1] and Matthew J. Barrett [1]

[1] Neurology Department, Virginia Commonwealth University, VA, USA
[2] Faculty of Medicine, Cairo University, Egypt
[3] Faculty of Medicine, Al-Azhar University, Egypt
[4] Faculty of Medicine, Tantia University, Egypt
[5] Dnipro State Medical University, Dnipro, Ukraine
* Correspondence: ahmed.negida@vcuhealth.org or ahmed.said.negida@gmail.com; Tel.: +1-347-251-3472

Abstract

Background: Parkinson's disease (PD) is clinically heterogeneous, yet the genetic architecture underlying this heterogeneity remains incompletely understood. We examined the genetic correlates of four complementary PD subtyping frameworks: clinical motor subtype (tremor-dominant [TD] vs. postural instability/gait difficulty [PIGD]), alpha-synuclein seed amplification assay status (SAA+ vs. SAA−), pathological subtype (brain-first vs. body-first, based on REM sleep behavior disorder), and data-driven subtype (diffuse malignant [DM] vs. mild-motor predominant [MMP] vs. intermediate [IM]). **Methods:** We analyzed 1,597 PD patients from the Parkinson's Progression Markers Initiative (PPMI) with genetic testing for seven PD-associated genes (LRRK2, GBA, SNCA, PRKN, PINK1, PARK7, VPS35), including specific variant resolution (LRRK2 G2019S, R1441G/C/H; GBA N409S, severe variants; SNCA A53T), and APOE genotyping (ε2/ε3/ε4 alleles). Genetic variant frequencies were compared across subtypes using chi-square or Fisher's exact tests with Benjamini–Hochberg false discovery rate (FDR) correction. Effect sizes were quantified using Cramér's V. Multivariable logistic regression (statsmodels) estimated adjusted odds ratios with Wald-based 95% confidence intervals. **Results:** Among 1,390 genotyped PD patients, LRRK2 carriers constituted 13.7% (190/1,390; 170 G2019S, 18 R1441G/C/H), GBA 8.6% (119/1,390; 96 N409S, 23 severe), and SNCA 2.0% (28/1,390; all A53T). APOE ε4 carriers comprised 23.4% (323/1,380). SAA-negative patients were markedly enriched for LRRK2 variants (37.1% vs. 10.2%, $P = 3.7 \times 10^{-19}$, $q < 0.001$, $V = 0.25$), driven by G2019S (28.5% vs. 9.6%, $P = 4.9 \times 10^{-11}$, $q < 0.001$) and R1441G/C/H (7.9% vs. 0.5%, $P = 2.7 \times 10^{-12}$, $q < 0.001$). Body-first PD was enriched for GBA carriers (12.3% vs. 6.7%, $P = 0.004$, $q = 0.021$) and depleted for LRRK2 (7.9% vs. 15.0%, $P = 0.002$, $q = 0.013$). The DM subtype carried the highest GBA frequency (14.0% vs. MMP 5.9%, $P < 0.001$, $q = 0.003$). After FDR correction, 10 of 48 univariate tests remained significant. Clinical subtypes (TD vs. PIGD) showed only nominal LRRK2 differences that did not survive FDR correction. APOE genotype did not differ across any framework. **Conclusions:** PD subtypes defined by alpha-synuclein pathology (SAA), pathological onset pattern (brain-first/body-first), and data-driven classification (DM/MMP/IM) show distinct genetic profiles that survive multiple comparison correction. LRRK2 variants strongly associate with SAA-negativity ($V = 0.25$); GBA variants associate with the severe body-first onset and the diffuse malignant subtype.

Keywords: Parkinson's disease; subtypes; genetics; LRRK2; GBA; SNCA; APOE; alpha-synuclein; seed amplification assay; brain-first; body-first; diffuse malignant

1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder, affecting over 10 million people worldwide [1–3]. Despite a shared core pathology of dopaminergic neuronal loss

 doi:10.20944/preprints202603.0734.v1

and alpha-synuclein aggregation, PD exhibits remarkable clinical heterogeneity in motor presentation, non-motor burden, and disease trajectory [4–8]. This heterogeneity has motivated the development of multiple subtyping frameworks, each capturing different aspects of disease biology [8–19].

Four major subtyping approaches have emerged. First, the clinical motor subtype classification distinguishes tremor-dominant (TD) from postural instability/gait difficulty (PIGD) subtypes based on the ratio of tremor to PIGD scores from the MDS-UPDRS [10,20]. Second, the alpha-synuclein seed amplification assay (SAA) status dichotomizes patients based on cerebrospinal fluid alpha-synuclein seeding activity, with approximately 5–10% of clinically diagnosed PD patients testing SAA-negative [21–23]. Third, the pathological subtype classification, proposed by Borghammer [18,19], uses baseline REM sleep behavior disorder (RBD) severity to infer whether PD pathology originated in the brainstem (brain-first) or peripheral autonomic nervous system (body-first). Fourth, the data-driven subtype classification of Fereshtehnejad et al.[15,16] integrates motor and non-motor measures into a data-driven taxonomy of diffuse malignant (DM), mild-motor predominant (MMP), and intermediate (IM) subtypes.

The genetic architecture of PD is increasingly well characterized. Pathogenic variants in LRRK2 (most commonly G2019S and R1441G/C/H) and GBA (most commonly N409S, with severe variants such as L483P) are the most prevalent monogenic risk factors. Less common causes include SNCA (A53T), PRKN, PINK1, PARK7, and VPS35. Additionally, the APOE ε4 allele, a major risk factor for Alzheimer's disease, has been investigated as a potential modifier of cognitive decline in PD.

Whether these genetic variants differentially distribute across PD subtypes has important implications for understanding disease mechanisms and for stratifying patients in clinical trials. In this study, we systematically examine the genetic correlates of all four subtyping frameworks in a large, well-characterized PD cohort from the Parkinson's Progression Markers Initiative (PPMI), with specific variant-level resolution.

2. Materials and Methods

Study Population

We analyzed data from 1,597 PD patients enrolled in the PPMI, a multicenter, longitudinal observational study. All patients met established diagnostic criteria for PD and had baseline clinical assessments and genetic testing available. The study was approved by institutional review boards at all participating sites, and all participants provided written informed consent.

Genetic Testing

Genetic data were derived from the Indiana University genetic consensus file (January 2026), which integrates results from CLIA-certified testing, genome-wide association studies, whole-exome sequencing, whole-genome sequencing, and Sanger sequencing. We analyzed seven PD-associated genes: LRRK2, GBA, SNCA, PRKN, PINK1, PARK7, and VPS35.

For LRRK2, we separately identified the G2019S and R1441G/C/H variants. For GBA, we distinguished the N409S (mild) variant from severe variants (L483P, IVS2+1G>A, etc.). For SNCA, we identified the A53T variant. Binary carrier status was defined as any pathogenic or likely pathogenic variant.

APOE genotyping was available for 1,380 PD patients. ε4 carrier status (at least one ε4 allele), ε4 homozygosity (ε4/ε4), and ε2 carrier status were computed. Full genotype distributions (ε2/ε2, ε2/ε3, ε2/ε4, ε3/ε3, ε3/ε4, ε4/ε4) were tabulated across subtypes.

Subtype Classification

- *Clinical Motor Subtypes (TD/PIGD).*

 doi:10.20944/preprints202603.0734.v1

Classification followed Stebbins et al. (2013). The tremor score was the mean of 10 MDS-UPDRS Part III tremor items (NP3PTRMR, NP3PTRML, NP3KTRMR, NP3KTRML, NP3RTARU, NP3RTALU, NP3RTALJ, NP3RTARL, NP3RTALL, NP3RTCON). The PIGD score was the mean of 5 items (NP3GAIT, NP3FRZGT, NP3PSTBL, NP2WALK, NP2FREZ). A tremor/PIGD ratio ≥ 1.15 classified TD; ≤ 0.90 classified PIGD; intermediate ratios were classified as Indeterminate [10,20].

- *SAA Status (SAA+/SAA−).*

Classification was based on cerebrospinal fluid alpha-synuclein seed amplification assay results using the Amprion platform. Positive (SAA+) and negative (SAA−) results were used as reported.

- *Pathological Subtype (Brain-first/Body-first).*

Classification followed the Borghammer (2021) model using baseline RBD Screening Questionnaire (RBDSQ) total score, computed as the sum of 12 binary items from questions Q1–Q9 (range 0–12; Q6 contributes 4 sub-items, Q10 is excluded in PD cohorts). A score ≥ 6 classified body-first PD; ≤ 3 classified brain-first PD; scores of 4–5 were classified as Indeterminate [18,19,24].

- *Data-driven Subtype (DM/MMP/IM).*

Classification followed Fereshtehnejad et al.[15,16]. Baseline Z-scores were computed for three motor composites (PIGD/tremor ratio, MDS-UPDRS III OFF total, MDS-UPDRS II total) and three non-motor measures (MoCA [inverted], RBDSQ total, SCOPA-AUT total). The 75th percentile of each defined the worst quartile. Patients with a motor composite Z-score ≥ P75 and ≥ 1 non-motor Z ≥ P75, or with all 3 non-motor Z-scores ≥ P75, were classified as diffuse malignant (DM). Those with motor composite Z < P75 and all 3 non-motor Z < P75 were classified as mild-motor predominant (MMP). All others with complete data were classified as intermediate (IM).

Statistical Analysis

Genetic variant carrier frequencies were compared across subtypes using Pearson's chi-square test or Fisher's exact test when any expected cell count was < 5. All P-values are two-sided. To account for multiple testing across 48 univariate comparisons (12 genetic variants × 4 subtyping schemes), we applied the Benjamini–Hochberg procedure to control the false discovery rate (FDR) at 5%. FDR-adjusted P-values (q-values) are reported alongside nominal P-values [25]. Effect sizes for categorical associations were quantified using Cramér's V [26].

Multivariable logistic regression was performed using maximum-likelihood estimation (statsmodels Logit) to estimate adjusted odds ratios (OR) with Wald-based 95% confidence intervals (CI) and P-values. Models included age at visit (standardized), LRRK2 carrier status, GBA carrier status, APOE ε4 carrier status, and sex as covariates. Model fit was assessed using McFadden's pseudo-R^2 and Akaike information criterion (AIC). For predictors exhibiting perfect or quasi-complete separation, results are reported as not estimable.

APOE genotype distributions were compared across subtypes using chi-square tests on the full 6-level genotype table. Cross-scheme agreement between subtyping frameworks was assessed using Cramér's V on cross-tabulations of subtype assignments. All analyses were performed in Python 3.13 using pandas, scipy, statsmodels, and matplotlib. Code and data processing pipelines are available upon request.

3. Results

Study Cohort

The PD cohort comprised 1,597 patients (mean age 63.1 ± 9.8 years, 34.4% male). Genetic data were available for 1,390 patients (87.0%). Of these, 190 (13.7%) carried LRRK2 variants (170 G2019S, 18 R1441G/C/H, 2 other), 119 (8.6%) carried GBA variants (96 N409S, 23 severe), 28 (2.0%) carried SNCA A53T, 17 (1.2%) carried PRKN variants, and no patients carried PINK1, PARK7, or VPS35 variants. Overall, 287 patients (20.6%) carried at least one pathogenic variant.

 doi:10.20944/preprints202603.0734.v1

APOE genotyping was available for 1,380 PD patients. APOE ε4 carriers comprised 23.4% (323/1,380), including 24 ε4/ε4 homozygotes (1.7%). APOE ε2 carriers comprised 15.1% (209/1,380).

Subtype Distributions

- Clinical (TD/PIGD): Among 1,220 patients with evaluable tremor/PIGD scores, 793 (65.0%) were classified as TD, 296 (24.3%) as PIGD, and 131 (10.7%) as Indeterminate.
- SAA Status: Among 1,268 patients with SAA results, 1,112 (87.7%) were SAA+ and 156 (12.3%) were SAA−.
- Pathological Subtype: Among 1,560 patients with RBDSQ data, 985 (63.1%) were classified as brain-first, 342 (21.9%) as body-first, and 233 (14.9%) as Indeterminate.
- Data-driven Subtype: Among 1,272 patients with complete baseline data, 322 (25.3%) were classified as DM, 441 (34.7%) as MMP, and 509 (40.0%) as IM.

The four subtyping frameworks showed low inter-scheme agreement (Supplementary Figure S2), with the highest concordance between Data-driven and Pathological (Cramér's V = 0.29) and Data-driven and Clinical (V = 0.25) classifications.

Genetic Correlates by Subtyping Framework

- *Clinical Motor Subtype (TD vs. PIGD)*

Genetic variant frequencies showed only marginal differences between TD and PIGD subtypes (Table 2, Figure 3a). LRRK2 carrier frequency was higher in PIGD (7.0% [19/270]) than TD (3.4% [25/739], P = 0.024, q = 0.095, V = 0.07), but this did not survive FDR correction. No significant differences were observed for GBA (PIGD 3.0% vs. TD 3.1%, P = 1.0), SNCA (PIGD 0.9% vs. TD 0.0%, P = 0.07), PRKN, APOE ε4 (PIGD 22.7% vs. TD 24.5%, P = 0.54), or any other tested variant. APOE genotype distribution did not differ between TD and PIGD ($\chi^2 = 4.28$, df = 5, P = 0.51).

In adjusted logistic regression (n = 586, pseudo-R^2 = 0.014, AIC = 669.5), no predictor reached significance for PIGD classification: LRRK2 (OR = 2.41 [0.63–9.17], P = 0.20), GBA (not estimable due to separation), APOE ε4 (OR = 0.76 [0.48–1.21], P = 0.24), age (OR = 0.97 [0.80–1.16], P = 0.72), male sex (OR = 1.26 [0.84–1.88], P = 0.26).

- *SAA Status (SAA+ vs. SAA−)*

SAA-negative patients showed dramatically higher rates of LRRK2 variants (Table 3, Figure 1, Figure 3b). Overall LRRK2 carrier frequency was 37.1% in SAA− vs. 10.2% in SAA+ ($P = 3.7 \times 10^{-19}$, q < 0.001, V = 0.25). This was driven by both G2019S (28.5% vs. 9.6%, $P = 4.9 \times 10^{-11}$, q < 0.001, V = 0.18) and R1441G/C/H (7.9% vs. 0.5%, $P = 2.7 \times 10^{-12}$, q < 0.001, V = 0.20). Any pathogenic variant carrier status was significantly higher in SAA− (43.0% vs. 19.2%, $P = 6.4 \times 10^{-11}$, q < 0.001, V = 0.18). GBA did not differ between SAA groups (SAA− 4.6% vs. SAA+ 7.5%, P = 0.28). Neither APOE ε4 (P = 0.89) nor APOE genotype distribution ($\chi^2 = 6.34$, df = 5, P = 0.27) differed between SAA groups.

In adjusted logistic regression (n = 600, pseudo-R^2 = 0.018, AIC = 440.3), LRRK2 carrier status was the only significant predictor of SAA+ status (OR = 0.22 [0.06–0.78], P = 0.02), reflecting the strong enrichment of LRRK2 variants in the SAA-negative group. GBA was not estimable due to quasi-complete separation.

- *Pathological Subtype (Brain-first vs. Body-first)*

Body-first PD was enriched for GBA carriers (12.3% [37/302] vs. 6.7% [59/879] in brain-first, P = 0.004, q = 0.021, V = 0.08, Table 4), with GBA N409S showing a nominal enrichment (12.0% vs. 6.6%, P = 0.015) that did not survive FDR correction (q = 0.067). Conversely, body-first PD was depleted for LRRK2 carriers (7.9% [24/302] vs. 15.0% [132/879], P = 0.002, q = 0.013, V = 0.08), driven by G2019S (7.0% vs. 13.3%, P = 0.004, q = 0.020, V = 0.08). APOE ε4 did not differ (body-first 26.9% vs. brain-first 22.8%, P = 0.20). APOE genotype distribution was also non-significant ($\chi^2 = 7.98$, df = 5, P = 0.16).

In adjusted logistic regression (n = 595, pseudo-R^2 = 0.039, AIC = 671.7), male sex was the strongest predictor of body-first classification (OR = 2.14 [1.38–3.31], $P = 6.3 \times 10^{-4}$), followed by age

 doi:10.20944/preprints202603.0734.v1

(OR = 1.37 [1.13–1.67], P = 0.002). LRRK2 (OR = 0.58 [0.13–2.69], P = 0.49), GBA (OR = 1.13 [0.28–4.49], P = 0.86), and APOE ε4 (OR = 1.48 [0.96–2.27], P = 0.08) were non-significant.

- *Data-driven Subtype (DM vs. MMP vs. IM)*

The DM subtype showed significantly elevated GBA carrier frequency (14.0% [39/279] vs. IM 6.3% [28/447] vs. MMP 5.9% [24/407], $P < 0.001$, $q = 0.003$, $V = 0.11$), driven by GBA N409S (DM 13.6% vs. MMP 4.6%, $P < 0.001$, $q = 0.003$, $V = 0.11$). Any pathogenic variant carrier status was also highest in DM (32.3% vs. IM 21.3% vs. MMP 18.2%, $P < 0.001$, $q = 0.003$, $V = 0.11$, Table 5). LRRK2 carrier frequency showed a non-significant trend (DM 15.1% vs. IM 12.8% vs. MMP 9.1%, $P = 0.11$, $q = 0.31$). APOE ε4 did not differ (DM 20.8% vs. IM 26.9% vs. MMP 23.1%, $P = 0.18$). APOE genotype was non-significant ($\chi^2 = 12.94$, df = 10, $P = 0.23$).

In adjusted logistic regression (n = 560, pseudo-$R^2 = 0.040$, AIC = 603.7), age was the strongest predictor of DM classification (OR = 1.63 [1.31–2.03], $P = 1.2 \times 10^{-5}$). LRRK2 (OR = 1.46 [0.37–5.78], P = 0.59), GBA (OR = 1.18 [0.30–4.72], P = 0.81), APOE ε4 (OR = 0.90 [0.56–1.45], P = 0.66), and male sex (OR = 1.37 [0.85–2.21], P = 0.20) were non-significant in the multivariable model, likely due to the small number of carriers and limited power.

GBA and LRRK2 Carrier Subtype Profiles

GBA carriers showed worse baseline motor scores (MDS-UPDRS III: P = 0.003, rank-biserial r = −0.17), lower cognitive performance (MoCA: P = 0.04, r = −0.11), and greater non-motor burden (MDS-UPDRS I: $P = 7.5 \times 10^{-4}$, r = −0.19) compared with non-carriers (Figure 4a–c). LRRK2 carriers similarly showed significant differences in motor scores (P = 0.02, r = 0.15), MoCA (P = 0.001, r = −0.15), and MDS-UPDRS I (P = 0.004, r = −0.13) (Figure 4d–f).

APOE Analysis

APOE ε4 carrier frequency was consistent across all four subtyping frameworks, ranging from 20.8% to 26.9% without statistically significant differences. Full APOE genotype distributions (ε2/ε2, ε2/ε3, ε2/ε4, ε3/ε3, ε3/ε4, ε4/ε4) were compared using chi-square tests and showed no significant associations with any subtype classification (all P > 0.16). These findings suggest that APOE genotype does not meaningfully differentiate PD subtypes in this cohort.

 doi:10.20944/preprints202603.0734.v1

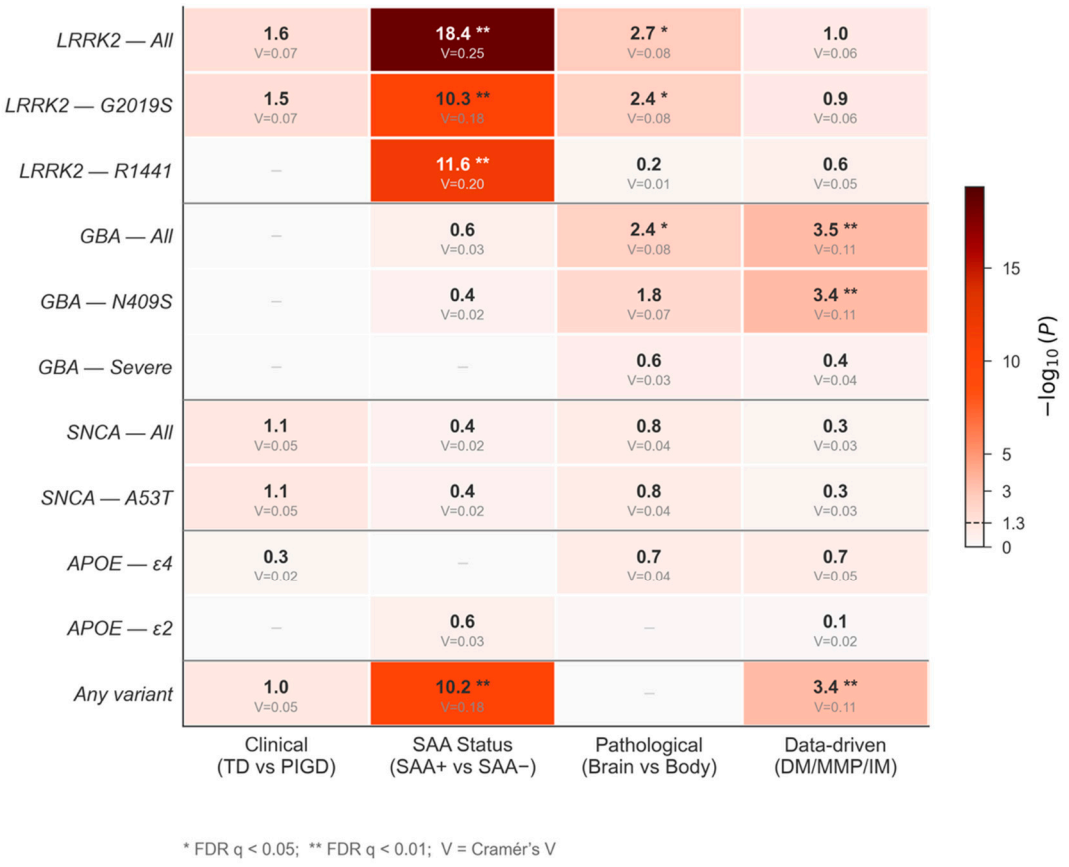

Figure 1. Genetic–Subtype Association Heatmap. Color intensity represents $-\log_{10}(P)$; Cramér's V effect sizes are annotated within cells. Stars indicate FDR-significant associations ($q < 0.05$).

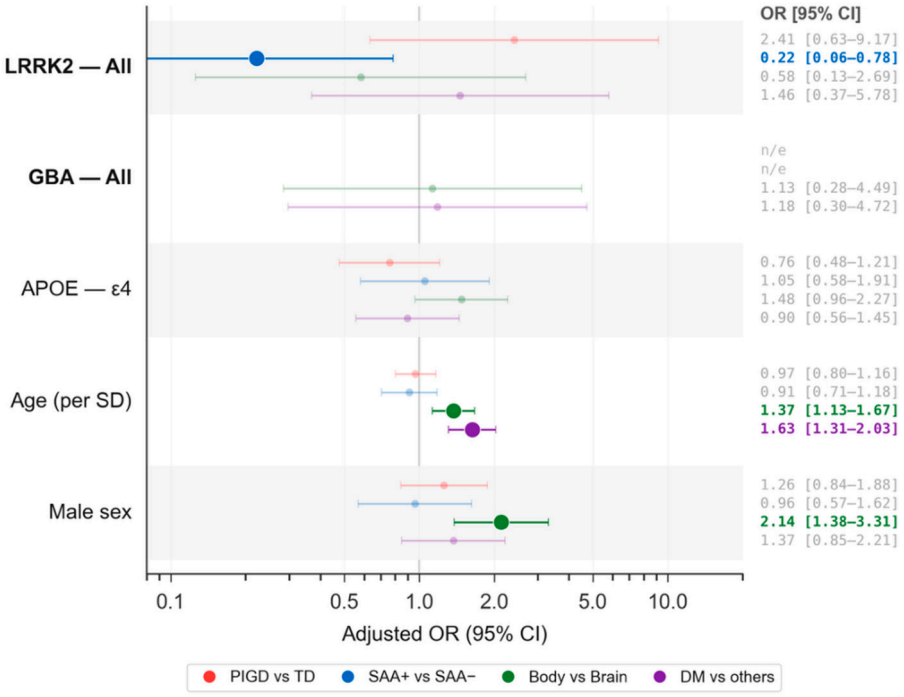

Figure 2. Adjusted Odds Ratios from Multivariable Logistic Regression. Table–forest hybrid showing adjusted ORs (95% Wald CI) for five predictors across four subtyping comparisons. Colored, bold entries indicate P <

 doi:10.20944/preprints202603.0734.v1

0.05. Notable associations: LRRK2 protective for SAA+ (OR = 0.22), male sex associated with body-first (OR = 2.14), age with body-first (OR = 1.37) and DM (OR = 1.63). Log scale.

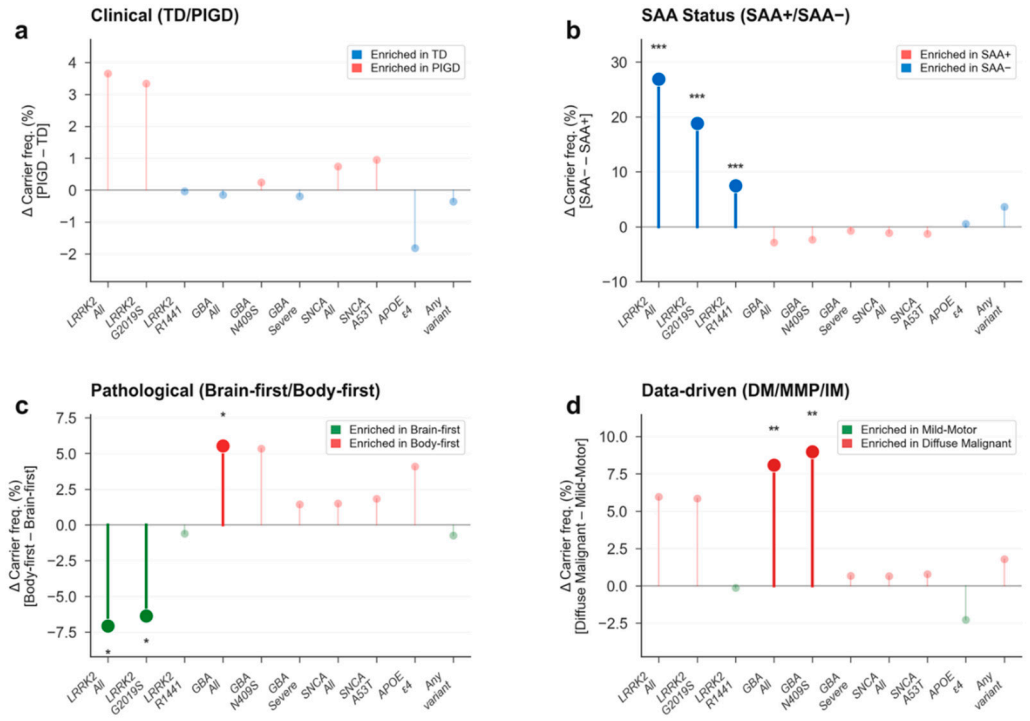

Figure 3. Diverging Carrier Frequency Differences Between Subtypes. Vertical diverging lollipop plots showing the difference in carrier frequency (%) between subtype groups. (a) Clinical (PIGD − TD). (b) SAA (SAA− − SAA+). (c) Pathological (Body-first − Brain-first). (d) Data-driven (DM − MMP). Stars indicate FDR significance ($q < 0.05$).

 doi:10.20944/preprints202603.0734.v1

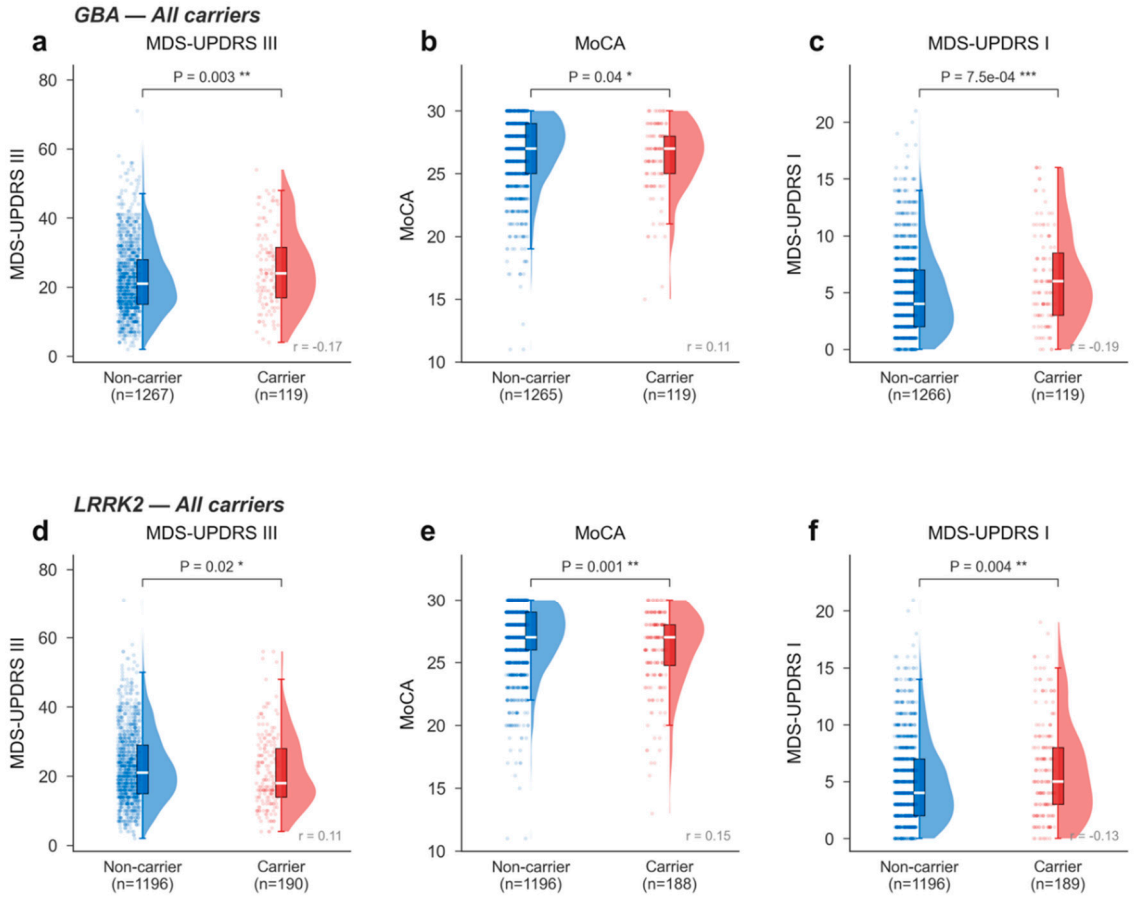

Figure 4. GBA and LRRK2 Carrier Clinical Subtype Profiles. Raincloud plots (half-violin + jitter strip + box) comparing baseline clinical measures between genetic carriers and non-carriers. (a–c) GBA carriers vs. non-carriers. (d–f) LRRK2 carriers vs. non-carriers.

Table 1. Baseline Demographics and Clinical Characteristics of the Study Population Categorized into Subtypes.

	Clinical Motor Subtypes		**SAA status**		**Pathological Subtypes**		**Data-driven Subtypes**		
	PIGD (n = 296)	**TD (n = 793)**	**SAA+ (n = 1,112)**	**SAA− (n = 156)**	**Body-first (n = 342)**	**Brain-first (n = 985)**	**DM (n = 322)**	**IM (n = 509)**	**MMP (n = 441)**
Age, years	63.2 ± 9.5	63.5 ± 9.5	62.6 ± 9.6	66.1 ± 9.6	63.3 ± 9.5	62.6 ± 9.9	64.0 ± 10.3	62.7 ± 9.1	59.8 ± 10.1
Male sex, n (%)	117 (40%)	317 (40%)	377 (34%)	49 (31.4%)	148 (43%)	338 (34%)	133 (41%)	208 (41%)	160 (36%)
Education, years	15.9 ± 3.4	16.2 ± 3.0	16.2 ± 3.2	15.1 ± 4.2	15.7 ± 3.4	16.1 ± 3.4	15.9 ± 3.6	16.1 ± 3.3	16.2 ± 3.2
MDS-UPDRS III	21.6 ± 9.9	22.6 ± 9.7	22.7 ± 9.9	20.6 ± 9.0	23.7 ± 11.4	21.8 ± 10.0	28.0 ± 12.1	21.3 ± 9.6	19.5 ± 8.3
MoCA	26.9 ± 2.4	26.9 ± 2.5	26.9 ± 2.5	26.0 ± 2.8	26.4 ± 2.9	26.8 ± 2.7	25.8 ± 2.9	26.3 ± 2.6	28.1 ± 1.3
H&Y stage	2 [1–2]	2 [1–2]	2 [1–2]	2 [1–2]	2 [1–2]	2 [1–2]	2 [2–2]	2 [1–2]	2 [1–2]
SAA Positive, n (%)	207 (83%)	647 (91%)	1112 (100%)	0 (0%)	240 (89%)	704 (86%)	218 (87%)	370 (88%)	351 (91%)
LRRK2 carrier, n (%)	19 (7.0%)	25 (3.4%)	109 (10.2%)	56 (37.1%)	24 (7.9%)	132 (15.0%)	42 (15.1%)	57 (12.8%)	37 (9.1%)
GBA carrier, n (%)	8 (3.0%)	23 (3.1%)	80 (7.5%)	7 (4.6%)	37 (12.3%)	59 (6.7%)	39 (14.0%)	28 (6.3%)	24 (5.9%)
SNCA carrier, n (%)	2 (0.7%)	0 (0.0%)	12 (1.1%)	0 (0.0%)	9 (3.0%)	13 (1.5%)	8 (2.9%)	7 (1.6%)	9 (2.2%)
PRKN carrier, n (%)	3 (1.1%)	12 (1.6%)	11 (1.0%)	3 (2.0%)	4 (1.3%)	10 (1.1%)	4 (1.4%)	4 (0.9%)	6 (1.5%)

 doi:10.20944/preprints202603.0734.v1

PINK1 carrier, n (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
PARK7 carrier, n (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
VPS35 carrier, n (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
LRRK2-G2019S, n (%)	17 (6.3%)	22 (3.0%)	103 (9.6%)	43 (28.5%)	21 (7.0%)	117 (13.3%)	40 (14.4%)	51 (11.4%)	35 (8.6%)
LRRK2-R1441G/C/H, n (%)	1 (0.4%)	3 (0.4%)	5 (0.5%)	12 (7.9%)	3 (1.0%)	14 (1.6%)	1 (0.4%)	6 (1.3%)	2 (0.5%)
GBA-N409S, n (%)	6 (2.8%)	16 (2.6%)	66 (7.2%)	6 (4.8%)	30 (12.0%)	49 (6.6%)	32 (13.6%)	23 (6.2%)	16 (4.6%)
GBA severe, n (%)	2 (0.9%)	7 (1.1%)	14 (1.5%)	1 (0.8%)	7 (2.8%)	10 (1.3%)	7 (3.0%)	5 (1.3%)	8 (2.3%)
SNCA-A53T, n (%)	2 (0.9%)	0 (0.0%)	12 (1.3%)	0 (0.0%)	9 (3.6%)	13 (1.8%)	8 (3.4%)	7 (1.9%)	9 (2.6%)
APOE ε4 carrier, n (%)	61 (22.7%)	179 (24.5%)	254 (24.0%)	37 (24.5%)	81 (26.9%)	199 (22.8%)	58 (20.8%)	119 (26.9%)	93 (23.1%)
APOE ε2 carrier, n (%)	41 (15.2%)	114 (15.6%)	151 (14.3%)	27 (17.9%)	43 (14.3%)	131 (15.0%)	41 (14.7%)	69 (15.6%)	52 (12.9%)

Values are mean ± SD, median [IQR], or n (%). TD, tremor-dominant; PIGD, postural instability/gait difficulty; SAA, alpha-synuclein seed amplification assay; Pathological subtype based on Borghammer (2021) model using RBDSQ total score (12 items, Q1–Q9): ≥ 6 = body-first; ≤ 3 = brain-first. Indeterminate patients (n = 233) excluded; DM, diffuse malignant; IM, intermediate; MMP, mild-motor predominant. Classification per Fereshtehnejad et al. (2017). Patients with incomplete data excluded; MDS-UPDRS, Movement Disorder Society–Unified Parkinson's Disease Rating Scale; MoCA, Montreal Cognitive Assessment; H&Y, Hoehn & Yahr; SAA, seed amplification assay. Indeterminate patients (n = 131) excluded.

Table 2. Genetic Variant Carrier Frequencies by Clinical Motor Subtype (TD vs. PIGD).

Genetic Variant	PIGD, n/N (%)	TD, n/N (%)	Test	P-value	q-value	Cramér's V
LRRK2 — All	19/270 (7.0)	25/739 (3.4)	χ^2	0.024	0.095	0.07
GBA — All	8/270 (3.0)	23/739 (3.1)	χ^2	1.0	1.0	0.00
SNCA — All	2/270 (0.7)	0/739 (0.0)	Fisher	0.074	0.236	0.05
PRKN	3/270 (1.1)	12/739 (1.6)	Fisher	0.771	0.949	0.01
PINK1	0/270 (0.0)	0/739 (0.0)	—	—	—	—
PARK7	0/270 (0.0)	0/739 (0.0)	—	—	—	—
VPS35	0/270 (0.0)	0/739 (0.0)	—	—	—	—
Any variant	32/270 (11.9)	59/739 (8.0)	χ^2	0.096	0.288	0.05
APOE — ε4	61/269 (22.7)	179/731 (24.5)	χ^2	0.539	0.762	0.02
APOE — ε2	41/269 (15.2)	114/731 (15.6)	χ^2	0.902	1.0	0.00
LRRK2 — G2019S	17/269 (6.3)	22/739 (3.0)	χ^2	0.031	0.113	0.07
LRRK2 — R1441	1/269 (0.4)	3/739 (0.4)	Fisher	1.0	1.0	0.00
GBA — N409S	6/211 (2.8)	16/616 (2.6)	χ^2	1.0	1.0	0.00
GBA — Severe	2/211 (0.9)	7/616 (1.1)	Fisher	1.0	1.0	0.00
SNCA — A53T	2/211 (0.9)	0/616 (0.0)	Fisher	0.074	0.236	0.05

Carrier frequencies reported as n/N (%). χ^2, Pearson's chi-square test; Fisher, Fisher's exact test. q-values are Benjamini–Hochberg FDR-adjusted P-values across all 48 comparisons. Cramér's V quantifies effect size. — indicates not estimable (zero carriers in both groups). Bold q-values < 0.05 indicate FDR-significant associations.

Table 3. Genetic Variant Carrier Frequencies by SAA Status (SAA+ vs. SAA−).

Genetic Variant	SAA+, n/N (%)	SAA−, n/N (%)	Test	P-value	q-value	Cramér's V
LRRK2 — All	109/1069 (10.2)	56/151 (37.1)	χ^2	3.7×10^{-19}	< 0.001	0.25
GBA — All	80/1069 (7.5)	7/151 (4.6)	χ^2	0.279	0.514	0.03
SNCA — All	12/1069 (1.1)	0/151 (0.0)	Fisher	0.381	0.614	0.02
PRKN	11/1069 (1.0)	3/151 (2.0)	Fisher	0.400	0.620	0.02
PINK1	0/1069 (0.0)	0/151 (0.0)	—	—	—	—
PARK7	0/1069 (0.0)	0/151 (0.0)	—	—	—	—

 doi:10.20944/preprints202603.0734.v1

VPS35	0/1069 (0.0)	0/151 (0.0)	—	—	—	—
Any variant	205/1069 (19.2)	65/151 (43.0)	χ^2	6.4×10^{-11}	< 0.001	0.18
APOE — ε4	254/1059 (24.0)	37/151 (24.5)	χ^2	0.887	1.0	0.00
APOE — ε2	151/1059 (14.3)	27/151 (17.9)	χ^2	0.257	0.494	0.03
LRRK2 — G2019S	103/1069 (9.6)	43/151 (28.5)	χ^2	4.9×10^{-11}	< 0.001	0.18
LRRK2 — R1441	5/1069 (0.5)	12/151 (7.9)	χ^2	2.7×10^{-12}	< 0.001	0.20
GBA — N409S	66/916 (7.2)	6/124 (4.8)	χ^2	0.384	0.614	0.02
GBA — Severe	14/916 (1.5)	1/124 (0.8)	Fisher	1.0	1.0	0.01
SNCA — A53T	12/916 (1.3)	0/124 (0.0)	Fisher	0.381	0.614	0.02

Carrier frequencies reported as n/N (%). χ^2, Pearson's chi-square test; Fisher, Fisher's exact test. q-values are Benjamini–Hochberg FDR-adjusted P-values across all 48 comparisons. Cramér's V quantifies effect size. — indicates not estimable.

Table 4. Genetic Variant Carrier Frequencies by Pathological Subtype (Body-first vs. Brain-first).

Genetic Variant	Body-first, n/N (%)	Brain-first, n/N (%)	Test	P-value	q-value	Cramér's V
LRRK2 — All	24/302 (7.9)	132/879 (15.0)	χ^2	0.002	0.013	0.08
GBA — All	37/302 (12.3)	59/879 (6.7)	χ^2	0.004	0.021	0.08
SNCA — All	9/302 (3.0)	13/879 (1.5)	χ^2	0.164	0.394	0.04
PRKN	4/302 (1.3)	10/879 (1.1)	Fisher	0.764	0.949	0.00
PINK1	0/302 (0.0)	0/879 (0.0)	—	—	—	—
PARK7	0/302 (0.0)	0/879 (0.0)	—	—	—	—
VPS35	0/302 (0.0)	0/879 (0.0)	—	—	—	—
Any variant	72/302 (23.8)	210/879 (23.9)	χ^2	0.978	1.0	0.00
APOE — ε4	81/301 (26.9)	199/872 (22.8)	χ^2	0.200	0.436	0.04
APOE — ε2	43/301 (14.3)	131/872 (15.0)	χ^2	0.803	0.963	0.01
LRRK2 — G2019S	21/302 (7.0)	117/878 (13.3)	χ^2	0.004	0.020	0.08
LRRK2 — R1441	3/302 (1.0)	14/878 (1.6)	Fisher	0.583	0.799	0.01
GBA — N409S	30/251 (12.0)	49/741 (6.6)	χ^2	0.015	0.067	0.07
GBA — Severe	7/251 (2.8)	10/741 (1.3)	χ^2	0.237	0.494	0.03
SNCA — A53T	9/251 (3.6)	13/741 (1.8)	χ^2	0.164	0.394	0.04

Carrier frequencies reported as n/N (%). χ^2, Pearson's chi-square test; Fisher, Fisher's exact test. q-values are Benjamini–Hochberg FDR-adjusted P-values across all 48 comparisons. Cramér's V quantifies effect size. — indicates not estimable.

Table 5. Genetic Variant Carrier Frequencies by Data-Driven Subtype (DM vs. IM vs. MMP).

Genetic Variant	DM, n/N (%)	IM, n/N (%)	MMP, n/N (%)	Test	P-value	q-value	Cramér's V
LRRK2 — All	42/279 (15.1)	57/447 (12.8)	37/407 (9.1)	χ^2	0.108	0.305	0.06
GBA — All	39/279 (14.0)	28/447 (6.3)	24/407 (5.9)	χ^2	3.4×10^{-4}	0.003	0.11
SNCA — All	8/279 (2.9)	7/447 (1.6)	9/407 (2.2)	χ^2	0.497	0.723	0.03
PRKN	4/279 (1.4)	4/447 (0.9)	6/407 (1.5)	χ^2	0.672	0.895	0.03
PINK1	0/279 (0.0)	0/447 (0.0)	0/407 (0.0)	—	—	—	—
PARK7	0/279 (0.0)	0/447 (0.0)	0/407 (0.0)	—	—	—	—
VPS35	0/279 (0.0)	0/447 (0.0)	0/407 (0.0)	—	—	—	—
Any variant	90/279 (32.3)	95/447 (21.3)	74/407 (18.2)	χ^2	3.7×10^{-4}	0.003	0.11
APOE — ε4	58/279 (20.8)	119/443 (26.9)	93/403 (23.1)	χ^2	0.182	0.417	0.05
APOE — ε2	41/279 (14.7)	69/443 (15.6)	52/403 (12.9)	χ^2	0.718	0.932	0.02
LRRK2 — G2019S	40/277 (14.4)	51/447 (11.4)	35/407 (8.6)	χ^2	0.122	0.325	0.06
LRRK2 — R1441	1/277 (0.4)	6/447 (1.3)	2/407 (0.5)	χ^2	0.255	0.494	0.05
GBA — N409S	32/235 (13.6)	23/371 (6.2)	16/345 (4.6)	χ^2	3.6×10^{-4}	0.003	0.11
GBA — Severe	7/235 (3.0)	5/371 (1.3)	8/345 (2.3)	χ^2	0.356	0.614	0.04
SNCA — A53T	8/236 (3.4)	7/371 (1.9)	9/345 (2.6)	χ^2	0.497	0.723	0.03

 doi:10.20944/preprints202603.0734.v1

Carrier frequencies reported as n/N (%). DM, diffuse malignant; IM, intermediate; MMP, mild-motor predominant. χ^2, Pearson's chi-square test; Fisher, Fisher's exact test. q-values are Benjamini–Hochberg FDR-adjusted P-values across all 48 comparisons. Cramér's V quantifies effect size. — indicates not estimable.

4. Discussion

This study provides a comprehensive examination of the genetic correlates of four complementary PD subtyping frameworks in the PPMI cohort, with specific variant-level resolution for the major PD genes and inclusion of APOE genotype analysis. Our findings reveal that certain subtyping schemes capture distinct genetic architectures more effectively than others, with 10 of 48 univariate tests surviving FDR correction.

The SAA-based classification yielded the most robust genetic associations. SAA-negative PD patients showed a strikingly high prevalence of LRRK2 variants (37.1%), driven by both G2019S (28.5%) and R1441G/C/H (7.9%). The effect size was substantial (Cramér's V = 0.25), and all SAA–LRRK2 associations survived FDR correction with $q < 0.001$. This is consistent with the known biology of LRRK2-PD, which often shows less alpha-synuclein pathology compared to idiopathic PD. Autopsy studies have established that an appreciable subset of LRRK2-PD patients lack Lewy body pathology despite having dopaminergic neuron loss indistinguishable from sporadic PD, instead exhibiting alternative proteinopathies such as tau or TDP-43 aggregation [27,28]. In the PPMI cohort, Siderowf et al.[29] reported that only 68% of LRRK2-PD participants were alpha-synuclein SAA-positive, mirroring the frequency of typical Lewy pathology observed at autopsy, while 96% of GBA-PD and 93% of sporadic PD cases tested positive. More recently, a Brain Communications study demonstrated that one-third of LRRK2 parkinsonism cases had no in vivo evidence of alpha-synuclein aggregates based on CSF SAA testing, contrasting sharply with only 7–9% SAA-negativity in sporadic PD [30].

The very high enrichment of R1441G/C/H in SAA− (7.9% vs. 0.5%, $P = 2.7 \times 10^{-12}$, $q < 0.001$) suggests these rarer variants may have an even stronger association with SAA-negativity than G2019S. In adjusted analysis, LRRK2 carrier status was the only significant predictor of SAA status (OR = 0.22 [0.06–0.78], P = 0.02). This differential effect by variant type aligns with reports that R1441G/C/H carriers present a more homogeneous subtype with higher rates of preserved olfaction and potentially lower alpha-synuclein burden than G2019S carriers [31].

GBA variants, predominantly N409S, were enriched in body-first PD (12.3% vs. 6.7%, $q = 0.021$) and DM (14.0% vs. MMP 5.9%, $q = 0.003$). This aligns with the known association of GBA with more aggressive disease, including faster cognitive decline, earlier autonomic dysfunction, and greater non-motor burden. Indeed, GBA carriers in our cohort showed worse baseline motor scores, cognitive performance, and non-motor burden (Figure 4). GBA-PD has been consistently associated with earlier disease onset, more frequent cognitive impairment, neuropsychiatric symptoms, and faster motor–cognitive progression compared with idiopathic PD, with a genotype-severity dependence whereby severe GBA variants confer more aggressive subtypes [32–36]. Additionally, GBA carriers with PD show more widespread Lewy body distribution across cortical and brainstem regions [37,38], and GBA mutation severity correlates with CSF alpha-synuclein seeding kinetics [39,40].

Body-first PD, characterized by prominent RBD and autonomic dysfunction, may represent a subtype where peripheral alpha-synuclein spread predominates, and GBA may facilitate this spread through impaired lysosomal function. Mechanistically, GBA mutations cause loss of glucocerebrosidase activity and downstream lysosomal-autophagic dysfunction, leading to glycosphingolipid accumulation that promotes alpha-synuclein aggregation and impairs its clearance [41,42]. This pathway provides a plausible biological link between GBA carrier status and body-first PD, where peripheral alpha-synuclein propagation through the autonomic nervous system—potentially facilitated by impaired lysosomal processing—may underlie the prominent RBD and autonomic dysfunction that characterize this subtype [17,19].

LRRK2 variants showed a mirror-image pattern to GBA: enrichment in brain-first (15.0% vs. 7.9% body-first, $q = 0.013$) and SAA-negative (37.1% vs. 10.2%, $q < 0.001$) PD. This suggests LRRK2-

 doi:10.20944/preprints202603.0734.v1

PD represents a distinct biological entity with less alpha-synuclein pathology and predominant nigral-striatal degeneration, consistent with the brain-first model. This interpretation is supported by the observation that LRRK2-PD patients without alpha-synuclein aggregates (SAA-negative) show a pattern of subcortical atrophy with relatively preserved olfaction, distinct from the limbic and brainstem pathology typical of body-first synucleinopathy [22]. It also aligns with neuropathological evidence that LRRK2 dysfunction may drive neurodegeneration through alpha-synuclein-independent mechanisms involving kinase-mediated disruption of autophagy and vesicle trafficking [43]. The variant-level analysis revealed that G2019S accounts for the majority of this association, but R1441G/C/H variants show an even stronger per-carrier effect (Figure 1, Figure 3b).

The TD/PIGD classification showed only nominal LRRK2 enrichment in PIGD (7.0% vs. 3.4%, P = 0.024), but this did not survive FDR correction (q = 0.095). No other genetic markers differentiated these subtypes. The low cross-scheme agreement between Clinical and other frameworks (V = 0.05–0.25; Supplementary Figure S2) suggests that clinical motor subtype may not cleanly map onto distinct genetic etiologies. This is consistent with prior work showing that the TD/PIGD classification is unstable over time, with a substantial proportion of patients switching subtypes during longitudinal follow-up, suggesting it may capture a disease-stage-dependent dimension rather than a fixed biological entity [11]. Dulski et al. [5] similarly found only modest genome-wide associations with clinical motor subtypes compared with more robust genetic signals for non-motor-based classifications.

APOE ε4 carrier frequency and full genotype distributions were remarkably consistent across all four frameworks. While APOE ε4 is a major risk factor for Alzheimer's disease, our findings do not support its role in defining PD subtypes. This is consistent with prior studies showing that APOE ε4's role in PD is primarily in modifying cognitive outcomes rather than defining motor or pathological subtypes. Multiple studies have linked APOE ε4 to faster cognitive decline in PD [44,45], while associations with motor severity and motor subtypes have been largely inconclusive [46–50]. A recent PPMI-based longitudinal study confirmed that APOE ε4 accelerates cognitive decline specifically in sporadic PD but not in GBA1-PD or LRRK2-PD, underscoring the subtype-specific effects of APOE on cognition rather than motor subtype [51]. Taken together, these findings confirm that APOE is a modifier of cognitive trajectory rather than a determinant of PD subtype membership.

The four subtyping frameworks showed generally low concordance, indicating they capture different aspects of PD biology. The highest agreement was between Data-driven and Pathological subtypes (V = 0.29), consistent with the overlap in non-motor features (particularly RBD) used by both approaches. The Clinical framework showed particularly low agreement with all others (V = 0.05–0.25), reinforcing that motor subtype classification captures a distinct—and genetically less informative—dimension of heterogeneity. This low concordance between subtyping frameworks has been observed previously by Chen et al.[52], who compared multiple data-driven PD subtyping methods and found that agreement was highly dependent on the clinical domains incorporated. Our genetic data provide a complementary perspective: subtyping frameworks that incorporate non-motor features and biological markers (SAA, RBD, multi-domain composites) capture more genetically informative dimensions of PD heterogeneity than purely motor-based classifications.

Our findings have direct implications for clinical trial design. First, SAA status should be considered a primary stratification variable, given its strong association with LRRK2 carrier status and the largest effect sizes among all tested associations. This recommendation aligns with the recent FDA letter of support endorsing alpha-synuclein SAA as a biomarker for PD clinical trials and the proposed Neuronal Synuclein Disease Integrated Staging System (NSD-ISS), which uses SAA as a foundational biological criterion for enrollment and staging [53]. Second, GBA-targeted therapies may be most relevant for body-first and DM subtypes, which show the highest carrier frequencies. Several GBA-targeted disease-modifying therapies are under development, including small-molecule GCase activators, substrate reduction therapies, and gene therapy approaches [32,34,38,54]. Our data suggest that body-first and DM subtypes, where GBA carriers are most concentrated, may represent optimal enrichment populations for such trials. Third, the lack of APOE-subtype

 doi:10.20944/preprints202603.0734.v1

associations suggests APOE need not be a stratification variable for PD motor subtype trials, though it remains relevant for cognitive outcomes. Fourth, the weak genetic differentiation of TD/PIGD suggests this classification may be less useful for genotype-stratified trials.

This study has several limitations. The PPMI is an enrichment cohort with overrepresentation of genetic PD relative to the general PD population. The cross-sectional design limits inference about temporal relationships between genetic status and subtype evolution. Some subtyping algorithms require specific clinical data that may introduce selection bias. The pathological subtype classification uses RBDSQ as a proxy for RBD, which has imperfect sensitivity and specificity for polysomnography-confirmed RBD. The logistic regression models for some comparisons exhibited quasi-complete separation due to low carrier frequencies in certain subtype groups, preventing estimation of some odds ratios. Despite this, the primary findings (SAA–LRRK2, GBA–body-first, GBA–DM) survived FDR correction across all 48 univariate tests.

5. Conclusions

We demonstrate that SAA status, pathological subtype, and data-driven subtype classifications capture distinct genetic architectures. LRRK2 variants (G2019S, R1441G/C/H) strongly associate with SAA-negativity ($V = 0.25$, $q < 0.001$) and brain-first PD subtype ($q = 0.013$). GBA variants (N409S, severe) associate with body-first ($q = 0.021$) and diffuse malignant subtypes ($q = 0.003$). Clinical motor subtypes (TD/PIGD) shows only nominal genetic differences that do not survive multiple comparison correction. APOE genotype does not differentiate PD subtypes. These findings support the biological validity of pathology-based and data-driven PD subtypes and have implications for genotype-stratified clinical trial design.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, A.N.; methodology, A.N. and M.A.; software, A.N.; validation, A.N.; formal analysis, A.N., M.A., B.M.H, Y.H.; investigation, A.N.; resources, A.N.; data curation, A.N.; writing—original draft preparation, A.N., M.A., B.M.H, and Y.H.; writing—review and editing, A.N., M.A., B.M.H, Y.H., A.D., Y.N., M.B., and B.B.; visualization, A.N.; supervision, M.B., and B.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was approved by institutional review boards at all participating sites in the PPMI studies, and all participants provided written informed consent.

Informed Consent Statement: As per the PPMI protocols, informed consent was obtained from all participants involved in the study.

Data Availability Statement: PPMI data are available here upon request: https://www.ppmi-info.org/.

Acknowledgments: Data used in the preparation of this article were obtained on 2024-09-28 from the Parkinson's Progression Markers Initiative (PPMI) database (https://www.ppmi-info.org/access-data-specimens/download-data), RRID:SCR_006431. For up-to-date information on the study, visit http://www.ppmi-info.org. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including 4D Pharma, Abbvie, AcureX, Allergan, Amathus Therapeutics, Aligning Science Across Parkinson's, AskBio, Avid Radiopharmaceuticals, BIAL, BioArctic, Biogen, Biohaven, BioLegend, BlueRock Therapeutics, Bristol-Myers Squibb, Calico Labs, Capsida Biotherapeutics, Celgene, Cerevel Therapeutics, Coave Therapeutics, DaCapo Brainscience, Denali, Edmond J. Safra Foundation, Eli Lilly, Gain Therapeutics, GE HealthCare, Genentech, GSK, Golub Capital, Handl Therapeutics, Insitro, Jazz Pharmaceuticals, Johnson & Johnson Innovative Medicine, Lundbeck, Merck, Meso Scale Discovery, Mission Therapeutics, Neurocrine Biosciences, Neuron23, Neuropore, Pfizer, Piramal, Prevail Therapeutics, Roche,

 doi:10.20944/preprints202603.0734.v1

Sanofi, Servier, Sun Pharma Advanced Research Company, Takeda, Teva, UCB, Vanqua Bio, Verily, Voyager Therapeutics, the Weston Family Foundation and Yumanity Therapeutics.

Conflicts of Interest: Ahmed Negida and Matthew J. Barrett disclose support for the research of this work from NIH/NIA 1R01NS142622 and 1R21AG077469-01. Ahmed Negida has received a grant from the International Parkinson and Movement Disorder Society. Brian D. Berman has received research grant support from the Dystonia Coalition (receives the majority of its support through NIH grant NS065701 from the Office of Rare Diseases Research in the National Center for Advancing Translational Science and National Institute of Neurological Disorders and Stroke), the Parkinson's Foundation, the VCU School of Medicine, the Administration for Community Living, and the Dystonia Medical Research Foundation. He has served as a consultant for the Dystonia Medical Research Foundation and has received honoraria from the International Parkinson and Movement Disorder Society. He currently serves on the Medical and Scientific Advisory Council of the Dystonia Medical Research Foundation as well as the director of the Medical Advisory Board of the Benign Essential Blepharospasm Research Foundation and member of the National Spasmodic Torticollis Association.

Abbreviations

The following abbreviations are used in this manuscript:

PD — Parkinson's disease
TD —Tremor-dominant
PIGD — Postural instability/gait difficulty
SAA — Seed amplification assay
DM — Diffuse malignant
MMP — Mild-motor predominant
IM — Intermediate
PPMI — Parkinson's Progression Markers Initiative
LRRK2 — Leucine-rich repeat kinase 2
GBA — Glucocerebrosidase
SNCA— Synuclein alpha
APOE — Apolipoprotein E
MDS-UPDRS — Movement Disorder Society–Unified Parkinson's Disease Rating Scale
MoCA — Montreal Cognitive Assessment
RBD — Rapid eye movement sleep behavior disorder
RBDSQ — RBD Screening Questionnaire
CSF — Cerebrospinal fluid
CLIA — Clinical Laboratory Improvement Amendments
OR — Odds ratio
CI — Confidence interval
FDR — False discovery rate
AIC — Akaike information criterion
IQR — Interquartile range
SD — Standard deviation
SCOPA-AUT — Scales for Outcomes in Parkinson's Disease–Autonomic

References

1. de Lau LML, Breteler MMB. Epidemiology of Parkinson's disease. Lancet Neurol 2006;5:525–35. https://doi.org/10.1016/S1474-4422(06)70471-9.
2. Grotewold N, Albin RL. Update: Descriptive epidemiology of Parkinson disease. Parkinsonism & Related Disorders 2024:106000. https://doi.org/10.1016/j.parkreldis.2024.106000.

 doi:10.20944/preprints202603.0734.v1

3. Negida A, Altamimi FQ, Elsayed M, Ebuid YRI, Nabil Y, Parker MO, et al. Chapter 14 - Epidemiology and risk factors of Parkinson's disease. In: Mohamed WMY, editor. Essential Guide to Neurodegenerative Disorders, Academic Press; 2025, p. 225–34. https://doi.org/10.1016/B978-0-443-15702-8.00014-2.
4. Albrecht F, Poulakis K, Freidle M, Johansson H, Ekman U, Volpe G, et al. Unraveling Parkinson's disease heterogeneity using subtypes based on multimodal data. Parkinsonism Relat Disord 2022;102:19–29. https://doi.org/10.1016/j.parkreldis.2022.07.014.
5. Dulski J, Uitti RJ, Beasley A, Hernandez D, Ramanan VK, Cahn EJ, et al. Genetics of Parkinson's disease heterogeneity: A genome-wide association study of clinical subtypes. Parkinsonism Relat Disord 2024;119:105935. https://doi.org/10.1016/j.parkreldis.2023.105935.
6. Wüllner U, Borghammer P, Choe C, Csoti I, Falkenburger B, Gasser T, et al. The heterogeneity of Parkinson's disease. J Neural Transm (Vienna) 2023;130:827–38. https://doi.org/10.1007/s00702-023-02635-4.
7. von Coelln R, Shulman LM. Clinical subtypes and genetic heterogeneity: of lumping and splitting in Parkinson disease. Curr Opin Neurol 2016;29:727–34. https://doi.org/10.1097/WCO.0000000000000384.
8. Mastenbroek SE, Vogel JW, Collij LE, Serrano GE, Tremblay C, Young AL, et al. Disease progression modelling reveals heterogeneity in trajectories of Lewy-type α-synuclein pathology. Nat Commun 2024;15:5133. https://doi.org/10.1038/s41467-024-49402-x.
9. Adams C, Suescun J, Haque A, Block K, Chandra S, Ellmore TM, et al. Updated Parkinson's disease motor subtypes classification and correlation to cerebrospinal homovanillic acid and 5-hydroxyindoleacetic acid levels. Clin Park Relat Disord 2023;8:100187. https://doi.org/10.1016/j.prdoa.2023.100187.
10. Jankovic J, McDermott M, Carter J, Gauthier S, Goetz C, Golbe L, et al. Variable expression of Parkinson's disease. Neurology 1990;40:1529–1529. https://doi.org/10.1212/WNL.40.10.1529.
11. Ren J, Pan C, Li Y, Li L, Hua P, Xu L, et al. Consistency and Stability of Motor Subtype Classifications in Patients With de novo Parkinson's Disease. Front Neurosci 2021;15:637896. https://doi.org/10.3389/fnins.2021.637896.
12. Chen Q, Scherbaum R, Gold R, Pitarokoili K, Mosig A, Zella S, et al. Data-driven subtyping of Parkinson's disease: comparison of current methodologies and application to the Bochum PNS cohort. J Neural Transm (Vienna) 2023;130:763–76. https://doi.org/10.1007/s00702-023-02627-4.
13. Deng X, Saffari SE, Xiao B, Ng SYE, Chia N, Choi X, et al. Disease Progression of Data-Driven Subtypes of Parkinson's Disease: 5-Year Longitudinal Study from the Early Parkinson's Disease Longitudinal Singapore (PALS) Cohort. J Parkinsons Dis 2024;14:1051–9. https://doi.org/10.3233/JPD-230209.
14. Negida A, Vohra HZ, Lageman SK, Mukhopadhyay N, Berman BD, Weintraub D, et al. Parkinson's Disease Mild Cognitive Impairment with MRI evidence of Cholinergic Nucleus 4 Degeneration: A New Subtype? Parkinsonism & Related Disorders 2025;0. https://doi.org/10.1016/j.parkreldis.2025.108072.
15. Fereshtehnejad S-M, Zeighami Y, Dagher A, Postuma RB. Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression. Brain 2017;140:1959–76. https://doi.org/10.1093/brain/awx118.
16. Fereshtehnejad S-M, Romenets SR, Anang JBM, Latreille V, Gagnon J-F, Postuma RB. New Clinical Subtypes of Parkinson Disease and Their Longitudinal Progression: A Prospective Cohort Comparison With Other Phenotypes. JAMA Neurol 2015;72:863–73. https://doi.org/10.1001/jamaneurol.2015.0703.
17. Bohnen NI, Postuma RB. Body-first versus brain-first biological subtyping of Parkinson's disease. Brain 2020;143:2871–3. https://doi.org/10.1093/brain/awaa293.
18. Borghammer P, Horsager J, Andersen K, Van Den Berge N, Raunio A, Murayama S, et al. Neuropathological evidence of body-first vs. brain-first Lewy body disease. Neurobiology of Disease 2021;161:105557. https://doi.org/10.1016/j.nbd.2021.105557.
19. Borghammer P, Van Den Berge N. Brain-First versus Gut-First Parkinson's Disease: A Hypothesis. Journal of Parkinson's Disease 2019;9:S281–95. https://doi.org/10.3233/JPD-191721.
20. Stebbins GT, Goetz CG, Burn DJ, Jankovic J, Khoo TK, Tilley BC. How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale. Mov Disord 2013;28:668–70. https://doi.org/10.1002/mds.25383.

 doi:10.20944/preprints202603.0734.v1

21. Brockmann K, Lerche S, Baiardi S, Rossi M, Wurster I, Quadalti C, et al. CSF α-synuclein seed amplification kinetic profiles are associated with cognitive decline in Parkinson's disease. Npj Parkinsons Dis 2024;10:1–9. https://doi.org/10.1038/s41531-023-00627-5.
22. Brooker SM, Pasquini J, Choi SH, Lafontant D-E, Fereshtehnejad S-M, Zeighami Y, et al. Clinical and Imaging Characteristics of Parkinson's Disease with Negative Alpha-Synuclein Seed Amplification Assay. Mov Disord 2026. https://doi.org/10.1002/mds.70197.
23. Coughlin DG, Shifflett B, Farris CM, Ma Y, Galasko D, Edland SD, et al. α-Synuclein Seed Amplification Assay Amplification Parameters and the Risk of Progression in Prodromal Parkinson Disease. Neurology 2025;104:e210279. https://doi.org/10.1212/WNL.0000000000210279.
24. Xu Z, Hu T, Xu C, Liang X, Li S, Sun Y, et al. Disease progression in proposed brain-first and body-first Parkinson's disease subtypes. NPJ Parkinsons Dis 2024;10:111. https://doi.org/10.1038/s41531-024-00730-1.
25. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 1995;57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
26. Feller W. Review: Harald Cramer, Mathematical Methods of Statistics. The Annals of Mathematical Statistics 1947;18:136–9. https://doi.org/10.1214/aoms/1177730503.
27. O'Hara DM, Pawar G, Kalia SK, Kalia LV. LRRK2 and α-Synuclein: Distinct or Synergistic Players in Parkinson's Disease? Frontiers in Neuroscience 2020;14:577. https://doi.org/10.3389/fnins.2020.00577.
28. Kalia LV, Lang AE, Hazrati L-N, Fujioka S, Wszolek ZK, Dickson DW, et al. Clinical correlations with Lewy body pathology in LRRK2-related Parkinson disease. JAMA Neurology 2015;72:100–5. https://doi.org/10.1001/jamaneurol.2014.2704.
29. Siderowf A, Concha-Marambio L, Lafontant D-E, Farris CM, Ma Y, Urenia PA, et al. Assessment of heterogeneity among participants in the Parkinson's Progression Markers Initiative cohort using α-synuclein seed amplification: a cross-sectional study. Lancet Neurol 2023;22:407–17. https://doi.org/10.1016/S1474-4422(23)00109-6.
30. Chahine LM, Lafontant D-E, Choi SH, Iwaki H, Blauwendraat C, Singleton AB, et al. LRRK2-associated parkinsonism with and without in vivo evidence of alpha-synuclein aggregates: longitudinal clinical and biomarker characterization. Brain Commun 2025;7:fcaf103. https://doi.org/10.1093/braincomms/fcaf103.
31. Vinagre-Aragón A, Campo-Caballero D, Mondragón-Rezola E, Pardina-Vilella L, Hernandez Eguiazu H, Gorostidi A, et al. A More Homogeneous Phenotype in Parkinson's Disease Related to R1441G Mutation in the LRRK2 Gene. Front Neurol 2021;12. https://doi.org/10.3389/fneur.2021.635396.
32. Höglinger G, Schulte C, Jost WH, Storch A, Woitalla D, Krüger R, et al. GBA-associated PD: chances and obstacles for targeted treatment strategies. Journal of Neural Transmission 2022;129:1219–33. https://doi.org/10.1007/s00702-022-02511-7.
33. Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. New England Journal of Medicine 2009;361:1651–61. https://doi.org/10.1056/NEJMoa0901281.
34. Huh YE, Usnich T, Scherzer CR, Klein C, Chung SJ. GBA1 Variants and Parkinson's Disease: Paving the Way for Targeted Therapy. J Mov Disord 2023;16:261–78. https://doi.org/10.14802/jmd.23023.
35. Maple-Grødem J, Dalen I, Tysnes O-B, Macleod AD, Forsgren L, Counsell CE, et al. Association of GBA Genotype With Motor and Functional Decline in Patients With Newly Diagnosed Parkinson Disease. Neurology 2021;96:e1036–44. https://doi.org/10.1212/WNL.0000000000011411.
36. Menozzi E, Schapira AHV. Exploring the Genotype–Phenotype Correlation in GBA-Parkinson Disease: Clinical Aspects, Biomarkers, and Potential Modifiers. Front Neurol 2021;12. https://doi.org/10.3389/fneur.2021.694764.
37. Walton RL, Koga S, Beasley AI, White LJ, Griesacker T, Murray ME, et al. Role of GBA variants in Lewy body disease neuropathology. Acta Neuropathol 2024;147:54. https://doi.org/10.1007/s00401-024-02699-w.
38. Smith L, Schapira AHV. GBA Variants and Parkinson Disease: Mechanisms and Treatments. Cells 2022;11:1261. https://doi.org/10.3390/cells11081261.

 doi:10.20944/preprints202603.0734.v1

39. Brockmann K, Lerche S, Baiardi S, Rossi M, Wurster I, Quadalti C, et al. CSF α-synuclein seed amplification kinetic profiles are associated with cognitive decline in Parkinson's disease. NPJ Parkinsons Dis 2024;10:24. https://doi.org/10.1038/s41531-023-00627-5.
40. Brockmann K, Quadalti C, Lerche S, Rossi M, Wurster I, Baiardi S, et al. Association between CSF alpha-synuclein seeding activity and genetic status in Parkinson's disease and dementia with Lewy bodies. Acta Neuropathol Commun 2021;9:175. https://doi.org/10.1186/s40478-021-01276-6.
41. Stoker TB, Torsney KM, Barker RA. Pathological Mechanisms and Clinical Aspects of GBA1 Mutation-Associated Parkinson's Disease. In: Stoker TB, Greenland JC, editors. Parkinson's Disease: Pathogenesis and Clinical Aspects, Brisbane (AU): Codon Publications; 2018.
42. Granek Z, Barczuk J, Siwecka N, Rozpędek-Kamińska W, Kucharska E, Majsterek I. GBA1 Gene Mutations in α-Synucleinopathies—Molecular Mechanisms Underlying Pathology and Their Clinical Significance. International Journal of Molecular Sciences 2023;24:2044. https://doi.org/10.3390/ijms24032044.
43. Lamonaca G, Volta M. Alpha-Synuclein and LRRK2 in Synaptic Autophagy: Linking Early Dysfunction to Late-Stage Pathology in Parkinson's Disease. Cells 2020;9:1115. https://doi.org/10.3390/cells9051115.
44. Aarsland D, Batzu L, Halliday GM, Geurtsen GJ, Ballard C, Ray Chaudhuri K, et al. Parkinson disease-associated cognitive impairment. Nat Rev Dis Primers 2021;7:1–21. https://doi.org/10.1038/s41572-021-00280-3.
45. Kim I, Shin N-Y, Yunjin Bak, Hyu Lee P, Lee S-K, Mee Lim S. Early-onset mild cognitive impairment in Parkinson's disease: Altered corticopetal cholinergic network. Sci Rep 2017;7:2381. https://doi.org/10.1038/s41598-017-02420-w.
46. Dadu A, Satone V, Kaur R, Hashemi SH, Leonard H, Iwaki H, et al. Identification and prediction of Parkinson's disease subtypes and progression using machine learning in two cohorts. NPJ Parkinsons Dis 2022;8:172. https://doi.org/10.1038/s41531-022-00439-z.
47. Pu J-L, Jin C-Y, Wang Z-X, Fang Y, Li Y-L, Xue N-J, et al. Apolipoprotein E Genotype Contributes to Motor Progression in Parkinson's Disease. Mov Disord 2022;37:196–200. https://doi.org/10.1002/mds.28805.
48. Conti M, Mascioli D, Simonetta C, Ferrari V, Bissacco J, Bagetta S, et al. Clinical, Biological, and Functional Connectivity Profile of Patients With De Novo Parkinson Disease Who Are APOE ε4 Carriers. Neurology 2026;106:e214449. https://doi.org/10.1212/WNL.0000000000214449.
49. Xu X, Zhang S, Xu C, Zhang W, Zhao H, Liu Y, et al. Identifying subtypes of longitudinal motor symptom severity trajectories in early Parkinson's disease patients. Front Neurol 2025;16:1597132. https://doi.org/10.3389/fneur.2025.1597132.
50. Soraya GV, Fitrah YA, Bintang AK, Akbar M, Jannah AR, Ulhaq ZS, et al. Elucidating the role of APOE ε4 gene variants in the clinical manifestation of Parkinson's disease. Front Aging Neurosci 2025;17:1632480. https://doi.org/10.3389/fnagi.2025.1632480.
51. Botta R, Locascio JJ, Ye R, Goodheart AE, Gomperts SN. APOE, Aβ42, and tau differentially impact cognitive decline in Sporadic, GBA1 and LRRK2 Parkinson's disease. Npj Parkinsons Dis 2026. https://doi.org/10.1038/s41531-026-01290-2.
52. Chen Q, Scherbaum R, Gold R, Pitarokoili K, Mosig A, Zella S, et al. Data-driven subtyping of Parkinson's disease: comparison of current methodologies and application to the Bochum PNS cohort. J Neural Transm (Vienna) 2023;130:763–76. https://doi.org/10.1007/s00702-023-02627-4.
53. Simuni T, Chahine LM, Poston K, Brumm M, Buracchio T, Campbell M, et al. A biological definition of neuronal α-synuclein disease: towards an integrated staging system for research. The Lancet Neurology 2024;23:178–90. https://doi.org/10.1016/S1474-4422(23)00405-2.
54. Zhang X, Wu H, Tang B, Guo J. Clinical, mechanistic, biomarker, and therapeutic advances in GBA1-associated Parkinson's disease. Transl Neurodegener 2024;13:48. https://doi.org/10.1186/s40035-024-00437-6.