

<https://doi.org/10.1038/s41531-025-01037-5>

Clinical correlates of data-driven subtypes of deep gray matter atrophy and dopamine availability in early Parkinson's disease

Yoonsang Oh¹, Joong-Seok Kim^{1,3}✉, Gilsoon Park^{2,3}✉, Sang-Won Yoo¹, Dong-Woo Ryu¹ & Hosung Kim²

Recent machine-learning techniques may be useful to identify subtypes with distinct spatial patterns of biomarker abnormality in the various neurodegenerative diseases. Using the Subtype and Stage Inference (SuStaln) technique, we categorized data-driven subtypes of PD by examining the deep gray matter volume and dopamine availability and compared cardiac denervation, cognition, and motor symptoms between these subtypes. The SuStaln algorithm revealed two distinctive subtypes, which were well replicated in an external dataset. Subtype 1 was characterized by lower dopamine availability apparent at early inferred stages, severe cardiac denervation, mild cognitive dysfunction in the early stage, and patterns suggesting accelerated motor and cognitive dysfunction associated with later stages. In contrast, subtype 2 showed patterns indicative of earlier brain atrophy, mild cardiac denervation, and severe cognitive dysfunction apparent at early inferred stages, with no significant correlation between motor and cognitive status and SuStaln stage. These findings suggest that the machine-learning model can identify heterogeneity in PD biomarker profiles, offering insights into potential region and stage-specific patterns of biomarker abnormality and their clinical implications.

Parkinson's disease (PD) is a complex neurodegenerative disease that is defined by a combination of core motor features, such as bradykinesia, rest tremor, rigidity, and postural instability. The clinical manifestations of PD vary by individuals according to age, sex, accompanied medical conditions, constellation of motor/nonmotor features, as well as dopaminergic treatment response¹.

Substantial phenotypic heterogeneity is well recognized within the disease, complicating the design and interpretation of clinical trials and limiting patients' counseling about their prognosis. Most attempts for the characterization of disease subtypes have been associated with categorization based on the most observable features, such as age at onset or motor/nonmotor features². This dichotomous separation, while intuitive, does not faithfully represent the clinical features of the disease, which are quantitative, complex, and interrelated. A more realistic representation of the disease and disease course requires a transition to a data-driven, multi-dimensional schema that encapsulates the constellation of interrelated features and allows tracking (and ultimately predicting) change^{3,4}.

A new phenotype to explain PD based on the α -Synuclein pathology of origin has been proposed as the dichotomy between "brain-first" versus "body-first" based on rapid-eye-movement sleep behavior disorder (RBD) presence⁵. In addition, a great deal of evidence suggests parallel degeneration

in the central nervous system and peripheral nervous system in PD. However, it may be premature to use this phenotype without pathological confirmation.

Recent advances in artificial intelligence and machine learning have great promise in classifying patients who share similar pathobiological mechanisms rather than common clinical features. Subtype and Stage Inference (SuStaln) is an advanced event-based model designed to infer distinct subtypes characterized by unique sequential patterns of biomarker abnormality. This data-driven approach also positions individuals along those sequences ('stage') and allows for modeling heterogeneity in progressive conditions using only cross-sectional data⁶. The different subtypes identified by SuStaln represent distinct inferred sequences of abnormality in biomarkers such as clinical or imaging features obtained from cross-sectional data. This algorithm has been applied to numerous neurodegenerative diseases such as Alzheimer's disease and frontotemporal dementia⁶, tau deposition in Alzheimer's disease⁷, and non-motor symptoms in PD⁸.

We have previously used the SuStaln algorithm to identify PD-related cognitive impairment across disease stages⁹. We demonstrated the utility of data-driven approaches in the region and stage dissection and have also led us to the next logical step towards identifying PD subtypes defined by distinct patterns of disease staging. While previous SuStaln studies in PD

¹Department of Neurology, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea. ²USC Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. ³These authors contributed equally: Joong-Seok Kim, Gilsoon Park. ✉e-mail: neuronet@catholic.ac.kr; gp_446@usc.edu



have focused primarily on either non-motor symptom progression linked with specific magnetic resonance imaging (MRI) markers⁸ or cognitive decline based on brain volumetry alone¹⁰, the interplay between dopaminergic degeneration, deep gray matter atrophy, and their combined relationship with peripheral denervation and core clinical symptoms remains less explored. This study aims to bridge this gap by applying SuStaln to multimodal imaging data encompassing both dopamine availability and deep gray matter volumes and correlating the identified subtypes and stages with cardiac sympathetic innervation, motor severity, and comprehensive cognitive assessments.

The objective of this study is to leverage a data-driven approach to identify distinct PD subtypes characterized by region and stage-specific patterns of abnormality in dopamine availability and deep gray matter volume per region of interest, as measured by dopamine transporter positron emission tomography (PET) and T1-weighted MRI, respectively. In addition, for the identified subtypes, we aim to analyze the differences in cognitive function and motor symptoms between the subtypes and assess their alignment with established PD phenotypes based on the degree of cardiac sympathetic denervation. We expect this study to provide new insights into the heterogeneity of PD progression and its clinical and phenotypic correlates.

Results

The primary analyses presented here were conducted on the site 1 cohort, comprising 287 newly diagnosed patients with PD and 46 healthy controls (Fig. 1). Detailed demographic and clinical characteristics of this cohort are provided in Table 1. Our core analyses utilized multimodal neuroimaging data, specifically dopamine transporter availability quantified by dopamine transporter PET standardized uptake value ratios (SUVRs) and regional deep gray matter volumes derived from T1-weighted MRI. First, we present the comparison of these baseline imaging biomarkers between the overall PD group and the control group. Subsequently, we describe the data-driven identification of PD subtypes using the SuStaln algorithm based on the biomarker profiles within the site 1 PD cohort. The robustness of these subtypes is then evaluated through external validation using an independent cohort (site 2; Fig. 1). Finally, we detail the comparison of associated clinical characteristics, including motor severity, cognitive function, and cardiac sympathetic denervation, between the identified subtypes and examine their relationship with the inferred SuStaln stage.

Comparison between PD patients and controls

Considering the exclusion criteria described previously, 46 normal controls and 287 PD patients were included in this study. Normal controls were significantly younger than PD patients ($P = 0.014$) and had higher smoking history ($P = 0.003$). The distribution of sex, diabetes mellitus, and hypertension was similar between controls and the PD group (Table 1).

Compared to healthy controls, PD patients in the primary cohort (site 1) exhibited significantly lower volumes in the globus pallidus ($P < 0.001$), thalamus ($P = 0.014$), and hippocampus ($P < 0.001$), after adjusting for age and sex (Table 2). No significant differences were found in caudate or putamen volumes between the controls and site 1 PD patients overall. In the external validation cohort (site 2), PD patients also showed significantly lower volumes compared to the same control group in the globus pallidus ($P < 0.001$), thalamus ($P = 0.029$), and caudate ($P = 0.003$, Table 2). Regarding dopamine availability, SUVRs were significantly lower in PD patients compared to controls across all measured striatal subregions and the thalamus in both the site 1 and site 2 cohorts (all $P < 0.001$, Table 2).

SuStaln subtyping of PD patients

Of 287 PD patients, 13 patients were not classified into either subtype 1 or 2 as they were assigned to stage 0 (“complete normal”). The remaining 274 patients were divided into two distinct subtypes. Two hundred and fifteen patients were assigned to subtype 1, and the other 59 patients were assigned to subtype 2.

Each subtype displayed distinctive spatial patterns of biomarker abnormality associated with inferred stage in dopamine transport and deep gray matter volume (Fig. 2). In subtype 1, lower SUVRs were observed initially in posterior putamen at early stages, with lower SUVRs of anterior and ventral putamen, globus pallidus and ventral striatum typically associated with subsequent stages (Fig. 2A). Deep gray matter volume loss was typically observed first in the globus pallidus and hippocampus at stage 18, while lower SUVRs across various other striatal and thalamic regions were generally associated with later stages (Fig. 2A). However, in subtype 2, lower globus pallidus and hippocampal volumes were observed at early stages, with lower thalamic volume and SUVR typically associated with subsequent stages (Fig. 2B). Lower volume of the caudate and lower SUVRs of the posterior caudate and posterior putamen were generally observed at later stages in this subtype (Fig. 2B).

When applying the SuStaln algorithm to the independent site 2 validation cohort ($n = 95$), we observed the pattern similar with that in the primary analysis. The model selection process (based on CVIC and log-likelihood) identified a two-subtype solution as optimal. The resulting biomarker progression patterns for subtype 1 and subtype 2 in the site 2 cohort are depicted in Supplementary Fig. 3 and show patterns broadly consistent with those observed in the primary cohort (site 1). Specifically, subtype 1 ($n = 80$) in site 2 again demonstrated patterns suggestive of earlier SUVR reductions preceding volume loss, while subtype 2 ($n = 15$) in site 2 showed patterns indicative of earlier volume loss preceding SUVR reductions.

Comparison of clinical characteristics between sustain subtypes

Mean age (subtype 1; 72.7 ± 9.7 , subtype 2; 72.2 ± 8.8 , $P = 0.680$) and sex distribution ($P = 0.502$) between two subtypes did not differ. Disease

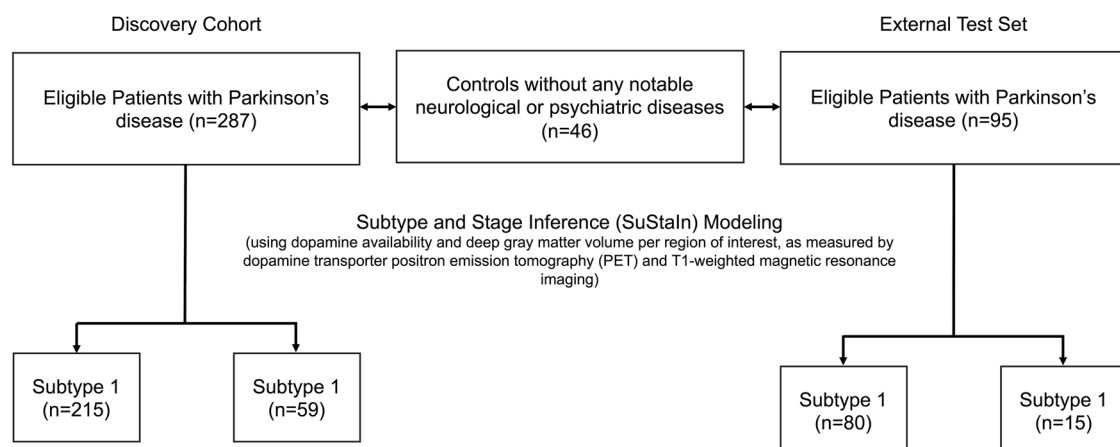


Fig. 1 | Flowchart of subjects.

Table 1 | Demographics and clinical characteristics of controls, patients with Parkinson’s disease (PD), and PD subgroups by subtype

	Control (n = 46)	Site 1			Site 2						
		PD (n = 287)	P-value (vs control)	Subtype 1 (n = 215)	Subtype 2 (n = 59)	P-value (S1 vs S2)	PD (n = 95)	P-value (vs control)	Subtype 1 (n = 80)	Subtype 2 (n = 15)	P-value (S1vs S2)
Age, years	69.3 ± 2.8	72.8 ± 9.5	0.014	72.7 ± 9.7	72.2 ± 8.8	0.680	68.7 ± 9.2	0.669	67.9 ± 9.6	72.8 ± 5.6	0.060
Sex, male ^a	24 (52.2)	144 (50.2)	0.801	106 (49.3)	32 (54.2)	0.502	45 (47.4)	0.593	35 (43.8)	10 (66.7)	0.103
Diabetes mellitus ^a	11 (23.9)	78 (27.2)	0.642	52 (24.2)	24 (40.7)	0.012	18 (18.9)	0.511	11 (13.8)	7 (46.7)	0.007
Hypertension ^a	22 (47.8)	144 (50.2)	0.767	105 (48.8)	31 (52.5)	0.614	39 (41.1)	0.473	32 (40.0)	7 (46.7)	0.776
Current or ex-smoking ^a	18 (39.1)	52 (18.1)	0.003	35 (16.3)	14 (23.7)	0.407	28 (29.5)	0.258	21 (26.2)	7 (46.7)	0.130
Disease duration, years	-	1.2 ± 1.7	-	1.1 ± 1.3	1.5 ± 2.6	0.104	1.2 ± 1.1	-	1.1 ± 1.0	1.4 ± 1.6	0.375
UPDRS part III score	-	16.6 ± 12.6	-	16.3 ± 12.3	18.0 ± 13.4	0.362	16.6 ± 8.2	-	15.8 ± 7.8	20.5 ± 9.3	0.039
Total UPDRS score	-	28.2 ± 20.3	-	27.2 ± 19.9	31.6 ± 21.1	0.142	25.8 ± 12.0	-	24.8 ± 11.7	31.1 ± 12.5	0.063
Hoehn and Yahr stage	-	1.9 ± 0.7	-	1.8 ± 0.7	2.1 ± 0.9	0.007	1.8 ± 0.6	-	1.7 ± 0.6	2.2 ± 0.6	0.007
Early H/M ratio	-	1.6 ± 0.4	-	1.6 ± 0.3	1.8 ± 0.3	<0.001	1.6 ± 0.3	-	1.6 ± 0.3	1.8 ± 0.3	0.005
Delayed H/M ratio	-	1.6 ± 0.4	-	1.5 ± 0.4	1.8 ± 0.4	<0.001	1.6 ± 0.4	-	1.5 ± 0.4	1.9 ± 0.4	0.002
Cognitive status ^a	-	-	-	-	-	<0.001	-	-	-	-	<0.001
Normal cognition	-	-	-	26 (12.1)	5 (8.5)	-	-	-	14 (17.5)	1 (6.7)	-
MCI	-	-	-	155 (72.1)	25 (42.4)	-	-	-	63 (78.8)	7 (46.7)	-
PDD	-	-	-	34 (15.8)	29 (49.1)	-	-	-	3 (3.8)	7 (46.7)	-
MMSE	-	-	-	24.0 ± 5.1	21.0 ± 6.0	<0.001	-	-	27.4 ± 2.4	22.8 ± 4.1	<0.001
CDR	-	-	-	0.5 ± 0.4	0.8 ± 0.5	<0.001	-	-	0.3 ± 0.3	0.8 ± 0.7	<0.001
GDS	-	-	-	2.7 ± 1.1	3.5 ± 1.3	<0.001	-	-	2.6 ± 0.7	3.8 ± 1.0	<0.001
Attention/working memory function	-	-	-	0.17 ± 1.23	-0.22 ± 0.97	0.076	-	-	0.12 ± 1.07	-0.74 ± 1.14	0.023
Language function	-	-	-	-0.81 ± 2.11	-1.16 ± 1.80	0.356	-	-	-0.12 ± 2.17	-3.53 ± 5.39	0.002
Visuospatial function	-	-	-	-1.61 ± 1.99	-1.98 ± 1.85	0.304	-	-	-0.79 ± 2.29	-1.47 ± 1.18	0.393
Memory function	-	-	-	-1.56 ± 1.02	-2.17 ± 1.03	0.001	-	-	-0.74 ± 0.86	-1.95 ± 1.36	<0.001
Frontal/executive function	-	-	-	-1.50 ± 1.76	-2.56 ± 1.73	<0.001	-	-	-0.62 ± 1.75	-2.92 ± 1.98	<0.001

Values represent mean with standard deviation or numbers of patients (percentage).

Analyses were performed by independent sample t-test.

UPDRS Unified Parkinson’s Disease Rating Scale, H/M heart-to-mediastinum, MCI mild cognitive impairment, PDD Parkinson’s disease dementia, MMSE Mini-Mental status examination, CDR clinical dementia rating, GDS geriatric deterioration scale, S1 subtype 1, S2 subtype 2.

^aPearson’s χ^2 test.

Table 2 | Comparison of regional deep gray matter volumes and SUVRs of controls and patients with Parkinson’s disease (PD)

	Control (n = 46)	Site 1		Site 2	
		PD (n = 287)	P-value (vs control)	PD (n = 95)	P-value (vs control)
Subregional volume (mm³)					
Caudate	8178.6 (876.1)	7926.8 (1247.1)	0.765	7752.1 (990.6)	0.003
Putamen	10191.8 (913.0)	9807.0 (1621.3)	0.556	10106.9 (1699.2)	0.400
Globus pallidus	4481.4 (361.1)	4023.8 (611.5)	<0.001	3736.3 (504.3)	<0.001
Thalamus	15765.2 (1249.6)	14703.5 (2073.9)	0.014	15385.3 (1845.5)	0.029
Hippocampus	9428.9 (944.7)	8375.5 (1537.7)	<0.001	9243.9 (1059.8)	0.075
Subregional SUVR					
Caudate	4.85 (1.35)	3.61 (1.52)	<0.001	3.85 (1.48)	<0.001
Anterior	5.15 (1.55)	3.76 (1.72)	<0.001	4.07 (1.71)	<0.001
Posterior	3.86 (1.20)	2.64 (1.20)	<0.001	3.15 (1.26)	<0.001
Putamen	7.01 (1.27)	4.42 (1.37)	<0.001	4.08 (1.25)	<0.001
Anterior	7.31 (1.49)	4.51 (1.52)	<0.001	4.19 (1.49)	<0.001
Posterior	6.76 (1.40)	3.52 (1.41)	<0.001	3.21 (1.19)	<0.001
Ventral	5.64 (0.93)	4.04 (1.03)	<0.001	3.85 (0.90)	<0.001
Globus pallidus	5.01 (1.04)	3.65 (1.06)	<0.001	3.48 (1.02)	<0.001
Thalamus	1.62 (0.13)	1.46 (0.18)	<0.001	1.43 (0.14)	<0.001
Ventral striatum	6.50 (1.53)	5.30 (1.38)	<0.001	4.88 (1.30)	<0.001

Values represent mean with standard deviation.
Analyses were performed by analysis of covariance (ANCOVA) test controlling for age and sex between control and PD.

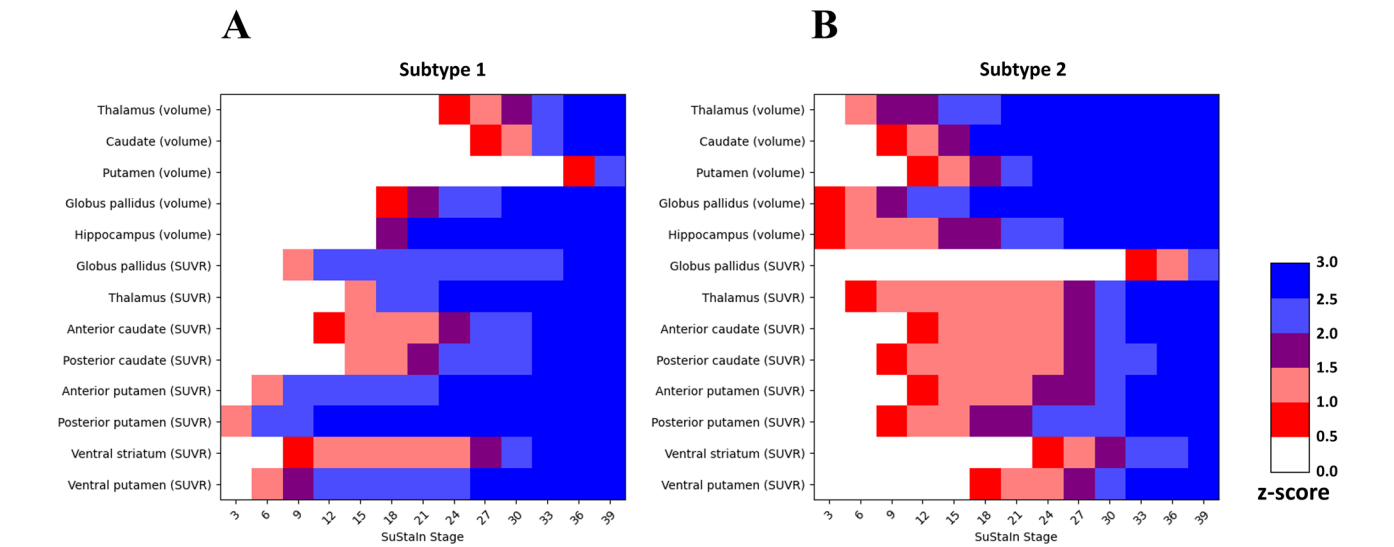
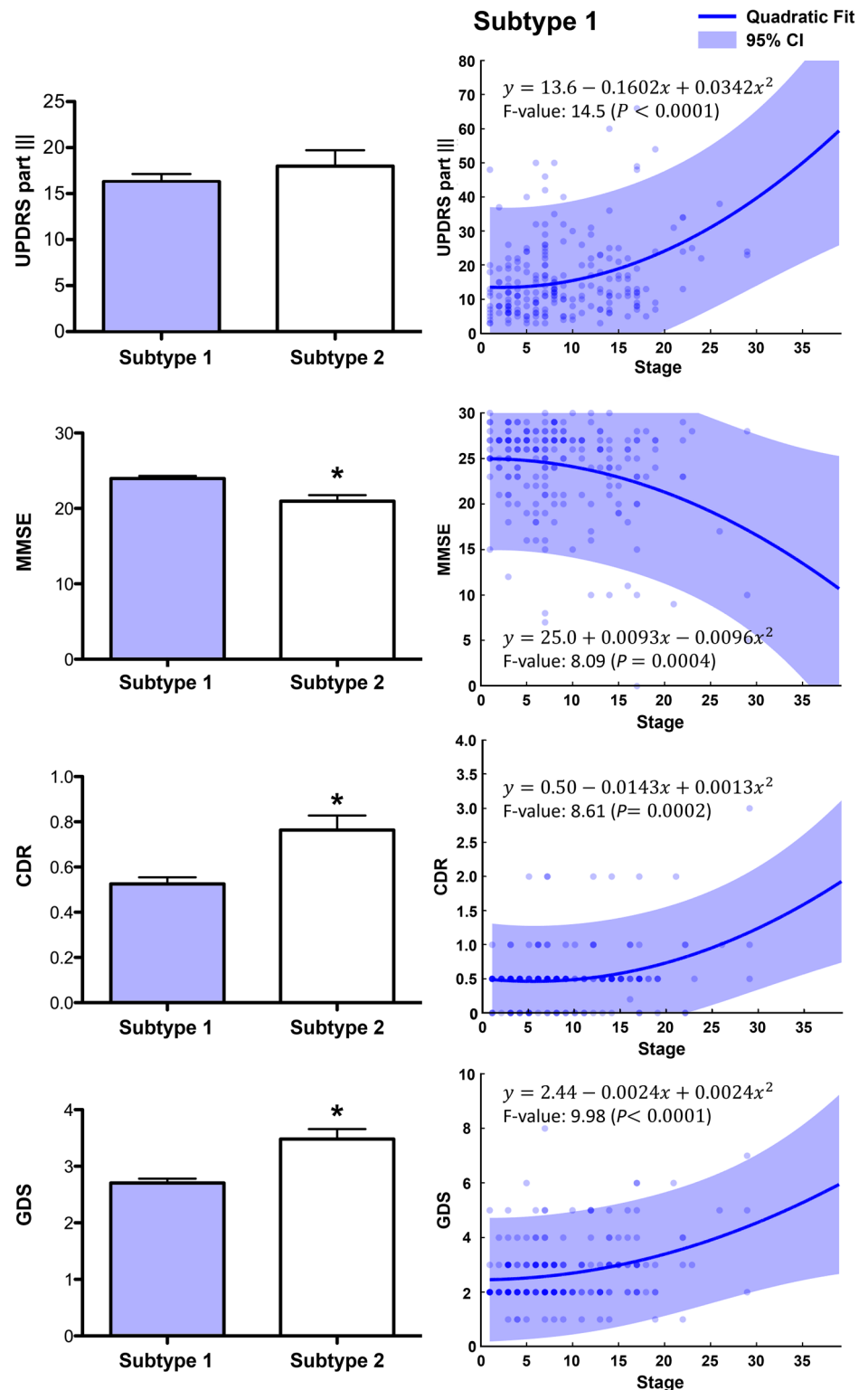


Fig. 2 | Results of the two-subtype model generated by SuStaln using deep gray matter volume and dopamine availability data (in the primary cohort, site 1). **A** In subtype 1, patterns suggest lower Standardized Uptake Value Ratio (SUVR) of the posterior putamen is typically observed at early stages, with lower SUVRs in other regions and subsequently lower volumes of deep gray matter associated with later stages. **B** In subtype 2, patterns indicate earlier lower volumes of the globus pallidus and hippocampus. Lower SUVR and volume of the thalamus, as well as lower volumes and SUVRs of other regions, are generally associated with subsequent stages. The stage (x-axis) indicates different phases of biomarker abnormality patterns relative to controls. The color bar indicates the z-score level for regional volume and SUVR loss: white (no effect) to blue (severe effect).

duration (subtype 1; 1.1 ± 1.3 , subtype 2; 1.5 ± 2.6 , $P = 0.104$) and UPDRS part III score (subtype 1; 16.3 ± 12.3 , subtype 2; 18.0 ± 13.4 , $P = 0.362$) between two groups were not different. H&Y stage score in subtype 2 (2.1 ± 0.9) was higher than that in subtype 1 (1.8 ± 0.7 , $P = 0.007$). In subtype 2, the delayed H/M ratio in subtype 2 (1.8 ± 0.4) was higher compared to subtype 1 (1.5 ± 0.4 , $P < 0.001$). General cognition status, including MMSE (subtype 1: 24.0 ± 5.1 vs. subtype 2: 21.0 ± 6.0 , $P < 0.001$), CDR (0.5 ± 0.4 vs. 0.8 ± 0.5 , $P < 0.001$), and GDS (2.7 ± 1.1 vs. 3.5 ± 1.3 , $P < 0.001$), was better in subtype 1 than subtype 2. In subtype 1, memory

function ($P = 0.001$) and frontal/executive function ($P < 0.001$) were better than subtype 2. The proportion of three cognitive function statuses (normal, MCI, and PDD) was different between groups, and more patients with PDD (49.1 vs. 15.8%) and fewer patients with normal cognition (8.5 vs. 12.1%) and MCI (42.4 vs. 72.1%) were found in subtype 2 compared with subtype 1 ($P < 0.001$, Table 1).
In subtype 1, the mean scores of MMSE, CDR, and GDS indicated better functional performance compared to those in subtype 2 ($P < 0.05$). Furthermore, the quadratic model examining motor symptoms and general

Fig. 3 | Comparison of two subtypes and correlations between SuStaiN stage and motor and cognitive function in subtype 1 (in the primary cohort, site 1). Mean Unified Parkinson's Disease Rating Scale (UPDRS) Part III scores are comparable between the two subtypes, whereas mean Mini-Mental Status Examination (MMSE), Clinical Dementia Rating (CDR), and Global Deterioration Scale (GDS) scores show significant differences. Quadratic regression models were fitted to examine the relationship between SuStaiN stage and clinical scores in Subtype 1. Significant quadratic relationships were observed for UPDRS Part III ($F = 14.50$, $P < 0.0001$), MMSE ($F = 8.09$, $P = 0.0004$), CDR ($F = 8.61$, $P = 0.0002$), and GDS ($F = 9.98$, $P < 0.0001$), indicating an accelerating decline in motor and cognitive function at later stages within this subtype. Individual data points are overlaid on the plots. Conversely, in subtype 2, UPDRS part III, CDR, and GDS scores do not show a statistically significant correlation with stage ($P > 0.1$, data not shown). * indicates P value < 0.05 .



cognition in relation to SuStaiN stages revealed distinct patterns of association between SuStaiN stage and clinical scores between the two subtypes. Subtype 1 exhibited a slowly worsening trend in UPDRS part III, MMSE, CDR, and GDS from early to mid-stage, followed by a rapid deterioration ($P < 0.001$, Fig. 3). Conversely, in subtype 2, UPDRS part III, MMSE, CDR, and GDS scores remained unchanged across stages ($P > 0.1$).

We then compared the clinical characteristics between the two subtypes identified within the site 2 cohort (Table 1). Several key findings from

the primary cohort were replicated: subtype 2 in site 2 patients had significantly lower MMSE scores ($P < 0.001$), higher CDR ($P < 0.001$), and higher GDS scores ($P < 0.001$), compared to subtype 1 in site 2. Furthermore, subtype 2 in site 2 had a higher prevalence of diabetes mellitus ($P = 0.007$), more advanced Hoehn and Yahr stage ($P = 0.007$), and significantly higher delayed H/M ratios ($P = 0.002$). Additionally, UPDRS Part III scores were significantly higher in subtype 2 within the site 2 cohort ($P = 0.039$).

Discussion

In the current study, a data-driven machine learning technique revealed two distinct phenotypic heterogeneities based on patterns of biomarker abnormality inferred by the model in dopamine availability and deep gray matter volume. In primary cohort (site 1), subtype 1 was characterized by patterns indicating earlier involvement of dopamine availability in the posterior putamen, with deep gray matter atrophy more prominent at later inferred stages, whereas subtype 2 exhibited earlier deep gray matter atrophy patterns (especially in the hippocampus and globus pallidus), with dopamine availability reductions more apparent at later inferred stages. Subtype 1 displayed severe cardiac denervation, mild cognitive dysfunction in the early stage, and rapid decline in motor and cognitive function in the later stage, whereas subtype 2 showed mild cardiac denervation, severe cognitive dysfunction in the early stage, but no significant correlation between motor and cognitive status and SuStAn stage. This novel finding expands upon previous results, supporting the hypothesis that body-first PD initially shows cardiac sympathetic denervation followed by secondary dopaminergic denervation, while brain-first PD shows the opposite pattern¹¹. Cardiac sympathetic denervation occurred in Braak stage I, and dopamine depletion in substantia nigra occurred in Braak stage III¹². Brain-first PD, corresponding to subtype 2 in our study, is characterized by nigrostriatal dopaminergic depletion prior to PNS involvement, while body-first PD, corresponding to subtype 1, is characterized by an involvement of the sympathetic nervous system earlier than the dopaminergic system¹². PD patients with RBD showed more decreased delayed H/M ratios than those without RBD⁵. As RBD is a representative marker for body-first PD, body-first PD has shown more severe cardiac sympathetic denervation, and cardiac sympathetic innervation is more preserved in brain-first PD^{11,13}.

Importantly, the two distinctive subtypes, one characterized by earlier dopaminergic deficiency and the other by earlier structural atrophy and their clinical characteristics were reproduced in an independent validation cohort (site 2). This replication achieved despite differences in MRI field strength and participants demographics, supporting the robustness and generalizability of the SuStAn-inferred subtypes and their potential relevance to distinct neuropathological trajectories within early-stage PD.

In our study, patients with PD presented significantly lower dopamine availability than normal controls in all the subregions analyzed, as expected. In addition, PD patients exhibited deep gray matter atrophy of the globus pallidus, thalamus, and hippocampus compared to normal controls in our study. Recent studies confirm the presence of deep gray matter atrophy in PD. One study revealed that PD patients showed decreased gray matter volume of cortex and striatum compared with healthy controls¹⁴. A more recent study found smaller gray matter volumes of caudate nucleus, putamen, globus pallidus, thalamus, and hippocampus in PD patients than healthy controls¹⁵. Our result outstands that subtype 2 showed more deep gray matter atrophy in the caudate nucleus, putamen, and thalamus, compared to subtype 1. This finding, however, conflicts with a recent study that there was no different brain atrophy pattern between brain-first PD and body-first PD¹⁵. This inconsistency could be attributed mainly to the subtyping method (data-driven disease staging model vs. hypothesis-driven clinical biomarker model) and other minor combined effects of different ethnicity (ours: Asian vs. prior study: predominantly white American) and imaging markers (ours: using PET and MRI vs. MRI only).

The UPDRS part III scores were similar between SuStAn subtype 1 and subtype 2. However, the relationship between the UPDRS part III score and SuStAn stage differed between the two subtypes. Patients assigned to subtype 1 exhibited a significant quadratic relationship between SuStAn stage and UPDRS part III score, with scores suggesting relatively milder motor impairment at early-to-mid stages and markedly greater impairment associated with later stages. In contrast, patients with subtype 2 remained unchanged across SuStAn stages. These findings suggest that subtype 1 has faster disease progression than subtype 2, which is consistent with previous studies showing faster motor progression in PD patients with decreased ¹²³I-MIBG uptake^{16,17}.

In our study, the subtype 1 and subtype 2 groups showed comparable cognitive performance. To note, patients with body-first PD are typically associated with worse cognitive status. PD patients with RBD, a marker for body-first PD, often exhibit severe cognitive impairment and a rapid progression to dementia^{5,18}. In our analysis, patients with subtype 2 (considered brain-first PD) had poorer performance in general cognitive scores, memory function, and frontal/executive function compared to those with subtype 1 (considered body-first PD). Further, the proportion of dementia was higher in subtype 2. This discrepancy may be explained by the small number of patients and different proportion of subjects with decreased ¹²³I-MIBG in subtype 2. Moreover, the early-stage volume reduction of the hippocampus and striatum in subtype 2 could directly contribute to declines in general cognition and memory function^{19,20}. The patterns of association between SuStAn stage and clinical measures of motor symptoms and cognitive function differed between the two subtypes. Subtype 1, with a higher prevalence of patients with decreased ¹²³I-MIBG uptake, significant quadratic relationships, suggesting relatively gradual differences in motor and cognitive function across early stages followed by apparently steeper differences associated with late stages. This is in line with previous findings that the progression to dementia becomes faster in brain-first PD²¹. However, this pattern of association between stage and clinical scores was not observed in subtype 2.

This study has several strengths. The innovative SuStAn machine learning technique allowed us to assign every patient into different disease progression stages based on patterns of neuroimaging abnormalities (e.g., reduced dopamine availability and decreased volume) and identify two or more subtypes explaining heterogeneity of the given disease spectrum. Indeed, our SuStAn analysis revealed two distinct subtypes of PD. Also, our study included a relatively large number of newly diagnosed and non-medicated patients. Since all subjects were not taking any antiparkinsonian and anti-dementia medications, the possible confounding effects of these medications on motor symptoms and cognition in PD could be excluded^{22,23}. In addition, we conducted comprehensive neuropsychological tests and classified patients based on reliable criteria. Moreover, the standardized quantitative analyses of MRI and PET imaging could reduce inherent subjectivity and errors stemming from visual or semiquantitative methods.

However, several limitations also exist in our study. First, differentiation between PD and dementia with Lewy bodies is challenging in the early disease stages, although we applied strict diagnostic criteria. This task is especially more difficult in body-first PD, where the progression to dementia is shorter and faster^{21,24}. Fortunately, as the number of PDD in subtype 1 was small, this potential confusion was likely minimal. Second, the PD patients in our primary cohort (site 1) were significantly older than healthy controls. However, the external validation cohort (site 2) was well-matched for age and sex with the control group (Table 1, Supplementary Fig. 3). Our key findings regarding the two distinct subtypes and their clinical characteristics were discovered consistently in both site 1 and 2. This finding supports the robustness and generalizability of our model despite the age difference between cohorts. Third, as the study was cross-sectional in nature, longitudinal alterations of deep gray matter volume and dopamine availability in the two groups were not examined. The SuStAn model could assign each patient into one of different PD progression stages. However, it is difficult to claim that the model can infer a pattern of temporal progression from cross-sectional data.

In conclusion, we revealed associations among alterations in brain volume and dopamine availability, cardiac denervation patterns, cognition, and motor symptoms according to identified two distinct PD subtypes characterized by different inferred biomarker abnormality patterns. One subtype exhibited patterns indicative of earlier dopamine availability reduction, severe cardiac sympathetic denervation, mild cognitive dysfunction associated with early stages, and accelerated decline in motor and cognitive function associated with later stages, whereas the other subtype showed patterns suggesting earlier deep gray matter atrophy, mild cardiac sympathetic denervation, severe cognitive dysfunction associated with early

stages, and no significant correlation found between motor and cognitive status and stage. These novel findings, derived from multimodal imaging and data-driven machine learning, provide insights into the heterogeneity of PD biomarker profiles and their clinical and phenotypic correlates.

Methods

Subjects

This study protocol was approved by the Institutional Review Board at Catholic Medical Center, The Catholic University of Korea, and all subjects provided written informed consent. All experiments were performed in accordance with relevant guidelines and regulations.

Newly diagnosed patients with PD who visited the movement disorder clinic in Yeouido St. Mary's Hospital were included. PD was diagnosed based on the UK PD Society Brain Bank clinical diagnostic criteria and Movement Disorder Society clinical diagnostic criteria for PD^{25,26}. Physical and neurological examinations were performed on all subjects. Demographics of age, sex, disease duration, smoking status, and medical history of diabetes mellitus and hypertension were collected. MRI of the brain, ¹⁸F-N-(3-fluoropropyl)-2beta-carbon ethoxy-3beta-(4-iodophenyl) nortropane (¹⁸F-FP-CIT) PET, and ¹²³I-meta-iodomethylguanodine (¹²³I-MIBG) myocardial scintigraphy were performed on all patients at the time of diagnosis. All included patients had decreased dopamine transporter uptake in the striatum, mainly in the posterior putamen.

Excluded subjects had: (1) normal dopamine transporter scan based on the Movement Disorder Society clinical diagnostic criteria for PD²⁶; (2) neurological abnormalities related to atypical or secondary parkinsonism; (3) structural or space-occupying lesions on the basal ganglia; (4) taking anti-dementia medications or other medications to influence cognitive function; (5) a history of diabetic neuropathy or other peripheral/autonomic neuropathy; (6) a history of relevant cardiac disease (such as ischemic heart disease, heart failure, or cardiomyopathy) or any abnormalities on routine chest radiography or electrocardiography; or (7) medication regimens known to influence myocardial ¹²³I-MIBG uptake or striatal dopamine uptake.

Forty-six healthy subjects without any notable neurological or psychiatric diseases were recruited and included as controls. All control subjects underwent brain MRI and did not demonstrate any abnormalities beyond mild white matter changes.

MR imaging acquisition and processing

Three-dimensional T1-weighted MR images with Magnetization Prepared Rapid Gradient Echo (MP-RAGE) sequence with following parameters: repetition time (TR)/echo time (TE) = 2040/2.43 ms, flip angle (FA) = 9°, slice thickness = 1 mm, acquisition matrix = 256 × 224, and voxel size = 1 × 1 × 1 mm³, using a 3 T MRI scanner (Skyra, Siemens, Munich, Germany). For a more accurate volumetric analysis of deep gray matter structures, we used deep neural networks to segment these structures²⁷. This network structure with highlighting foreground (HF) modules was proposed to alleviate the commonly observed imbalanced data problems in medical image segmentation, where the foreground (= target brain structure to segment) is smaller than the background (= nontarget brain and non-brain structures). This HF network demonstrated the best performance in the white matter hyperintensity segmentation challenge (team pgs, <https://wmh.isi.uu.nl/results/>). To adapt the model to the segmentation of deep gray matter, we chose, as the training data, the Parkinson's Progression Markers Initiative (PPMI) that included healthy and PD subjects. We randomly chose 42 subjects from the dataset. The label information of the selected dataset was delineated by a neuroanatomist. Using the external dataset, we performed 5-fold cross-validation to evaluate the segmentation performance. As a result, the HF network showed remarkable segmentation performance, which was comparable to experts' intra-rater reproducibility (mean Dice score = 0.93). The mean Dice score for automated segmentation was 0.9211 (Thalamus: 0.9477, Caudate: 0.9417, Putamen: 0.9435, Pallidum: 0.9045, Hippocampus: 0.9261). Finally, we applied the ensemble of five models resulting from the 5-fold cross-validation to the present study.

PET imaging acquisition and processing

Computed tomography (CT) and ¹⁸F-FP-CIT PET images were acquired using a Biograph 40 True Point PET CT scanner (Siemens, Munich, Germany). At 3-h after the intravenous injection of an average of 3.7 MBq/kg of ¹⁸F-FP-CIT, a brain CT scan was acquired for attenuation correction, followed by a 10-min ¹⁸F-FP-CIT emission PET scan. The PET image was reconstructed into a 256 × 256 × 148 matrix using an ordered-subsets expectation maximization algorithm. The voxel size was 1.336 × 1.336 × 1.500 mm³. An MR-guided conventional spatial normalization method was used for spatial normalization of ¹⁸F-FP-CIT PET images²⁸. Then, PET images were co-registered to patients' individual MR images and spatially normalized to the Montreal Neurological Institute (MNI) template space with parameter normalizing skull-stripped MR images. Subject-specific striatal volumes of interest (VOI) templates were derived from FreeSurfer 5.1 (Massachusetts General Hospital, Harvard Medical School; <http://surfer.nmr.mgh.harvard.edu>) after automated subcortical segmentation and partial volume correction. We measured the regional standardized uptake value (SUV) of the cerebellum and each side of the caudate, putamen, globus pallidus, thalamus, and ventral striatum in the PET images using these VOI templates. We also used an in-house MATLAB 2020a (MathWorks, Natick, MA, USA) for simple arithmetic operations on images and measuring the regional uptake values. The mean SUV ratio (SUVr) was calculated as the target SUV divided by the cerebellum SUV.

Data-driven subtyping and progression patterns modeling

The SuStaIn was applied to discover data-driven subtypes that were developed with progression patterns of dopamine availability and regional deep gray matter volume. To achieve this, we used all the measurements available, including 8 SUVrs (anterior and posterior caudate, anterior, posterior, and ventral putamen, globus pallidus, thalamus, and ventral striatum) and 5 regional deep gray matter volumes (caudate, putamen, globus pallidus, thalamus, and hippocampus), consisting of 13 regions of interest in total. Each regional brain volume and SUVr was transformed into a z-score relative to the normal control group. This z-score transformation allowed us to assess each individual's deviation from the norm in the normal control group. The progression pattern of the disease was described as a negative linear z-score model, and z-scores increased as the SUVr and regional brain volume decreased. SuStaIn machine learning allowed multiple sequences of progression, each representing a subtype. These stages represented different phases of disease progression, allowing a detailed understanding of how the disease evolves over time, i.e., which biomarkers (volume vs. dopamine availability) and which regions play pivotal roles at various stages of the disease. Each subject was assigned to a specific subtype and stage using maximum likelihood based on an uncertainty approximation derived from Markov Chain Monte Carlo sampling. The number of subtypes was determined by 10-fold cross-validation. For each fold, the model fit was evaluated by the Cross-Validation Information Criterion (CVIC) and out-of-sample log-likelihood. We observed that CVIC markedly decreased in the two-subtype model from the model without subtyping and then slowly and gradually decreased towards the seven-subtype model (Supplementary Fig. 1A). In cases of little improvement in model fit, simpler models should be favored. Two- to seven-subtype models displayed similarly high log likelihood values, and there was no significant difference of log likelihood among these models (Supplementary Fig. 1B). Considering the goal of model simplicity (favoring simpler models without a substantial loss in fit), and the practical requirement of adequate sample sizes within subtypes for meaningful clinical correlation analyses (e.g., the four-subtype model yielded groups with *n* = 181, 37, 26, and 31, Supplementary Fig. 2), the two-subtype model was chosen as providing the best balance of model fit, interpretability, and statistical robustness for subsequent analyses. Finally, we have implemented our SuStaIn based on pySuStaIn²⁹.

Motor and cognitive function measurements

Parkinsonian motor status was measured by the Unified Parkinson's Disease Rating Scale (UPDRS) and modified Hoehn and Yahr (H&Y) stage

scores. General cognitive status and dementia severity were evaluated using the Korean version of the Mini-Mental Status Examination (MMSE), the Clinical Dementia Rating (CDR), and the Global Deterioration Scale (GDS). Cognitive function was assessed in five cognitive domains, which comprised of attention/working memory, memory, language and related function, visuospatial and frontal executive functions³⁰, and subtests for each domain were selected. Each quantifiable neuropsychological test score was converted into a previously determined standardized score (z-score) based on age-, sex-, and education-specific normal controls, and the scores were classified as abnormal when they were below 1.5 standard deviations (SD) from the mean of normal controls. The z-scores of each test were calculated separately. In domains with multiple z-scores, the average scores of each domain were defined as the representative values³⁰. Mild cognitive impairment (MCI) was diagnosed by scores <1.5 SD of normative data on at least two measures within at least one of the five cognitive domains and unimpaired functional activities of daily living assessed by ADL, according to the Movement Disorder Society Task Force Level II criteria (comprehensive assessment)³¹. Parkinson's disease dementia (PDD) was diagnosed using clinical diagnostic criteria for dementia associated with PD³².

¹²³I-meta-iodobenzylguanidine myocardial scintigraphy

¹²³I-MIBG scintigraphy was performed with a dual-head camera equipped with a low-energy, high-resolution collimator (Siemens, Munich, Germany/Infina, GE Healthcare, Chicago, IL, USA). Data were collected 30 min (early) and 2 h (delayed) after injecting 111 MBq of ¹²³I-MIBG. A static image with a 128 × 128 matrix was obtained. Regions of interest were drawn manually around the whole heart and mediastinum. The heart-to-mediastinum (H/M) ratio was calculated from the average counts per pixel in the heart and mediastinum. We used the delayed H/M ratio because the delayed phase H/M reflected the active neuronal uptake of ¹²³I-MIBG without passive transfer and was recommended for diagnostic studies in PD, characterized by postganglionic presynaptic sympathetic failure³³.

External validation cohort and analysis

To assess the robustness and generalizability of our findings, we included an independent external validation cohort (hereafter referred to as site 2) consisting of 95 patients with newly diagnosed PD recruited from the movement disorder clinic at Seoul St. Mary's Hospital in Seoul, Korea. All patients at the site 2 met the same diagnostic criteria (UK Parkinson's Disease Society Brain Bank and MDS criteria) and exclusion criteria as the primary cohort (site 1)^{25,26}. They underwent ¹⁸F-FP-CIT PET (Discovery PET/CT 710, General Electric Healthcare, Waukesha, WI, USA) and ¹²³I-MIBG myocardial scintigraphy (Infina, GE Healthcare, Chicago, IL, USA) using identical protocols to site 1. T1-weighted Brain MRI was acquired on a 3 T scanner (Magnetom Vida, Siemens, Forchheim, Germany). Demographic and clinical data were collected similarly (Table 1). Notably, this cohort was well-matched for age ($P = 0.669$) and sex ($P = 0.593$) with the healthy control group (Table 1). MRI and PET image processing, including deep gray matter segmentation and SUVR calculation, followed the same pipelines used for the site 1 cohort. The SuStaIn algorithm was applied independently to the site 2 biomarker data (age-adjusted z-scores relative to the same control group) using the same procedures described above to identify subtypes and stages within this separate cohort.

Statistical analysis

Statistical analyses were performed using SPSS software version 24.0 for Mac (IBM Corporation, New York, NY, USA). Pearson's χ^2 test was used to compare the frequencies of categorical variables. An independent sample t-test or one-way analysis of variance was used to compare means among the groups. The subregional dopamine transporter availability and deep gray matter volume of the groups were analyzed using analysis of covariance (ANCOVA). Statistical significance was set at $P < 0.05$ after adjusting p -values using Bonferroni post-hoc correction. Furthermore, to elucidate the changes in the four scores (UPDRS part III, MMSE, CDR, and GDS) as the disease stage increases within each subtype, we conducted an analysis of

the associations employing a quadratic model. This analysis was performed using the 'fitlm' function in MATLAB R2023a (MathWorks, Natick, MA, USA), which provided the estimated equation, correlation coefficient, and p -value for each fitting model.

Data availability

Anonymized data generated during this study are available from the corresponding author on request from individuals affiliated with research or health care institutions.

Code availability

Statistical analyses were performed using SPSS Statistics 24. SPSS codes are available upon reasonable request to the corresponding authors.

Received: 17 November 2024; Accepted: 2 June 2025;

Published online: 12 June 2025

References

- Dadu, A. et al. Identification and prediction of Parkinson's disease subtypes and progression using machine learning in two cohorts. *Npj. Parkinsons Dis.* **8**, 172 (2022).
- Stebbins, G. T. et al. 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.* **28**, 668–670 (2013).
- Faghri, F. et al. Identifying and predicting amyotrophic lateral sclerosis clinical subgroups: a population-based machine-learning study. *Lancet Digit. Health* **4**, e359–e369 (2022).
- Satone, K. et al. Predicting Alzheimer's disease progression trajectory and clinical subtypes using machine learning. *bioRxiv*. 792432 <https://doi.org/10.1101/792432>
- Horsager, J. et al. Brain-first versus body-first Parkinson's disease: a multimodal imaging case-control study. *Brain* **143**, 3077–3088 (2020).
- Young, A. L. et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nat. Commun.* **9**, 4273 (2018).
- Vogel, J. W. et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nat. Med.* **27**, 871–881 (2021).
- Zhou, C. et al. Two distinct trajectories of clinical and neurodegeneration events in Parkinson's disease. *Npj. Parkinsons Dis.* **9**, 111 (2023).
- Oh, Y., Kim, J. S., Lyoo, C. H., Park, G. & Kim, H. Spatiotemporal progression patterns of dopamine availability and deep gray matter volume in Parkinson disease-related cognitive impairment. *Neurology* **103**, e209498 (2024).
- Sakato, Y. et al. Delineating three distinct spatiotemporal patterns of brain atrophy in Parkinson's disease. *Brain* **147**, 3702–3713 (2024).
- Horsager, J., Knudsen, K. & Sommerauer, M. Clinical and imaging evidence of brain-first and body-first Parkinson's disease. *Neurobiol. Dis.* **164**, 105626 (2022).
- Borghammer, P. & Van Den Berge, N. Brain-first versus gut-first Parkinson's disease: a hypothesis. *J. Parkinsons Dis.* **9**, S281–S295 (2019).
- Ryu, D. W. et al. Comparison of disease progression between brain-predominant Parkinson's disease versus Parkinson's disease with body-involvement phenotypes. *Neurobiol. Dis.* **174**, 105883 (2022).
- Xu, X. et al. Grey matter abnormalities in Parkinson's disease: a voxel-wise meta-analysis. *Eur. J. Neurol.* **27**, 653–659 (2020).
- Banwinkler, M. et al. Gray matter volume loss in proposed brain-first and body-first Parkinson's disease subtypes. *Mov. Disord.* **37**, 2066–2074 (2022).
- Tsujikawa, K. et al. Chronological changes of ¹²³I-MIBG myocardial scintigraphy and clinical features of Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **86**, 945–951 (2015).

17. Pitton Rissardo, J. & Fornari Caprara, A. L. Cardiac 123I-Metaiodobenzylguanidine (MIBG) scintigraphy in Parkinson's disease: a comprehensive review. *Brain Sci.* **13**, 1471 (2023).
18. Fereshtehnejad, S. M. et al. New clinical subtypes of Parkinson disease and their longitudinal progression: a prospective cohort comparison with other phenotypes. *Jama. Neurol.* **72**, 863–873 (2015).
19. Kandiah, N. et al. Hippocampal volume and white matter disease in the prediction of dementia in Parkinson's disease. *Parkinsonism Relat. Disord.* **20**, 1203–1208 (2014).
20. Brown, G. et al. Frontostriatal and limbic contributions to cognitive decline in Parkinson's disease. *J. Neuroimaging* **33**, 121–133 (2023).
21. Borghammer, P. The alpha-Synuclein Origin and Connectome Model (SOC Model) of Parkinson's Disease: Explaining Motor Asymmetry, Non-Motor Phenotypes, and Cognitive Decline. *J. Parkinsons Dis.* **11**, 455–474 (2021).
22. Roy, M. A., Doiron, M., Talon-Croteau, J., Dupre, N. & Simard, M. Effects of Antiparkinson medication on cognition in Parkinson's disease: a systematic review. *Can. J. Neurol. Sci.* **45**, 375–404 (2018).
23. Seppi, K. et al. Update on treatments for nonmotor symptoms of Parkinson's disease—an evidence-based medicine review. *Mov. Disord.* **34**, 180–198 (2019).
24. Borghammer, P. et al. Neuropathological evidence of body-first vs. brain-first Lewy body disease. *Neurobiol. Dis.* **161**, 105557 (2021).
25. Gibb, W. R. & Lees, A. J. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **51**, 745–752 (1988).
26. Postuma, R. B. et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord.* **30**, 1591–1601 (2015).
27. Park, G., Hong, J., Duffy, B. A., Lee, J. M. & Kim, H. White matter hyperintensities segmentation using the ensemble U-Net with multi-scale highlighting foregrounds. *Neuroimage* **237**, 118140 (2021).
28. Baek, M. S., Cho, H., Ryu, Y. H. & Lyoo, C. H. Customized FreeSurfer-based brain atlas for diffeomorphic anatomical registration through exponentiated lie algebra tool. *Ann. Nucl. Med.* **34**, 280–288 (2020).
29. Aksman, L. M. et al. pySuStaln: a Python implementation of the Subtype and Stage Inference algorithm. *SoftwareX* **16**, 100811 (2021).
30. Ryu, H. J. & Yang, D. W. The Seoul Neuropsychological Screening Battery (SNSB) for comprehensive neuropsychological assessment. *Dement. Neurocogn. Disord.* **22**, 1–15 (2023).
31. Litvan, I. et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: movement disorder society task force guidelines. *Mov. Disord.* **27**, 349–356 (2012).
32. Emre, M. et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov. Disord.* **22**, 1689–1707 (2007).
33. Chung, E. J. & Kim, S. J. (123I)-Metaiodobenzylguanidine Myocardial Scintigraphy in Lewy body-related disorders: a literature review. *J. Mov. Disord.* **8**, 55–66 (2015).

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the

Ministry of Science and ICT (NRF-2017R1D1A1B06028086 and NRF-2019R1G1A1099554). This research was also supported by the “Korea National Institute of Health” research project (2024ER100200). H.K. was supported by the National Institutes of Health grants (U01AG024904) and Michael J Fox Foundation funds (MJFF-023385).

Author contributions

Y.O., J.-S.K., G.P., and H.K. contributed the conception and design of the study; Y.O., S.-W.Y., D.-W.R., and J.-S.K. contributed to the acquisition and analysis of data; Y.O., J.-S.K., G.P., and H.K. contributed to the interpretation of results, drafting the text and preparing figure; Y.O. drafted the manuscript; J.-S.K., G.P., S.-W.Y., D.-W.R., and H.K. revised the manuscript. Y.O., J.-S.K., and H.K. obtained funding. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41531-025-01037-5>.

Correspondence and requests for materials should be addressed to Joong-Seok Kim or Gilsoon Park.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025