

## ARTICLE OPEN



# Two distinct trajectories of clinical and neurodegeneration events in Parkinson's disease

Cheng Zhou<sup>1,8</sup>, Linbo Wang<sup>2,3,4,5,8</sup>, Wei Cheng<sup>2,3,4,5,6</sup>✉, JinChao Lv<sup>2</sup>, Xiaoujun Guan<sup>1</sup>, Tao Guo<sup>1</sup>, Jingjing Wu<sup>1</sup>, Wei Zhang<sup>2,3,4,5</sup>, Ting Gao<sup>7</sup>, Xiaocao Liu<sup>1</sup>, Xueqin Bai<sup>1</sup>, Haoting Wu<sup>1</sup>, Zhengye Cao<sup>1</sup>, Luyan Gu<sup>7</sup>, Jingwen Chen<sup>1</sup>, Jiaqi Wen<sup>1</sup>, Peiyu Huang<sup>1</sup>, Xiaojun Xu<sup>1</sup>, Baorong Zhang<sup>1</sup>, Jianfeng Feng<sup>2,3,4,5,6</sup>✉ and Minming Zhang<sup>1</sup>✉

Increasing evidence suggests that Parkinson's disease (PD) exhibits disparate spatial and temporal patterns of progression. Here we used a machine-learning technique—Subtype and Stage Inference (SuStaln)—to uncover PD subtypes with distinct trajectories of clinical and neurodegeneration events. We enrolled 228 PD patients and 119 healthy controls with comprehensive assessments of olfactory, autonomic, cognitive, sleep, and emotional function. The integrity of substantia nigra (SN), locus coeruleus (LC), amygdala, hippocampus, entorhinal cortex, and basal forebrain were assessed using diffusion and neuromelanin-sensitive MRI. SuStaln model with above clinical and neuroimaging variables as input was conducted to identify PD subtypes. An independent dataset consisting of 153 PD patients and 67 healthy controls was utilized to validate our findings. We identified two distinct PD subtypes: subtype 1 with rapid eye movement sleep behavior disorder (RBD), autonomic dysfunction, and degeneration of the SN and LC as early manifestations, and cognitive impairment and limbic degeneration as advanced manifestations, while subtype 2 with hyposmia, cognitive impairment, and limbic degeneration as early manifestations, followed later by RBD and degeneration of the LC in advanced disease. Similar subtypes were shown in the validation dataset. Moreover, we found that subtype 1 had weaker levodopa response, more GBA mutations, and poorer prognosis than subtype 2. These findings provide new insights into the underlying disease biology and might be useful for personalized treatment for patients based on their subtype.

npj Parkinson's Disease (2023)9:111; <https://doi.org/10.1038/s41531-023-00556-3>

## INTRODUCTION

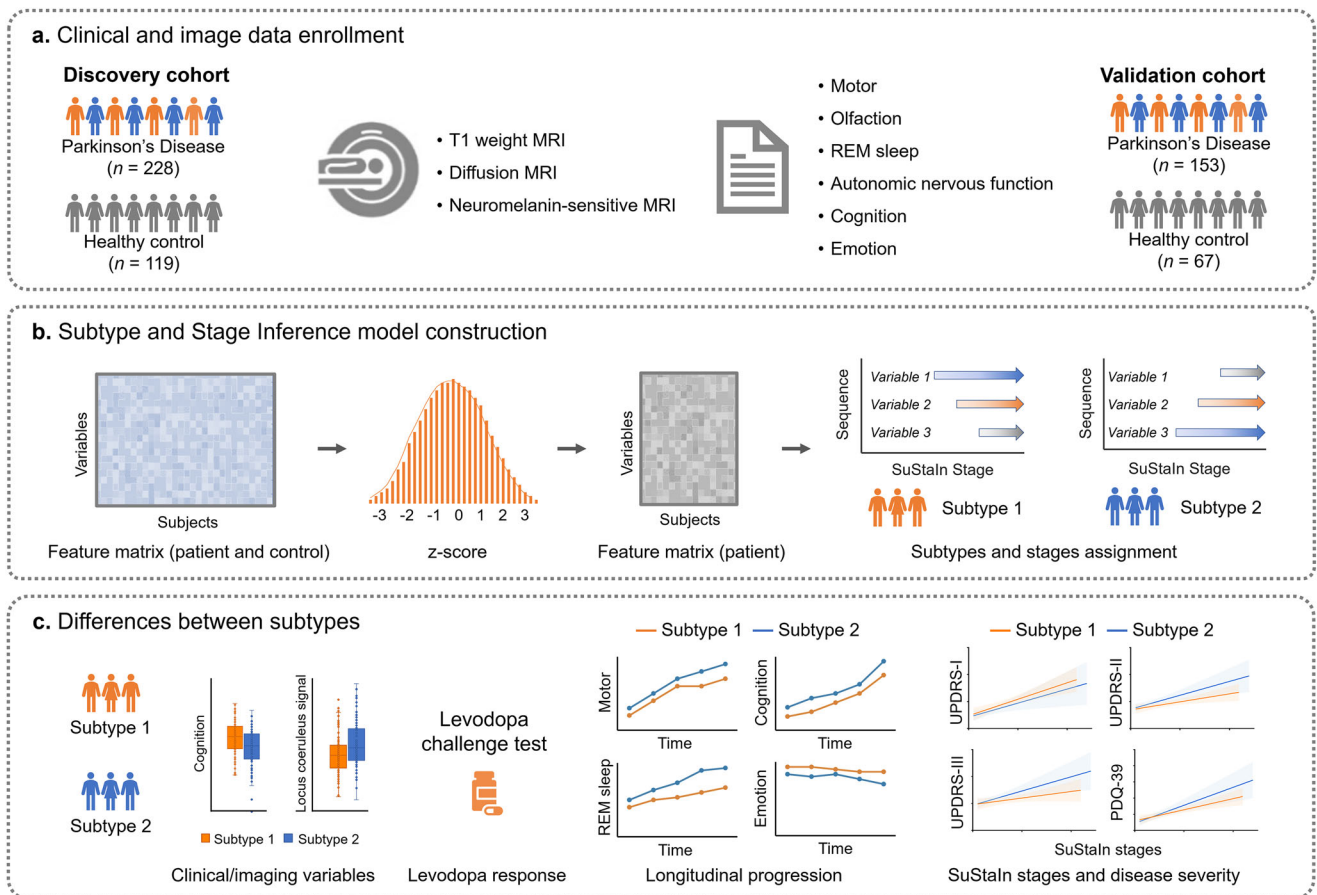
Parkinson's disease (PD) is the second most common neurodegenerative disease, and is characterized by loss of dopaminergic neurons in the substantia nigra (SN) and progressive and irreversible aggregation of misfolded  $\alpha$ -synuclein in multiple brain regions<sup>1</sup>. PD varies dramatically in its manifestations, response to therapy, and long-term prognosis, suggesting the existence of PD subtypes<sup>2</sup>. Identification of these subtypes is crucial for better understanding the mechanisms of the disease, better prognosis, and ultimately, for delivering personalized medicine<sup>3</sup>.

Great efforts to identify subtypes of PD have been made, from the initial classification based on a single motor symptom, to classification based on multidomain phenotypes including motor and non-motor symptoms, neuroimaging data, and biochemical markers<sup>3–8</sup>. Owing to the poor reliability and difficulty in characterizing disease heterogeneity, empirical classification methods have lost favor and have been replaced by data-driven methods without a priori hypotheses<sup>9–14</sup>. These data-driven studies provide invaluable information toward the understanding of complex mechanisms, but have limitations. Typical data-driven subtyping methods are limited by the confound of disease stage, as the samples included in such studies are often at different timepoints in the course of the disease<sup>12,15</sup>. Restricted by feasibility issues such as sample size, it is difficult to stratify patients by disease duration or stage. Therefore, patients can be misclassified by data-driven subtyping methods if individuals at different

timepoints or stages are included. In addition, a typical clustering analysis might not work well if the patients have different progression trajectories<sup>16</sup>. We speculate that this is also one of the reasons why many of these data-driven PD subtype classification systems lack reproducibility<sup>17</sup>. Recently, “brain-first” and “body-first” subtypes of PD have been proposed, which suggest two contrasting spatiotemporal propagation routes of  $\alpha$ -synuclein pathology<sup>17–19</sup>. In the body-first subtype, the initial pathology may first appear in the enteric or peripheral autonomic nervous system. It then propagates to the medulla oblongata via the vagus nerve. Ascending pathology affects pons giving rise to Rapid Eye Movement (REM) sleep behavior disorder (RBD) before the SN is involved<sup>19</sup>. In the brain-first subtype, the pathology may appear in the amygdala or olfactory bulb, and then spreads to the brainstem and cortex<sup>19,20</sup>. In this sense, uncovering both temporal and phenotype heterogeneity when conducting a data-driven clustering analysis is somewhat meaningful.

An emerging model called Subtype and Stage Inference (SuStaln) has been successfully used to reveal the heterogeneity and temporal complexity of neurodegenerative diseases, including Alzheimer's disease, multiple sclerosis, and frontotemporal dementia<sup>15,16,21–23</sup>. The SuStaln model is a machine-learning method that disentangles temporal and phenotypic heterogeneity to identify subtypes with distinct progression trajectories from readily accessible cross-sectional patients. Meanwhile, the model assigns each patient with the most appropriate subtype and

<sup>1</sup>Department of Radiology, The Second Affiliated Hospital, Zhejiang University School of Medicine, 310000 Hangzhou, China. <sup>2</sup>Institute of Science and Technology for Brain-inspired Intelligence, Fudan University, 200433 Shanghai, China. <sup>3</sup>Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, Shanghai, China. <sup>4</sup>MOE Frontiers Center for Brain Science, Fudan University, Shanghai, China. <sup>5</sup>Zhangjiang Fudan International Innovation Center, Shanghai, China. <sup>6</sup>Department of Computer Science, University of Warwick, Coventry CV4 7AL, United Kingdom. <sup>7</sup>Department of Neurology, The Second Affiliated Hospital, Zhejiang University School of Medicine, 310000 Hangzhou, China. <sup>8</sup>These authors contributed equally: Cheng Zhou, Linbo Wang. ✉email: [wcheng.fdu@gmail.com](mailto:wcheng.fdu@gmail.com); [jffeng@fudan.edu.cn](mailto:jffeng@fudan.edu.cn); [zhangminming@zju.edu.cn](mailto:zhangminming@zju.edu.cn)



**Fig. 1** A flowchart of characterization of clinical and imaging heterogeneity in Parkinson's disease. **a** A total of 567 individuals (347 in discovery dataset and 220 in validation dataset) were enrolled in this study; five clinical variables and 6 imaging variables were used for model construction; these imaging variables were extracted from 6 brain structures including substantia nigra, locus coeruleus, amygdala, hippocampus, entorhinal cortex, and basal forebrain by using neuromelanin-sensitive imaging and diffusion tensor imaging. **b** These variables were normalized relative to healthy control subjects using z-score; the SuStaln model was conducted to identify distinct patterns of progression (progression trajectories of clinical and neurodegenerative events) using these variables and to cluster patients into distinct patterns (subtypes) and stages. **c** The differences of clinical/imaging variables, levodopa response, and longitudinal progression were assessed between subtypes; the consistency between the inferred SuStaln stage and the disease severity measured by the UPDRS, PDQ-39, and other measures was evaluated. The box plot and line chart were used for visualization purposes only. Box plots with center line indicating median, bounds of boxes showing upper and lower quartile, whiskers illustrating 1.5 \* interquartile range, and dots representing the distribution of raw data (minima and maxima are included). MRI magnetic resonance imaging, SuStaln Subtype and Stage Inference, REM rapid eye movement. UPDRS Unified Parkinson's Disease Rating Scale, PDQ-39 Parkinson's Disease Questionnaire-39 items.

calculates each patient's stage within that subtype. In this model, progression trajectory refers to the sequence of abnormalities in clinical or imaging features, and stage refers to the cumulative degree of abnormalities in clinical or imaging features. If all patients of a certain type have atrophy in subcortical nuclei, but only some have atrophy in the neocortex, the SuStaln model would infer with high confidence that subcortical atrophy occurs before neocortical atrophy in this type of patient<sup>24</sup>. A patient with both subcortical and neocortical atrophy would therefore have an advanced disease stage compared with those with pure subcortical atrophy.

Here, we aimed to identify PD subtypes with distinct trajectories based on the above-mentioned data-driven method with multi-dimensional data, including comprehensive clinical evaluations and MRI features (diffusion tensor imaging and neuromelanin-sensitive MRI). These MRI measures are considered as sensitive imaging biomarkers for evaluating neurodegeneration of PD patients<sup>6</sup>. Furthermore, we fitted analogous multidimensional data from the Parkinson's Progression Markers Initiative (PPMI) dataset into the SuStaln model to test the repeatability of PD subtypes.

Finally, we investigated the differences in clinical characteristics, imaging features, response to levodopa treatment, and longitudinal prognosis between these PD subtypes (Fig. 1).

## RESULTS

Healthy controls and PD patients were not significantly different in terms of gender, age and education level. PD patients showed significant hyposmia, autonomic dysfunction, RBD symptoms, cognitive decline, and depression compared to the healthy control group. Additionally, PD patients showed significant decreased contrast-to-noise ratio (CNR) of substantia nigra (SN) and locus coeruleus (LC), and increased free-water of hippocampus and entorhinal cortex. Detailed demographics, clinical and imaging characteristics were shown in Table 1.

### Clinical and neurodegeneration trajectories of subtypes

Of the 228 PD patients, 114 were assigned to subtype 1 and 102 to subtype 2. Only 12 patients were categorized as stage 0, and not assigned to any subtype. Each subtype was defined by a

**Table 1.** Demographics, clinical and imaging characteristics of healthy controls and Parkinson's disease patients in discovery dataset.

	Health control	Parkinson's disease	<i>p</i> value
Sex (male)	119(68)	228(148)	0.157
Age (year)	61.71 ± 7.28	60.53 ± 8.56	0.20
Education (year)	9.56 ± 4.20	9.06 ± 3.85	0.268
Duration (year)	–	4.29 ± 2.82	–
LEDD	–	346.25 ± 272.41	–
UPDRS-I	–	1.54 ± 1.51	–
UPDRS-II	–	9.10 ± 5.33	–
UPDRS-III	–	22.69 ± 12.97	–
UPDRS-IV	–	0.83 ± 1.54	–
HY stage	–	2.09 ± 0.62	–
PDQ-39	–	19.76 ± 18.8	–
ADL	–	22.14 ± 5.17	–
RBDQ-HK	9.64 ± 5.77	20.75 ± 16.75	<0.001***
SCOPA-AUT	4.76 ± 3.44	9.57 ± 6.17	<0.001***
BSIT	8.50 ± 1.63	5.21 ± 2.47	<0.001***
MoCA	24.53 ± 3.36	23.11 ± 4.24	0.002**
HAMD	2.91 ± 2.99	3.80 ± 3.95	0.032*
CNR <sub>SN</sub>	2.50 ± 0.42	2.01 ± 0.44	<0.001***
CNR <sub>LC</sub>	2.26 ± 0.78	2.07 ± 0.74	0.030*
Free-water of basal forebrain	0.196 ± 0.021	0.197 ± 0.017	0.721
Free-water of entorhinal cortex	0.20 ± 0.022	0.209 ± 0.024	0.001**
Free-water of amygdala	0.172 ± 0.010	0.174 ± 0.012	0.073
Free-water of hippocampus	0.180 ± 0.016	0.185 ± 0.018	0.019*

LEDD levodopa equivalent daily dose, UPDRS Unified Parkinson's Disease Rating Scale, HY Hoehn–Yahr, PDQ-39 Parkinson's Disease Questionnaire-39 items, ADL activity of daily life scale, RBDQ-HK Rapid Eye Movement Sleep Behavior Disorder Questionnaire (Chinese University of Hong Kong version), SCOPA-AUT Scales for Outcomes in Parkinson's Disease-Autonomic, BSIT Brief Smell Identification Test modified for Chinese, MoCA Montreal Cognitive Assessment, HAMD Hamilton Rating Scale for Depression, CNR<sub>SN</sub> contrast-to-noise ratio of substantia nigra, CNR<sub>LC</sub> contrast-to-noise ratio of locus coeruleus.

\**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001.

different sequence of abnormalities in clinical symptoms and brain degeneration (Fig. 2a). Distribution of two subtypes across SuStaln stages were shown in Fig. 2b. Boxplots of probability of maximum likelihood subtype suggested that all patients have high probability (>50%, Fig. 2c). Figure 2d, e showed the Log Likelihood across Markov Chain Monte Carlo (MCMC) iterations and 10-fold cross-validation, respectively. Figure 2f showed the cross-validation information criterion (CVIC) under different number of subtypes. A decreased CVIC illustrated an improved model fit. The CVIC showed clear decrease from one to two, but less obvious decrease from two to three subtypes. In case of very little improvement in model fit, a simpler model should be favored. Thus, the SuStaln model with two subtypes was selected.

The SuStaln model estimated that progression starts with classic prodromal features of PD, such as RBD, hyposmia, autonomic dysfunction, and depression, which precede cognitive impairment later in the disease course. Degeneration of the SN and LC were the earliest manifestations of the disease captured by imaging in subtype 1. These were followed by degeneration of the entorhinal cortex. Consistent with the sequence of cognitive decline, the degeneration of limbic systems such as hippocampus, amygdala, and basal forebrain

occurred in the advanced stages. The first symptoms occurred in subtype 2 patients were olfaction, autonomic function, and cognition. In contrast to subtype 1, RBD and depression were confined to the latter stages of the disease. SN degeneration was the earliest sign of the disease to be imaged. Early degeneration of hippocampus and amygdala were also characteristic of this subtype. These were followed by degeneration of entorhinal cortex and basal forebrain. LC degeneration were confined to the latter stages of the disease.

Additionally, we constructed a new SuStaln model using diffusion MRI data only (specifically measuring the free-water of substantia nigra, locus coeruleus, basal forebrain, entorhinal cortex, amygdala, and hippocampus) in the discovery dataset. The resulting progress patterns of two subtypes were consistent with above results (Supplementary Fig. 1).

### Differences of demographics, clinical variables, imaging parameters, and response to levodopa treatment between subtypes

The two subtypes were not significantly different in their gender, age, disease duration, and levodopa equivalent daily dose (LEDD). Subtype 1 patients had lower education level (*p* = 0.005) and higher SuStaln stage than subtype 2 (*p* < 0.001). Therefore, education and SuStaln stage were regressed as covariates when comparing the difference of clinical variables, imaging parameters, and levodopa response between subtypes. General linear model suggested that no significant difference was found for part I, II, and III of Unified Parkinson's Disease Rating Scale (UPDRS), and Hoehn and Yahr (HY) scores between the two subtypes. In addition, more severe RBD symptoms, autonomic dysfunction, and LC degeneration were observed in subtype 1 compared with subtype 2. By contrast, more significant declines in cognition and olfaction were found for subtype 2. More serious degeneration in SN, hippocampus, amygdala, and basal forebrain were also found for subtype 2 [Fig. 3, False discovery rate (FDR) corrected].

We compared the difference of levodopa response between the two subtypes. One hundred and twelve patients (subtype 1 = 55, subtype 2 = 57) undertook the levodopa challenge test. The two subgroups were well matched in age, gender, education level, disease duration, and LEDD. Subtype 1 patients showed significant poor levodopa responses than that of subtype 2 patients (*p* = 0.026). The difference kept significant in regression of LEDD between two subtypes (*p* = 0.025). Supplementary Table 1 shares detailed demographic information of the patients who took the levodopa challenge test.

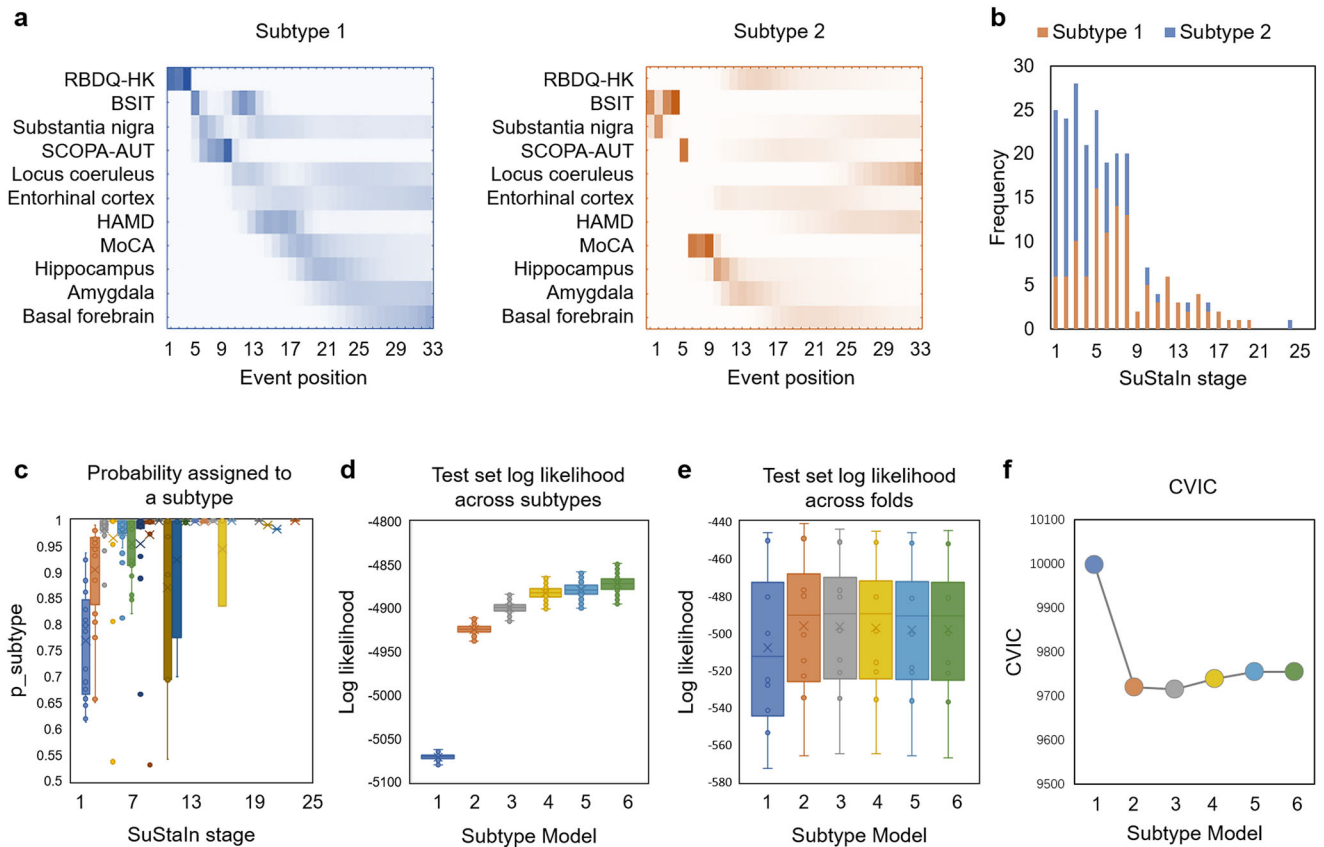
No significant difference was observed between the two subtypes for tremor, rigidity, bradykinesia, and axial symptoms, histories of smoking, alcohol consumption, pesticide and toxic exposure, and asymmetry index of imaging variables (see Supplementary Materials for details).

### Correlations between inferred SuStaln stage and disease severity

Significant correlations were found between SuStaln stage and UPDRS-I, UPDRS-II, UPDRS-III, UPDRS-IV, Parkinson's Disease Questionnaire-39 items (PDQ-39), and activity of daily life scale (ADL) scores in subtype 1, and UPDRS-I, UPDRS-II, UPDRS-IV, PDQ-39, and ADL scores in subtype 2 (Fig. 4, FDR corrected). These results suggest that the features used in the SuStaln model were related to PD pathology, and that the SuStaln stages assigned to patients were reliable.

### Subtypes in validation dataset

We validated our findings in the PPMI dataset. Healthy controls and PD patients were well matched in gender, age and education level. Detailed demographics, clinical and imaging characteristics



**Fig. 2** The progression patterns of Parkinson's disease subtypes in the discovery dataset. **a** The positional density maps of two Parkinson's disease subtypes. x axis indicates the number of events (11 variables \* 3 grads). The color intensity reflects the row-wise positional density and confidence in the ordering. The location of highest color intensity means the most probable stage/sequence of this feature. **b** Distribution of two subtypes across SuStaln stages. **c** Boxplots of probability of maximum likelihood subtype. **d** Log Likelihood across Markov chain Monte Carlo iterations. **e** Log Likelihood across 10-fold cross-validation. Log Likelihood increased dramatically from 1 subtype to 2 subtypes, but not increased slowly or even decreased from 2 subtypes to 3 or more subtypes. **f** The CVIC under different number of subtypes. Decreased CVIC illustrated improved model fit. In case of little improvement in model fit, a simpler model should be favored. Box plots with center line indicating median, bounds of boxes showing upper and lower quartile, whiskers illustrating 1.5 \* interquartile range, and dots representing the distribution of raw data (minima and maxima are included). The contrast-to-noise ratio of substantia nigra and locus coeruleus, the free-water of basal forebrain, entorhinal cortex, amygdala, and hippocampus were enrolled. RBDQ-HK Rapid Eye Movement Sleep Behavior Disorder Questionnaire (Chinese University of Hong Kong version), BSIT Brief Smell Identification Test modified for Chinese, SCOPA-AUT Scales for Outcomes in Parkinson's Disease-Autonomic, HAMD Hamilton Rating Scale for Depression, MoCA Montreal Cognitive Assessment, SuStaln Subtype and Stage Inference, CVIC cross-validation information criterion.

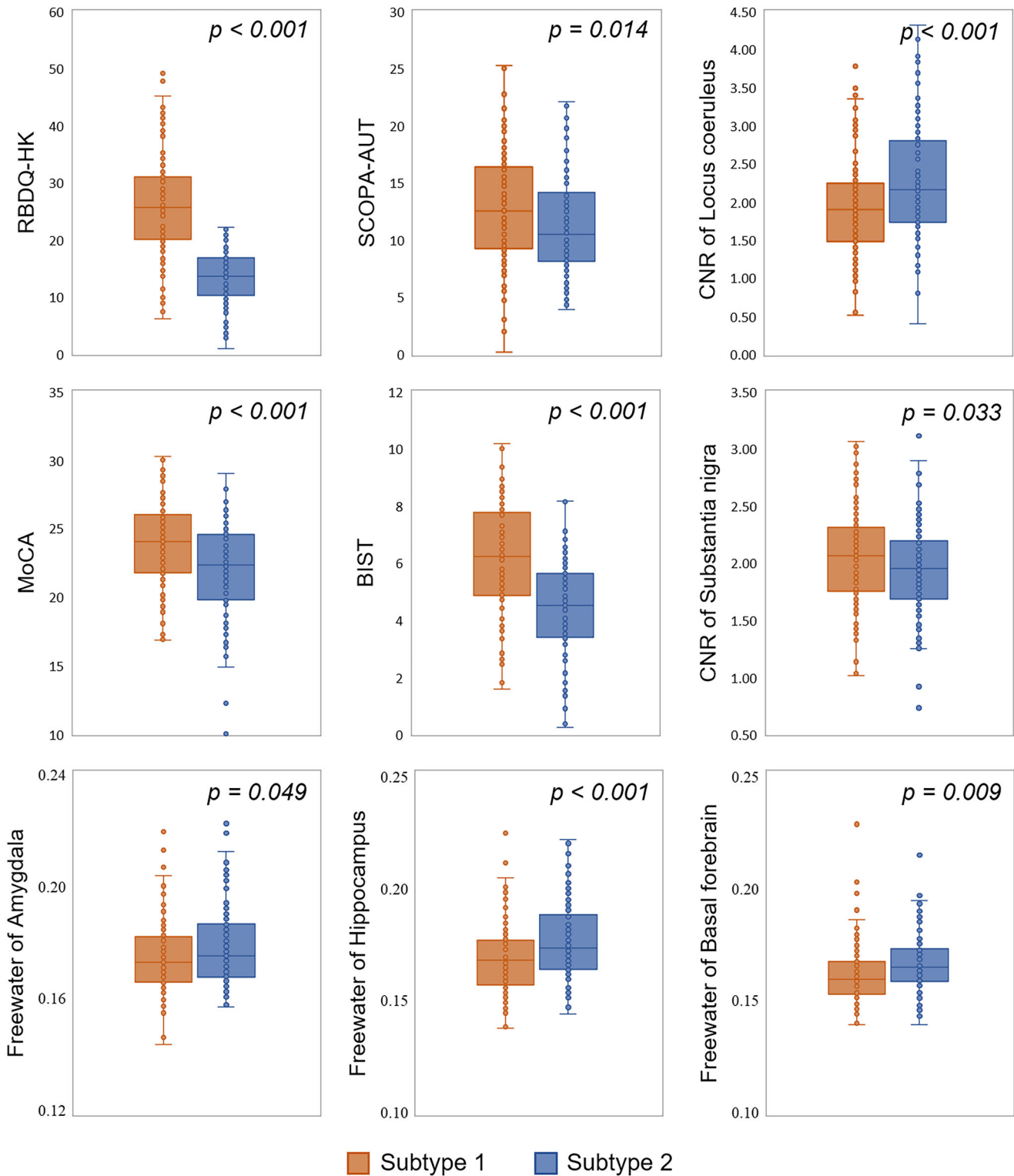
were shown in Supplementary Table 2. One hundred and twenty-eight patients were divided into two distinct progression subtypes. The first dysfunctional signs in subtype 1 patients ( $n = 102$ ) were in olfaction and REM sleep. These were followed by depression and autonomic dysfunction, while cognitive decline developed in the advanced stages. The initial symptoms in subtype 2 patients ( $n = 26$ ) were cognitive decline and hyposmia. Autonomic dysfunction, RBD symptoms, and depression were limited to the latter stages of the disease. Alterations in imaging features were restricted to the end of disease stage in both subtypes. The remaining 25 patients in that dataset were categorized as stage 0. Distribution of two subtypes across SuStaln stages, boxplots of probability of maximum likelihood subtype, MCMC iterations and 10-fold cross-validation, and the CVIC under different number of subtypes were shown in Supplementary Fig. 2. Two subtypes were determined as the optional number of subtypes.

Consistent with the discovery dataset, subtype 1 patients displayed more severe RBD ( $p = 0.044$ ), but not survived after FDR correction, while subtype 2 patients displayed significant lower Montreal Cognitive Assessment (MoCA) score ( $p < 0.001$ ). Three PD patients with GBA mutations and 1 PD patient with LRRK2

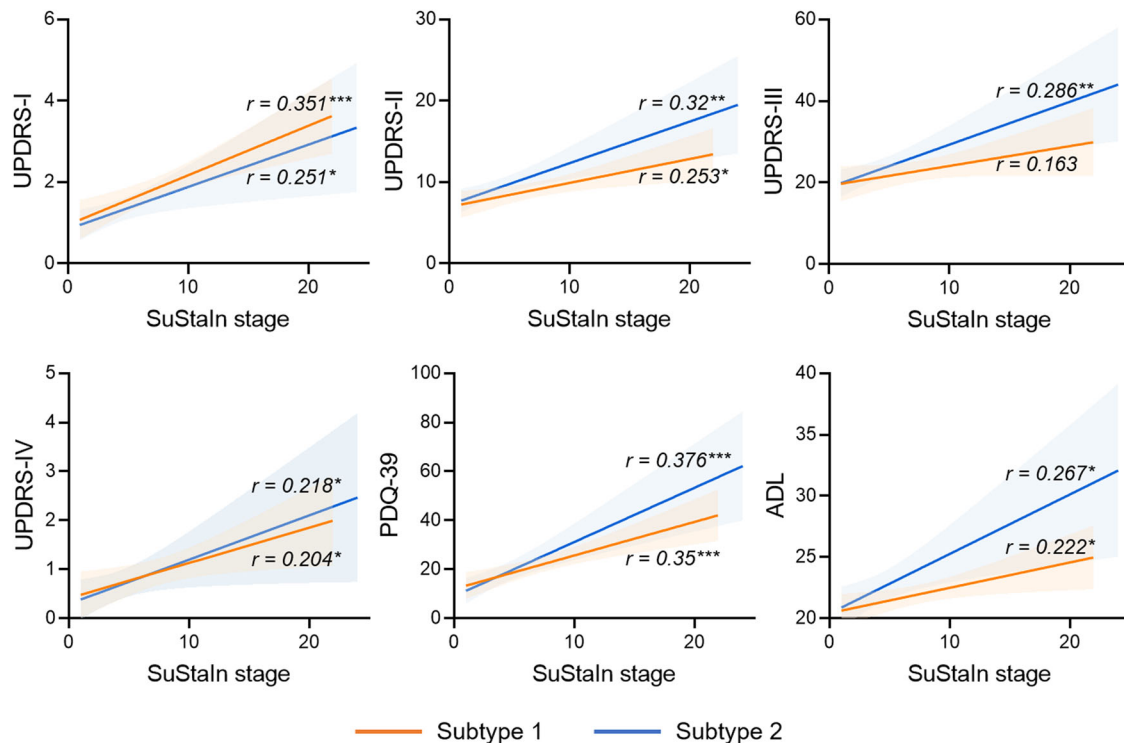
mutations were identified. Three patients with GBA mutations were grouped to subtype 1, and one patient with LRRK2 mutations was grouped to subtype 2. We also assessed the differences of longitudinal progression between subtypes. Linear mixed-effects model analysis suggested that subtype 1 had faster rate of progression in MDS-UPDRS-III and MDR-UPDRS-Total scores during 5 years follow up (group \* time,  $p = 0.030$  and  $p = 0.040$ , respectively, uncorrected).

No significant difference was found between the two subtypes for gender, age, education level, disease duration, MDS-UPDRS scores and motor subscales of tremor, rigidity, bradykinesia, axial symptoms, and postural instability and gait difficulty (PIGD). No significant difference was found in environmental factors, asymmetry index, dopamine transporter (DAT) binding ratios, levels of amyloid- $\beta$ 42,  $\alpha$ -synuclein, and phosphorylated tau protein in cerebrospinal fluid (Supplementary Table 4).

In addition, we found significant correlations between SuStaln stage and MDS-UPDRS-I, MDS-UPDRS-II, and MDS-UPDRS-III scores for subtype 1 ( $p = 0.004$ ,  $p < 0.001$ , and  $p = 0.015$ , respectively) but not for subtype 2, which might be due to the small sample size of subtype 2 (FDR corrected).



**Fig. 3** The differences of clinical and imaging features between two Parkinson's disease subtypes in the discovery dataset. Lower CNR of locus coeruleus or substantia nigra means more severe degeneration, while higher free-water measure means more severe degeneration. Box plots with center line indicating median, bounds of boxes representing upper and lower quartile, whiskers illustrating 1.5 \* interquartile range, and dots representing the distribution of raw data (minima and maxima are included). The  $p$  values were FDR corrected. RBDQ-HK Rapid Eye Movement Sleep Behavior Disorder Questionnaire—Chinese University of Hong Kong version, SCOPA-AUT Scales for Outcomes in Parkinson's Disease-Autonomic, CNR contrast to noise ratio, MoCA Montreal Cognitive Assessment scale, BIST modified Brief Smell Identification Test for Chinese. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .



**Fig. 4 Correlations between SuStaln stage and disease severity in two Parkinson's disease subtypes in the discovery dataset.** FDR corrected. UPDRS Unified Parkinson's Disease Rating Scale, PDQ-39 Parkinson's Disease Questionnaire-39 items, ADL activity of daily life scale. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

## DISCUSSION

We identified two distinct patterns of disease progression in PD patients: subtype 1, in which RBD, hyposmia, and autonomic dysfunction are early symptoms, and subtype 2, in which hyposmia and cognitive impairment are early manifestations. Furthermore, subtype 1 was characterized by early degeneration in the SN and LC, followed by degeneration in amygdala, hippocampus, and basal forebrain as the disease progresses. Subtype 2 was characterized by early degeneration in the SN, hippocampus, amygdala, and basal forebrain, with spread to the LC as the disease progressed. The validation dataset showed patterns of progression similar to the discovery dataset. Moreover, subtype 1 patients had poor levodopa response, more GBA mutations, and worse longitudinal prognosis compared with subtype 2 patients. These findings provide new insights into the heterogeneity of PD from the perspective of dynamic disease evolution, and are also potentially useful for the stratification of patients in clinical trials and practice.

According to the Braak model, PD typically begins with RBD, autonomic dysfunction symptoms, and hyposmia, which then progress to emotional and classical motor disorders before ultimately leading to cognitive impairment<sup>24,25</sup>. In this study, we identified a subtype of PD with an identical pattern of progress from multidimensional data by using a novel data-driven approach. Meanwhile, subtype 1 was characterized by early degeneration in SN and LC, which was followed by entorhinal cortex degeneration, and eventually spreading to the limbic system such as hippocampus, amygdala, and basal forebrain. This trajectory of brain degeneration in subtype 1 is similar with the Braak staging hypothesis which describes the intracerebral formation of abnormal proteinaceous Lewy bodies and the spread of Lewy neurites from the brainstem, and midbrain to limbic cortex<sup>25,26</sup>. Braak also pointed that the pathology of PD may develop in the olfactory bulb and extend to adjacent regions, but

these lesions do not appear to advance substantially beyond those areas<sup>25</sup>. Therefore, Braak concluded that although PD may have a dual-etiology (originating from the gut and olfactory bulb), it is the propagation of pathology through the brainstem that ultimately drives the disease<sup>25</sup>. This viewpoint is similar with the recently proposed "body-first" hypothesis<sup>19</sup>.

However, neuropathological and clinical evidence suggests that Braak's hypothesis does not apply to all PD patients. We identified a subtype characterized by early hyposmia and cognitive impairment, which in the later stages is accompanied by REM sleep dysfunction. Furthermore, degeneration in the hippocampus, amygdala, and basal forebrain occurred long before it did in the LC. This subtype was consistent with the "brain-first" hypothesis<sup>19</sup>. A recent postmortem study further supports our findings, reporting that pathology in "brain-first" patients is initially triggered in the olfactory bulb or amygdala<sup>20</sup>. In addition, it has been recognized that hyposmia and cognitive impairment can occur in the prodromal stage of PD, and is included as part of the MDS research criteria for prodromal PD<sup>27</sup>. Our findings further suggest that early occurrence of hyposmia and cognitive impairment might not merely be a prodromal marker for increased risk of PD, instead it might herald a specific PD subtype.

The clinical and imaging profiles differed significantly between the two subtypes. Subtype 1 showed more severe RBD symptoms, autonomic dysfunction-related manifestations, and LC degeneration. By contrast, subtype 2 showed more severe cognitive impairment, hyposmia, and limbic degeneration. These characteristics of the two subtypes were well aligned with that of "body-first" and "brain-first" subtypes, respectively<sup>18,19,28,29</sup>, and the neuroanatomic biotypes identified using a data-driven clustering approach<sup>14</sup>. These authors reported two subtypes of PD, one of which was characterized by more severe RBD and autonomic dysfunction<sup>14,18,19,28,29</sup>. According to the "body-first and brain-first" hypothesis,  $\alpha$ -synuclein aggregates originate in the peripheral or enteric nervous systems and gradually invade the

brainstem and cortex in “body-first” subtype patients. In “brain-first” subtype patients,  $\alpha$ -synuclein aggregation starts in the limbic system and gradually spreads to the brainstem and peripheral and enteric nervous systems. Consistent with these two subtypes, autopsies demonstrate two dominant patterns of Lewy body pathology distribution in the PD population: a brainstem-predominant pattern and a limbic-predominant pattern<sup>29–32</sup>. Additionally, recent study reviewed postmortem datasets and further concluded that limbic-predominant patients had initial olfactory bulb pathology and then spreads to amygdala and neighborhood regions<sup>20</sup>. This viewpoint could explain why hyposmia was more severe in subtype 2 patients in our current findings.

It is worth noting that previous studies also found that PD patients with RBD symptoms have worse cognitive performance<sup>33</sup>. Traditional clustering analysis also tends to divide patients into mild (fewer non-motor symptoms) or diffuse malignant (more severe RBD symptoms and worse cognitive performance) subtypes<sup>10</sup>. We considered that disease stage as a confounding factor may contribute to differences in our conclusions. For instance, a PD subtype with RBD as an early manifestation might develop cognitive impairment at an advanced stage. In contrast, a PD subtype with cognitive impairment as an early manifestation could also show the symptoms of RBD at an advanced stage. Both types of patients will reach a similar end point characterized by severe disability followed by death<sup>7</sup>. PD patients with RBD symptoms usually implies that the  $\alpha$ -synuclein pathology has been present for more than 10 years, which is much earlier than the onset of motor symptoms<sup>3</sup>. These patients might be in the advanced pathological stages, and consequently can have more severe clinical manifestations. However, this does not imply that PD patients with RBD symptoms suffered an early cognitive decline. Accordingly, a clinical study with a large sample size suggested that subjects (but not PD patients) with RBD symptoms do not undergo a cognitive decline during a 5-year follow-up period<sup>34</sup>. Therefore, the clinical characteristics of the two subtypes identified in our study do not conflict with previous studies. In addition, because the subjects included in clinical studies are at different stages of the disease, most clustering algorithms will partition individuals into early-stage or advanced-stage subtypes, or mild or diffuse malignant subtypes. This is not conducive to resolving the heterogeneous patterns of progression subtypes, which are, in theory, orthogonal to disease progression itself<sup>16</sup>. Therefore, the SuStaln algorithm which combining clustering with disease progression modeling was conducted to overcome this limitation.

We found that subtype 1 had a significantly poor response to levodopa replacement therapy compared to subtype 2, which suggests different pathobiological mechanisms. In line with this finding, recent animal and human studies have shown that a deficiency of the LC noradrenergic system can complicate PD symptoms and diminishes the therapeutic efficacy of levodopa<sup>35,36</sup>. We found that subtype 2 patients have slight severe SN degeneration when compared with subtype 1 patients, probably because “brain-first” subtype is characterized by earlier dopamine deficiency than “body-first” subtype<sup>18</sup>. Nonetheless, the difference is minor and necessitates a prudent explanation. The UPDRS-III scores and subscores for tremor, rigidity, bradykinesia, axial symptoms did not differ between two PD subtypes, which indicate a comparable burden of nigrostriatal dopaminergic deficit between two PD subtypes. This finding was supported by previous study, which suggested that motor symptom are not the core features for distinguishing PD subtypes<sup>19</sup>. Furthermore, no significant association was observed between substances such as alcohol, tobacco, and coffee and PD subtypes in our current study. And pesticide and toxic exposure may not directly affect the subtype of PD. However, to address this issue more comprehensively, further specialized epidemiological studies are

required, including meticulous quantification of factors such as alcohol content, precise amount of drinking, number of cigarettes smoked, and years of smoking.

In addition to classifying patients into different subtypes, we assigned each patient for a SuStaln stage. We found that SuStaln stage was significantly correlated with UPDRS-I, UPDRS-II, UPDRS-III, PDQ-39, and ADL scores. Considering that these clinical variables are mostly used for evaluating disease severity from different dimensions, these results confirmed the rationality of using the SuStaln model to assign patients into specific subtypes and stages.

The two subtypes identified from the validation dataset were consistent with the subtypes determined from the discovery dataset. Subtype 1 patients had severe RBD symptom, while subtype 2 patients had significant poorer cognitive performance. It should be noted that the results obtained from free-water measures in the validation dataset were not entirely congruent with those obtained from the discovery dataset. We speculated that the lack of additional b0 with reverse phase-encode polarity contributed to the low sensitivity in detecting the neurodegeneration of PD. It is known that susceptibility-induced distortions could significantly reduce the accuracy of signal in small brain regions<sup>37</sup>. The findings from validation dataset also provided new knowledge about PD subtypes. Subtype 1 patients had faster rate of progression in motor symptom and global disease severity compared to subtype 2. This was consistent with previous studies that reported a poor prognosis in PD patients with RBD symptoms<sup>3,29</sup>. We also found that 3 patients with GBA mutations were grouped to subtype 1, and 1 patient with LRRK2 mutations was grouped to subtype 2. This finding was consistent the proposed “body-first” and “brain-first” phenotypes<sup>3,18</sup>. They considered that the pathogenic mechanism introduced by LRRK2 mutations may facilitate a “brain-first” subtype, whereas the GBA mutation carriers seem to be characterized by early, marked sympathetic denervation resembling more the “body-first” phenotype. Nonetheless, we acknowledged that the sample size for this analysis was small, and further validation is needed.

No significant difference was found in DAT binding ratios between PD subtypes, which might because of the flooring effect of DAT imaging. In addition, no difference was found in levels of amyloid- $\beta$ 42 and pathological tau protein in cerebrospinal fluid between the two subtypes. These findings suggest that the cognitive differences may not be caused by the burden of Alzheimer’s disease-related pathology, but rather by the different pattern of progression. Taken together, these findings expand our understanding of PD heterogeneous by showing two subtypes with different temporal and phenotype progression.

We acknowledge that the current findings are not completely aligned with the pathological evidence of the “body-first and brain-first” hypothesis<sup>31</sup>. Measurement errors in clinical and imaging assessments, patient heterogeneity, or the different pathological susceptibility of regions of the brain, might contribute to this inconsistency. For example, objective measures, such as polysomnography and colonic transit study, are more effective in quantifying the severity of RBD and constipation compared to subjective questionnaires. Although free-water measurements may provide better pathophysiological information regarding PD, diffusion MRI scanned with single-shell ( $b = 1000$ ) is not the optimal choice for estimating free-water<sup>38</sup>. Further studies based on multi-shell diffusion MRI scans are necessary. Future studies to investigate this inconsistency are warranted. Moreover, future studies involving peripheral system examinations such as meta-[<sup>123</sup>I]iodobenzylguanidine (<sup>123</sup>I-MIBG) scintigraphy and intestinal biopsies might improve the understanding of distinct progression subtypes. Incorporating a vast number of samples should increase the confidence of subtype and stage inference, and therefore, future studies with larger sample sizes will be valuable for validating the current findings. Finally, considering

the complexity of multidimensional data, having a subtype and stage inference model better suited to integrating multidimensional data in the near future would be a significant development<sup>39</sup>.

In conclusion, we defined two PD subtypes with distinct sequences of clinical and neurodegeneration events. These subtypes exhibit differing clinical and imaging profiles and responses to treatment. These findings corroborated the classical view on disease progression reveal new insights into the non-classical trajectory of clinical symptoms and neurodegeneration. This classification might also be useful for stratifying patients entering clinical trials, and might eventually shape more individualized treatments.

## METHODS

### Participants

The discovery dataset included 242 PD patients and 133 healthy control subjects. The PD patients were recruited from the Second Affiliated Hospital of Zhejiang University School of Medicine and diagnosed by experienced neurologists according to the United Kingdom Parkinson's Disease Society Brain Bank criteria. Healthy controls were recruited from the community. Participants with missing values in clinical assessments or MRI scans that used the SuStaln model were excluded. In addition, all images were visually inspected, and images showing intracranial mass, cerebrovascular disorders, and obvious artifacts were excluded. In total, 228 PD patients and 119 healthy controls were included.

The protocol, consent form, and other relevant documentations were approved by the ethics committee of the Second Affiliated Hospital of Zhejiang University School of Medicine before the study commenced. The study was performed in accordance with the Declaration of Helsinki and consistently with Good Clinical Practice. Before enrollment, all patients provided their written informed consent.

### Clinical and neuropsychological assessments

The UPDRS, HY, PDQ-39, and ADL scales were evaluated in PD patients. The subscores of tremors, rigidity, bradykinesia, and axial symptoms were calculated: subscore for tremor is obtained by adding the UPDRS-III items 20–21; subscore for rigidity is equal to the UPDRS-III item 22; subscore for bradykinesia is obtained by adding the UPDRS-III items 23–26 and 31; subscore for axial symptoms is obtained by adding the UPDRS-III items 27–30. All participants underwent evaluation for Brief Smell Identification Test (BSIT) modified for Chinese, Scales for Outcomes in Parkinson's Disease-Autonomic (SCOPA-AUT) survey, Rapid Eye Movement Sleep Behavior Disorder Questionnaire—Chinese University of Hong Kong version (RBDQ-HK), MoCA test, Hamilton Rating Scale for Depression (HAM-D). For PD patients, all of the above assessments were conducted during the "OFF state" (a period at least 12 h after withholding PD medications).

The self-reported lifetime histories of smoking, alcohol consumption, pesticide and toxic exposure of participants were recorded. Participants were defined as alcohol drinkers if they reported drinking at least 1 drink per week, and nondrinkers were defined as those reporting drinking 0 or <1 alcohol drink per week. Participants were defined as smokers if they reported smoking at least 1 cigarette per week, and nonsmokers were defined as those reporting smoking 0 or <1 cigarette per week. Pesticide exposure history refers to having worked in a job that required mixing, applying, or being exposed to any form of pesticide, such as herbicides, insecticides, fungicides, rodenticides, or acaricides (>6 months). Toxic exposure history was defined as exposed to some of toxic, including heavy metals, and organic pollutants (>6 months).

Additionally, 112 patients voluntarily underwent a levodopa challenge test. The UPDRS-III score was assessed during the OFF state and repeated one hour after administration of 200 mg levodopa and 50 mg benserazide (the "ON state")<sup>40</sup>. The levodopa response was expressed as the rate of change in the UPDRS-III score  $[(UPDRS-III_{OFF} - UPDRS-III_{ON})/UPDRS-III_{OFF}]$ .

### Magnetic resonance imaging data acquisition and analysis

**Image acquisition.** All participants were scanned on a GE Discovery MR750 3.0T MRI scanner. Earplugs and foam pads were used to reduce noise and head motion. High-resolution 3D T1-weighted structural MRI, diffusion tensor imaging (DTI), and neuromelanin-sensitive imaging (NM-MRI) were performed.

High-resolution 3D T1-weighted image was acquired using a fast spoiled gradient-recalled sequence: echo time (TE) = 3.036 ms; repetition time (TR) = 7.336 ms; inversion time = 450 ms; flip angle (FA) = 11°; field of view (FOV) = 260 × 260 mm<sup>2</sup>; matrix = 256 × 256; slice thickness = 1.2 mm; number of slices = 196 (sagittal). DTI was acquired using a spin echo-echo planar imaging sequence: TR = 8000 ms; TE = 80 ms; flip angle = 90 degrees; field of view = 256 × 256 mm<sup>2</sup>; matrix = 128 × 128; slice thickness = 2 mm; slice gap = 0 mm; number of slices = 67 (axial). Diffusion images were acquired from 30 gradient directions ( $b = 1000$  s/mm<sup>2</sup>), and included five acquisitions without diffusion weighting ( $b = 0$ ). An additional  $b = 0$  acquisition with reverse phase-encode polarity was acquired for distortion correction. Neuromelanin image was acquired using a T1-weighted fast spin echo sequence: TE = 18.6 ms; TR = 600 ms; FA = 77°; FOV = 220 × 220 mm<sup>2</sup>; matrix = 512 × 512; slice thickness = 3 mm; slice gap = 0 mm; number of slices = 17 (axial). Scanning coverage was set from the top of basal ganglia to the bottom of the medulla oblongata. The acquisition plane was orthogonal to the brainstem.

### Image processing

**Neuromelanin-sensitive MRI.** SN and LC are vital nodes in the trajectory of PD progression. With the advent of NM-MRI, the SN and LC are visible and measurable in vivo (Supplementary Fig. 3). Therefore, we used NM-MRI to measure the integrity of SN and LC. An author (C.Z.), who was blinded to the subjects' information, performed two manual measurements with a time interval of one month. These measurements were conducted using ITK-SNAP (<https://sourceforge.net/projects/itk-snap/>). The mean and standard deviation (SD) of the signal intensity (SI) in SN, LC, cerebral peduncle (CP), and pontine (PT) were calculated. The  $CNR_{SN/LC}$  was calculated as  $(SI_{SN/LC} - SI_{CP/PT})/SD_{CP/PT}$ . The averaged  $CNR_{SN/LC}$  value from twice assessments were used for final analysis<sup>41</sup>. A lower CNR value indicates more serious degeneration. Detailed steps about manual measurements could be found in our previous study<sup>35</sup>.

**Diffusion MRI.** According to previous studies focusing on the trajectory of PD progression<sup>18,20,28</sup>, we measure the microstructural integrity of amygdala, hippocampus, entorhinal cortex, and basal forebrain. In this study, we calculated the free-water fractional volumes of them, which is an advanced diffusion MRI measure that estimates the volume fraction of extracellular space<sup>42</sup>. Previous studies suggested that free-water increased under the conditions of neuroinflammatory and atrophy-based neurodegeneration<sup>43</sup>. Free-water is considered as a more sensitive measure than traditional diffusion and volume measures in neurodegeneration<sup>44</sup>.

DTI data were processed using the FMRIB Software Library (FSL, <http://www.fmrib.ox.ac.uk/fsl>) and MRtrix3 ([www.mrtrix.org](http://www.mrtrix.org)). First, skulls were stripped from the DTI data for each participant. Then, topup and eddy were executed to correct susceptibility-induced distortions, eddy currents, and movements in diffusion data<sup>45</sup>. According to previous study, free-water was calculated by fitting a



bi-tensor model<sup>46</sup>. The signal attenuation of the water molecules caused by and extracellular and intracellular water was predicted by this model, respectively. The derived free-water image reflects the volume fraction of extracellular space water. The steps to extract the free-water measure of above brain regions were as follows: 1) the native T1-weighted images were linearly and non-linearly aligned into Montreal Neurological Institute (MNI) standard space, and the outputs were visually inspected for errors and bad registration (4 PD patients and 1 healthy control were excluded in the following analysis); 2) the b0 images were linearly aligned to native T1-weighted images, and the free-water images were projected onto the T1-weighted images; 3) the free-water images were then linearly and non-linearly aligned into MNI space using the transformation matrix derived from T1-weighted images registration; 4) the free-water of four regions were then extracted: amygdala, hippocampus, entorhinal cortex, and basal forebrain. Amygdala, hippocampus, and entorhinal cortex atlas were acquired from the Human Brainnetome Atlas (<http://atlas.brainnetome.org/>). The basal forebrain (ch1-4 were combined) was defined according to the stereotaxic probabilistic maps imbedded in the Statistical Parametric Mapping (SPM, <https://www.fil.ion.ucl.ac.uk/spm/>) toolbox Anatomy v22c<sup>47</sup>. To alleviate partial volume effects, internal erosion with one voxel were implemented in all atlas.

In summary,  $CNR_{SN}$ ,  $CNR_{LC}$ , free-water of amygdala, hippocampus, entorhinal cortex, and basal forebrain were calculated for next analysis. We also evaluated the asymmetry of imaging features included in the SuStaln model. The formula used to calculate the asymmetry index was as follows:  $Asymmetry\ index = \frac{feature_{left} - feature_{right}}{(\text{mean}(feature_{left} + feature_{right}))}$ <sup>48</sup>.

### Construction of the subtype and stage inference model

**Feature selection.** Five continuous clinical features, namely RBDQ-HK, SCOPA-AUT, BSIT, MoCA, and HAMD, were included in the analysis. Motor symptoms were only assessed on PD patients, and were therefore excluded. In addition, as we aim to model the progression in established PD (rather than prodromal disease), which is not suitable for UPDRS scores (all patients have motor dysfunction) in the SuStaln model. The inclusion of UPDRS scores in the model could result in motor dysfunction being identified as the earliest event in PD progression. At the individual patient level, this may be misleading as research conducted on populations in preclinical stages of Parkinson's disease has revealed that changes in smell and sleep tend to appear prior to motor symptoms<sup>24</sup>.

MRI features included  $CNR_{SN}$ ,  $CNR_{LC}$ , free-water of amygdala, hippocampus, entorhinal cortex, and basal forebrain, each of them was averaged over the left and right hemispheres. As a result, a final subject\*feature matrix (347\*11) was obtained for further analysis. The confounding variables including age, gender, and years of education were adjusted for as covariates of no interest. The adjusted features were then converted into z scores relative to the healthy controls (by first subtracting the means of healthy controls, then dividing by the standard deviation of healthy controls). As the measure of BSIT, MoCA,  $CNR_{SN}$ , and  $CNR_{LC}$  decrease indicates impairment, we multiplied their z scores by  $-1$ , so that a larger z score indicates more severe impairment for all features. Finally, the normalized patient\*feature matrix (228\*11) was input into the model.

**Modeling.** pySuStaln, a Python implementation of the SuStaln algorithm<sup>49</sup>, was used to identify PD subtypes with distinct progression patterns and to grade a disease stage for each participant. Each feature was assigned three z scores (1, 2, and 3), which represented three continuous events (stages). In other words, each stage corresponds to a feature achieving a new z score. The advantage of this method is that it represents the

continuous linear accumulation of changes caused by pathologic damage, rather than an instantaneous switch from normal to abnormal. Therefore, the pattern of progression for each disease subtype was characterized as a piecewise linear z score model, which consisted of a series of stages. Individuals were typed according to the likelihood they belonged to each SuStaln subtype<sup>49</sup>, and the SuStaln stage was determined by choosing the most probable stage<sup>16</sup>. The estimation of model uncertainty involved conducting 1,000,000 MCMC iterations. SuStaln utilizes this method calculated probability for each individual into a specific stage and subtype based on their maximum likelihood<sup>16</sup>. Individuals with no biomarker abnormality were defined as being at stage 0 and no subtype was assigned. Finally, by iteratively increased the number of subtypes in the SuStaln model, a less complex model (smaller number of subtypes) with lower CVIC calculated through 10-fold cross-validation was considered as the optional number of subtypes<sup>21</sup>. The algorithms and analytical procedures used here have been described in detail for recent studies<sup>15,16,23</sup>.

**Validation dataset.** We validated the reproducibility of the PD subtypes in the PPMI dataset ([www.ppmi-info.org/access-data-specimens/download-data](http://www.ppmi-info.org/access-data-specimens/download-data)), RRID:SCR\_006431<sup>50</sup>. For up-to-date information on the PPMI, visit [www.ppmi-info.org](http://www.ppmi-info.org). Subjects were scanned on Siemens 3.0T TIM Trio scanners with the same scanning parameters. The scanning parameters of 3D T1 image were as follows: TR = 2300 ms; TE = 2.98 ms; inversion time = 900 ms; flip angle = 90°; slice number = 176; acquisition matrix = 240 × 256 and voxel size = 1 × 1 × 1 mm<sup>3</sup>. Diffusion tensor image was obtained with the following parameters: TR = 900 ms, TE = 88 ms, flip angle = 90°, voxel size = 2 × 2 × 2 mm<sup>3</sup>, slice number = 72, 64 gradient directions with a b-value of 1000 s/mm<sup>2</sup>. One non-gradient volume (b = 0 s/mm<sup>2</sup>) was also acquired. As the PPMI dataset does not contain NM-MRI data, we only calculated the free-water of SN and LC in MNI space. The SN label was obtained from the Automated anatomical labeling atlas 3<sup>51</sup>, and the LC label were acquired based on a LC probability atlas with peak signal coordinates observed at two SDs<sup>52</sup>. Consequently, the free-water of SN, LC, amygdala, hippocampus, entorhinal cortex, and basal forebrain were input into the SuStaln model. Clinical assessments including University of Pennsylvania Smell Identification Test (UPSIT), SCOPA-AUT, RBDQ, MoCA, and Geriatric Depression Scale (GDS) were used enrolled for SuStaln model construction. Finally, the baseline clinical and MRI data of 153 PD patients and 67 healthy controls were enrolled for validation analysis. The model construction procedures were same with the discovery dataset.

The MDS-UPDRS and HY scale were conducted during "OFF" state. Subscore for tremor is obtained by adding the MDS-UPDRS-III items 15–18. Subscore for rigidity is equal to the MDS-UPDRS-III item 3. Subscore for bradykinesia is obtained by adding the MDS-UPDRS-III items 4–9. Subscore for axial symptoms is obtained by adding the MDS-UPDRS-III items 1 and 10–13. In addition, subscore for PIGD was calculated as the mean of 5 items, including walking, balance and freezing in MDS-UPDRS-II items 12–13, and gait, postural stability and freezing of gait in MDS-UPDRS-III items 10–12<sup>53,54</sup>. The DAT binding rate (caudate and putamen) and cerebrospinal fluid proteins ( $\alpha$ -synuclein, amyloid- $\beta$ 42, tau, and phosphorylated tau protein) were collected. The details of the DAT processing and cerebrospinal fluid biomarker measurements could be found in our previous study<sup>55,56</sup>. The gene mutation information is only available in the PPMI dataset. According to the methods used in recent studies, the LRRK2 genetic testing included G2019S and R1441G/C, N1437H mutations; GBA genetic testing included N370S (also called N409S) in all, and L483P, L444P, IVS2 + 1, and 84GG mutations; SNCA genetic testing included A53T mutation;

dual mutation carriers for both LRRK2 and GBA were excluded<sup>57–59</sup>.

Part of the PPMI's participants collected the information about environmental factors tied to PD risk. This is conducted through a single site, University of California, San Francisco, and PPMI participants provide consent to be followed in PPMI FOUND separately. Alcohol drinkers were defined as drunk 100 or more alcoholic drinks during lifetime; smokers were defined as smoked 100 or more cigarettes during lifetime; coffee drinkers were defined as regularly drunk caffeinated coffee, that is, at least once per week for 6 months or longer; pesticides exposure were defined as had or ever had a job in which you mixed, applied, or were exposed in some other way to any type of pesticide, including herbicides, fungicides, insecticides, rodenticides or fumigants; toxicants exposure were defined as used glues or adhesives 100 or more days at work or at home during lifetime.

The PPMI project was approved by the Institutional Review Board or Independent Ethics Committee of all participating sites in Europe, including Attikon University Hospital (Greece), Hospital Clinic de Barcelona and Hospital Universitari Donostia (Spain), Innsbruck University (Austria), Paracelsus-Elena-Klinik Kassel/University of Marburg (Germany), Imperial College London (UK), Pitié-Salpêtrière Hospital (France), University of Salerno (Italy), and in the USA, including Emory University, Johns Hopkins University, University of Alabama at Birmingham, PD and Movement Disorders Center of Boca Raton, Boston University, Northwestern University, University of Cincinnati, Cleveland Clinic Foundation, Baylor College of Medicine, Institute for Neurodegenerative Disorders, Columbia University Medical Center, Beth Israel Medical Center, University of Pennsylvania, Oregon Health and Science University, University of Rochester, University of California at San Diego, and University of California, San Francisco. Informed consent was provided according to the Declaration of Helsinki.

**Statistical analysis.** Statistical Package for the Social Sciences (version 23) was used for statistical analysis. Comparisons of demographics, clinical and imaging variables between PD patients and healthy control were performed using two-sample *t*-test and chi-squared test as appropriate. General liner model was conducted to assess the difference of clinical/imaging variables, asymmetry index, levodopa response, DAT binding rate (validation dataset), and cerebrospinal fluid proteins (validation dataset) between PD subtypes. Confounding variables including gender, age, education, and SuStaln stage were adjusted as fixed effects in linear mixed-effects model. Comparisons of environmental factors and gene mutation (validation dataset) between PD subtypes were performed using chi-squared test. Pearson correlation analysis was conducted to assess the consistency between the SuStaln-model-inferred disease stage and UPDRS, PDQ-39, and ADL scores, which reflect the severity of the disease. FDR correction was performed for multiple comparison correction, and a two-tailed *p*-value < 0.05 was considered significant. Additionally, the linear mixed-effects model was conducted to assessed the difference of longitudinal progression between PD subtype (group\*time effect, validation dataset, Supplementary Table 5).

#### DATA AVAILABILITY

The dataset used in this study, obtained from The Second Affiliated Hospital, Zhejiang University School of Medicine, is available for qualified researchers to request for the purpose of reproducibility of the research. To access the dataset, interested researchers can contact the corresponding author directly. The PPMI dataset (<https://www.ppmi-info.org/>) used for this study is available for qualified researchers through application.

#### CODE AVAILABILITY

FSL (v6.0, <http://www.fmrib.ox.ac.uk/fsl/>), MRtrix3 ([www.mrtrix.org/](http://www.mrtrix.org/)), SPM12 (<https://www.fil.ion.ucl.ac.uk/spm/>), and the code for SuStaln model construction (<https://github.com/ucl-pond/pySuStaln>) are open sources.

Received: 22 January 2023; Accepted: 4 July 2023;

Published online: 13 July 2023

#### REFERENCES

- Kalia, L. V. & Lang, A. E. Parkinson's disease. *Lancet* **386**, 896–912 (2015).
- Lee, S. H. et al. Parkinson's disease subtyping using clinical features and biomarkers: literature review and preliminary study of subtype clustering. *Diagnostics* **12** <https://doi.org/10.3390/diagnostics12010112> (2022).
- Berg, D. et al. Prodromal Parkinson disease subtypes - key to understanding heterogeneity. *Nature reviews. Neurology* **17**, 349–361 (2021).
- Myers, P. S. et al. Distinct progression patterns across Parkinson disease clinical subtypes. *Ann. Clin. Transl. Neurol.* **8**, 1695–1708 (2021).
- Uribe, C. et al. Progression of Parkinson's disease patients' subtypes based on cortical thinning: 4-year follow-up. *Parkinsonism Relat. Disord.* **64**, 286–292 (2019).
- Mitchell, T. et al. Emerging neuroimaging biomarkers across disease stage in Parkinson disease: a review. *JAMA Neurol.* **78**, 1262–1272 (2021).
- De Pablo-Fernández, E., Lees, A. J., Holton, J. L. & Warner, T. T. Prognosis and neuropathologic correlation of clinical subtypes of Parkinson disease. *JAMA Neurol.* **76**, 470–479 (2019).
- Guo, T. et al. Clinically relevant connectivity features define three subtypes of Parkinson's disease patients. *Hum. Brain Mapp.* **41**, 4077–4092 (2020).
- Inguanzo, A. et al. Hierarchical cluster analysis of multimodal imaging data identifies brain atrophy and cognitive patterns in Parkinson's disease. *Parkinsonism Relat. Disord.* **82**, 16–23 (2021).
- Fereshtehnejad, S. M., Zeighami, Y., Dagher, A. & Postuma, R. B. Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression. *Brain* **140**, 1959–1976 (2017).
- Zhang, X. et al. Data-driven subtyping of Parkinson's disease using longitudinal clinical records: a cohort study. *Sci. Rep.* **9**, 797 (2019).
- Mestre, T. A. et al. Parkinson's disease subtypes: critical appraisal and recommendations. *J. Parkinsons Dis.* **11**, 395–404 (2021).
- Rodríguez-Sánchez, F. et al. Identifying Parkinson's disease subtypes with motor and non-motor symptoms via model-based multi-partition clustering. *Sci. Rep.* **11**, 23645 (2021).
- Wang, L. et al. Association of specific biotypes in patients with Parkinson disease and disease progression. *Neurology* **95**, e1445–e1460 (2020).
- Young, A. L. & Marinescu, R. V. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nat. Commun.* **9**, 4273 (2018).
- Vogel, J. W., Young, A. L. & Oxtoby, N. P. Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nat. Med.* **27**, 871–881 (2021).
- Mestre, T. A. et al. Reproducibility of data-driven Parkinson's disease subtypes for clinical research. *Parkinsonism Relat. Disord.* **56**, 102–106 (2018).
- Borghammer, P. & Van Den Berge, N. Brain-first versus gut-first Parkinson's disease: a hypothesis. *J. Parkinsons Dis.* **9**, S281–s295 (2019).
- Horsager, J. et al. Brain-first versus body-first Parkinson's disease: a multimodal imaging case-control study. *Brain* **143**, 3077–3088 (2020).
- Borghammer, P. et al. A postmortem study suggests a revision of the dual-hit hypothesis of Parkinson's disease. *NPJ Parkinson's Dis.* **8**, 166 (2022).
- Collij, L. E. et al. Spatial-temporal patterns of  $\beta$ -amyloid accumulation: a subtype and stage inference model analysis. *Neurology* **98**, e1692–e1703 (2022).
- Young, A. L. et al. Characterizing the clinical features and atrophy patterns of MAPT-related frontotemporal dementia with disease progression modeling. *Neurology* **97**, e941–e952 (2021).
- Eshaghi, A., Young, A. L. & Wijeratne, P. A. Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. *Nat. Commun.* **12**, 2078 (2021).
- Oxtoby, N. P. et al. Sequence of clinical and neurodegeneration events in Parkinson's disease progression. *Brain* **144**, 975–988 (2021).
- Braak, H. et al. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* **24**, 197–211 (2003).
- Braak, H., Ghebremedhin, E., Rüb, U., Bratzke, H. & Del Tredici, K. Stages in the development of Parkinson's disease-related pathology. *Cell Tissue Res.* **318**, 121–134 (2004).
- Heinzel, S. et al. Update of the MDS research criteria for prodromal Parkinson's disease. *Mov. Disord. Off. J. Mov. Disord. Soc.* **34**, 1464–1470 (2019).

28. Pyatigorskaya, N. et al. Parkinson disease propagation using MRI biomarkers and partial least squares path modeling. *Neurology* **96**, e460–e471 (2021).
29. Horsager, J., Knudsen, K. & Sommerauer, M. Clinical and imaging evidence of brain-first and body-first Parkinson's disease. *Neurobiol. Dis.* 105626 <https://doi.org/10.1016/j.nbd.2022.105626> (2022).
30. Raunio, A. et al. Lewy-related pathology exhibits two anatomically and genetically distinct progression patterns: a population-based study of Finns aged 85. *Acta Neuropathol.* **138**, 771–782 (2019).
31. Borghammer, P. et al. Neuropathological evidence of body-first vs. brain-first Lewy body disease. *Neurobiol. Dis.* **161**, 105557 (2021).
32. Halliday, G., Hely, M., Reid, W. & Morris, J. The progression of pathology in longitudinally followed patients with Parkinson's disease. *Acta Neuropathol.* **115**, 409–415 (2008).
33. 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).
34. Pagano, G. et al. REM behavior disorder predicts motor progression and cognitive decline in Parkinson disease. *Neurology* **91**, e894–e905 (2018).
35. Zhou, C. et al. Locus coeruleus degeneration correlated with levodopa resistance in parkinson's disease: a retrospective analysis. *J. Parkinsons Dis.* **11**, 1631–1640 (2021).
36. Ostock, C. Y., Lindenbach, D., Goldenberg, A. A., Kampton, E. & Bishop, C. Effects of noradrenergic denervation by anti-DBH-saporin on behavioral responsiveness to L-DOPA in the hemi-parkinsonian rat. *Behav. Brain Res.* **270**, 75–85 (2014).
37. Qiao, Y. & Shi, Y. Unsupervised deep learning for FOD-based susceptibility distortion correction in diffusion MRI. *IEEE Trans. Med. Imaging* **41**, 1165–1175 (2022).
38. Golub, M., Neto Henriques, R. & Gouveia Nunes, R. Free-water DTI estimates from single b-value data might seem plausible but must be interpreted with care. *Magn. Reson. Med.* **85**, 2537–2551 (2021).
39. Young, A. L. et al. Ordinal SuStaln: subtype and stage inference for clinical scores, visual ratings, and other ordinal data. *Front. Artif. Intell.* **4**, 613261 (2021).
40. Dirks, M. F. et al. Cerebral differences between dopamine-resistant and dopamine-responsive Parkinson's tremor. *Brain* **142**, 3144–3157 (2019).
41. Zhou, C. et al. Locus coeruleus degeneration is associated with disorganized functional topology in Parkinson's disease. *NeuroImage Clin.* **32**, 102873 (2021).
42. Burciu, R. G. et al. Progression marker of Parkinson's disease: a 4-year multi-site imaging study. *Brain* **140**, 2183–2192 (2017).
43. Wang, Y. et al. Quantification of increased cellularity during inflammatory demyelination. *Brain* **134**, 3590–3601 (2011).
44. Andica, C. et al. Free-water imaging in white and gray matter in Parkinson's disease. *Cells* **8** <https://doi.org/10.3390/cells8080839> (2019).
45. Andersson, J. L. R. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* **125**, 1063–1078 (2016).
46. Zhou, L. et al. Increased free water in the substantia nigra in idiopathic REM sleep behaviour disorder. *Brain* **144**, 1488–1497 (2021).
47. Mesulam, M. M., Mufson, E. J., Wainer, B. H. & Levey, A. I. Central cholinergic pathways in the rat: an overview based on an alternative nomenclature (Ch1-Ch6). *Neuroscience* **10**, 1185–1201 (1983).
48. Banwinkler, M., Dzialas, V., Hoenig, M. C. & van Eimeren, T. Gray matter volume loss in proposed brain-first and body-first Parkinson's disease subtypes. *Mov. Disord.* **37**, 2066–2074 (2022).
49. Aksman, L. M. et al. pySuStaln: a Python implementation of the subtype and stage inference algorithm. *SoftwareX* **16** <https://doi.org/10.1016/j.softx.2021.100811> (2021).
50. Kenneth Marek, D. J. & Lasch, S. The Parkinson progression marker initiative (PPMI). *Prog. Neurobiol.* **95**, 629–635 (2011).
51. Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J. & Joliot, M. Automated anatomical labelling atlas 3. *Neuroimage* **206**, 116189 (2020).
52. Keren, N. I., Lozar, C. T., Harris, K. C., Morgan, P. S. & Eckert, M. A. In vivo mapping of the human locus coeruleus. *Neuroimage* **47**, 1261–1267 (2009).
53. Qin, Y. et al. Neuroimaging uncovers distinct relationships of glymphatic dysfunction and motor symptoms in Parkinson's disease. *J. Neurol.* **270**, 2649–2658 (2023).
54. Taniuchi, R. et al. The power of instruction on retropulsion: a pilot randomized controlled trial of therapeutic exercise focused on ankle joint movement in Parkinson's disease. *Clin. Park Relat. Disord.* **7**, 100151 (2022).
55. Wang, L. et al. Association of cortical and subcortical microstructure with clinical progression and fluid biomarkers in patients with Parkinson disease. *Neurology* <https://doi.org/10.1212/wnl.000000000207408> (2023).
56. Zhou, C. et al. Progressive brain atrophy in Parkinson's disease patients who convert to mild cognitive impairment. *CNS Neurosci. Ther.* **26**, 117–125 (2020).
57. Simuni, T. et al. Clinical and dopamine transporter imaging characteristics of leucine rich repeat kinase 2 (LRRK2) and glucosylceramidase beta (GBA) Parkinson's disease participants in the Parkinson's progression markers initiative: a cross-sectional study. *Mov. Disord.* **35**, 833–844 (2020).
58. Nalls, M. A. et al. Baseline genetic associations in the Parkinson's Progression Markers Initiative (PPMI). *Mov. Disord.* **31**, 79–85 (2016).
59. Koros, C. et al. Serum uric acid level as a biomarker in idiopathic and genetic (p.A53T alpha-synuclein carriers) Parkinson's disease: data from the PPMI study. *J. Parkinsons Dis.* **10**, 481–487 (2020).

## ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos. 82271935, 81971577, 82171888, 82202091, 82001767, 91630314, 82071997, and 81771216), the Natural Science Foundation of Zhejiang Province (Grant No. LY22H180002 and LQ21H180008), the 13th Five-year Plan for National Key Research and Development Program of China (Grant No. 2016YFC1306600), the Shanghai Municipal Science and Technology Major Project (No. 2018SHZDZX01), ZJ Lab, and Shanghai Center for Brain Science. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research funding partners 4D Pharma, Abbvie, Acurex Therapeutics, Allergan, Amathus Therapeutics, ASAP, Avid Radiopharmaceuticals, Bial Biotech, Biogen, BioLegend, Bristol-Myers Squibb, Calico, Celgene, Dacapo Brain Science, Denali, The Edmond J. Safra Foundation, GE Healthcare, Genentech, GlaxoSmithKline, Golub Capital, Handl Therapeutics, Insitro, Janssen Neuroscience, Lilly, Lundbeck, Merck, Meso Scale Discovery, Neurocrine Biosciences, Pfizer, Piramal, Preval, Roche, Sanofi Genzyme, Servier, Takeda, Teva, UCB, Verily, and Voyager Therapeutics.

## AUTHOR CONTRIBUTIONS

C.Z. and L.W.: Research project—Conception and Execution, Statistical analysis—Design and Execution, Manuscript—Writing of the first draft; W.C., J.L.: Statistical analysis—Review, Manuscript—Review; X.G., T.Guo., J.W., W.Z., T.Gao., X.L., X.B., H.W., Z.C., L.G., J.C., J.W., P.H., X.X., and B.Z.: Data—Collection and Manuscript—Review; J.F. and M.Z.: Research project—Conception and Organization; Manuscript—Review.

## 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/s41531-023-00556-3>.

**Correspondence** and requests for materials should be addressed to Wei Cheng, Jianfeng Feng or Mingming Zhang.

**Reprints and permission 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 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023