Temporal Progression Patterns of Brain Atrophy in Corticobasal Syndrome and Progressive Supranuclear Palsy Revealed by Subtype and Stage Inference (SuStaIn)

Yuya Saito[1], Koji Kamagata[1]*, Peter A. Wijeratne[2], Christina Andica[1], Wataru Uchida[1], Kaito Takabayashi[1], Shohei Fujita[1,3], Toshiaki Akashi[1], Akihiko Wada[1], Keigo Shimoji[4], Masaaki Hori[5], Yoshitaka Masutani[6], Daniel C. Alexander[2] and Shigeki Aoki[1]

[1] Department of Radiology, Juntendo University Graduate School of Medicine, Tokyo, Japan, [2] Centre for Medical Image Computing, Department of Computer Science, University College London, London, United Kingdom, [3] Department of Radiology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan, [4] Department of Radiology, Tokyo Metropolitan Geriatric Hospital and Institute of Gerontology, Tokyo, Japan, [5] Department of Radiology, Toho University Omori Medical Center, Tokyo, Japan, [6] Department of Biomedical Information Sciences, Hiroshima City University Graduate School of Information Sciences, Hiroshima, Japan

Differentiating corticobasal degeneration presenting with corticobasal syndrome (CBD-CBS) from progressive supranuclear palsy with Richardson's syndrome (PSP-RS), particularly in early stages, is often challenging because the neurodegenerative conditions closely overlap in terms of clinical presentation and pathology. Although volumetry using brain magnetic resonance imaging (MRI) has been studied in patients with CBS and PSP-RS, studies assessing the progression of brain atrophy are limited. Therefore, we aimed to reveal the difference in the temporal progression patterns of brain atrophy between patients with CBS and those with PSP-RS purely based on cross-sectional data using Subtype and Stage Inference (SuStaIn)—a novel, unsupervised machine learning technique that integrates clustering and disease progression modeling. We applied SuStaIn to the cross-sectional regional brain volumes of 25 patients with CBS, 39 patients with typical PSP-RS, and 50 healthy controls to estimate the two disease subtypes and trajectories of CBS and PSP-RS, which have distinct atrophy patterns. The progression model and classification accuracy of CBS and PSP-RS were compared with those of previous studies to evaluate the performance of SuStaIn. SuStaIn identified distinct temporal progression patterns of brain atrophy for CBS and PSP-RS, which were largely consistent with previous evidence, with high reproducibility (99.7%) under cross-validation. We classified these diseases with high accuracy (0.875) and sensitivity (0.680 and 1.000, respectively) based on cross-sectional structural brain MRI data; the accuracy was higher than that reported in previous studies. Moreover, SuStaIn stage correctly reflected disease severity without the label of disease stage, such as disease duration. Furthermore, SuStaIn also showed the genialized performance of differentiation

and reflection for CBS and PSP-RS. Thus, SuStaIn has potential for improving our understanding of disease mechanisms, accurately stratifying patients, and providing prognoses for patients with CBS and PSP-RS.

INTRODUCTION

Corticobasal degeneration (CBD) and progressive supranuclear palsy (PSP) are sporadic atypical parkinsonian disorders associated with the accumulation of insoluble deposits of predominantly four microtubule-binding domain repeat (4R) tau protein in specific central nervous system neurons and glia (1–4). CBD has a variety of phenotypes, and pathological symptoms depend on tau lesions in the frontoparietal cortex, particularly in the primary motor and somatosensory cortices (5, 6). The most common clinical syndrome of CBD is corticobasal syndrome (CBS), which is characterized by Parkinsonism, rigidity, unilateral dystonia, myoclonus, alien limb, and ideomotor apraxia (7). In contrast, PSP is a classical Richardson's syndrome with symptoms of postural instability and vertical supranuclear gaze palsy (8); it is pathologically characterized by tau lesions mainly in the midbrain and superior cerebellar peduncle (SCP), especially in the substantia nigra and dentate nucleus (9). Currently, effective treatments for CBS or PSP-RS are not available. Moreover, the etiology and onset mechanism remain poorly understood. Furthermore, the neurodegenerative conditions closely overlap in terms of clinical information, pathology, biochemistry, and genetic risk factors; thus, differentiating CBD presenting with CBS (CBD-CBS) from PSP-RS, particularly in early stages, is often difficult (10). Several morphological markers on MRI, including the "hummingbird" sign of midbrain atrophy compared with pons and the "morning glory" sign of midbrain tegmentum atrophy (11–13), can indicate PSP-RS. However, in the study of MR findings before autopsy confirmation, these signs have high specificity but low sensitivity ["hummingbird" sign (sensitivity: ∼51.0%; specificity: ∼99.5%] and "morning glory" sign (sensitivity: ∼37.0%; specificity: ∼97.0%)] for distinguishing PSP-RS from CBD-CBS (14). However, accurate differentiation between PSP-RS and CBS is important to facilitate the early diagnosis of PSP-RS and CBS for the accurate prognostication and stratification of patients for clinical trials.

Numerous studies have evaluated MRI-based brain volumetry data to distinguish CBS from PSP-RS based on the specific patterns of brain atrophy. Brain atrophy in CBS primarily involves the frontoparietal lobe, especially the pre- and postcentral gyri (15–21). In contrast, in patients with PSP-RS, brain atrophy is most prominent in the brainstem regions, particularly in the midbrain tegmentum and SCP (15–19, 22, 23). Although revealing the temporal progression patterns of brain atrophy may help understand disease mechanisms and enable more accurate patient stratification and prognostication, most studies evaluating brain MRI data in patients with CBS and PSP-RS were cross-sectional in nature. A significant

amount of longitudinal data is required to evaluate disease progression, and the collection of such data adds a considerable burden in terms of time, effort, and money. Moreover, tracking large populations is challenging. Although a few studies have investigated longitudinal brain atrophy using longitudinal brain structural MRI at some time points (e.g., baseline, over 6 months and 1 year) based on the assumption that volumetric changes are linear (24–26), they failed to identify temporal atrophy progression with over a few disease stages or provide more information on the longitudinal atrophy of patients with CBS and PSP-RS.

The accuracy of MRI-based brain volumetry has not always been higher than that of clinical criteria; sample sizes were often small, and its use for single-subject decision-making was limited. Therefore, various classifiers based on MRI-based brain volumetry have been proposed for differentiating patients with PSP-RS from those with CBS. In a previous study, Correia et al. used a support vector machine (SVM)—a statistical classifier—and gray matter volume data to classify 19 patients with PSP-RS and 19 patients with CBS; however, the classification accuracy was only 62.2% (27). Gröschel et al. (16) used a mathematical model for brain MR volumetry, including the midbrain, parietal white matter, temporal gray matter, brainstem, frontal white matter, and pons, in patients with PSP-RS ($n = 33$) and CBS ($n = 18$) and achieved a classification accuracy of 79.5%. Therefore, it is challenging to precisely differentiate between CBS and PSP-RS even using the latest methods. The low accuracy of differentiation between patients with CBS and those with PSP-RS might have resulted from developing the model with an aim to only predict disease subtypes, without considering the disease-stage heterogeneity (28).

Despite the development of disease progression and classifier models for CBS and PSP-RS using MRI-based brain volumetric data, to our knowledge, no model has been developed that can integrates and simultaneously estimates disease progression as well as can differentiate CBS and PSP-RS based on the longitudinal and low amount of cross-sectional brain structural MRI regional brain volume data. Recently, an unsupervised machine learning technique called Subtype and Stage Inference (SuStaIn) (28) was developed to identify data-driven disease phenotypes with distinct temporal progression patterns (**Figure 1**). The technique integrates clustering and disease progression modeling based on widely available cross-sectional data (**Figures 1a–c**). Thus, compared with models that only predict disease subtypes, SuStaIn can model disease progression using only cross-sectional, but not longitudinal, data and disease-stage heterogeneity to allow better stratification of patients with CBS and those with PSP-RS (28). If the sample size of the input cross-sectional data (**Figure 1b**) is insufficient
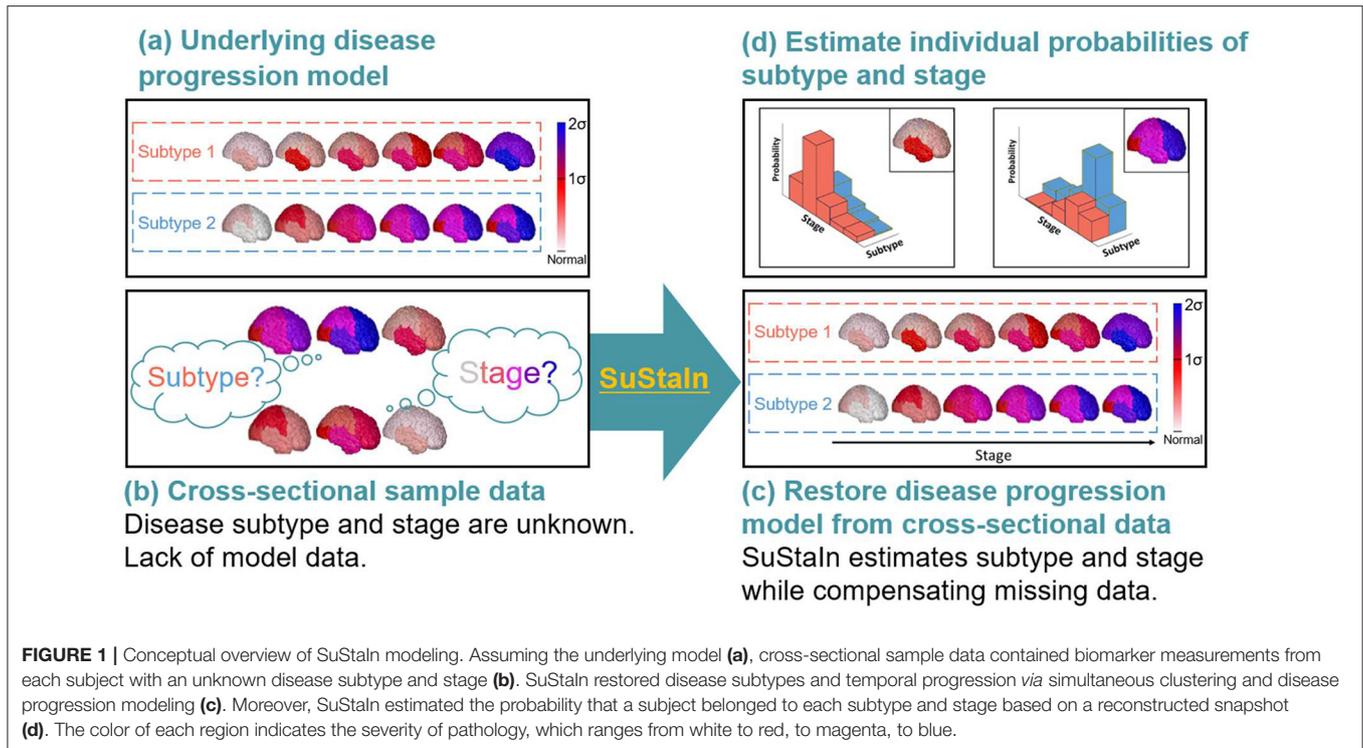
FIGURE 1 | Conceptual overview of SuStaIn modeling. Assuming the underlying model **(a)**, cross-sectional sample data contained biomarker measurements from each subject with an unknown disease subtype and stage **(b)**. SuStaIn restored disease subtypes and temporal progression *via* simultaneous clustering and disease progression modeling **(c)**. Moreover, SuStaIn estimated the probability that a subject belonged to each subtype and stage based on a reconstructed snapshot **(d)**. The color of each region indicates the severity of pathology, which ranges from white to red, to magenta, to blue.

for reconstructing the underlying disease progression model (**Figure 1a**), SuStaIn can restore disease progression model from the insufficient input data (**Figure 1c**) and estimate individual probabilities of disease subtypes and stages (**Figure 1d**). Young et al. (28) applied SuStaIn to reveal disease subtypes and distinct trajectories of regional brain atrophy-based MR volumetry in genetic frontotemporal dementia (FTD) and Alzheimer's disease (AD). As a validation, SuStaIn correctly identifies distinct genetic subtypes of FTD without seeing the genetic information. In sporadic AD, the algorithm identified three subtypes that correspond to end-stage patterns observed in postmortem pathology. Moreover, SuStaIn provided good separation between presymptomatic and symptomatic mutation carriers of FTD, cognitively normal patients, and patients with AD. Thus, SuStaIn may enable the simultaneous identification of disease subtypes and allow inferences on the progression of each subtype.

In this study, we applied SuStaIn to cross-sectional brain structural MRI regional brain volume data to identify differences in the temporal progression patterns of brain atrophy between patients with CBS and PSP-RS. We hypothesized that SuStaIn is useful in assessing and differentiating the progressions of CBS and PSP-RS.

METHODS

Study Cohorts

The data used in the preparation of this manuscript were obtained from the 4-Repeat Tauopathy Neuroimaging Initiative (4RTNI) and the Frontotemporal Lobar Degeneration Neuroimaging Initiative (FTLDNI) database. The primary goal of

the 4RTNI is to identify neuroimaging and biomarker indicators for disease progression in 4R tauopathy neurodegenerative diseases, PSP-RS, and CBD. The primary goals of FTLDNI are to identify neuroimaging modalities and analysis methods for tracking frontotemporal lobar degeneration and assess the value of imaging vs. other biomarkers for diagnosis. Both initiatives are managed by the University of California San Francisco (UCSF) and follow the same principal study design and protocol. Detailed information on 4RTNI and FTLDNI is available at http://4rtni-ftldni.ini.usc.edu/. In this study, we used two datasets of different sites and MR acquisition parameters as (internal dataset: 114 participants, external dataset: 17 participants).

As internal dataset, data of 25 patients with CBS and 39 with PSP-RS were obtained from the 4RTNI database. Data of 50 healthy controls (HCs) were obtained from the FTLDNI database. This study included the data of age-, sex-, and disease duration-matched participants recruited at the UCSF during their first visit and who were scanned using the same MRI acquisition parameters (see Methods: MRI acquisition) as those used for participants to adjust for any effects due to the differences in MRI acquisition site and parameters (29). To evaluate generalization performance of SuStaIn, as external dataset, data of 5 patients with CBS and 12 with PSP-RS, who were recruited at 3 sites [University of Toronto (Toronto); University of California, San Diego (UCSD); Massachusetts General Hospital [MGH]] during their first visit and who were scanned using the same MRI acquisition parameters, were obtained from the 4RTNI database.

All participants underwent a comprehensive neurological examination, which included the following: the PSP Rating Scale

(PSPRS) ranging from 0 (best) to 100 (worst) (30), four motor subscores of total PSPRS, including PSPRS Bulbar, PSPRS Ocular Motor, PSPRS Limb Motor, and PSPRS Gait/Midline (30) and Unified Parkinson's Disease Rating Scale (UPDRS) ranging from 0 (best) to 108 (worst) (31) for evaluating motor function; the Mini-Mental State Examination (MMSE) ranging from 0 (worst) to 30 (best) (32) and Montreal Cognitive Assessment (MoCA) ranging from 0 (worst) to 30 (best) (33) for evaluating cognitive function; the Clinical Dementia Rating (CDR) Box ranging from 0 (best) to 18 (worst) (34), the Schwab and England Activities of Daily Living (SEADL) ranging from 0% (worst) to 100% (best) (35), and the Functional Activities Questionnaire (FAQ) ranging from 0 (best) to 30 (worst) (36) for evaluating abilities of daily living activities. Patients with PSP-RS met the National Institute of Neurological Disorders and Stroke for PSP (NINDS-SPSP) criteria for PSP-Richardson's syndrome (AL-108-231) (8, 37). Patients with CBS met the Armstrong's criteria for possible or probable CBD-CBS subtype (7). Patients with CBS and PSP-RS who had no motor symptoms were excluded. All HCs were cognitively normal with an MMSE score of ≥27 and a MoCA score of ≥25 and had no impairments in the activities of daily living with a CDR Box score of ≤1, FAQ score of ≤1, and a SEADL score of 100%. Autopsy-confirmed diagnosis was available for two patients with CBD-CBS and two patients with PSP-RS. Furthermore, T1-weighted images of all participants were evaluated by a neuroradiologist (K.K.) with >10 years of experience in MRI of neurodegenerative disease to assess the classical MR findings of the "asymmetrical frontoparietal atrophy" in CBS (15–21) and "hummingbird" sign in PSP-RS (11–13). Population demographics and clinical characteristics are shown in **Table 1**.

MRI Acquisition

For internal dataset, three-dimensional T1-weighted images were acquired at the UCSF using the same MRI scanner and acquisition parameters to remove effects caused by differences in MRI acquisition site and parameters. MRI data were acquired on a Siemens Tim Trio (Siemens Healthcare Inc., Erlangen, Germany) 3 Tesla MRI scanner with a 12-channel receiver head coil. Whole brain images were acquired using a volumetric magnetization-prepared rapid gradient-echo sequence using the following parameters: repetition time/echo time/inversion time = 2300/2.98/900 ms; α = 9°; sagittal orientation, 256 × 240 × 160 matrix size; and 1 mm³ isotropic voxel resolution. MRI data quality was centrally evaluated at the UCSF.

For external dataset, three-dimensional T1-weighted images were acquired at the Toronto, UCSD, MGH using the same MRI scanner and acquisition parameters. MRI data were acquired on a GE DISCOVERY MR 750 (GE Healthcare Inc., Milwaukee, WI, USA) 3 Tesla MRI scanner with a 8-channel receiver head coil. Whole brain images were acquired using a volumetric magnetization-prepared rapid gradient-echo sequence using the following parameters: repetition time/echo time/inversion time = 7340/3.04/400 ms; α = 11°; sagittal orientation, 196 × 256 × 256 matrix size; and 1.20 × 1.02 × 1.02 mm³ isotropic voxel resolution. These MRI data qualities were also centrally evaluated at the UCSF.

MRI Processing and Volumetry

In total, 10 subregions of gray matter volume were calculated using FreeSurfer (38) version 6.0.0, including the frontal, temporal, parietal, and occipital lobes as well as the cingulate gyrus, basal ganglia, SCP, midbrain, pons, and medulla. To remove the effect of different head sizes, all regional volumes were normalized by dividing the volumes by total intracranial volume. Each regional volume was converted into a z-score relative to a control population so that the control population had a mean of 0 and a standard deviation of 1. Because z-scores become negative as regional brain volumes decrease with disease progression, for simplicity, we multiplied z-scores by minus one so that they increased as regional brain volumes decreased.

Brain Atrophy Progression Modeling

We applied SuStaIn (28) to the cross-sectional regional brain volumes of internal dataset to estimate the two disease subtypes and trajectories of CBS and PSP-RS, which have distinct atrophy patterns (15, 18, 39, 40). A conceptual overview of SuStaIn modeling is shown in **Figure 1**. SuStaIn evaluated the optimal clustering of individuals for each disease subtype (CBS or PSP-RS) that reflected distinct patterns of brain atrophy progression based on cross-sectional regional brain volumes. Simultaneously, each progression pattern was inferred as a sequential transition of individual subregions of gray matter volume from one z-score to another, relative to a control population. Subtype and disease stage were determined by those with the highest likelihood of participants being assigned to each disease subtype and disease stage. The progression trajectory of each subtype was described as a linear z-score model (28), which was composed of a sequence of stages in which each regional brain volume followed a piecewise linear trajectory. In practice, the number of progression patterns of brain atrophy is too large to evaluate all possible progression patterns. So, we use Markov Chain Monte Carlo (MCMC) sampling, which is able to indirectly obtain inference on the posterior distribution using computer simulations (41), to provide an approximation to this uncertainty (42, 43). We used 1,000,000 MCMC samples initialized from the maximum likelihood solution.

The linear z-score model underlying SuStaIn is based on the event-based model (42, 43). The event-base model identifies disease progression as a series of events corresponding to a regional brain atrophy from a normal to an abnormal level in this study. For biomarker $i = 1 \ldots l$ (i.e., brain region), the occurrence of an event E_i (i.e., brain atrophy) is informed by the measurements x_{ij} of biomarker i in subject $j, j = 1 \ldots J$ (i.e., regional brain volume). The most likely ordering of the event is the sequence S (i.e., brain atrophy progression) that maximizes the data likelihood is described using as the whole data set $X = \{x_{ij} | i = 1 \ldots I, j \ldots J\}$ as:

$$P(X|S) = \prod_{j=1}^{J} \left[\sum_{k=0}^{I} \left(P(k) \prod_{i=1}^{k} P(x_{ij}|E_i) \prod_{i=k+1}^{I} P(x_{ij}|\neg E_i) \right) \right],$$

where $P(x_{ij}|E_i)$ and $P(x_{ij}|\neg E_i)$ are the likelihoods of measurement x given that biomarker i has or has not

TABLE 1 | Demographic characteristics of participants.

Information	HC	CBS	PSP-RS	HC vs. CBS/PSP-RS *P*-value	CBS vs. PSP-RS *P*-value
(a) Internal dataset					
No. of MRI scans	50	25	39	—	—
Age, y	68.7 ± 4.0	67.8 ± 5.2	68.8 ± 7.1	*n.s.*	*n.s.*
Sex, % male	52.0	52.0	53.8	*n.s.*	*n.s.*
Disease duration, y	—	5.3 ± 2.8	4.8 ± 3.1	—	*n.s.*
Autopsy-confirmed subjects	—	2	2	—	—
PSPRS total	—	25.6 ± 9.5	36.0 ± 16.0	—	<0.05*
PSPRS bulbar	—	1.1 ± 1.0	2.6 ± 1.5	—	<0.001****
PSPRS ocular motor	—	1.9 ± 1.9	7.4 ± 3.6	—	<0.001****
PSPRS limb motor	—	8.7 ± 3.2	4.8 ± 2.5	—	<0.001****
PSPRS gait/midline	—	6.0 ± 4.9	9.8 ± 5.3	—	<0.05*
UPDRS-III total	—	35.4 ± 16.0	30.8 ± 16.5	—	*n.s.*
CDR box	0.1 ± 0.2	3.4 ± 2.9	3.7 ± 3.0	<0.001****	*n.s.*
MMSE total	29.4 ± 0.8	23.8 ± 5.5	25.2 ± 3.6	<0.001****	*n.s.*
MoCA total	27.4 ± 1.3	19.7 ± 6.5	21.3 ± 3.8	<0.001****	*n.s.*
FAQ total	0.0 ± 0.2	11.2 ± 7.0	14.1 ± 7.7	<0.001****	*n.s.*
SEADL, %	100.0 ± 0.0	56.8 ± 16.9	55.8 ± 27.2	<0.001****	*n.s.*
MRI findings					
Asymmetrical frontoparietal atrophy, % (*n*)	0% (0/50)	84.0% (21/25)	23.1% (9/39)	—	—
Hummingbird sign, % (*n*)	0% (0/50)	16.0% (4/25)	64.1% (25/39)	—	—

Information		CBS	PSP-RS		CBS vs. PSP-RS *P*-value
(b) External dataset					
No. of MRI scans		5	12		—
Age, y		68.8 ± 9.2	71.0 ± 7.5		*n.s.*
Sex, % male		60.0	50.0		*n.s.*
Disease duration, y		*none*	*none*		—
Autopsy confirmed subjects		0	0		—
PSPRS Total		24.6 ± 6.6	37.2 ± 19.1		*n.s.*
PSPRS Bulbar		2.2 ± 1.2	2.4 ± 1.3		*n.s.*
PSPRS Ocular–Motor		1.8 ± 1.5	7.3 ± 4.0		<0.01**
PSPRS Limb–Motor		7.8 ± 2.9	4.9 ± 3.3		*n.s.*
PSPRS Gait/Midline		4.0 ± 10.5	10.5 ± 6.1		*n.s.*
UPDRS-III Total		32.8 ± 7.4	26.8 ± 13.0		*n.s.*
CDR Box		4.8 ± 4.5	5.8 ± 4.5		*n.s.*
MMSE		18.8 ± 10.6	18.8 ± 7.0		*n.s.*
MoCA Total		21.3 ± 11.6	24.5 ± 7.9		*n.s.*
FAQ		11.8 ± 9.5	12.0 ± 10.0		*n.s.*
SEADL, %		52.0 ± 19.4	60.9 ± 32.9		*n.s.*
MRI findings					
Asymmetrical frontoparietal atrophy, % (*n*)		80.0% (4/5)	27.3% (3/12)		—
Hummingbird sign, % (*n*)		20.0% (1/5)	63.6% (8/12)		—

HC, healthy control; CBS, corticobasal syndrome; PSP, progressive supranuclear palsy; PSPRS, PSP Rating Scale; Four motor subscores from total PSPRS, PSPRS Bulbar, PSPRS Ocular–Motor, PSPRS Limb–Motor, and PSPRS Gait/Midline; UPDRS-III, Part-III (motor exams) of the Unified Parkinson's Disease Rating Scale; CDR Box, Clinical Dementia Rating Sum of Boxes; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; FAQ, Functional Activities Questionnaire; SEADL, Schwab and England Activities of Daily Living; n.s., not significant. *: P < 0.05, **: P < 0.01, ***: P < 0.005, ****: P < 0.001.

become abnormal, respectively. $P(k)$ is the prior likelihood of being at stage k and, at which the events $E_1, \ldots, E_k$ have occurred, and the events $E_{k+1}, \ldots, E_l$ have yet to occur. The model uses a uniform prior on the stage, so that a priori individuals are equally likely to belong to any stage along the progression pattern. The likelihoods $P(x_{ij}|E_i)$ and $P(x_{ij}|\neg E_i)$ are modeled as normal distributions from input data. The linear *z*-score model in this study reformulated the event-based model by replacing the instantaneous normal to abnormal events with events that

represent the linear accumulation of a biomarker from zero to two z-score.

Model fitting requires simultaneously optimizing disease subtype (i.e., CBS and PSP-RS), subtype trajectory and the posterior distributions of both. The cost function depends on the sequence ordering based on the well-established methods developed for the event-based model (42–45) with convergence and optimality in simulation and in our datasets. The SuStaIn model was fitted hierarchically with the number of clusters estimated *via* model selection criteria obtained from cross-validation. The hierarchical fitting initializes the fitting of each C-cluster (subtype) model from the previous C-1-cluster model, that is, the clustering problem was solved sequentially from $C = 1, 2$ in this study, initializing each model using the previous model.

Cross-Validation

Fivefold cross-validation was performed to evaluate the consistency of the disease progression patterns of each subtype by splitting the data into five folds and fitting the model to each subset (i.e., four of folds). The remaining fold was retained to test the model so that the model was fitted to each fold repeatedly, and ultimately, it was fitted to the whole internal dataset.

Model cross-validation similarity was measured using the Bhattacharyya coefficient (46) to assess the consistency of models across cross-validation folds.

The Bhattacharyya coefficient of the event position between each subregion of the gray matter volume of the two subtype progression patterns was measured by averaging across the events of subregion volume and MCMC method samples. The Bhattacharyya coefficient ranged from 0 (no similarity) to 1 (maximum similarity) for distribution similarity of event positions of subregion volumes of the subtype sequences.

Classification of Disease Subtypes

To evaluate the classification performance of SuStaIn for CBS and PSP-RS, we calculated the classification accuracy and sensitivity of the fivefold cross-validation for CBS and PSP-RS using internal dataset. Additionally, to evaluate generalization performance of SuStaIn classification, we also calculated the classification accuracy and sensitivity of the external dataset for CBS and PSP-RS with parameters that were fitted using the internal dataset.

Relationship Between SuStaIn Stage and Disease Severity

To evaluate the clinical utility of SuStaIn for CBS and PSP-RS, the correlation analysis was performed between the disease stage of internal dataset estimated by SuStaIn and disease severity including disease duration, total PSPRS score, PSPRS subscores, SEADL, UPDRS-III, MoCA, MMSE, CDR Box FAQ scores using Spearman rank correlation test. Additionally, to evaluate generalization performance of SuStaIn staging, we also analyzed these correlations of the external dataset that individual subject was assigned to a SuStaIn stage with parameters that were fitted using the internal dataset. Statistical significance for the correlation tests was set at 0.05.

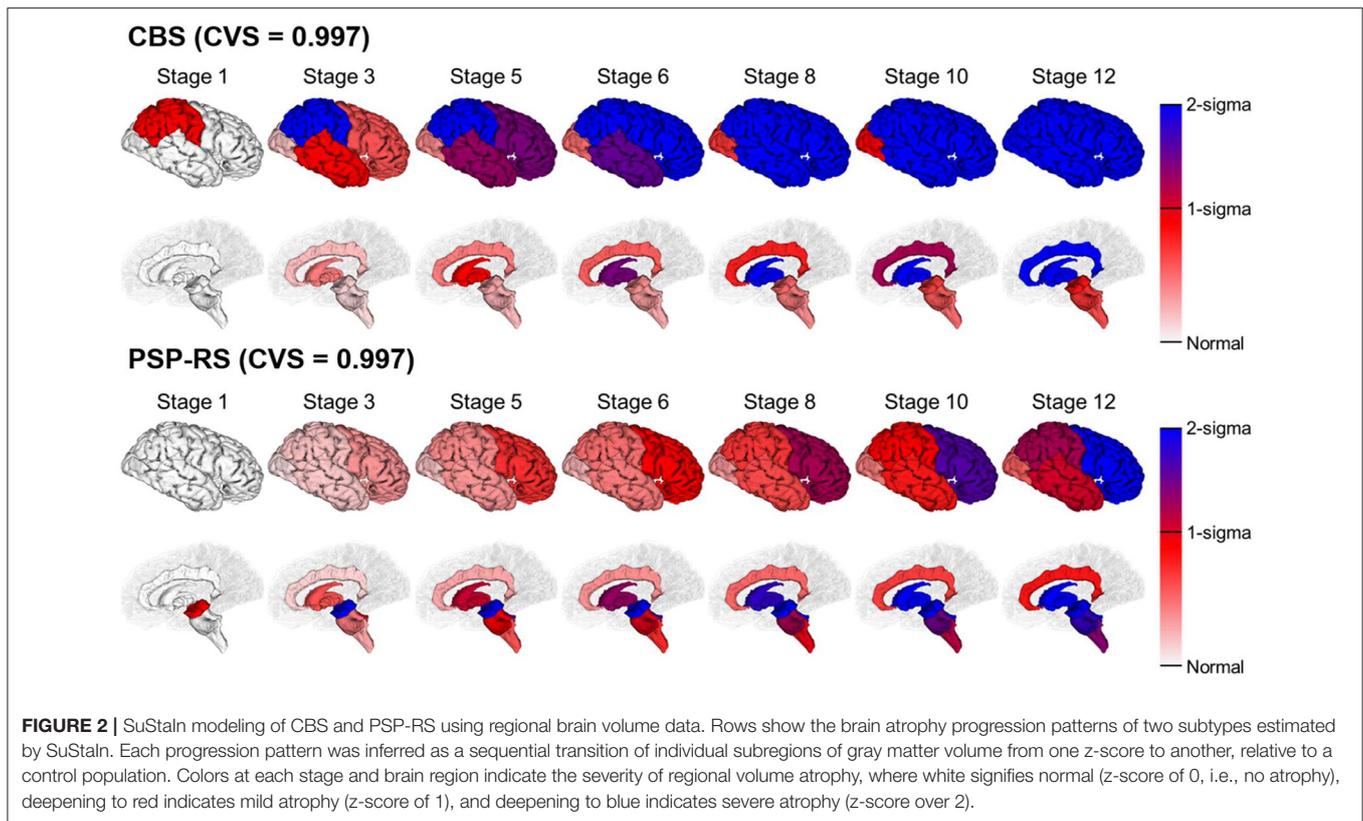# RESULTS

Demographic and Clinical Data Findings

Table 1 summarizes the demographic, clinical data and MRI findings. Groups were well-matched for age and sex distributions. The patient group in the internal dataset showed significant impairments in both motor and cognitive functions compared with the control group as shown using CDR Box data, MMSE total, MoCA total, FAQ, and SEADL scores. The PSP-RS group had significantly higher total PSPRS score and PSPRS subscores than the CBS group, with the exception of the PSPRS limb motor subscore, which showed that the CBS group was more impaired than the PSP-RS group.

In the internal dataset, overall, 84.0% (21/25) of patients with CBS had asymmetrical frontoparietal atrophy and 16.0% (4/25) showed the hummingbird sign. Four patients with CBS with the hummingbird sign also had asymmetrical frontoparietal atrophy, so the percentage of patients with CBS with only asymmetrical frontoparietal atrophy was 68.0% (17/25). In contrast, 23.1% (9/39) of patients with PSP-RS had asymmetrical frontoparietal atrophy and 64.1% (25/39) showed the hummingbird sign. Seven of 25 patients with PSP-RS with the hummingbird sign also had asymmetrical frontoparietal atrophy, so the percentage of patients with PSP-RS with only the hummingbird sign was 46.2% (18/39). In summary, the classification accuracy between CBS and PSP-RS based on classical MR signs was 54.7% (35/64) and the sensitivities for CBS and PSP-RS were 68.0% (17/25) and 46.2% (18/39), respectively.

In the external dataset, overall, 80.0% (4/5) of patients with CBS had asymmetrical frontoparietal atrophy and 20.0% (1/5) showed the hummingbird sign. One patient with CBS with the hummingbird sign also had asymmetrical frontoparietal atrophy, so the percentage of patients with CBS with only asymmetrical frontoparietal atrophy was 60.0% (3/5). In contrast, 27.3% (3/12) of patients with PSP-RS had asymmetrical frontoparietal atrophy and 63.6% (8/12) showed the hummingbird sign. Three of eight patients with PSP-RS with the hummingbird sign also had asymmetrical frontoparietal atrophy, so the percentage of patients with PSP-RS with only the hummingbird sign was 41.7% (5/12). In summary, the classification accuracy between CBS and PSP-RS based on classical MR signs was 47.1% (8/17) and the sensitivities for CBS and PSP-RS were 60.0% (3/5) and 41.7% (5/12), respectively. Thus, the external dataset had similar characters of MR findings to the internal dataset.

Disease Subtype Progression Patterns

Figure 2 shows the temporal progression patterns of atrophy of the CBS and PSP-RS groups obtained using SuStaIn using the internal dataset. SuStaIn identified that atrophy in patients with CBS started from the frontoparietal lobe, followed by the temporo-occipital lobe and basal ganglia, and finally reached the cingulate gyrus and brain stem. In contrast, in the PSP-RS group, atrophy started from the midbrain and SCP, followed by the pons, medulla, basal ganglia, cingulate gyrus, and frontoparietal lobe, and eventually reached the temporo-occipital lobe. These progression patterns for the two subtypes of CBS and PSP-RS were highly reproducible under cross-validation, which showed

FIGURE 2 | SuStaIn modeling of CBS and PSP-RS using regional brain volume data. Rows show the brain atrophy progression patterns of two subtypes estimated by SuStaIn. Each progression pattern was inferred as a sequential transition of individual subregions of gray matter volume from one z-score to another, relative to a control population. Colors at each stage and brain region indicate the severity of regional volume atrophy, where white signifies normal (z-score of 0, i.e., no atrophy), deepening to red indicates mild atrophy (z-score of 1), and deepening to blue indicates severe atrophy (z-score over 2).

an average similarity between cross-validation folds of >99.7% for each disease subtype.

Disease Classification

Table 2 shows the classification accuracy of SuStaIn for identifying disease subtype (CBS and PSP-RS) from cross-sectional MRI data. In the internal dataset, the classification accuracy of SuStaIn was 0.875 (56/64), and the sensitivity for CBS and PSP-RS were 0.680 (17/25) and 1.000 (39/39), respectively. In this study, the accuracy of the SuStaIn classification was superior to that of the classification based on MR findings (accuracy: 0.547 (35/64), CBS sensitivity: 0.680 (17/25), PSP-RS sensitivity: 0.462 (18/39). Moreover, four autopsy-confirmed CBD and patients with PSP-RS were correctly classified as CBS and PSP-RS, respectively. In the internal dataset, the classification accuracy of SuStaIn was 0.875 (56/64), and the sensitivity for CBS and PSP-RS were 0.600 (3/5) and 0.833 (10/12), respectively. In this study, the accuracy of the SuStaIn classification was superior to that of the classification based on MR findings (accuracy: 0.471 (8/12), CBS sensitivity: 0.600 (3/5), PSP-RS sensitivity: 0.417 (5/12). Thus, the classification accuracy of SuStaIn was superior to that of classic MR findings.
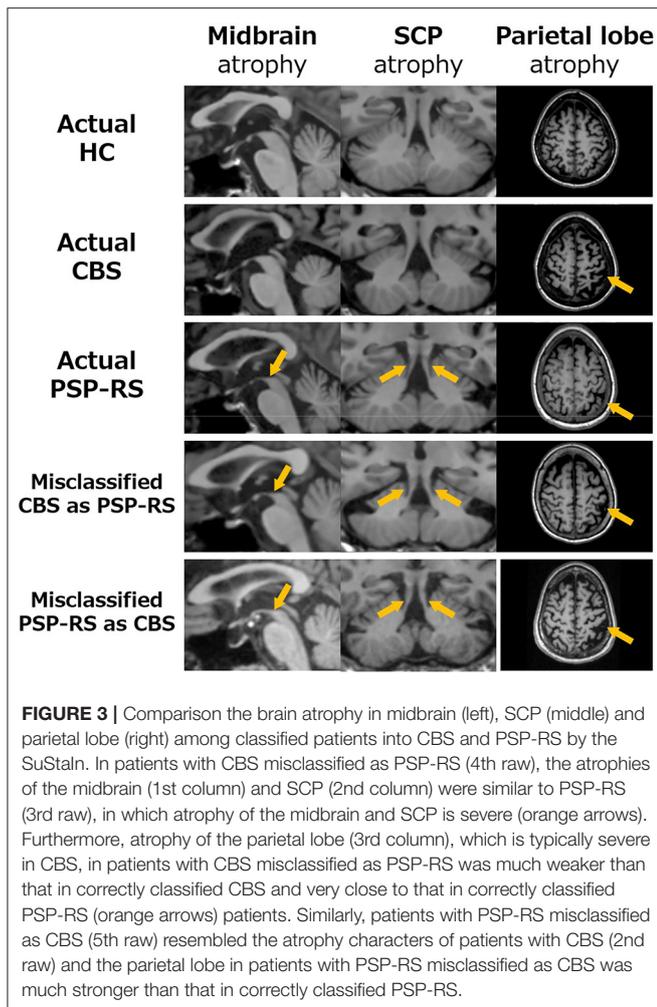
Although SuStaIn classified CBS and PSP-RS as individual disease subtypes with high accuracy, it misclassified CBS as PSP-RS and PSP-RS as CBS, in some patients. To investigate the reason for SuStaIn misclassification, the remarkable regional parietal lobe atrophy in patients with CBS and midbrain and SCP atrophy in patients with PSP-RS were compared between the patient groups classified by SuStaIn (i.e., correctly classified

TABLE 2 | Confusion matrix for the classification of CBS and PSP using SuStaIn.

	Predicted CBS	Predicted PSP-RS	Sensitivity
(a) Internal dataset			
Actual CBS	17	8	0.680 (17/25)
Actual PSP-RS	0	39	1.000 (39/39)
Precision	1.000 (17/17)	0.830 (39/48)	**Accuracy 0.875** (56/64)

	Predicted CBS	Predicted PSP-RS	Sensitivity
(b) External dataset			
Actual CBS	3	2	0.600 (3/5)
Actual PSP-RS	2	10	0.833 (10/12)
Precision	0.600 (3/5)	0.833 (10/12)	**Accuracy 0.765** (13/17)

HC, healthy control; CBS, corticobasal syndrome; PSP-RS, progressive supranuclear palsy Richardson's syndrome.

CBS and PSP-RS vs. misclassified CBS as PSP-RS) based on the z-score, relative to HCs with a mean of 0 and a standard deviation of 1. The PSPRS total scores were also compared among these patient populations. **Figure 3** and **Table 3** show the differences in regional brain atrophy among the HC, correctly

FIGURE 3 | Comparison the brain atrophy in midbrain (left), SCP (middle) and parietal lobe (right) among classified patients into CBS and PSP-RS by the SuStaIn. In patients with CBS misclassified as PSP-RS (4th raw), the atrophies of the midbrain (1st column) and SCP (2nd column) were similar to PSP-RS (3rd raw), in which atrophy of the midbrain and SCP is severe (orange arrows). Furthermore, atrophy of the parietal lobe (3rd column), which is typically severe in CBS, in patients with CBS misclassified as PSP-RS was much weaker than that in correctly classified CBS and very close to that in correctly classified PSP-RS (orange arrows) patients. Similarly, patients with PSP-RS misclassified as CBS (5th raw) resembled the atrophy characters of patients with CBS (2nd raw) and the parietal lobe in patients with PSP-RS misclassified as CBS was much stronger than that in correctly classified PSP-RS.

classified CBS and PSP-RS, misclassified CBS as PSP-RS groups and misclassified PSP-RS as CBS groups. The characteristics of regional brain atrophy and the PSPRS total scores in misclassified patients with CBS were close to those of the correctly classified patients with PSP-RS, while the characteristics of regional brain atrophy and the PSPRS total scores in misclassified patients with PSP-RS also were close to those of the correctly classified patients with CBS. Notably, patients with CBS classified as PSP-RS showed the remarkable MR finding of the hummingbird sign which is characteristic of PSP-RS. Similarly, patients with PSP-RS classified as CBS showed the remarkable MR finding of the asymmetrical frontoparietal atrophy which is characteristic of CBS.

Relationship Between SuStaIn Stage and Severity

Table 4 shows the correlation (r_s) between the disease stage as estimated by SuStaIn and disease severity in CBS and PSP-RS. The results of the internal data set for CBS and PSP-RS showed a significant correlation between SuStaIn stage and disease severity, except for PSPRS Ocular Motor ($P = 0.939$). However, the correlation with disease duration only tended toward significance

TABLE 3 | Comparison of brain atrophy in the parietal lobe, midbrain, and SCP and PSPRS total score among patients classified as CBS and PSP-RS by SuStaIn.

	Actual CBS	Misclassified CBS	Actual PSP-RS	Misclassified PSP-RS
(a) Internal dataset				
n	17	8	39	*none*
Parietal lobe (*z*-score)	2.72 ± 1.18	0.74 ± 1.21	0.86 ± 1.11	*none*
Midbrain (*z*-score)	0.18 ± 1.02	1.47 ± 1.06	2.90 ± 0.84	*none*
SCP (*z*-score)	−0.08 ± 1.31	1.18 ± 1.85	2.16 ± 1.24	*none*
PSPRS Total	24.6 ± 10.4	27.8 ± 6.9	36.0 ± 16.0	*none*
(b) External dataset				
n	3	2	10	2
Parietal lobe (*z*-score)	2.96 ± 1.01	−1.55 ± 0.23	0.38 ± 0.86	2.23 ± 0.46
Midbrain (*z*-score)	−0.73 ± 1.05	0.91 ± 2.24	2.36 ± 0.66	3.03 ± 0.43
SCP (*z*-score)	−0.07 ± 1.27	1.27 ± 1.65	1.72 ± 1.86	−0.26 ± 1.49
PSPRS Total	20.3 ± 4.50	31.0 ± 3.00	31.4 ± 15.9	63.0 ± 7.00

HC, healthy control; CBS, corticobasal syndrome; PSP-RS, progressive supranuclear palsy Richardson's syndrome, PSPRS, PSP Rating Scale. Each regional volume was converted to a z-score relative to a control population so that the control population had a mean of 0 and a standard deviation of 1. Because z-scores become negative as regional brain volumes decrease with disease progression, for simplicity, we multiplied the z-scores by −1 so that they increased as regional brain volumes decreased.

($P < 0.058$). The significant relationship between the SuStaIn stage and disease severity in CBS was indicated in the FAQ, CDR Box, and SEADL score, whereas that in PSP-RS was indicated in all scores except for the FAQ and CDR Box. Although there were no significant correlations between cognitive function and MoCA or MMSE scores in CBS, the relationship tended toward significance (MoCA, $P = 0.071$; MMSE, $P = 0.066$). The total PSPRS score and subscore were significantly correlated with SuStaIn stage only in PSP-RS. The increment in disease duration, PSPRS, UPDRS-III, FAQ, and CDR Box scores as well as the decrement in SEADL, MoCA, and MMSE scores reflect high disease severity. This tendency was similar to the results of the external data. Thus, SuStaIn stage in CBS and PSP-RS was significantly related to disease severity and correctly reflected the same.

DISCUSSION

In this study, we used SuStaIn, a recently developed unsupervised machine learning technique for data-driven disease phenotype discovery, to clarify the differences in the temporal progression patterns of brain atrophy between CBS and PSP-RS for the first time. SuStaIn successfully revealed distinct brain atrophy progression patterns that corresponded to CBS and PSP-RS, with a high average similarity across cross-validation folds for each disease subtype. The progression model of CBS showed that brain atrophy started from the cerebral neocortex, especially the frontoparietal lobe, followed by the cingulate gyrus and basal ganglia, and finally reached the brain stem. In contrast, PSP-RS progression showed atrophy starting from the brainstem, specifically the midbrain and SCP, followed by the basal ganglia and cingulate gyrus, and eventually reaching the cerebral

neocortex. Furthermore, SuStaIn classified CBS and PSP-RS as individual disease subtypes with high accuracy. Additionally, the disease stage of individual patients with CBS and PSP-RS had significant relation with the disease severity and correctly reflect it.

Our estimated atrophy progression of CBS and PSP-RS using SuStaIn is largely consistent with previous histopathological studies. In a postmortem study investigating the distribution and severity of tau pathology in preclinical and end-stage CBD, changes in neuronal lesions were largest in the frontal and parietal cortices, moderate in the basal ganglia, and mild in the brainstem (47). Another postmortem study that investigated the distribution of tau pathology in PSP-RS showed sequential distribution patterns that suggested the accumulation of different cellular tau pathologies in PSP-RS (48). The sequence of PSP-RS related neuronal tau pathology started from the substantia nigra, followed by the midbrain tegmentum, medulla oblongata, pons base, and frontal lobe, and eventually reached the parietal, temporal, and occipital lobes (48). This suggests that temporal atrophy progression patterns identified using SuStaIn mirror the progression of tau pathology in CBD-CBS and PSP-RS.

Our finding that temporal atrophy progression in PSP-RS starts from the SCP (49) and that of CBS starts from the frontoparietal cortex (50) is consistent with histopathological reports on cerebral tau accumulation. The SCP comprises the dentate nucleus in the cerebellum that ascends to the ventrolateral thalamus through the SCP. Degeneration and activated microglia along this tract have been shown to be the specific characteristics of PSP-RS pathology (51, 52). In our atrophy progression model of PSP-RS, we found other prominent regions of atrophy in the early stages, such as the midbrain, followed by basal ganglia. These regions are consistent with neuropathological distributions of tau-positive astrocytic inclusions observed in patients with PSP-RS (53). Unlike the histopathological features of PSP-RS, which are characterized by neuronal loss, gliosis, and abundant neurofibrillary tangles in the midbrain and SCP (2), the extensive accumulation of tau-immunoreactive inclusions and astrocytic plaques in the gray and white matter is a distinct characteristic of CBD pathology (54).

In line with the postmortem studies, previous cross-sectional brain volumetric studies using structural MRI have revealed midbrain and SCP atrophy as the primary features of PSP-RS (15, 18, 39, 40, 55, 56). In contrast, greater atrophy is observed in the frontoparietal cortex in CBS (16, 19). A meta-analysis of voxel-based morphometry studies showed more prominent atrophy in the superior parietal lobe in CBS compared with PSP-RS, although there was significant overlap of atrophy between PSP-RS and CBS (57). In a pathology-proven sample, atrophy of the midbrain and SCP was strongly associated with PSP-RS, whereas frontoparietal degeneration without significant brainstem atrophy was more implicative of CBD-CBS (18). Atrophy progression estimated using SuStaIn reflected regional changes in brain areas that are known to be most severely affected by CBD and PSP-RS pathology, which included the frontoparietal cortex and midbrain/SCP, respectively. These regions showed the earliest longitudinal changes in patients with CBS and PSP-RS.

The temporal brain atrophy progression patterns of patients with CBS and those with PSP-RS estimated by SuStaIn were considerably consistent with those of previous longitudinal studies. A 1-year longitudinal study, which used voxel-based morphometry of structural MRI data, (24) reported that the regional volumes of patients with CBS compared with HCs at baseline showed atrophy in the precentral gyrus, supplementary motor cortex, and postcentral gyrus. Regional brain atrophy extended to the putamen and pallidum of the basal ganglia over 6 months. Finally, atrophy reached the midbrain and pons at 12 months. For patients with PSP-RS compared with HCs at baseline, the most remarkable regional atrophy occurred in the midbrain, pontine tegmentum, and SCP. Over 1-year, cortical atrophy extended over the frontoparietal and occipital lobes, accompanied by atrophy of the basal ganglia, which included the putamen and pallidum. Additionally, Zhang et al. (56) investigated the progression of microstructural degeneration in CBS and PSP-RS using diffusion tensor imaging, which is sensitive to more subtle brain pathology. Results showed that the most prominent changes in PSP-RS were in the SCP, followed by the basal ganglia, whereas those in CBS were in the basal ganglia as well as widespread white matter regions, which included the frontal, parietal, occipital, and temporal lobes.

We compared the classification accuracies of SuStaIn and classic MR signs evaluated by an expert neuroradiologist in this study with those of a previous study (11–14, 16, 27, 58, 59) in the differentiation of CBS and PSP-RS. In the present study, the classification accuracy of internal dataset between CBS and PSP-RS based on the expert neuroradiologist's diagnosis from MR findings was 0.547 (35/64) and the sensitivities for CBS and PSP-RS were 0.680 (17/25) and 0.462 (18/39), respectively, and additionally, the classification accuracy of internal dataset was similar to that of external dataset (accuracy, 47.1% (8/17); sensitivity for CBS, 60.0% (3/5); sensitivity for PSP-RS, 41.7% (5/12). It was reported that the sensitivity and specificity according to classical morphological markers of PSP-RS on MRI were approximately 0.510 and 0.995 based on the "hummingbird" sign and a∼0.370 and 0.970 based on the "morning glory" sign, respectively (11–13). Thus, these two signs had high specificity but low sensitivity to distinguish PSP-RS and CBS (14). Overall, the classification performance of SuStaIn was superior to that of classic MR signs in terms of accuracy, sensitivity, and specificity. There were some reasons for this finding. A previous study using MRI volumetry reported that midbrain atrophy was associated with the clinical presentation of PSP-RS but not with the pathological diagnosis of PSP in the absence of clinical PSP-RS (23, 60). Similarly, asymmetric atrophy of the frontal lobe, including the premotor cortex, which is common in CBS, was not associated with background neuropathologies. In fact, previous studies reported midbrain atrophy in patients with pathologically proven CBD and asymmetric frontal atrophy in patients with pathologically proven PSP (23, 61, 62). Thus, the classification performance of classic MR signs could be lower than that of SuStaIn, because there is some heterogeneity in typical brain atrophy based on MR findings in CBS and PSP-RS.

In a previous study using a classifier, Gröschel et al. used the mathematical model for brain MR volumetry, which included the midbrain, parietal white matter, temporal gray matter, brainstem, frontal white matter, and pons in patients with PSP-RS ($n = 33$)

and CBD-CBS ($n = 18$) (16). The model correctly predicted PSP-RS and CBD-CBS with 76 and 83% sensitivity, respectively. Correia et al. used a SVM method on gray matter volume data to classify 19 patients with PSP-RS and 19 patients with CBS (27). Using a leave-two-out cross-validation approach, the mean classification accuracy of the SVM was found to be 62.2%. Also using a leave-two-out cross-validation approach, Correia et al. also observed 79.8% mean accuracy using diffusion tensor image data, which included measures of fractional anisotropy and mean diffusivity. There are also studies that have used other modalities. One study used a semiquantitative 123I-N-ω-fluoropropyl-2β-carbomethoxy-3β-(4-iodophenyl) nortropane (123I-FP-CIT) single-photon emission computed tomography striatal evaluation combined with SVM (58) that differentiated 41 patients with PSP-RS and 28 patients with CBS with 73.9% accuracy (PSP-RS sensitivity and specificity of 82.6 and 72.7%, respectively). Another study used an assay of cerebrospinal fluid (CSF) tau (59, 63), which reported 84.2% sensitivity and 66.7% specificity for separating 21 patients with PSP-RS and 12 patients with CBS using the p-tau/t-tau ratio. Urakami et al. obtained 81.5% sensitivity and 80% specificity for classifying CBS and PSP-RS based on tau protein levels in the CSF using sandwich enzyme linked immunosorbent assay. The classification accuracy of our study using SuStaIn based on only structural MRI was higher than previous studies and suggested that SuStaIn modeling for disease-stage heterogeneity allows better stratification of CBS and PSP-RS compared with models that only predict disease subtypes (28). Furthermore, in this study, the classification accuracy of external dataset was similar to that of internal dataset and it suggested that SuStaIn had the generalization performance of classification for CBS and PSP-RS. We will look at including other modalities in future work to uncover disease progression of CBS and PSP-RS and stratify these diseases with higher accuracy and reliability.

Although SuStaIn classified CBS and PSP-RS as individual disease subtypes with high accuracy, it misclassified CBS as PSP-RS in some patients. Notably, half of these patients with CBS misclassified as PSP-RS showed the hummingbird sign on MRI. As shown in **Figure 3** and **Table 3**, in patients with CBS misclassified as PSP-RS, atrophy of the midbrain and SCP was similar to that in correctly classified PSP-RS cases, where atrophy of the midbrain and SCP is severe (15, 18, 39, 40), and the PSP-RSRS total score approached that of correctly classified patients with PSP-RS. Furthermore, atrophy of the parietal lobe in patients with CBS misclassified as PSP-RS (15, 16) was much weaker than that in correctly classified patients with CBS and very close to that in correctly classified patients with PSP-RS. Thus, the above results suggest that the eight patients with CBS misclassified as PSP-RS, by SuStaIn, had atrophy characteristics strongly associated with PSP-RS, which may have resulted in the misclassification. Importantly, CBD pathology is observed not only in CBS (∼50%) but also in PSP-RS, Alzheimer's disease (AD), frontal temporal lobar degeneration (FTLD), and Creutzfeldt–Jakob disease (64). Although CBD pathology is determined by the Armstrong criteria and has four phenotypes in including CBS, frontal behavioral-spatial syndrome (FBS), nonfluent/agrammatic variant of primary progressive aphasia

TABLE 4 | Relationship between disease severity and SuStaIn stage.

Information	CBS and PSP-RS		CBS		PSP-RS	
	r_s	P-value	r_s	P-value	r_s	P-value
(a) Internal dataset						
Disease duration	0.25	0.058	0.02	0.934	0.41	<0.05
PSPRS total	0.50	<0.001	0.11	0.658	0.55	<0.005
PSPRS bullbar	0.37	<0.01	−0.04	0.864	0.41	<0.05
PSPRS ocular–motor	0.44	<0.005	−0.17	0.471	0.41	<0.05
PSPRS limb–motor	−0.01	0.939	−0.25	0.285	0.50	<0.005
PSPRS gait/midline	0.51	<0.001	0.41	0.075	0.49	<0.005
UPDRS-III total	0.31	<0.05	0.36	0.089	0.41	<0.05
FAQ total	0.41	<0.005	0.45	<0.05	0.29	0.111
CDR box	0.39	<0.005	0.57	<0.005	0.29	0.083
SEADL	−0.61	<0.001	−0.56	<0.05	−0.67	<0.001
MoCA total	−0.33	<0.05	−0.41	0.071	−0.39	<0.05
MMSE total	−0.31	<0.05	−0.41	0.066	−0.33	<0.05
(b) External dataset						
Disease duration	none	none	none	none	none	none
PSPRS total	0.63	<0.01	0.74	0.155	0.70	<0.05
PSPRS bullbar	0.46	0.065	0.49	0.406	0.47	0.120
PSPRS ocular–motor	0.36	0.169	0.87	0.058	0.72	<0.05
PSPRS limb–motor	0.72	<0.005	−0.16	0.794	0.86	<0.001
PSPRS gait/midline	0.42	0.104	0.38	0.530	0.62	<0.05
UPDRS-III total	0.95	<0.001	0.97	<0.005	0.95	<0.001
FAQ total	0.60	<0.05	0.45	0.453	0.65	<0.05
CDR box	0.54	<0.05	0.94	0.057	0.52	0.082
SEADL	−0.59	<0.05	−0.97	<0.005	−0.50	0.116
MoCA total	−0.64	<0.05	−0.94	0.057	−0.79	<0.01
MMSE total	−0.71	<0.005	−0.87	0.333	−0.79	<0.005

HC, healthy control; CBS, corticobasal syndrome; PSP-RS, progressive supranuclear palsy Richardson's syndrome; PSPRS, PSP Rating Scale; Four motor subscores from total PSPRS, PSPRS Bulbar, PSPRS Ocular Motor, PSPRS Limb Motor, and PSPRS Gait/Midline; UPDRS-III, Part-III (motor exams) of the Unified Parkinson's Disease Rating Scale; CDR Box, Clinical Dementia Rating Sum of Boxes; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; FAQ, Functional Activities Questionnaire; SEADL, Schwab and England Activities of Daily Living.

(naPPA), and PSP syndrome (PSPS) (7), a postmortem pathology study using the antemortem Armstrong criteria (65)—reported that the accuracy of diagnosis was 68% (13 of 19 CBD patients) and that these patients with CBS met the criteria for probable CBD; however, all patients (14/14) with CBS without CBD pathology, such as AD and FTLD pathologies, met the criteria for probable or possible CBD. These results imply that the Armstrong criteria lack the necessary specificity for accurate antemortem clinical diagnosis of CBD.

As shown in **Table 4**, the disease stage of individual patients with CBS and PSP-RS was significantly related to disease severity. The increment in disease duration, PSPRS, UPDRS-III, FAQ, and CDR Box scores as well as the decrement in SEADL, MoCA, and MMSE scores reflect high disease severity. The significant relationship between SuStaIn stage and disease severity in CBS was indicated in the FAQ, CDR Box, and SEADL scores, whereas that in PSP-RS was indicated in all scores except for the FAQ and CDR Box. In the internal data set, the correlations

between cognitive function and MoCA and MMSE scores in CBS only tended toward significance (MoCA, $P = 0.071$; MMSE, $P = 0.066$). Similarly, the correlation between motor function and PSPRS Gait/Midline and UPDRS-III scores in PSP-RS only tended toward significance (PSPRS Gait/Midline, $P = 0.075$; UPDRS-III, $P = 0.089$). In this study, the sample size of the CBS group ($n = 25$) was smaller than that of the PSP-RS group ($n = 39$). Therefore, this insufficient sample size might have affected the significance of the correlation test. Moreover, taking into consideration that the correlation coefficients (r_s) of the clinical scores related to cognitive and motor function were quite similar between CBS (MoCA, $r_s = -0.41$; MMSE, $r_s = -0.41$; PSPRS Gait/Midline, $r_s = 0.41$; UPDRS-III, $r_s = 0.36$) and PSP-RS (MoCA, $r_s = -0.39$; MMSE, $r_s = -0.33$; PSPRS Gait/Midline, $r_s = 0.49$; UPDRS-III, $r_s = 0.41$) and because the overlap in CBS and PSP-RS between clinical symptoms including cognitive and motor function is well-known, the SuStaIn stage could reflect exactly the disease severity of cognitive function. The total PSPPR score and subscore were significantly correlated with SuStaIn stage in only PSP-RS. A previous study on clinical correlations between annual change in PSPRS score and loss in regional brain volume demonstrated a tendency toward a significant correlation in the PSP-RS group but not in the CBS group ([24], [66]). The results of this study provide support for the findings of previous research. Thus, SuStaIn stage in CBS and PSP-RS was significantly associated with disease severity and correctly reflected it without the label of disease stage, such as disease duration. Moreover, in this study, the above tendency of the internal data set was similar to the result of the external data set, and it suggests that SuStaIn generally reflects the disease severity in patients with CBS and PSP-RS. Furthermore, SuStaIn could be used to stratify patients with CBS or PSP-RS and predict patient prognoses using only brain structural characters based on cross-sectional MRI.

The present study is limited by the fact that the diagnoses of our patients were based on clinical symptoms without autopsy confirmation. Although clinical criteria accurately stratified PSP-RS pathology, they have been shown to be insufficiently specific for CBD pathology and there is a heterogeneous condition consisted of not only CBD but also various neurodegenerative disorders such as PSP, AD, TDP-43 proteinopathy, Pick disease, and dementia with Lewy bodies ([7]). So, there was possible to inclusion of a mixture of various neuropathologies in clinically-diagnosed CBS. Therefore, the inclusion of patients with CBS who may have had various neuropahologies except for CBD pathology results may have biased our findings of the CBS group. We did not confirm whether our patients with CBS had these various neuropahologies based on negative results from amyloid, dopamine transporter or 18F-fluorodeoxyglucose imaging or CSF analysis and did not exclude patients with CBS who had PSP, AD, TDP-43 proteinopathy, Pick disease, and dementia with Lewy bodies pathology. Therefore, additional studies that exclude positive cases of these imaging will be necessary to avoid the effect of the heterogeneity of CBS neuropathologies. Lastly, further studies of autopsy-confirmed CBD and PSP-RS cases are needed to definitively determine differences in brain atrophy progression between these conditions.

While pioneering studies investigated longitudinal brain atrophy using longitudinal brain structural MRI at a few time points based on the assumption that volumetric changes are linear ([24]–[26]), this study performed SuStaIn, a recent unsupervised machine learning innovation that integrates clustering and disease progression modeling based on widely available cross-sectional data, for cross-sectional brain structural MRI to identify temporal atrophy progressions with over a few stages as **Figure 1** shows and thus had more information on the longitudinal atrophy of CBS and PSP-RS. Moreover, we estimated the temporal atrophy progression patterns of each disease subtype based on widely available cross-sectional data without disease labels, and the results were largely consistent with previous studies. Therefore, SuStaIn has the potential to be a useful tool for longitudinal studies to uncover data-driven disease phenotypes with distinct temporal progression patterns. Our findings advance our understanding of the temporal atrophy progression patterns of PSP-RS and CBS and are the first to apply such a data-driven modeling approach to study CBS and PSP-RS.

CONCLUSION

The present study is the first report that applied SuStaIn to clarify brain atrophy progression of CBS and PSP-RS. SuStaIn identified two disease subtypes and trajectories of CBS and PSP-RS with distinct atrophy patterns using cross-sectional regional brain volume data. Furthermore, the temporal progression patterns of brain atrophy in CBS and PSP-RS estimated by SuStaIn were largely consistent with previous reports, and the classification accuracy of CBS and PSP-RS was higher than those of previous studies. Although the disease mechanisms remain poorly understood, SuStaIn may be a promising tool to achieve accurate patient stratification and prognostication to develop effective treatment methods for PSP-RS and CBS.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YS, KK, and KS conceived the presented idea. YS developed the theory and performed the computations. PW and DA verified the analytical methods. KK encouraged YS to investigate brain atrophy progression in PSP and CBS and supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

FUNDING

REFERENCES

1. Boeve BF. Parkinson-related dementias. *Neurol Clin*. (2007) 25:761–81, vii. doi: 10.1016/j.ncl.2007.04.002

2. Litvan I, Bhatia KP, Burn DJ, Goetz CG, Lang AE, McKeith I, et al. Movement disorders society scientific issues committee report: SIC task force appraisal of clinical diagnostic criteria for Parkinsonian disorders. *Mov Disord*. (2003) 18:467–86. doi: 10.1002/mds.10459

3. Litvan I, Hauw JJ, Bartko JJ, Lantos PL, Daniel SE, Horoupian DS, et al. Validity and reliability of the preliminary NINDS neuropathologic criteria for progressive supranuclear palsy and related disorders. *J Neuropathol Exp Neurol*. (1996) 55:97–105. doi: 10.1097/00005072-199601000-00010

4. Williams DR, Lees AJ. Progressive supranuclear palsy: clinicopathological concepts and diagnostic challenges. *Lancet Neurol*. (2009) 8:270–9. doi: 10.1016/S1474-4422(09)70042-0

5. Dickson DW, Bergeron C, Chin SS, Duyckaerts C, Horoupian D, Ikeda K, et al. Office of rare diseases neuropathologic criteria for corticobasal degeneration. *J Neuropathol Exp Neurol*. (2002) 61:935–46. doi: 10.1093/jnen/61.11.935

6. Kouri N, Murray ME, Hassan A, Rademakers R, Uitti RJ, Boeve BF, et al. Neuropathological features of corticobasal degeneration presenting as corticobasal syndrome or Richardson syndrome. *Brain*. (2011) 134:3264–75. doi: 10.1093/brain/awr234

7. Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B, et al. Criteria for the diagnosis of corticobasal degeneration. *Neurology*. (2013) 80:496–503. doi: 10.1212/WNL.0b013e31827f0fd1

8. Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, et al. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. *Neurology*. (1996) 47:1–9. doi: 10.1212/WNL.47.1.1

9. Williams DR, Holton JL, Strand C, Pittman A, de Silva R, Lees AJ, et al. Pathological tau burden and distribution distinguishes progressive supranuclear palsy-parkinsonism from Richardson's syndrome. *Brain*. (2007) 130:1566–76. doi: 10.1093/brain/awm104

10. Ling H, Macerollo A. Is it useful to classify PSP and CBD as different disorders? *Yes Mov Disord Clin Pract*. (2018) 5:145–8. doi: 10.1002/mdc3.12581

11. Adachi M, Kawanami T, Ohshima H, Sugai Y, Hosoya T. Morning glory sign: a particular MR finding in progressive supranuclear palsy. *Magn Reson Med Sci*. (2004) 3:125–32. doi: 10.2463/mrms.3.125

12. Mueller C, Hussl A, Krismer F, Heim B, Mahlknecht P, Nocker M, et al. The diagnostic accuracy of the hummingbird and morning glory sign in patients with neurodegenerative parkinsonism. *Parkinsonism Relat Disord*. (2018) 54:90–4. doi: 10.1016/j.parkreldis.2018.04.005

13. Whitwell JL, Höglinger GU, Antonini A, Bordelon Y, Boxer AL, Colosimo C, et al. Radiological biomarkers for diagnosis in PSP: where are we and where do we need to be? *Mov Disord*. (2017) 32:955–71. doi: 10.1002/mds.27038

14. Massey LA, Micallef C, Paviour DC, O'Sullivan SS, Ling H, Williams DR, et al. Conventional magnetic resonance imaging in confirmed progressive supranuclear palsy and multiple system atrophy. *Mov Disord*. (2012) 27:1754–62. doi: 10.1002/mds.24968

15. Boxer AL, Geschwind MD, Belfor N, Gorno-Tempini ML, Schauer GF, Miller BL, et al. Patterns of brain atrophy that differentiate corticobasal degeneration syndrome from progressive supranuclear palsy. *Arch Neurol*. (2006) 63:81–6. doi: 10.1001/archneur.63.1.81

16. Gröschel K, Hauser TK, Luft A, Patronas N, Dichgans J, Litvan I, et al. Magnetic resonance imaging-based volumetry differentiates progressive supranuclear palsy from corticobasal degeneration. *Neuroimage*. (2004) 21:714–24. doi: 10.1016/j.neuroimage.2003.09.070

17. Josephs KA, Whitwell JL, Boeve BF, Knopman DS, Petersen RC, Hu WT, et al. Anatomical differences between CBS-corticobasal degeneration and CBS-Alzheimer's disease. *Mov Disord*. (2010) 25:1246–52. doi: 10.1002/mds.23062

18. Josephs KA, Whitwell JL, Dickson DW, Boeve BF, Knopman DS, Petersen RC, et al. Voxel-based morphometry in autopsy proven PSP and CBD. *Neurobiol Aging*. (2008) 29:280–9. doi: 10.1016/j.neurobiolaging.2006.09.019

19. Soliveri P, Monza D, Paridi D, Radice D, Grisoli M, Testa D, et al. Cognitive and magnetic resonance imaging aspects of corticobasal degeneration and progressive supranuclear palsy. *Neurology*. (1999) 53:502–7. doi: 10.1212/WNL.53.3.502

20. Lee SE, Rabinovici GD, Mayo MC, Wilson SM, Seeley WW, DeArmond SJ, et al. Clinicopathological correlations in corticobasal degeneration. *Ann Neurol*. (2011) 70:327–40. doi: 10.1002/ana.22424

21. Whitwell JL, Jack CR Jr, Boeve BF, Parisi JE, Ahlskog JE, Drubach DA, et al. Imaging correlates of pathology in corticobasal syndrome. *Neurology*. (2010) 75:1879–87. doi: 10.1212/WNL.0b013e3181feb2e8

22. Gröschel K, Kastrup A, Litvan I, Schulz JB. Penguins and hummingbirds: midbrain atrophy in progressive supranuclear palsy. *Neurology*. (2006) 66:949–50. doi: 10.1212/01.wnl.0000203342.77115.bf

23. Whitwell JL, Jack CR Jr, Parisi JE, Gunter JL, Weigand SD, Boeve BF, et al. Midbrain atrophy is not a biomarker of progressive supranuclear palsy pathology. *Eur J Neurol*. (2013) 20:1417–22. doi: 10.1111/ene.12212

24. Dutt S, Binney RJ, Heuer HW, Luong P, Attygalle S, Bhatt P, et al. Progression of brain atrophy in PSP and CBS over 6 months and 1 year. *Neurology*. (2016) 87:2016–25. doi: 10.1212/WNL.0000000000003305

25. Josephs KA, Xia R, Mandrekar J, Gunter JL, Senjem ML, Jack CR Jr, et al. Modeling trajectories of regional volume loss in progressive supranuclear palsy. *Mov Disord*. (2013) 28:1117–24. doi: 10.1002/mds.25437

26. Paviour DC, Price SL, Jahanshahi M, Lees AJ, Fox NC. Longitudinal MRI in progressive supranuclear palsy and multiple system atrophy: rates and regions of atrophy. *Brain*. (2006) 129:1040–9. doi: 10.1093/brain/awl021

27. Correia MM, Rittman T, Barnes CL, Coyle-Gilchrist IT, Ghosh B, Hughes LE, et al. Towards accurate and unbiased imaging-based differentiation of Parkinson's disease, progressive supranuclear palsy and corticobasal syndrome. *Brain Commun*. (2020) 2:fcaa051. doi: 10.1093/braincomms/fcaa051

28. Young AL, Marinescu RV, Oxtoby NP, Bocchetta M, Yong K, Firth NC, et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nat Commun*. (2018) 9:4273. doi: 10.1038/s41467-018-05892-0

29. Chen J, Liu J, Calhoun VD, Arias-Vasquez A, Zwiers MP, Gupta CN, et al. Exploration of scanning effects in multi-site structural MRI studies. *J Neurosci Methods*. (2014) 230:37–50. doi: 10.1016/j.jneumeth.2014.04.023

30. Golbe LI, Ohman-Strickland PA. A clinical rating scale for progressive supranuclear palsy. *Brain*. (2007) 130:1552–65. doi: 10.1093/brain/awm032

31. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord*. (2008) 23:2129–70. doi: 10.1002/mds.22340

32. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. (1975) 12:189–98. doi: 10.1016/0022-3956(75)90026-6

33. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The montreal cognitive assessment, MoCA: a brief screening

tool for mild cognitive impairment. *J Am Geriatr Soc.* (2005) 53:695–9. doi: 10.1111/j.1532-5415.2005.53221.x

34. Morris JC. The clinical dementia rating (CDR): current version and scoring rules. *Neurology.* (1993) 43:2412–4. doi: 10.1212/WNL.43.11.2412-a

35. Schwab RS, editor. *Projection Technique for Evaluating Surgery in Parkinson's Disease.* Third symposium on Parkinson's disease: E&S Livingstone (1969).

36. Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol.* (1982) 37:323–9. doi: 10.1093/geronj/37.3.323

37. Boxer AL, Lang AE, Grossman M, Knopman DS, Miller BL, Schneider LS, et al. Davunetide in patients with progressive supranuclear palsy: a randomised, double-blind, placebo-controlled phase 2/3 trial. *Lancet Neurol.* (2014) 13:676–85. doi: 10.1016/S1474-4422(14)70088-2

38. Fischl B. FreeSurfer. *Neuroimage.* (2012) 62:774–81. doi: 10.1016/j.neuroimage.2012.01.021

39. Kurata T, Kametaka S, Ohta Y, Morimoto N, Deguchi S, Deguchi K, et al. PSP as distinguished from CBD, MSA-P and PD by clinical and imaging differences at an early stage. *Intern Med.* (2011) 50:2775–81. doi: 10.2169/internalmedicine.50.5954

40. Price S, Paviour D, Scahill R, Stevens J, Rossor M, Lees A, et al. Voxel-based morphometry detects patterns of atrophy that help differentiate progressive supranuclear palsy and Parkinson's disease. *Neuroimage.* (2004) 23:663–9. doi: 10.1016/j.neuroimage.2004.06.013

41. Geyer CJ. *Markov Chain Monte Carlo Maximum Likelihood.* Interface Foundation of North America, University of Minnesota Digital Conservancy, United States (1991). Available online at: https://hdl.handle.net/11299/58440

42. Fonteijn HM, Modat M, Clarkson MJ, Barnes J, Lehmann M, Hobbs NZ, et al. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. *Neuroimage.* (2012) 60:1880–9. doi: 10.1016/j.neuroimage.2012.01.062

43. Young AL, Oxtoby NP, Daga P, Cash DM, Fox NC, Ourselin S, et al. A data-driven model of biomarker changes in sporadic Alzheimer's disease. *Brain.* (2014) 137:2564–77. doi: 10.1093/brain/awu176

44. Oxtoby NP, Young AL, Cash DM, Benzinger TLS, Fagan AM, Morris JC, et al. Data-driven models of dominantly-inherited Alzheimer's disease progression. *Brain.* (2018) 141:1529–44. doi: 10.1093/brain/awy050

45. Wijeratne PA, Young AL, Oxtoby NP, Marinescu RV, Firth NC, Johnson EB, et al. An image-based model of brain volume biomarker changes in Huntington's disease. *Ann Clin Transl Neurol.* (2018) 5:570–82. doi: 10.1002/acn3.558

46. Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull Calcutta Math Soc.* (1943) 35:99–109.

47. Ling H, Kovacs GG, Vonsattel JP, Davey K, Mok KY, Hardy J, et al. Astrogliopathy predominates the earliest stage of corticobasal degeneration pathology. *Brain.* (2016) 139:3237–52. doi: 10.1093/brain/aww256

48. Kovacs GG, Lukic MJ, Irwin DJ, Arzberger T, Respondek G, Lee EB, et al. Distribution patterns of tau pathology in progressive supranuclear palsy. *Acta Neuropathol.* (2020) 140:99–119. doi: 10.1007/s00401-020-02158-2

49. Dickson DW. Neuropathology of progressive supranuclear palsy. *Handb Clin Neurol.* (2008) 89:487–91. doi: 10.1016/S0072-9752(07)01245-6

50. Sakai K, Piao YS, Kikugawa K, Ohara S, Hasegawa M, Takano H, et al. Corticobasal degeneration with focal, massive tau accumulation in the subcortical white matter astrocytes. *Acta Neuropathol.* (2006) 112:341–8. doi: 10.1007/s00401-006-0093-5

51. Ishizawa K, Dickson DW. Microglial activation parallels system degeneration in progressive supranuclear palsy and corticobasal degeneration. *J Neuropathol Exp Neurol.* (2001) 60:647–57. doi: 10.1093/jnen/60.6.647

52. Tsuboi Y, Slowinski J, Josephs KA, Honer WG, Wszolek ZK, Dickson DW. Atrophy of superior cerebellar peduncle in progressive supranuclear palsy. *Neurology.* (2003) 60:1766–9. doi: 10.1212/01.WNL.0000068011.21396.F4

53. Revesz T, Holton JL. Anatamopathological spectrum of tauopathies. *Mov Disord.* (2003) 18(Suppl. 6):S13–20. doi: 10.1002/mds.10558

54. Forman MS, Zhukareva V, Bergeron C, Chin SS, Grossman M, Clark C, et al. Signature tau neuropathology in gray and white matter of corticobasal degeneration. *Am J Pathol.* (2002) 160:2045–53. doi: 10.1016/S0002-9440(10)61154-6

55. Armstrong MJ. Progressive Supranuclear Palsy: an Update. *Curr Neurol Neurosci Rep.* (2018) 18:12. doi: 10.1007/s11910-018-0819-5

56. Zhang Y, Walter R, Ng P, Luong PN, Dutt S, Heuer H, et al. Progression of microstructural degeneration in progressive supranuclear palsy and corticobasal syndrome: a longitudinal diffusion tensor imaging study. *PLoS ONE.* (2016) 11:e0157218. doi: 10.1371/journal.pone.0157218

57. Yu F, Barron DS, Tantiwongkosi B, Fox P. Patterns of gray matter atrophy in atypical parkinsonism syndromes: a VBM meta-analysis. *Brain Behav.* (2015) 5:e00329. doi: 10.1002/brb3.329

58. Nicastro N, Wegrzyk J, Preti MG, Fleury V, Van de Ville D, Garibotto V, et al. Classification of degenerative parkinsonism subtypes by support-vector-machine analysis and striatal (123)I-FP-CIT indices. *J Neurol.* (2019) 266:1771–81. doi: 10.1007/s00415-019-09330-z

59. Urakami K, Wada K, Arai H, Sasaki H, Kanai M, Shoji M, et al. Diagnostic significance of tau protein in cerebrospinal fluid from patients with corticobasal degeneration or progressive supranuclear palsy. *J Neurol Sci.* (2001) 183:95–8. doi: 10.1016/S0022-510X(00)00480-9

60. Sakurai K, Morimoto S, Oguri T, Yuasa H, Uchida Y, Yamada K, et al. Multifaceted structural magnetic resonance imaging findings in demented patients with pathologically confirmed TDP-43 proteinopathy. *Neuroradiology.* (2019) 61:1333–9. doi: 10.1007/s00234-019-02289-8

61. Sakurai K, Tokumaru AM, Shimoji K, Murayama S, Kanemaru K, Morimoto S, et al. Beyond the midbrain atrophy: wide spectrum of structural MRI finding in cases of pathologically proven progressive supranuclear palsy. *Neuroradiology.* (2017) 59:431–43. doi: 10.1007/s00234-017-1812-4

62. Tokumaru AM, Saito Y, Murayama S, Kazutomi K, Sakiyama Y, Toyoda M, et al. Imaging-pathologic correlation in corticobasal degeneration. *AJNR Am J Neuroradiol.* (2009) 30:1884–92. doi: 10.3174/ajnr.A1721

63. Aerts MB, Esselink RA, Bloem BR, Verbeek MM. Cerebrospinal fluid tau and phosphorylated tau protein are elevated in corticobasal syndrome. *Mov Disord.* (2011) 26:169–73. doi: 10.1002/mds.23341

64. Cordato NJ, Halliday GM, McCann H, Davies L, Williamson P, Fulham M, et al. Corticobasal syndrome with tau pathology. *Mov Disord.* (2001) 16:656–67. doi: 10.1002/mds.1124

65. Alexander SK, Rittman T, Xuereb JH, Bak TH, Hodges JR, Rowe JB. Validation of the new consensus criteria for the diagnosis of corticobasal degeneration. *J Neurol Neurosurg Psychiatry.* (2014) 85:925–9. doi: 10.1136/jnnp-2013-307035

66. VandeVrede L, Dale ML, Fields S, Frank M, Hare E, Heuer HW, et al. Open-label phase 1 futility studies of salsalate and young plasma in progressive supranuclear palsy. *Mov Disord Clin Pract.* (2020) 7:440–7. doi: 10.1002/mdc3.12940