

Novel data-driven subtypes and stages of brain atrophy in the ALS-FTD spectrum

Ting Shen (✉ Ting.Shen@Pennmedicine.upenn.edu)

University of Pennsylvania Perelman School of Medicine <https://orcid.org/0000-0002-3937-4427>

Jacob W. Vogel

Lund University

Jeffrey Duda

University of Pennsylvania Perelman School of Medicine

Jeffrey S. Phillips

University of Pennsylvania Perelman School of Medicine

Philip A. Cook

University of Pennsylvania Perelman School of Medicine

James Gee

University of Pennsylvania Perelman School of Medicine

Lauren Elman

University of Pennsylvania Perelman School of Medicine

Colin Quinn

University of Pennsylvania Perelman School of Medicine

Defne A. Amado

University of Pennsylvania Perelman School of Medicine

Michael Baer

University of Pennsylvania Perelman School of Medicine

Lauren Massimo

University of Pennsylvania Perelman School of Medicine

Murray Grossman

University of Pennsylvania Perelman School of Medicine

David J. Irwin

University of Pennsylvania Perelman School of Medicine

Corey T. McMillan



University of Pennsylvania Perelman School of Medicine

Research Article

Keywords: amyotrophic lateral sclerosis, frontotemporal degeneration, disease heterogeneity, SuStaln model

Posted Date: August 10th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3183113/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

TDP-43 proteinopathies represents a spectrum of neurological disorders, anchored clinically on either end by amyotrophic lateral sclerosis (ALS) and frontotemporal degeneration (FTD). The ALS-FTD spectrum exhibits a diverse range of clinical presentations with overlapping phenotypes, highlighting its heterogeneity. This study aimed to use disease progression modeling to identify novel data-driven spatial and temporal subtypes of brain atrophy and its progression in the ALS-FTD spectrum.

Methods

We used a data-driven procedure to identify 13 anatomic clusters of brain volumes for 57 behavioral variant FTD (bvFTD; with either autopsy-confirmed TDP-43 or TDP-43 proteinopathy-associated genetic variants), 103 ALS, and 47 ALS-FTD patients with likely TDP-43. A Subtype and Stage Inference (SuStaln) model was trained to identify subtypes of individuals along the ALS-FTD spectrum with distinct brain atrophy patterns, and we related subtypes and stages to clinical, genetic, and neuropathological features of disease.

Results

SuStaln identified three novel subtypes: two disease subtypes with predominant brain atrophy either in prefrontal/somatomotor regions or limbic-related regions, and a normal-appearing group without obvious brain atrophy. The Limbic-predominant subtype tended to present with more impaired cognition, higher frequencies of pathogenic variants in *TBK1* and *TARDBP* genes, and a higher proportion of TDP-43 type B, E and C. In contrast, the Prefrontal/Somatomotor-predominant subtype had higher frequencies of pathogenic variants in *C9orf72* and *GRN* genes and higher proportion of TDP-43 type A. The normal-appearing brain group showed higher frequency of ALS relative to ALS-FTD and bvFTD patients, higher cognitive capacity, higher proportion of lower motor neuron onset, milder motor symptoms, and lower frequencies of genetic pathogenic variants. Overall SuStaln stages also correlated with evidence for clinical progression including longer disease duration, higher King's stage, and cognitive decline. Additionally, SuStaln stages differed across clinical phenotypes, genotypes and types of TDP-43 pathology.

Conclusions

Our findings suggest distinct neurodegenerative subtypes of disease along the ALS-FTD spectrum that can be identified *in vivo*, each with distinct brain atrophy, clinical, genetic and pathological patterns.

Introduction

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disorder characterized by loss of motor neurons in the brain and spinal cord, leading to muscle weakness, atrophy and ultimately paralysis [1].

Behavioral variant frontotemporal dementia (bvFTD) is the most common subtype of frontotemporal degeneration (FTD) characterized by impairments in behavior, personality, and/or executive function [2, 3]. ALS may additionally exhibit cognitive and behavioral symptoms overlapping with bvFTD [4, 5] and bvFTD can exhibit motor neuron dysfunction consistent with ALS [6]. These two phenotypes can either occur separately or simultaneously, and have shared clinical, neuropathological and genetic features, constituting two ends of a spectrum of disorders that ALS-FTD lies in between [7]. Defining subtypes and elaborating distinct characteristics in the ALS-FTD spectrum captures potential driving causes of neurodegeneration.

Neurodegenerative diseases display high degree of inter-individual variation in disease biomarkers, including neuropsychological profiles, neuroimaging features, and molecular biological indicators. Distinct patterns of brain atrophy have been observed along the ALS-FTD spectrum. Regarding clinical phenotypes, bvFTD patients exhibit greater grey matter atrophy in the frontotemporal cortex, insula, thalamus, striatum, hippocampus and amygdala, while ALS patients show more severe atrophy in the motor cortex, pons and brainstem [8, 9]. Different genetic pathogenic variants also result in distinct patterns of brain atrophy in individuals with ALS-FTD spectrum. These patterns vary in severity, progression rate, and affected brain regions. *C9orf72*-related FTD and ALS are associated with higher degree of atrophy extensively in frontal, parietal, occipital, cingulate and insula regions, thalamus and cerebellum compared to sporadic patients [10–17]. Progranulin (*GRN*)-frontotemporal lobar degeneration (FTLD) patients tend to exhibit greater grey matter volume loss in the frontal cortex [18, 19]. Longitudinal data suggest that patients with pathogenic variants in *GRN* experience faster brain atrophy progression than those with pathogenic variants in *C9orf72*, indicating different rates of pathological progression and fundamental mechanisms associated with different gene variants [10, 20]. Thus, distinct clinical phenotypes and genotypes may account for both spatial and temporal heterogeneity in brain atrophy patterns.

To better understand the spatial and temporal patterns of brain atrophy, an unsupervised machine-learning algorithm called Subtype and Stage Inference (SuStaln) was developed. This tool can identify distinct subtypes and extract their progression patterns simultaneously [21], unlike previous studies that applied either subtype-only [22–24] or stages-only [25–27] models. A recent study utilized the SuStaln algorithm to establish a data-driven pathological TDP-43 staging system in ALS, FTLD due to TDP-43 proteinopathies (FTLD-TDP), and limbic-predominant age-related TDP-43 encephalopathy neuropathologic change (LATE-NC) [28]. They identified two subtypes within FTLD-TDP that were cortical-predominant or brainstem-predominant, and two subtypes within ALS that were subcortical-predominant or corticolimbic-predominant. To date, this method has been applied to reconstruct different patterns of sequential disease progression trajectories in TDP-43 proteinopathies [28], FTD [21, 29] and Alzheimer's disease (AD) [30, 31], providing fundamental insights into the underlying biological processes of these diseases.

This study aimed to investigate the complex progression patterns and heterogeneity within earlier stages of the ALS-FTD spectrum, in contrast to late-stage neuropathological studies. To achieve this, we focused on individuals with high likelihood (clinical ALS) or definite (pathology confirmed or genetic variants) TDP-43 pathology, and we trained a SuStaln model on baseline cortical and subcortical volume data. Our prior study using the SuStaln model trained on TDP-43 proteinopathy data had limitations related to the focus on the end-stage of disease and reliance on ordinal pathology ratings [28]. In contrast, this study utilized more

quantitative data, the MRI-derived cortical and subcortical volumes that can identify earlier evidence of brain atrophy. We classified individuals into subtypes with different brain atrophy patterns and extracted a full trajectory for each subtype. Furthermore, we examined the differences in clinical phenotypes, genotypes and pathologies across subtypes. We also assessed the effectiveness of the fitted model by analyzing longitudinal brain volumetric data.

Methods

Participants

Participants were retrospectively selected from the Integrated NeuroDegenerative Disease (INDD) database at the University of Pennsylvania (Additional file 1: Fig.S1) [32]. This study included a cohort of individuals who met published clinical criteria for ALS (n = 103), ALS-FTD (n = 47), or bvFTD (n = 57) [33–35], diagnosed by board-certified neurologists. We also included 172 demographically-comparable (age, sex) healthy controls who self-reported a negative neurological and non-significant psychiatric history with a normal Mini-Mental Status Examination (MMSE) > 27 (out of 30). Individuals with bvFTD had either autopsy-confirmed TDP-43 proteinopathy or genetic evidence of pathogenic variants associated with TDP-43 proteinopathy including *C9orf72*, *GRN*, metalloendopeptidase (*MME*), TANK-binding kinase 1 (*TBK1*), and TAR DNA binding protein (*TARDBP*). Of the 207 individuals with ALS-FTD spectrum disorder, 62 (22 with ALS, 8 with ALS-FTD and 32 with bvFTD) had one follow-up MRI scan, which were used in secondary analyses to evaluate the longitudinal consistency of SuStaln subtype and stage assignments.

Neuroimaging Data & Processing

Structural T1-weighted MRI scans were acquired on a Siemens 3.0 Tesla scanner outfitted as a TIM Trio (n = 188) and subsequently as a Prisma Fit (n = 81). MRI scans were collected with similar magnetization-prepared rapid gradient-echo (MPRAGE) sequences as follows: 1) 3.0 Tesla Siemens TIM Trio scanner, 8 channel head coil, axial plane with repetition time (TR) ranging from 1620 ms to 1900 ms, echo time (TE) ranging from 3.09 ms to 4.38 ms, slice thickness = 1.0 mm or 1.5 mm, in-plane resolution = 0.98 × 0.98 mm. 2) 3.0 Tesla Siemens TIM Trio scanner, 64 channel head coil, sagittal plane with TR = 2200 ms or 2300 ms, TE ranging from 2.95 ms to 4.63 ms, slice thickness = 1.0 mm or 1.2 mm, in plane resolution = 1.0 × 1.0 mm. 3) 3.0 Tesla Siemens Prisma scanner, 64 channel head coil, sagittal plane with TR = 2400 ms, TE = 1.96 ms, slice thickness = 0.8 mm, in-plane resolution = 0.8 × 0.8 mm.

Images were processed using the Advanced Normalization Tools (ANTs) software package through standard preprocessing steps, as previously described [36]. Briefly, this procedure included N4 bias field correction, diffeomorphic and symmetric registration to a custom template, brain extraction, and segmentation into six-tissue classes (cortical grey matter, subcortical grey matter, deep white matter, CSF, brainstem, and cerebellum) using template-based priors [37]. The custom template was in turn aligned to the MNI152 2009c Asymmetric T1-weighted template. The Schaefer 17-network atlas with 100 cortical parcels [38] and the Melbourne subcortex atlas [39] were warped from MNI152 space through the custom template to individual space. From each label, volumetric measurement was extracted, normalized by age, sex, and intracranial volume and converted to w-scores relative to healthy controls [40].

Considering the relatively low dimensionality of input data required for SuStaln model, it is important to limit the number of features. We sought data reduction to enhance power of analysis, improve model identifiability, and reduce uncertainty. An unsupervised consensus-clustering algorithm, Bootstrap Analysis of Stable Clusters (BASC), was utilized to identify spatially stable clusters that consistently exhibited similar volumetric measurements of cortical and subcortical structures across subjects [41]. This algorithm performed k-means clustering on 1000 bootstrapped samples to reduce the dimensions of input data. A stability matrix was generated to represent the probabilities of each pair of brain regions falling into the same cluster. Based on the Silhouette index, an optimal number of data-driven clusters were identified. The volumetric measurements of BASC-identified clusters were then extracted and used as input biomarkers to the SuStaln model (Fig. 1a, Additional file 1: Table S1).

Clinical data

Clinical and neuropsychological assessments were conducted at the Penn Frontotemporal Degeneration Center and Penn Comprehensive ALS Clinic. Neuropsychological test scores were obtained from the testing visit that was closest to the MRI scan. Demographic information, including age, sex, years of education, disease duration from symptom onset to MRI scan, diagnostic delay (the time interval between symptom onset and confirmed disease diagnosis), and site of symptom onset were collected.

Motor Assessments. The Penn Upper Motor Neuron Score (PUMNS) measures upper motor neuron signs in individuals with ALS/ALS-FTD [42]. The Revised ALS Functional Rating Scale (ALSFERS-R) evaluates the severity of motor symptom functional impairment in ALS/ALS-FTD [43]. Disease progression was measured by the Progression index, which is calculated as (48 - ALSFRS-R score) divided by duration in months [44]. We also calculated King's stage, derived from the ALSFRS-R, to assess spreading of motor symptoms [45].

Cognitive Assessments. Cognitive and behavioral changes were evaluated using tests, including MMSE, Edinburgh Cognitive Assessment Scale (ECAS) [5, 46], Philadelphia Brief Assessment of Cognition (PBAC) [47], Boston Naming Test (BNT), semantically-guided category naming fluency for the number of animals generated in 60 seconds (Animal fluency score), letter guided category naming fluency for the number of 'F' words generated in 60 seconds (Letter fluency score), and digit-span for the longest number of digits repeated in forward and backward sequences (Digit forward span and Digit backward span).

Genetic screening

Genomic DNA was extracted from peripheral blood or frozen brain tissue collected from participants [48]. DNA was not available for 6 individuals. Genotyping for *C9orf72* hexanucleotide repeat expansions was performed using a modified repeat-primed polymerase-chain reaction, as previously described [49]. Pathogenic variants that associated with ALS-FTD spectrum were screened using either a targeted next-generation sequencing panel (MiND-Seq) [48] or whole exome/genome (WES/WGS) sequencing. Of 201 individuals who underwent genetic screening, 64 were found to have pathogenic variants. Specifically, 48 had repeat expansions in *C9orf72* (> 30 repeats), or known pathogenic variants including 11 in *GRN*, 1 in *MME*, 2 in *TBK1*, and 2 in *TARDBP*.

Neuropathological examination

Autopsy was performed on a subset of individuals (n = 55) including 21 ALS, 7 ALS-FTD, and 27 bvFTD. Neuropathological diagnosis of FTLD-TDP and ALS with TDP-43 proteinopathies (ALS-TDP) was performed by expert neuropathologists according to previously described protocols [50]. TDP-43 proteinopathies were classified into categories including types A, B, C and E [51]. Since type E is relatively rare and shows some biological overlap with type B [51], it has been proposed to combine these two types together. Of the 55 individuals, 16 were classified as type A cases, 18 as type B or E, 3 as type C cases, and the remaining 18 cases (1 bvFTD and 17 ALS) that could not be further subtyped were classified as TDP-43 non-specific type.

Subtype and Stage Inference modelling

We utilized the w-scored volumetric measurements of 13 BASC-identified clusters (Fig. 1a, Additional file 1: Table S1) as input biomarkers for training the SuStaln model (<https://github.com/ucl-pond/pySuStaln>). As the volumetric measurements were continuous variables, we employed the piecewise linear SuStaln model. This algorithm combines clustering and disease progression modelling to identify subtypes with different rates and patterns of disease progression [21]. To evaluate the performance of SuStaln model, we used 10-fold cross-validation, where the optimal number of subtypes was selected based on the out-of-sample log-likelihood and cross-validation information criterion (CVIC) [21] to better balance the model complexity with accuracy (Fig. 1b-c). Each subtype's disease progression pattern was described by a piecewise linear model, which reconstructed the trajectory of brain atrophy. Each event, alternatively referred to as stage, corresponded to a change in a specific biomarker, quantified by w-scores representing the severity of brain atrophy. We utilized w-score waypoints of 1, 2, and 3, with 3 set as the maximum value that represented the point at which the biomarker reached severe abnormality. To capture the progression pattern where each SuStaln stage corresponds to a new region reaching a new score, the number of stages was determined by multiplying the number of BASC-identified clusters (13) by the maximum w-score value (3), resulting in a total of 39 stages. The model uncertainty was estimated using 100,000 Markov chain Monte Carlo iterations (MCMC). For each subject, the SuStaln model assigned a probability value to each subtype and stage, enabling their assignment to a specific subtype and stage within the disease progression pattern of this subtype.

Longitudinal MRI scans were withheld from the SuStaln model calculations and then used in a secondary analysis to assess the stability of SuStaln subtypes and progression of SuStaln stages over time. At follow-up visits, the volumetric measurements were w-scored as described above using the same healthy control cohort for normalization. Subtype stability was determined as the proportion of individuals who were either assigned to the same subtype or progressed from normal-appearing group to a SuStaln subtype at follow-up visits. The advancement of SuStaln stage over time was evaluated in individuals with stable subtypes. The annualized change of SuStaln stage was calculated by dividing the change in SuStaln stage from baseline to follow-up visit by follow-up period.

Statistical analyses

The statistical analyses and plotting were conducted with R statistical software (version 4.2.0; R Foundation for Statistical Computing, Vienna, Austria) and GraphPad Prism (version 9.0; GraphPad Software, Inc., San Diego, CA). The brain heatmaps were visualized using BrainNet Viewer [52]. The normality of variable distribution was tested using the Shapiro-Wilk normality test. Continuous variables with normal distribution

were compared using two-sample t-test, while Mann-Whitney test was utilized for comparing variables with non-normal distribution. For comparison of categorical variables, chi-squared test or Fisher exact test was employed. We compared clinical features, frequencies of pathogenic variants, proportions of TDP-43 types, SuStaln stages and annualized change of SuStaln stage across subtypes. Additionally, subtype probability at baseline were compared between subtype-stable and unstable individuals. A significance level of $p < 0.05$ was considered significant. Cortical and subcortical volumes were compared between different groups using a generalized linear model, and a false discovery rate (FDR)-corrected $p < 0.05$ was used for multiple testing. Correlation analyses were conducted between the predicted SuStaln stages and clinical profiles, baseline and follow-up SuStaln stages, as well as the change in SuStaln stage and follow-up period. All correlation analyses were considered significant at a threshold of $p < 0.05$.

Results

Participants characteristics

The demographic, clinical, genetic and pathological characteristics are summarized in Table 1. Compared to ALS individuals, bvFTD individuals had longer disease duration. The diagnostic delay in individuals with ALS, ALS-FTD, and bvFTD is a multifactorial issue influenced by various elements, which increased in ascending order for these conditions. The ALS individuals were younger and had higher MMSE scores than ALS-FTD and bvFTD individuals. Individuals with bvFTD had higher frequencies of pathogenic variants in *C9orf72* and *GRN* genes than ALS/ALS-FTD, and two individuals had pathogenic *TARDBP* mutations were both bvFTD. Higher proportions of TDP-43 type A, B, and E cases were observed in ALS-FTD and bvFTD groups compared to ALS group. All three TDP-43 type C cases were bvFTD. Most of the ALS cases in our cohort were classified as TDP-43 non-specific type.

Table 1
Comparison of baseline characteristics between clinical phenotypes in all individuals with ALS-FTD spectrum disorder

	ALS (n = 103)	ALS-FTD (n = 47)	bvFTD (n = 57)	Missing data	P _{ALSvsALS-FTD}	P _{ALSvsbvFTD}	P _{ALS-FTDvsbvFTD}
Age at MRI (years)	58.9 (10.3)	62.6 (9.6)	62.9 (7.5)	0.0%	0.04	0.005	0.85
Gender (male%)	58 (56.3%)	28 (59.6%)	36 (63.2%)	0.0%	0.71	0.40	0.71
Education (years)	16.7 (12.0)	15.0 (2.8)	16.1 (2.8)	0.0%	0.72	0.09	0.09
Disease duration (months)	37.8 (36.4)	42.9 (33.8)	52.0 (38.1)	0.0%	0.20	0.0004	0.06
Diagnostic delay (months)	17.7 (17.6)	36.8 (34.6)	46.5 (33.2)	0.0%	< 0.0001	< 0.0001	0.007
MMSE	27.7 (2.8)	24.1 (5.7)	24.3 (4.9)	6.3%	< 0.0001	< 0.0001	0.87
Genetic pathogenic variants ^a	n = 99	n = 45	n = 57	2.9%	-	-	-
<i>C9orf72</i>	7 (7.1%)	10 (22.2%)	31 (54.4%)		0.009	< 0.0001	0.001
<i>GRN</i>	0 (0.0%)	0 (0.0%)	11 (19.3%)		1.00	< 0.0001	0.002
<i>MME</i>	1 (1.0%)	0 (0.0%)	0 (0.0%)		1.00	1.00	1.00
<i>TBK1</i>	0 (0.0%)	1 (2.2%)	1 (1.6%)		0.31	0.37	1.00
<i>TARDBP</i>	0 (0.0%)	0 (0.0%)	2 (3.5%)		1.00	0.13	0.50
FTLD/ALS-TDP ^b	n = 21	n = 7	n = 27	73.4%	0.0003	< 0.0001	0.13
Type A	1 (4.8%)	1 (14.3%)	14 (51.9%)	-	-	-	-
Type B/E	3 (14.3%)	6 (85.7%)	9 (33.3%)	-	-	-	-
Type C	0 (0.0%)	0 (0.0%)	3 (11.1%)	-	-	-	-

	ALS (n = 103)	ALS-FTD (n = 47)	bvFTD (n = 57)	Missing data	P _{ALSvsALS-FTD}	P _{ALSvsbvFTD}	P _{ALS-FTDvsbvFTD}
Non-specific	17 (81.0%)	0 (0.0%)	1 (3.7%)	-	-	-	-
SuStaln stage	3.3 (5.2)	12.1 (7.7)	15.6 (7.2)	0.0%	< 0.0001	< 0.0001	0.02

Data are presented as mean (standard deviation) for the continuous variables, and as number (frequency) for the categorical variables. Missing data indicates the percentage of individuals with missing data. ^a number of individuals underwent genetic screening; ^b number of individuals underwent neuropathological examination.

ALS amyotrophic lateral sclerosis, ALS-FTD amyotrophic lateral sclerosis-frontotemporal degeneration, bvFTD behavioral variant frontotemporal degeneration, MMSE Mini-Mental Status Examination, FTL/ALS-TDP frontotemporal lobar degeneration or amyotrophic lateral sclerosis with TDP-43 inclusions.

Subtype progression patterns

The SuStaln algorithm was applied to the baseline brain volumetric measurements, resulting in the identification of subtypes that exhibit distinct progression patterns of brain atrophy. Figure 2 illustrates the brain atrophy trajectory for each subtype, with the w-score ranging from 1 to 3, indicating the degree of brain atrophy from mild to moderate to severe. The most noticeable differences between the two subtypes with distinct brain atrophy patterns were observed in the initial sites of brain atrophy during the early SuStaln stages.

The first identified subtype, exhibited brain atrophy that initially appeared in the prefrontal cortex and subsequently in the somatomotor cortex at SuStaln stage 3, which we subsequently refer to as “Prefrontal/Somatomotor-predominant subtype”. By SuStaln stage 12–13, parts of the prefrontal cortex reached w-scores exceeding 3. Additionally, the volumetric loss of subcortical regions, including the thalamus, caudate, globus pallidus, putamen, and nucleus accumbens, was evident in early stages but developed more slowly than atrophy in the prefrontal cortex. This volume loss continued to progress and reaches a severe degree after SuStaln stage 17.

The second identified subtype displayed brain atrophy that was first observed in the temporal pole within the limbic network, hippocampus, and amygdala at SuStaln stage 1, which we subsequently referred to as “Limbic-predominant subtype”. The brain regions related to the limbic system experience a more rapid progression of atrophy. Specifically, the hippocampus and amygdala reached w-score 3 by SuStaln stage 8, while the temporal pole and insula reached w-score 3 by stage 12. The volumetric loss of subcortical regions also began in the early stages of atrophy progression, but it reached w-score 3 later than the Prefrontal/Somatomotor-predominant subtype, indicating a relatively slower rate of progression. It was worth noting that the 11th cluster, which included prefrontal regions, orbitofrontal cortex and insula, experienced significant volumetric loss in the early stages and ultimately reached a severe level of atrophy by

SuStaln stage 11 in both subtypes. In addition to these two subtypes with atrophy, individuals assigned to SuStaln stage 0 were labeled as “normal-appearing group”, which showed no detectable brain atrophy.

Subtype assignments

Of individuals with ALS, 48 (46.6%) were categorized as Prefrontal/Somatomotor-predominant subtype, 14 (13.6%) as Limbic-predominant subtype, and 41 (39.8%) as normal-appearing group. The ALS-FTD cohort consisted of 26 (55.3%) individuals classified as Prefrontal/Somatomotor-predominant subtype, 19 (40.4%) classified as Limbic-predominant subtype, and 2 (4.3%) categorized as normal-appearing group. Of individuals with bvFTD, 42 (73.7%) were assigned to Prefrontal/Somatomotor-predominant subtype, 14 (24.6%) assigned to Limbic-predominant subtype, and 1 (1.8%) categorized as normal-appearing group. Thus, individuals with ALS were more likely to be classified into the normal-appearing group, whereas the majority of the ALS-FTD and bvFTD individuals were assigned to atrophy subtypes.

Prefrontal/Somatomotor-predominant subtype was the most common assignment across clinical diagnoses, which has a higher prevalence compared to the Limbic-predominant subtype, occurring almost 2.5 times more frequently. The distributions of clinical phenotypes significantly differed across subtypes (Fig. 3a, Additional file 1: Table S2).

Comparison of cortical and subcortical volumes between subtypes

Comparing cortical and subcortical volumes across different groups (Fig. 4), we found that the normal-appearing group did not display any significant brain atrophy at their baseline MRI. As indicated by the name “normal-appearing group”, there was no noticeable reduction in brain volumes compared to healthy controls, which was in line with our expectations.

The two atrophy subtypes displayed extensive decreased brain volume in comparison to the normal-appearing group. The Prefrontal/Somatomotor-predominant subtype exhibited reduced volume in brain regions within several networks, including somatomotor, limbic, dorsal attention, salience/ventral attention, control, visual, and default mode networks. Additionally, this subtype had reduced volume in subcortical regions including thalamus, putamen, globus pallidus, caudate, nucleus accumbens, hippocampus, and amygdala. The Limbic-predominant subtype showed decreased volumes mainly in the limbic, dorsal attention, salience/ventral attention, control, and default mode networks, as well as subcortical regions including hippocampus, amygdala, thalamus, nucleus accumbens and putamen.

The two SuStaln subtypes exhibited distinct patterns of brain atrophy (Fig. 4). The Limbic-predominant subtype, as indicated by its name, demonstrated lower volumes in the limbic network including temporal pole, insula, parahippocampal cortex, hippocampus, and amygdala relative to the Prefrontal/Somatomotor-predominant subtype. The Prefrontal/Somatomotor-predominant subtype showed lower volumes in prefrontal and somatomotor cortices compared to the Limbic-predominant subtype.

Given the significant difference in SuStaln stage between subtypes, we conducted additional comparisons of volumetric measurements between subtypes while adjusting for SuStaln stage, to avoid attributing regional

atrophy differences solely to subtypes with more advanced atrophy due to disease progression (Additional file 1: Fig.S2). Similar findings were observed, with more concentrated in regions relevant to the respective subtypes. Specifically, the Prefrontal/Somatomotor-predominant subtype exhibited reduced volume primarily in the thalamus, prefrontal and somatomotor cortices, while the Limbic-predominant subtype showed decreased volumes mainly in the temporal lobe, insula, parahippocampal cortex, hippocampus, and amygdala.

Comparison of clinical, genetic, and neuropathological features between subtypes

Demographic, clinical, genetic and neuropathological characteristics for each subtype are summarized in Fig. 3 and Additional file 1: Table S2. Although the two SuStaln subtypes displayed different patterns of brain atrophy, there were substantial overlaps in clinical features across subtypes. This suggests that despite differences in neurodegenerative patterns, the clinical manifestations and symptomatology remain largely consistent between the subtypes. The Limbic-predominant subtype exhibited poorer performance on BNT, which assesses language and semantic memory, compared to the Prefrontal/Somatomotor-predominant subtype. In terms of genetic status, the Prefrontal/Somatomotor-predominant subtype had a significantly higher frequency of pathogenic variants in *GRN* compared to Limbic-predominant subtype. Notably, all 11 cases with *GRN* pathogenic variants were classified into the Prefrontal/Somatomotor-predominant subtype. Although not statistically significant, there was also a trend towards higher frequencies of repeat expansions in *C9orf72* in the Prefrontal/Somatomotor-predominant subtype. Additionally, it is worth highlighting that two individuals with bvFTD who had pathogenic variants in the *TARDBP* gene, as well as one individual with ALS-FTD and one with bvFTD who carried *TBK1* pathogenic variants, were all classified under the Limbic-predominant subtype. Distribution of TDP-43 types varied across SuStaln subtypes. The Prefrontal/Somatomotor-predominant subtype had a higher proportion of TDP-43 type A. The Limbic-predominant subtype was more prone to having TDP-43 type B or E, and all three bvFTD individuals with TDP-43 type C also belonged to this subtype. The TDP-43 non-specific type, predominantly observed in individuals with ALS-TDP, was more prevalent in the Prefrontal/Somatomotor-predominant subtype than Limbic-predominant subtype. Comparing to atrophy subtypes, the normal-appearing group had a significantly shorter diagnostic delay, and a higher proportion of individuals with ALS than ALS-FTD and bvFTD. Additionally, they had a lower frequency of cognitive onset in relation to lower and upper motor neuron onset. This group also showed higher cognitive scores, as evidenced by better performance on tests including MMSE, ECAS, PBAC, BNT, Animal and Letter fluency tasks, and Digit forward and backward span. Two cases in the normal-appearing group were found to have pathogenic variants in either *C9orf72* or *MME* genes. Additionally, most individuals in this group who underwent autopsy were classified as having TDP-43 non-specific type pathology.

Certain tests (including PUMNS, ALSFRS-R, Progression index, and King's stage) were specifically administered for individuals with ALS/ALS-FTD, as these tests were considered more relevant or sensitive in assessing motor impairments. Thus, we focused on ALS/ALS-FTD as a distinct subgroup to compare clinical profiles across subtypes (Fig. 3, Additional file 1: Table S3). Despite a smaller number of ALS-FTD cases in this cohort, the Limbic-predominant subtype still exhibited a higher percentage of individuals with ALS-FTD

compared to the Prefrontal/Somatomotor-predominant subtype. Likewise, individuals who experienced cognitive onset were more likely to be classified under the Limbic-predominant subtype, given that this subtype had more individuals with cognitive decline. Regarding the motor symptom scales, the normal-appearing group tended to have lower King's stages compared to atrophy subtypes. Moreover, by focusing solely on bvFTD (Additional file 1: Table S4), the research sample was relatively homogeneous, allowing for a comprehensive examination of cognitive function across subtypes. The Limbic-predominant subtype had longer disease duration and only showed worse performance on BNT.

To demonstrate that the differences between two subtypes were related to atrophy patterns rather than one subtype being in a more advanced stage, we further adjusted for SuStaln stage when comparing the clinical profiles. This adjustment allowed us to account for the potential confounding effect of disease progression. Even after adjusting for SuStaln stage, the Limbic-predominant subtype still showed poorer performance on BNT (t-statistic = -5.70, $p < 0.0001$) and language scale (t-statistic = -2.17, $p = 0.03$) of PBAC compared to the Prefrontal/Somatomotor-predominant subtype. This finding further supported the presence of language impairments in the Limbic-predominant subtype. Furthermore, the Limbic-predominant subtype showed longer diagnostic delay (t-statistic = 2.009, $p = 0.04$).

Relationship between SuStaln stage and clinical characteristics

Each individual was assigned a SuStaln stage, which reflected progression of brain atrophy. The distribution of individuals assigned to each SuStaln stage was illustrated in Fig. 5a. ALS individuals were predominantly assigned to earlier SuStaln stages of brain atrophy, while ALS-FTD and bvFTD individuals were more frequently assigned to later stages (Fig. 5b). Individuals in Limbic-predominant subtype had higher SuStaln stages than individuals in Prefrontal/Somatomotor-predominant subtype (Fig. 3b, Tables S2-S3).

We further investigated the relationship between SuStaln stages and clinical profiles, genotypes, and neuropathologies in all individuals. The SuStaln stage was positively correlated with disease duration ($r = 0.22$, $p = 0.002$; Fig. 5f) and diagnostic delay ($r = 0.46$, $p < 0.0001$; Fig. 5g), while negatively correlated with cognitive scales including MMSE ($r = -0.50$, $p < 0.0001$; Fig. 5h), ECAS scores, PBAC score, BNT, Animal and Letter fluency tasks, Digit forward and back span tasks (Additional file 1: Fig.S3). In terms of motor symptoms, individuals with ALS/ALS-FTD who had higher King's stages exhibited higher SuStaln stages compared to individuals in King's stage 1 (Fig. 5c). Furthermore, individuals carrying pathogenic variants in *C9orf72* and *GRN* had significantly higher SuStaln stages, compared to sporadic forms of the disease (Fig. 5d). Individuals with pathogenic variants in *C9orf72* exhibited higher SuStaln stages than those who had pathogenic variants in *GRN*. Furthermore, autopsy-confirmed TDP-43 typable cases including type A, B, C, and E, also showed significantly higher SuStaln stages than cases having TDP-43 non-specific type (Fig. 5e).

Longitudinal stability and reliability of SuStaln subtypes and stages

Subtyping stability. The mean follow-up period was 17.5 months, with a standard deviation of 13.1 months. The subtype assignments of follow-up visits are shown in Fig. 6a and Additional file 1: Table S5. Of the 62 follow-up visits, 55 (88.7%) remained consistent with their baseline subtype assignments. Additionally, 2 (3.2%) individuals initially assigned to the normal-appearing group progressed to Prefrontal/Somatomotor-predominant subtype, while 2 (3.2%) progressed to Limbic-predominant subtype. These 59 cases (95.2%) were deemed as “subtype stable” individuals. The remaining 3 (4.8%) follow-up visits resulted in inconsistent subtype assignments, and were considered as “subtype unstable”. The probability of subtype assignments at baseline was higher in subtype stable individuals than unstable individuals (Mann-Whitney U-statistic = 27, $p = 0.04$; Fig. 6b). Individuals assigned to Prefrontal/Somatomotor-predominant subtype exhibited more atrophy in its key regions, the BASC-identified clusters 1, 2, 5, and 10. The Limbic-predominant subtype showed more atrophy in its key regions, the BASC-identified clusters 9 and 12 (Fig. 2, Additional file 1: Fig.S4). During follow-up visits, brain atrophy showed slight progression. Specifically, the two normal-appearing cases that progressed to Limbic-predominant subtype exhibited significant atrophy progression, particularly in clusters 9 and 12. In contrast, the two normal-appearing cases progressed to Prefrontal/Somatomotor-predominant subtype showed more widespread atrophy progression, particularly in the prefrontal cortex, with less pronounced progression in limbic-related regions (Additional file 1: Fig.S4b). Cases displaying abnormal longitudinal changes were typically classified as “subtype unstable” or “stage unstable”.

Staging reliability. Among individuals with stable subtype, most of the follow-up visits were assigned to a more advanced SuStaln stage or remained at the same stage. Of the 59 subtype stable cases, 6 (10.2%) follow-up visits were retrogressed to an earlier stage and regarded as “stage unstable” individuals. The probability of stage assignments at baseline was significantly higher in stage stable individuals compared to unstable individuals (Mann-Whitney U-statistic = 45, $p = 0.003$; Fig. 6c). The annualized change in SuStaln stage may indicate rate of disease progression, with the normal-appearing group showing slower progression than the Prefrontal/Somatomotor-predominant subtype (Mann-Whitney U-statistic = 110, $p = 0.01$; Fig. 6d). In stage stable individuals, annualized change in SuStaln stage was significantly smaller in normal-appearing group compared to both atrophy subtypes (Mann-Whitney U-statistic = 86 and $p = 0.003$ for Prefrontal/Somatomotor-predominant subtype, and Mann-Whitney U-statistic = 14 and $p = 0.005$ for Limbic-predominant subtype). Additionally, the SuStaln stage at baseline was significantly correlated with stages at follow-up visits ($r = 0.89$, $p < 0.0001$; Fig. 6e). Furthermore, we observed a positive correlation between the follow-up period and the change of SuStaln stage ($r = 0.27$, $p = 0.04$; Fig. 6f).

Discussion

In this study, we utilized a data-driven SuStaln model approach to investigate diverse spatial and temporal patterns of brain atrophy in the ALS-FTD spectrum. By analyzing baseline cross-sectional volumetric imaging data, we identified distinct patterns of regional brain atrophy, which included a Prefrontal/Somatomotor-predominant subtype, a Limbic-predominant subtype and a normal-appearing group. These data-driven subtypes exhibited variations in clinical, genetic and neuropathological characteristics. Moreover, the data-driven SuStaln stages constructed progression trajectories of each subtype, which aligned with worsening clinical profiles. Together, our findings provided new insights into the

heterogeneity in progression patterns of brain atrophy in the ALS-FTD spectrum and highlighted the potential utility for patient stratification in precision medicine.

Supporting evidence has demonstrated that the ALS-FTD spectrum displays a high degree of clinical, genetic and neuropathological heterogeneity [8]. Although various biomarkers have been applied to subtype individuals and characterize their brain atrophy patterns within the ALS-FTD spectrum [20, 53, 54], there is still no ideal method to fully disentangle the heterogeneity of brain atrophy. Using the SuStaln model, we identified data-driven subtypes with distinct progression patterns of brain atrophy. The Prefrontal/Somatomotor-predominant and Limbic-predominant subtypes exhibited brain atrophy in shared and distinct brain regions. The two subtypes were characterized by their distinctive brain atrophy regions as their names suggest, with the Prefrontal/Somatomotor-predominant subtype exhibiting atrophy in prefrontal and somatomotor regions, while the Limbic-predominant subtype exhibited atrophy in limbic-related regions such as temporal regions, hippocampus and amygdala. In addition, both subtypes exhibited volumetric loss in several shared brain regions including prefrontal, paralimbic, and subcortical regions. The prefrontal regions were likely to be the vulnerable regions in the Prefrontal/Somatomotor-predominant subtype, while the orbitofrontal cortex and insula, as two major components of the paralimbic belt, were vulnerable regions in the Limbic-predominant subtype. Our findings were partly consistent with previous studies that have identified subtypes of brain atrophy in subsets of the ALS-FTD spectrum [24, 55]. Tan et al. utilized a subtype-only clustering algorithm and identified subtypes in ALS, one involving motor regions and the other involving orbitofrontal/temporal regions [55]. Bede et al. also identified two distinct subgroups in ALS, one with more motor involvement and one with more frontotemporal pathology [24]. Ranasinghe et al. focused on bvFTD and identified subgroups characterized by predominance in salience network, semantic appraisal network, and subcortical regions [56]. Young et al. applied the SuStaln model to genetic FTD, and identified a temporal subtype and a frontotemporal subtype of brain atrophy [29]. Our study trained the SuStaln model on a diverse range of clinical phenotypes within the ALS-FTD spectrum. Our approach benefited by considering both spatial and temporal progression of brain atrophy, setting it apart from previous subtype-only and stage-only studies. By incorporating spatial patterns of brain atrophy, we gained a more comprehensive understanding of the different subtypes within the ALS-FTD spectrum. Simultaneously, analysis of temporal progression allowed us to capture the dynamic nature of brain atrophy in the ALS-FTD spectrum and allowing determination of the progressive stage of an individual. As a result, the two subtypes we identified provide a comprehensive summary of the characteristics of previously identified subtypes.

The two brain atrophy subtypes showed distinct characteristics. The Limbic-predominant subtype captured a higher proportion of individuals with cognitive (rather than motor) symptom onset, with more pronounced cognitive decline, particularly in the language domain. This subtype resembled a semantic variant primary progressive aphasia pattern. The Prefrontal/Somatomotor-predominant subtype had higher frequencies of pathogenic variants in *C9orf72*. The *C9orf72* pathogenic variants carriers were demonstrated to exhibit prominent structural and functional disruptions in various brain regions, including prefrontal and motor cortices [57, 58]. Additionally, this subtype also covered all the *GRN* pathogenic variants carriers. FTD individuals with *GRN* pathogenic variants may exhibit asymmetric cortical atrophy involving frontal, temporal and parietal cortices [10, 59, 60]. Both two bvFTD individuals with the I383V variant in *TARDBP* gene fell into the Limbic-predominant subtype, consistent with previous observations that I383V variant was

associated with predominant atrophy of temporal lobes and hippocampus [61, 62]. The distribution of TDP-43 types was different between subtypes. The Prefrontal/Somatomotor-predominant subtype had a higher proportion of type A, which has been linked to atrophy in the dorsal frontotemporal, striatal, and thalamic regions [53, 57], all of which were predominant regions of this subtype. The Limbic-predominant subtype presented higher proportions of TDP-43 type B, and E. It has been reported that the TDP-43 type B was associated with relatively symmetric atrophy of the medial temporal, medial prefrontal, and orbitofrontal-insular cortices [53], which are regions involved in the Limbic-predominant subtype. TDP-43 type C was highly associated with neurodegeneration in the anterior temporal lobes including the temporal pole and amygdalo-hippocampal area [63]. It is notable that all three bvFTD individuals with confirmed TDP-43 type C pathology fell into the Limbic-predominant subtype, which aligns with a staging system of brain atrophy in TDP-43 type C with early involvement of amygdala, medial and lateral temporal cortex, and temporal pole, followed by later involvement of insula [64]. The normal-appearance group displayed better cognitive abilities in various domains including the executive functioning, language, visual skill, and memory, as well as milder behavioral symptoms, and a tendency towards shorter disease duration. This group mostly consisted of individuals with ALS, who exhibited better cognitive performance and were more likely to be lower motor neuron onset. This observation is in line with established knowledge, which suggests that ALS typically exhibits a lesser degree of cortical TDP-43 pathology and greater involvement of lower motor neurons [65]. The spread of TDP-43 pathology in ALS follows a sequential pattern, starting from motor neurons in the spinal cord, brainstem, and agranular motor cortex, then propagating to the frontotemporal and subcortical regions [50].

The SuStaln model further reconstructed the progression trajectories of brain atrophy of each subtype. The SuStaln stages represent ordered progression of brain atrophy from normal to a certain degree of abnormality. The Limbic-predominant subtype had higher SuStaln stages, indicating a more advanced degree of brain atrophy progression than Prefrontal/Somatomotor-predominant subtype. Individuals with genetic pathogenic variants were assigned to more advanced stages compared to sporadic individuals. Specifically, the individuals with pathogenic variants in *GRN* exhibited more advanced stages than those with the *C9orf72* repeat expansions. This aligns with previous works demonstrating a faster progression rate of brain atrophy in individuals with pathogenic variants in *GRN* than those in *C9orf72* [20, 66]. Furthermore, individuals with TDP-43 non-specific type exhibited higher SuStaln stages compared to those with typable TDP-43 pathology. This is because the TDP-43 non-specific type mainly consisted of ALS cases with less cortical pathology, making them unclassifiable into specific TDP-43 types. These individuals exhibited less brain atrophy, indicating an early-stage level of brain atrophy. As individuals entered advanced SuStaln stages, brain atrophy was increased in degree and spatial extent, accompanied by a subsequent progression of clinical symptoms. SuStaln stage showed good linear correlations with clinical progression measures including disease duration, motor symptoms severity and cognitive decline, making it a reliable representation of disease progression and could be used to evaluate the level of advancement of an individual's disease.

To test the reliability of the SuStaln model, we examined the consistency of subtype assignments on follow-up MRI data. The results supported the effectiveness of the disease progression model in subtyping and staging, as 95.2% of the individuals showed stable subtype assignments over time. This includes individuals

who were consistently assigned to the same subtype, and those who progressed from the normal-appearing group to corresponding atrophy subtypes as the brain atrophy initiated in either prefrontal/somatomotor or limbic-related regions. Overall, the model demonstrated a subtyping capability as high as 95.2%. Staging reliability refers to the proportion of follow-up visits where individuals either advanced to a higher SuStaln stage or remained at the same stage as baseline assessment. This model exhibited a staging reliability of 89.8%, which could be attributed to the lower probabilities of stage assignment in unstable stage cases, making them more prone to being retrogressed to an earlier stage. The reason for the “subtype unstable” or “stage unstable” assignments in longitudinal assessments could be attributed to various factors, including technical issues that may lead to inconsistencies in the measured imaging features used to classify subtypes or stages. Moreover, our finding revealed progressively worsening of brain atrophy over time, with longer follow-up periods associated with greater changes in SuStaln stage, reflecting more advanced disease progression.

There are several limitations to consider in future work. One limitation is the inherent heterogeneity of the ALS-FTD spectrum. Our clinical assessments were routinely collected clinical measures (e.g., ALSFRS-R, UMN) that largely did not differ across observed subtypes but more detailed clinical exam or finer-grained motor measures may better identify how our observed patterns may onto clinical heterogeneity in the future. Our study specifically focused on individuals associated with TDP-43 proteinopathies. This selective focus may restrict the generalizability of SuStaln model in capturing the full extent of heterogeneity within the ALS-FTD spectrum, including bvFTD due to a tauopathy or atypical form of AD. Another limitation is the lack of sampling from important regions including spinal cord and brainstem, which play crucial roles in pathophysiology of ALS. This limitation may partially explain why approximately 40% of ALS individuals were assigned to the normal-appearing group without apparent brain atrophy. The absence of data from these regions may mask important changes occurring specifically in spinal cord and brainstem, thereby restricting our ability to fully comprehend the underlying neurodegenerative processes in ALS. Future investigations should address these limitations to gain a more comprehensive understanding of the ALS-FTD spectrum.

Conclusions

In general, we utilized the SuStaln model to gain a deeper understanding of the heterogeneity within progressive processes of the ALS-FTD spectrum. We demonstrated two distinct spatiotemporal subtypes of cortical atrophy with varying clinical, genetic and neuropathological profiles, which shed light on the intricate progression patterns and heterogeneity of the ALS-FTD spectrum. This data-driven disease progression modelling method provided a valuable tool for individual classification and staging, paving the way for precision medicine in the field.

Abbreviations

AD Alzheimer’s disease

ALS Amyotrophic lateral sclerosis

ALSFERS-R Revised ALS Functional Rating Scale

BASC Bootstrap Analysis of Stable Clusters

BNT Boston Naming Test

bvFTD Behavioral variant frontotemporal dementia

CVIC Cross-validation information criterion

ECAS Edinburgh Cognitive Assessment Scale

FTLD Frontotemporal lobar degeneration

FTLD-TDP FTLD due to TDP-43 proteinopathies

ALS-TDP ALS with TDP-43 proteinopathies

GRN Progranulin

LATE-NC Limbic-predominant age-related TDP-43 encephalopathy neuropathologic change

MCMC Markov chain Monte Carlo iterations

MME Metalloendopeptidase

MMSE Mini-Mental Status Examination

PBAC Philadelphia Brief Assessment of Cognition

PUMNS Penn Upper Motor Neuron Score

SuStaln Subtype and Stage Inference

TBK1 TANK-binding kinase 1

TARDBP TAR DNA binding protein

Declarations

Acknowledgements

We would like to acknowledge the late Murray Grossman who participated in the design and interpretation of results.

Funding

JWV acknowledges funding from the NIH (T32MH019112) and the SciLifeLab & Wallenberg Data Driven Life Science Program (grant: KAW 2020.0239). Jeffrey S. Phillips was supported by NIH grant (R01-AG054519,

K01-AG061277). This work was supported by NIH funding (P30 AG072979, P01AG066597, R01NS109260), Penn Institute on Aging, Robinson Family Fund, Peisach Family Fund for FTD Research, and Arking Family Fund.

Author contributions

Study concept and design: TS, JWV, and CTM; Acquisition and analysis of MRI data: JD, JSP, PAC, and JG; Acquisition and analysis of clinical data: LE, CQ, DAA, MB, LM, and MG; Acquisition and analysis of genetic data: LM, and CTM. Acquisition and analysis of autopsy data: DJI; drafting of the manuscript: TS; critical revision of the manuscript: all authors.

Availability of data and materials

All data used in this study are available upon reasonable request and approval from the Penn Neurodegenerative Data Sharing Committee. Requests may be submitted using a webform request: <https://www.pennbindlab.com/data-sharing>.

Competing interests

All authors report that they have no competing interests to disclose.

Ethics approval and consent to participate

The protocols of this study were approved by an institutional review board convened at the University of Pennsylvania. All participants participated in an approved informed consent procedure in accordance with the Declaration of Helsinki.

Consent for publication

All authors read and approved the final manuscript.

References

1. van Es MA, Hardiman O, Chio A, Al-Chalabi A, Pasterkamp RJ, Veldink JH, et al. Amyotrophic lateral sclerosis. *Lancet*. 2017;390(10107):2084-2098.
2. Boeve BF, Boxer AL, Kumfor F, Pijnenburg Y, Rohrer JD. Advances and controversies in frontotemporal dementia: diagnosis, biomarkers, and therapeutic considerations. *Lancet Neurol*. 2022;21(3):258-272.
3. Piguet O, Hornberger M, Mioshi E, Hodges JR. Behavioural-variant frontotemporal dementia: diagnosis, clinical staging, and management. *Lancet Neurol*. 2011;10(2):162-172.
4. Goldstein LH, Abrahams S. Changes in cognition and behaviour in amyotrophic lateral sclerosis: nature of impairment and implications for assessment. *Lancet Neurol*. 2013;12(4):368-380.
5. McMillan CT, Wu J, Rascovsky K, Cosentino S, Grossman M, Elman L, et al. Defining cognitive impairment in amyotrophic lateral sclerosis: an evaluation of empirical approaches. *Amyotroph Lateral Scler Frontotemporal Degener*. 2022;23(7-8):517-526.

6. Lipton AM, White CL, 3rd, Bigio EH. Frontotemporal lobar degeneration with motor neuron disease-type inclusions predominates in 76 cases of frontotemporal degeneration. *Acta Neuropathol.* 2004;108(5):379-385.
7. Ferrari R, Kapogiannis D, Huey ED, Momeni P. FTD and ALS: a tale of two diseases. *Curr Alzheimer Res.* 2011;8(3):273-294.
8. Ahmed RM, Bocchetta M, Todd EG, Tse NY, Devenney EM, Tu S, et al. Tackling clinical heterogeneity across the amyotrophic lateral sclerosis-frontotemporal dementia spectrum using a transdiagnostic approach. *Brain Commun.* 2021;3(4):fcab257.
9. Luo C, Hu N, Xiao Y, Zhang W, Gong Q, Lui S. Comparison of Gray Matter Atrophy in Behavioral Variant Frontal Temporal Dementia and Amyotrophic Lateral Sclerosis: A Coordinate-Based Meta-Analysis. *Front Aging Neurosci.* 2020;12:14.
10. Whitwell JL, Weigand SD, Boeve BF, Senjem ML, Gunter JL, DeJesus-Hernandez M, et al. Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics. *Brain.* 2012;135(Pt 3):794-806.
11. Sha SJ, Takada LT, Rankin KP, Yokoyama JS, Rutherford NJ, Fong JC, et al. Frontotemporal dementia due to C9ORF72 mutations: clinical and imaging features. *Neurology.* 2012;79(10):1002-1011.
12. Byrne S, Elamin M, Bede P, Shatunov A, Walsh C, Corr B, et al. Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a C9orf72 repeat expansion: a population-based cohort study. *Lancet Neurol.* 2012;11(3):232-240.
13. Irwin DJ, McMillan CT, Brettschneider J, Libon DJ, Powers J, Rascovsky K, et al. Cognitive decline and reduced survival in C9orf72 expansion frontotemporal degeneration and amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry.* 2013;84(2):163-169.
14. Bede P, Bokde AL, Byrne S, Elamin M, McLaughlin RL, Kenna K, et al. Multiparametric MRI study of ALS stratified for the C9orf72 genotype. *Neurology.* 2013;81(4):361-369.
15. Devenney E, Hornberger M, Irish M, Mioshi E, Burrell J, Tan R, et al. Frontotemporal dementia associated with the C9ORF72 mutation: a unique clinical profile. *JAMA Neurol.* 2014;71(3):331-339.
16. Westeneng HJ, Walhout R, Straathof M, Schmidt R, Hendrikse J, Veldink JH, et al. Widespread structural brain involvement in ALS is not limited to the C9orf72 repeat expansion. *J Neurol Neurosurg Psychiatry.* 2016;87(12):1354-1360.
17. Schonecker S, Neuhofer C, Otto M, Ludolph A, Kassubek J, Landwehrmeyer B, et al. Atrophy in the Thalamus But Not Cerebellum Is Specific for C9orf72 FTD and ALS Patients - An Atlas-Based Volumetric MRI Study. *Front Aging Neurosci.* 2018;10:45.
18. Borroni B, Alberici A, Cercignani M, Premi E, Serra L, Cerini C, et al. Granulin mutation drives brain damage and reorganization from preclinical to symptomatic FTLD. *Neurobiol Aging.* 2012;33(10):2506-2520.
19. Bozzali M, Battistoni V, Premi E, Alberici A, Giulietti G, Archetti S, et al. Structural brain signature of FTLD driven by Granulin mutation. *J Alzheimers Dis.* 2013;33(2):483-494.
20. Staffaroni AM, Goh SM, Cobigo Y, Ong E, Lee SE, Casaletto KB, et al. Rates of Brain Atrophy Across Disease Stages in Familial Frontotemporal Dementia Associated With MAPT, GRN, and C9orf72

- Pathogenic Variants. *JAMA Netw Open*. 2020;3(10):e2022847.
21. 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(1):4273.
 22. Whitwell JL, Przybelski SA, Weigand SD, Ivnik RJ, Vemuri P, Gunter JL, et al. Distinct anatomical subtypes of the behavioural variant of frontotemporal dementia: a cluster analysis study. *Brain*. 2009;132(Pt 11):2932-2946.
 23. Noh Y, Jeon S, Lee JM, Seo SW, Kim GH, Cho H, et al. Anatomical heterogeneity of Alzheimer disease: based on cortical thickness on MRIs. *Neurology*. 2014;83(21):1936-1944.
 24. Bede P, Murad A, Lope J, Hardiman O, Chang KM. Clusters of anatomical disease-burden patterns in ALS: a data-driven approach confirms radiological subtypes. *J Neurol*. 2022;269(8):4404-4413.
 25. 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(3):1880-1889.
 26. Eshaghi A, Marinescu RV, Young AL, Firth NC, Prados F, Jorge Cardoso M, et al. Progression of regional grey matter atrophy in multiple sclerosis. *Brain*. 2018;141(6):1665-1677.
 27. 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(5):1529-1544.
 28. Young AL, Vogel JW, Robinson JL, McMillan CT, Ossenkoppele R, Wolk DA, et al. Data-driven neuropathological staging and subtyping of TDP-43 proteinopathies. *medRxiv*. 2023.
 29. Young AL, Bocchetta M, Russell LL, Convery RS, Peakman G, Todd E, et al. Characterizing the Clinical Features and Atrophy Patterns of MAPT-Related Frontotemporal Dementia With Disease Progression Modeling. *Neurology*. 2021;97(9):e941-e952.
 30. Vogel JW, Young AL, Oxtoby NP, Smith R, Ossenkoppele R, Strandberg OT, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nat Med*. 2021;27(5):871-881.
 31. Young AL, Vogel JW, Aksman LM, Wijeratne PA, Eshaghi A, Oxtoby NP, et al. Ordinal SuStaln: Subtype and Stage Inference for Clinical Scores, Visual Ratings, and Other Ordinal Data. *Front Artif Intell*. 2021;4:613261.
 32. Toledo JB, Van Deerlin VM, Lee EB, Suh E, Baek Y, Robinson JL, et al. A platform for discovery: The University of Pennsylvania Integrated Neurodegenerative Disease Biobank. *Alzheimers Dement*. 2014;10(4):477-484 e471.
 33. Brooks BR, Miller RG, Swash M, Munsat TL, World Federation of Neurology Research Group on Motor Neuron D. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord*. 2000;1(5):293-299.
 34. Strong MJ, Grace GM, Freedman M, Lomen-Hoerth C, Woolley S, Goldstein LH, et al. Consensus criteria for the diagnosis of frontotemporal cognitive and behavioural syndromes in amyotrophic lateral sclerosis. *Amyotroph Lateral Scler*. 2009;10(3):131-146.
 35. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*. 2011;134(Pt 9):2456-

2477.

36. Burke SE, Phillips JS, Olm CA, Peterson CS, Cook PA, Gee JC, et al. Phases of volume loss in patients with known frontotemporal lobar degeneration spectrum pathology. *Neurobiol Aging*. 2022;113:95-107.
37. Kinney NG, Bove J, Phillips JS, Cousins KAQ, Olm CA, Wakeman DG, et al. Social and leisure activity are associated with attenuated cortical loss in behavioral variant frontotemporal degeneration. *Neuroimage Clin*. 2021;30:102629.
38. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo XN, Holmes AJ, et al. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex*. 2018;28(9):3095-3114.
39. Tian Y, Margulies DS, Breakspear M, Zalesky A. Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nat Neurosci*. 2020;23(11):1421-1432.
40. Chung J, Yoo K, Lee P, Kim CM, Roh JH, Park JE, et al. Normalization of cortical thickness measurements across different T1 magnetic resonance imaging protocols by novel W-Score standardization. *Neuroimage*. 2017;159:224-235.
41. Bellec P, Rosa-Neto P, Lyttelton OC, Benali H, Evans AC. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. *Neuroimage*. 2010;51(3):1126-1139.
42. Quinn C, Edmundson C, Dahodwala N, Elman L. Reliable and efficient scale to assess upper motor neuron disease burden in amyotrophic lateral sclerosis. *Muscle Nerve*. 2020;61(4):508-511.
43. Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, et al. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J Neurol Sci*. 1999;169(1-2):13-21.
44. Kollwe K, Mauss U, Krampfl K, Petri S, Dengler R, Mohammadi B. ALSFRS-R score and its ratio: a useful predictor for ALS-progression. *J Neurol Sci*. 2008;275(1-2):69-73.
45. Balendra R, Jones A, Jivraj N, Knights C, Ellis CM, Burman R, et al. Estimating clinical stage of amyotrophic lateral sclerosis from the ALS Functional Rating Scale. *Amyotroph Lateral Scler Frontotemporal Degener*. 2014;15(3-4):279-284.
46. Niven E, Newton J, Foley J, Colville S, Swingler R, Chandran S, et al. Validation of the Edinburgh Cognitive and Behavioural Amyotrophic Lateral Sclerosis Screen (ECAS): A cognitive tool for motor disorders. *Amyotroph Lateral Scler Frontotemporal Degener*. 2015;16(3-4):172-179.
47. Libon DJ, Rascovsky K, Gross RG, White MT, Xie SX, Dreyfuss M, et al. The Philadelphia Brief Assessment of Cognition (PBAC): a validated screening measure for dementia. *Clin Neuropsychol*. 2011;25(8):1314-1330.
48. Caswell C, McMillan CT, Xie SX, Van Deerlin VM, Suh E, Lee EB, et al. Genetic predictors of survival in behavioral variant frontotemporal degeneration. *Neurology*. 2019;93(18):e1707-e1714.
49. Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. 2011;72(2):257-268.
50. Brettschneider J, Del Tredici K, Toledo JB, Robinson JL, Irwin DJ, Grossman M, et al. Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Ann Neurol*. 2013;74(1):20-38.
51. Neumann M, Lee EB, Mackenzie IR. Frontotemporal Lobar Degeneration TDP-43-Immunoreactive Pathological Subtypes: Clinical and Mechanistic Significance. *Adv Exp Med Biol*. 2021;1281:201-217.

52. Xia M, Wang J, He Y. BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One*. 2013;8(7):e68910.
53. Rohrer JD, Geser F, Zhou J, Gennatas ED, Sidhu M, Trojanowski JQ, et al. TDP-43 subtypes are associated with distinct atrophy patterns in frontotemporal dementia. *Neurology*. 2010;75(24):2204-2211.
54. Cividini C, Basaia S, Spinelli EG, Canu E, Castelnovo V, Riva N, et al. Amyotrophic Lateral Sclerosis-Frontotemporal Dementia: Shared and Divergent Neural Correlates Across the Clinical Spectrum. *Neurology*. 2021.
55. Tan HHG, Westeneng HJ, Nitert AD, van Veenhuijzen K, Meier JM, van der Burgh HK, et al. MRI clustering reveals three ALS subtypes with unique neurodegeneration patterns. *Ann Neurol*. 2022.
56. Ranasinghe KG, Rankin KP, Pressman PS, Perry DC, Lobach IV, Seeley WW, et al. Distinct Subtypes of Behavioral Variant Frontotemporal Dementia Based on Patterns of Network Degeneration. *JAMA Neurol*. 2016;73(9):1078-1088.
57. Yokoyama JS, Rosen HJ. Neuroimaging features of C9ORF72 expansion. *Alzheimers Res Ther*. 2012;4(6):45.
58. Lee SE, Khazenzon AM, Trujillo AJ, Guo CC, Yokoyama JS, Sha SJ, et al. Altered network connectivity in frontotemporal dementia with C9orf72 hexanucleotide repeat expansion. *Brain*. 2014;137(Pt 11):3047-3060.
59. Hakkinen S, Chu SA, Lee SE. Neuroimaging in genetic frontotemporal dementia and amyotrophic lateral sclerosis. *Neurobiol Dis*. 2020;145:105063.
60. Beck J, Rohrer JD, Campbell T, Isaacs A, Morrison KE, Goodall EF, et al. A distinct clinical, neuropsychological and radiological phenotype is associated with progranulin gene mutations in a large UK series. *Brain*. 2008;131(Pt 3):706-720.
61. Floris G, Borghero G, Cannas A, Di Stefano F, Murru MR, Corongiu D, et al. Clinical phenotypes and radiological findings in frontotemporal dementia related to TARDBP mutations. *J Neurol*. 2015;262(2):375-384.
62. Mol MO, Nijmeijer SWR, van Rooij JGJ, van Spaendonk RML, Pijnenburg YAL, van der Lee SJ, et al. Distinctive pattern of temporal atrophy in patients with frontotemporal dementia and the I383V variant in TARDBP. *J Neurol Neurosurg Psychiatry*. 2021;92(7):787-789.
63. Peet BT, Spina S, Mundada N, La Joie R. Neuroimaging in Frontotemporal Dementia: Heterogeneity and Relationships with Underlying Neuropathology. *Neurotherapeutics*. 2021;18(2):728-752.
64. Bocchetta M, Iglesias Espinosa MDM, Lashley T, Warren JD, Rohrer JD. In vivo staging of frontotemporal lobar degeneration TDP-43 type C pathology. *Alzheimers Res Ther*. 2020;12(1):34.
65. Geser F, Martinez-Lage M, Robinson J, Uryu K, Neumann M, Brandmeir NJ, et al. Clinical and pathological continuum of multisystem TDP-43 proteinopathies. *Arch Neurol*. 2009;66(2):180-189.
66. Whitwell JL, Boeve BF, Weigand SD, Senjem ML, Gunter JL, Baker MC, et al. Brain atrophy over time in genetic and sporadic frontotemporal dementia: a study of 198 serial magnetic resonance images. *Eur J Neurol*. 2015;22(5):745-752.

Figures

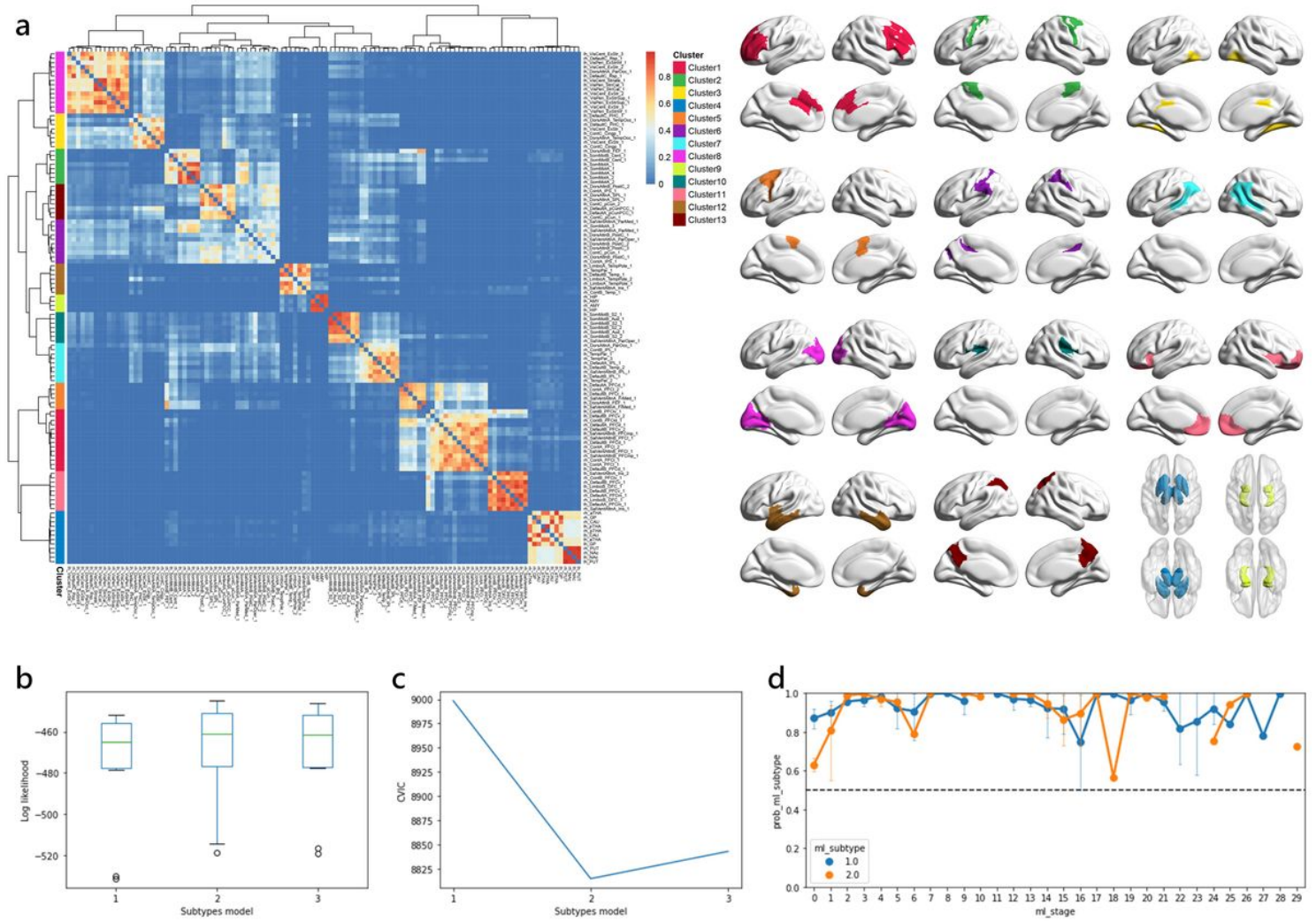


Figure 1

Methodology of selecting optimal number of brain clusters and subtypes. (a) Bootstrap analysis of stable clusters on cortical and subcortical volume. The stability matrix showed that partitions of the brain were classified into stable clusters. Cross-validation was employed and (b) out-of-sample log-likelihood and (c) CVIC were both calculated to select the optimal number of subtypes. (d) Subtype probability across SuStaln stage. *CVIC* cross-validation information criterion.

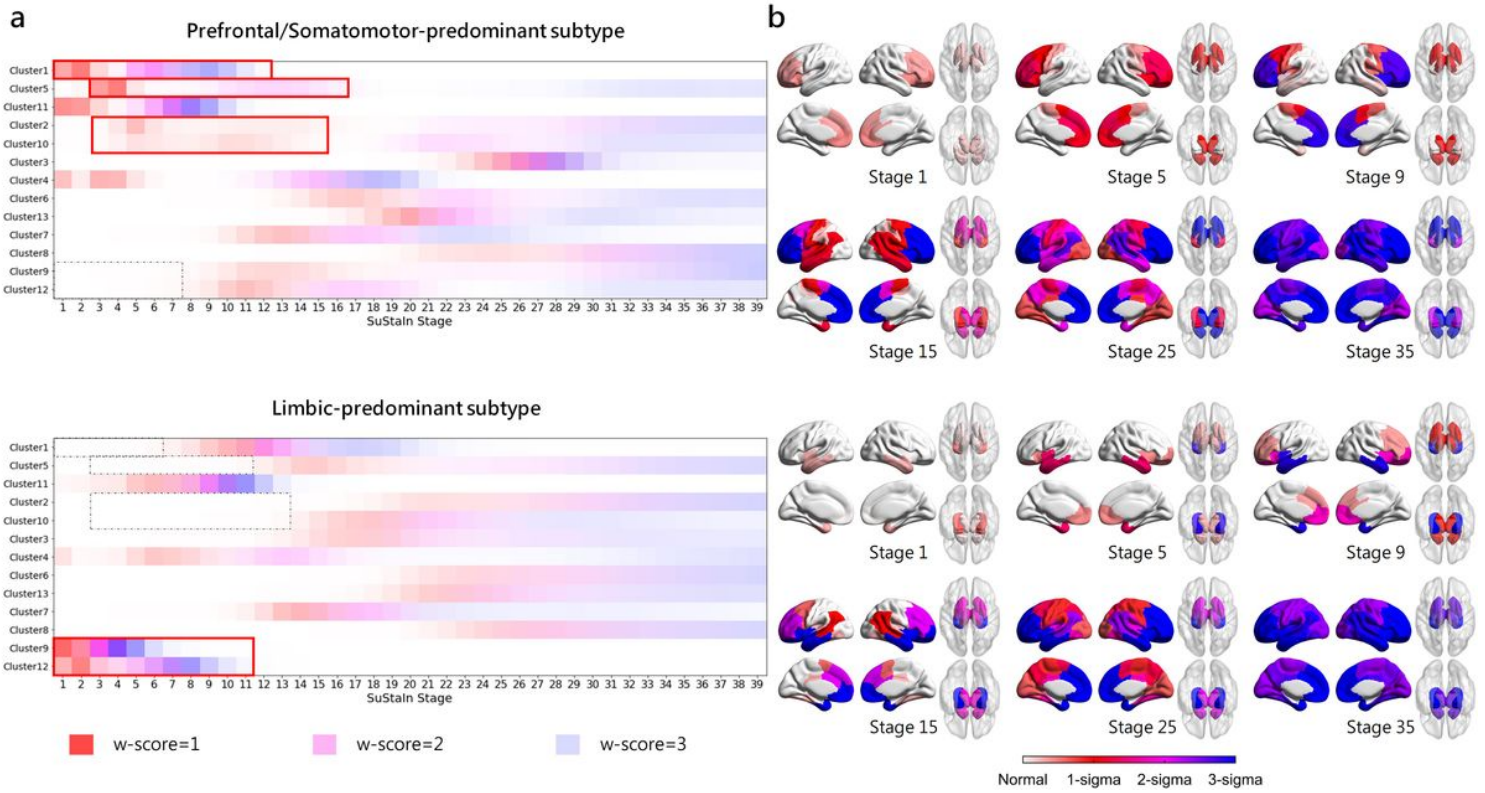


Figure 2

Subtypes progression patterns identified by SuStaln algorithm. (a) W-scores of subtype progression patterns for each region for each subtype. Color shade represents the probability that w-score in each region is reached at each SuStaln stage, with red for mild atrophy (w-score = 1), magenta for moderate atrophy (w-score = 2), and blue for severe atrophy (w-score = 3). (b) Spatial distribution and degree of cortical atrophy at each SuStaln stage. Color shades represent the cumulative sum of probabilities in each brain region.

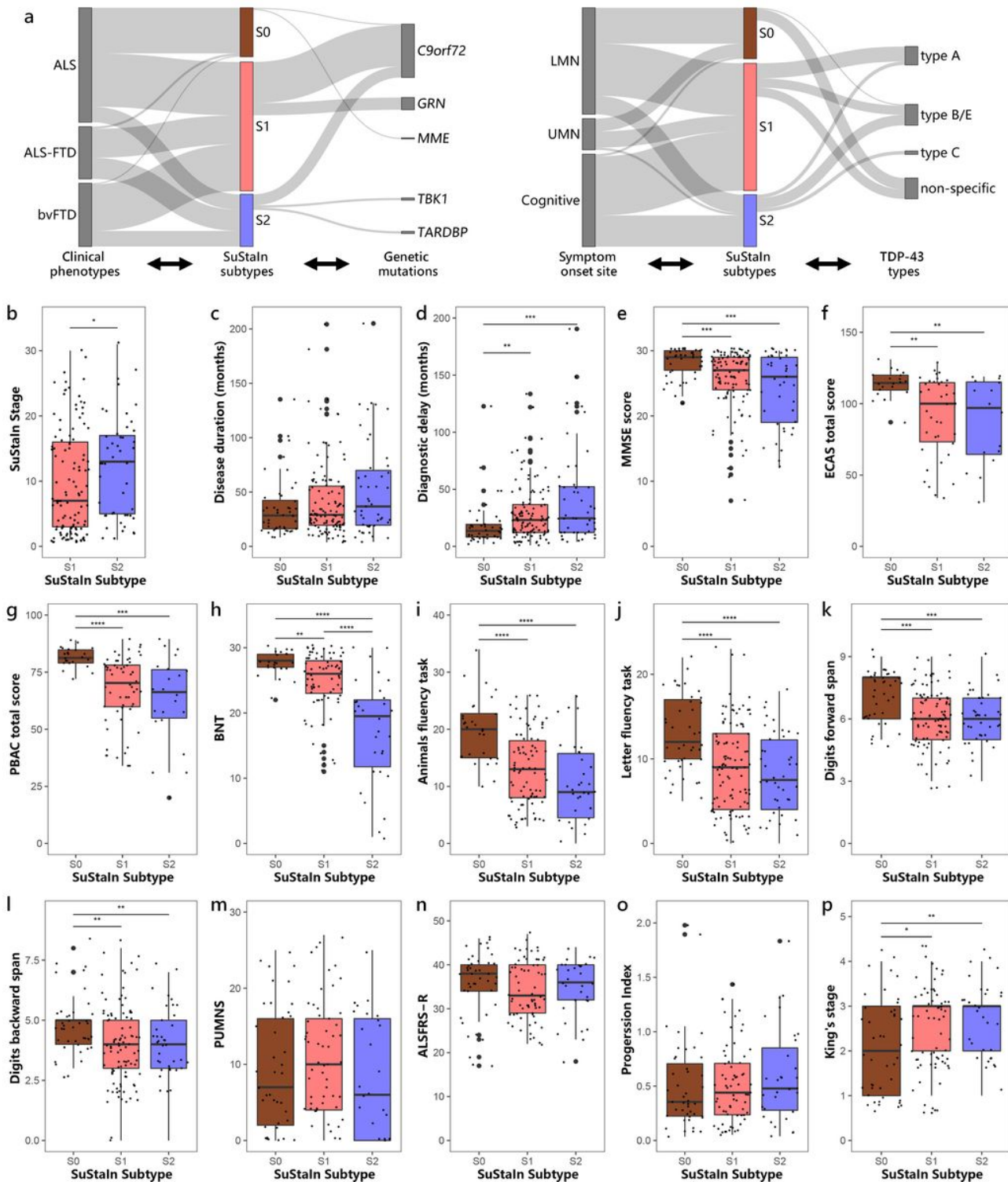


Figure 3

Comparison of clinical, genetic and pathological characteristics across subtypes. (a) number of clinical phenotypes, cases carrying genetic pathogenic variants, symptom onset sites and TDP-43 types assigned to each subtype. Comparison of (b) SuStain stage, (c) disease duration, (d) diagnostic delay, (e-l) cognitive scores across subtypes in all individuals, and (m) PUMNS, (n) ALSFRS-R, (o) progression index, (p) King's stage across subtypes in individuals with ALS/ALS-FTD. *p value < 0.05, **p value < 0.01, ***p value < 0.001, ****p value < 0.0001.

S0 Normal-appearing group, S1 Prefrontal/Somatomotor-predominant subtype, S2 Limbic-predominant subtype, MMSE Mini-Mental Status Examination, ECAS Edinburgh Cognitive Assessment Scale, PBAC Philadelphia Brief Assessment of Cognition, BNT Boston naming test, PUMNS Penn Upper Motor Neuron Score, ALSFRS-R Revised ALS Functional Rating Scale, LMN lower motor neuron, UMN upper motor neuron.

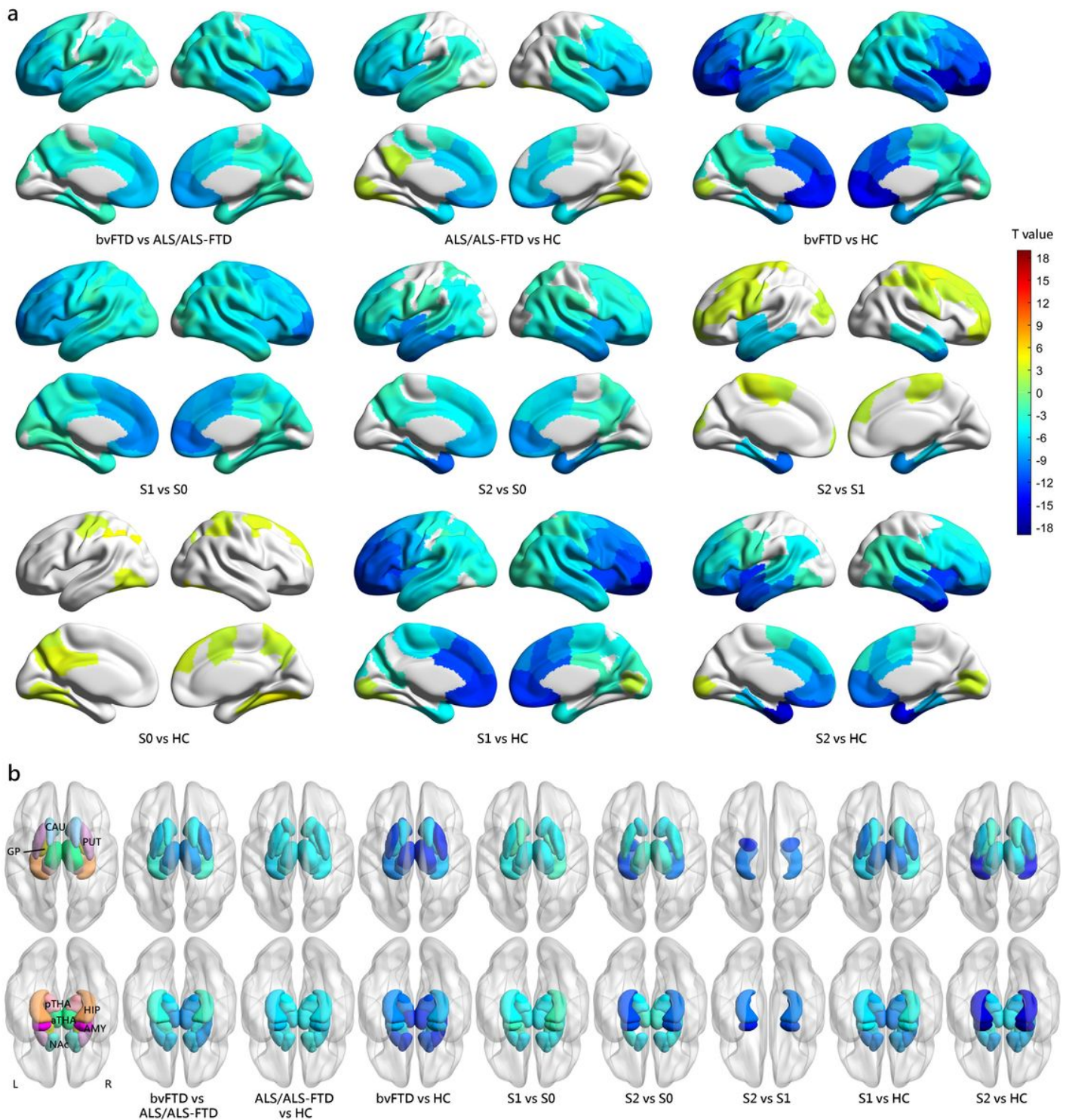


Figure 4

Comparison of volumetric measurements between groups at baseline. (a) Cortical volumetric differences between groups at baseline. (b) Subcortical volumetric differences between groups at baseline. Only results with a threshold at FDR-corrected p value < 0.05 were shown. Cool colors indicate more cortical atrophy in the former group than the latter one, while warm colors indicate more cortical atrophy in the latter group than the former one.

S0 Normal-appearing group, S1 Prefrontal/Somatomotor-predominant subtype, S2 Limbic-predominant subtype

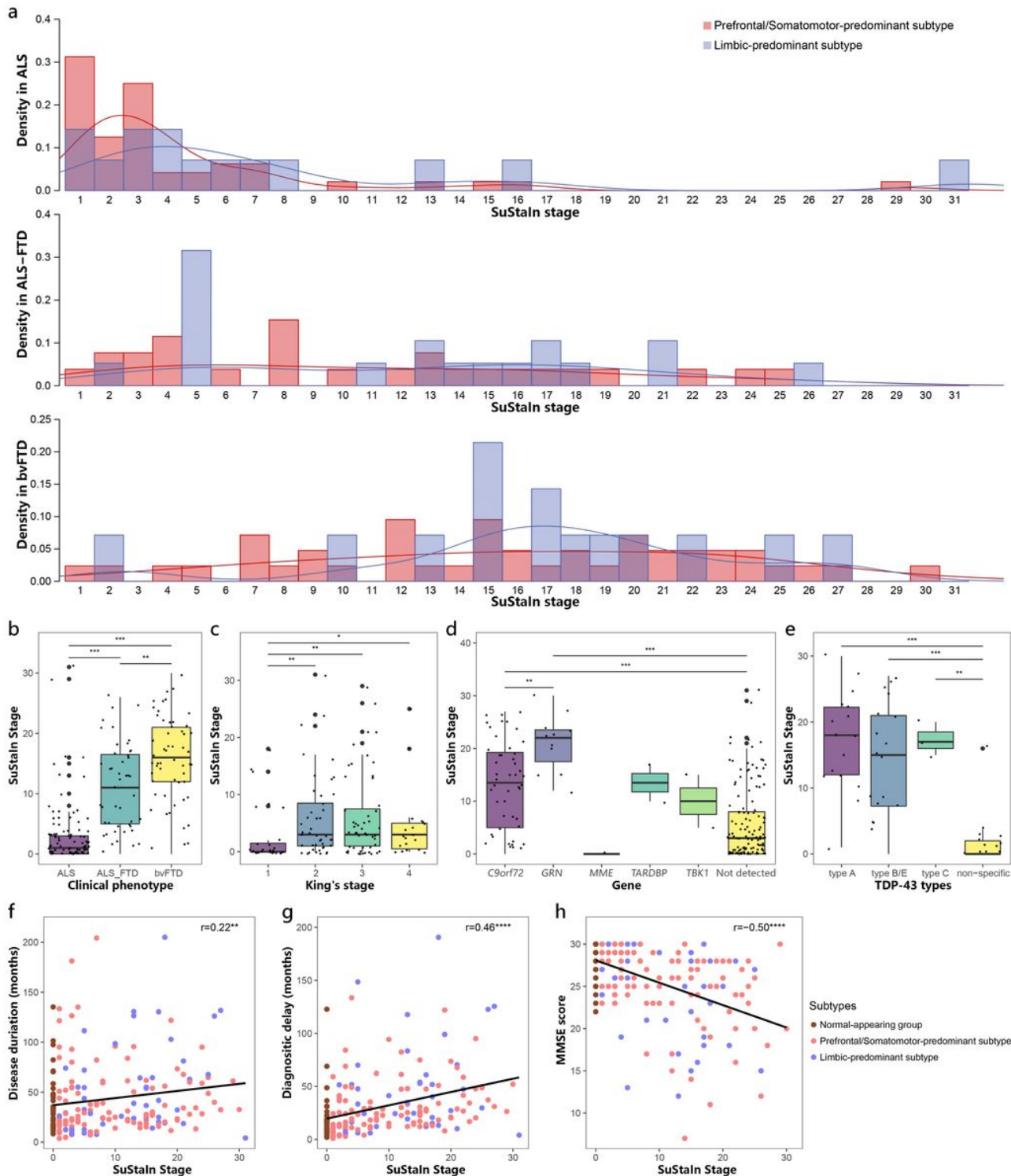


Figure 5

Progression of SuStaln subtypes. (a) Distribution of individuals assigned to each SuStaln stage in different clinical phenotypes. Comparison of SuStaln stages between different clinical phenotypes (b), King's stages (c), genetic pathogenic variants (d), and TDP-43 types (e). Increasing SuStaln stage was correlated with longer disease duration (f), longer diagnostic delay (g) and worse cognitive function (h) across all subtypes. *p value < 0.05, **p value < 0.01, ***p value < 0.001, ****p value < 0.0001.

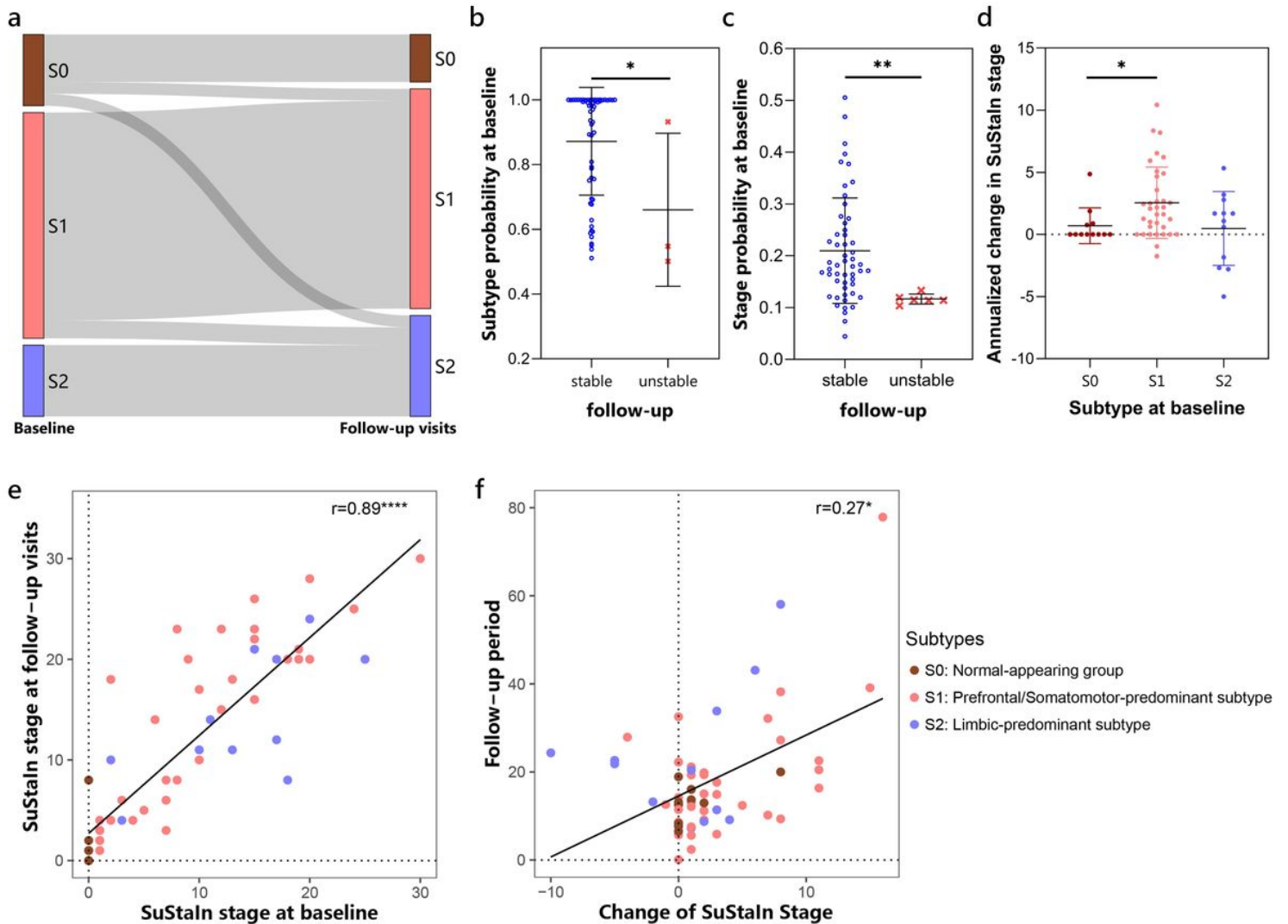


Figure 6

Stability of SuStaln subtypes. (a) Longitudinal subtype consistency. (b) Subtype probability at baseline in groups of stable or unstable longitudinal subtype assignments. (c) Stage probability at baseline in groups of stable or unstable longitudinal stage assignments. (d) Annualized change in SuStaln stage of each subtype in individuals with stable subtypes over time. (e) Correlations between SuStaln stages at baseline and follow-up visits. (f) Correlations between the follow-up period and change of SuStaln stages. *p value < 0.05, **p value < 0.01, ****p value < 0.0001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)