

Preoperative diffusion-perivascular MRI burden and postoperative delirium after subthalamic nucleus deep brain stimulation in Parkinson disease

Chaoshi Niu

niuchaoshi@ustc.edu.cn

The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China

Peng Chen

Chi Xiong

Manli Jiang

Shanshan Wang

Ce Zhang

Hu He

Jianmin Si

Article

Keywords: Parkinson disease, postoperative delirium, subthalamic nucleus deep brain stimulation, multimodal MRI, diffusion tensor imaging, DTI-ALPS, enlarged perivascular spaces

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11133556/v1>

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

Additional Declarations: The authors have declared there is **NO** conflict of interest to disclose

Abstract

Postoperative delirium after subthalamic nucleus deep brain stimulation may reflect pre-existing brain vulnerability, but the preoperative imaging correlates of this vulnerability remain uncertain. We examined whether a combined diffusion and perivascular imaging burden was associated with delirium after bilateral subthalamic nucleus deep brain stimulation. This single-centre observational cohort included 100 patients with Parkinson disease undergoing bilateral subthalamic nucleus deep brain stimulation and 32 healthy controls scanned with the same preoperative MRI protocol. Delirium was assessed during postoperative days using the Nursing Delirium Screening Scale with clinical adjudication. Age- and sex-adjusted reference models fitted in healthy controls were used to standardise tract diffusion metrics, diffusion tensor image analysis along the perivascular space, and basal ganglia enlarged perivascular space counts. Directionally aligned diffusion and perivascular scores were averaged with equal domain weights. Patients who developed postoperative delirium ($n = 30$) had higher combined burden than those who did not ($n = 70$; 95% CI, 0.584 to 1.526; $q < .001$). Using Firth penalised logistic regression, higher combined burden remained associated with postoperative delirium after adjustment for age, sex, education, Mini-Mental State Examination score, disease duration, and motor severity (95% CI, 1.49 to 4.25; Wald $P < .001$). Higher burden was also associated with longer delirium duration and higher peak screening scores among affected patients. Exploratory internal prediction yielded an area under the receiver-operating-characteristic curve of 0.701 after adding burden to clinical variables, compared with 0.538 for clinical variables alone. These findings support preoperative diffusion-perivascular burden as a candidate marker of perioperative neuropsychiatric vulnerability requiring multicentre validation.

Introduction

Postoperative delirium (POD) is an acute neuropsychiatric syndrome that can be followed by persistent cognitive decline [1, 2]. In Parkinson disease (PD), cognitive vulnerability is distributed across cortical–subcortical networks [3], whereas bilateral subthalamic nucleus deep brain stimulation (STN-DBS) is an established treatment for motor complications in appropriately selected patients [4].

Studies of delirium after DBS have identified associations with age, cognition, disease characteristics, and perioperative factors [5–8]. These clinical measures do not fully describe the cerebral changes that may underlie susceptibility. The interaction between baseline vulnerability and acute physiological stress provides a useful framework for studying delirium [9]. In older surgical populations, preoperative MRI studies have linked POD to regional morphometry, thalamic structure, white matter diffusion, and structural connectivity [10–14]. Whether a combined diffusion and perivascular imaging measure is associated with POD in patients undergoing STN-DBS remains uncertain.

Diffusion tensor imaging (DTI) measures tissue diffusion but does not identify a specific pathological process [15]. DTI analysis along the perivascular space (DTI-ALPS) and MRI-visible enlarged perivascular spaces provide additional information about local diffusion and perivascular morphology, although

neither directly measures brain clearance [16–20]. Studies in PD have associated these markers with cognitive impairment and subsequent clinical outcomes [21–24]. Expressing the measures relative to a healthy-control reference allows abnormalities with different units and directions to be summarised in a single score. Whether this summary carries information about POD beyond baseline clinical characteristics is an empirical question.

We therefore tested whether a preoperative healthy-control-referenced diffusion-perivascular MRI burden was associated with POD in patients with PD undergoing bilateral STN-DBS. The primary imaging endpoint in the present analysis was the combined diffusion-perivascular burden. Secondary analyses examined diffusion and perivascular components, delirium duration and severity, and exploratory grey matter localisation. We also evaluated whether the burden provided incremental information beyond clinical variables during internal prediction analyses.

Methods

Study design and participants

This single-centre observational multimodal MRI study was conducted at the First Affiliated Hospital of USTC between May 2024 and October 2025 and is reported following STROBE principles for observational studies, with prediction analyses described in line with TRIPOD-oriented reporting [25, 26]. The source population comprised patients with clinically diagnosed PD who were evaluated for bilateral STN-DBS during the recruitment period. Of 133 screened patients, 33 were excluded because of absence of preoperative T1-weighted imaging ($n = 11$), incomplete diffusion-weighted imaging data ($n = 13$), excessive head motion or imaging artefacts ($n = 6$), or missing prospective POD assessments ($n = 3$). The final surgical cohort included 100 patients with PD who underwent bilateral STN-DBS and had complete preoperative multimodal MRI and prospective POD assessments. Patients were classified as non-POD ($n = 70$) or POD ($n = 30$). Healthy controls ($n = 32$) provided the internal reference for imaging-derived z scores and were not surgical comparators. The study was approved by the local ethics committee (2024KY-099), and all participants provided written informed consent.

Eligible patients fulfilled established clinical indications for bilateral STN-DBS and had preoperative MRI of sufficient quality for morphometric, diffusion, and perivascular analyses. Exclusion criteria were incomplete preoperative MRI, motion or artefact precluding image analysis, a major structural brain lesion, previous intracranial surgery, or missing POD assessment. Baseline variables comprised age, sex, education, disease duration, levodopa equivalent daily dose, Hoehn and Yahr stage, Movement Disorder Society Unified Parkinson Disease Rating Scale Part III score, Mini-Mental State Examination (MMSE) score, Hamilton Depression Rating Scale score, Hamilton Anxiety Rating Scale score, Non-Motor Symptoms Scale score, Parkinson Disease Sleep Scale score, local and general anaesthesia durations, and total intracranial volume (TIV).

Postoperative delirium assessment

POD was monitored from postoperative day 1 through day 5, or until earlier discharge, using the Nursing Delirium Screening Scale (Nu-DESC) [27]. Assessments were performed once per nursing shift, with additional assessments when an acute change in attention, arousal, behaviour, or mental status was suspected. A Nu-DESC score of 2 or higher constituted a positive screen [27]. Positive screens were adjudicated clinically using bedside assessment and review of medical and nursing records.

POD occurrence was defined as at least one adjudicated episode during the monitoring period. POD duration was the cumulative number of postoperative days with at least one adjudicated episode; multiple positive assessments on the same day were counted once. POD severity was defined a priori as the peak Nu-DESC score during the monitoring period. Assessors were blinded to preoperative imaging results. Further details can be found in the Supplementary Methods.

Surgical procedure

All patients underwent bilateral STN-DBS implantation. Stereotactic planning used preoperative MRI, with target coordinates defined in anterior commissure-posterior commissure space. Intraoperative microelectrode recording was used to confirm electrophysiological passage through the STN before bilateral implantation of quadripolar electrodes. The implantable pulse generator was placed subcutaneously in the subclavian region under general anaesthesia. The same anaesthetic agents were used for both local (lidocaine and ropivacaine) and general anaesthesia (etomidate, propofol, cisatracurium, remifentanyl, sufentanyl and remimazolam), with dosages determined according to the prescribing information and adjusted for sex, body weight, and other relevant patient characteristics.

MRI acquisition

MRI was performed preoperatively on a 3.0-T GE Discovery MR750 scanner. The protocol included three-dimensional T1-weighted imaging for morphometry and EPVS assessment and diffusion-weighted imaging for white matter and DTI-ALPS analyses.

Three-dimensional T1-weighted images were acquired using a magnetisation-prepared rapid gradient echo sequence (repetition time, 8.6 ms; echo time, 3.2 ms; inversion time, 450 ms; flip angle, 12 degrees; field of view, 256 x 256 mm; matrix, 256 x 256; slice thickness, 1 mm; isotropic voxel size, 1 x 1 x 1 mm). Diffusion imaging included $b = 0$ and $b = 1000$ s/mm² acquisitions with 64 diffusion-encoding directions (repetition time, 15059 ms; echo time, 96 ms; matrix, 128 x 128; slice thickness, 2 mm).

Diffusion MRI preprocessing and tract-level analysis

Diffusion data were processed using FSL [28]. Raw images were inspected for signal dropout, motion, incomplete coverage, and distortion and were converted to NIfTI format with corresponding b values and gradient vectors. When reverse phase-encoding b = 0 images were available, susceptibility-induced distortion was corrected using TOPUP [29], and eddy-current distortion and head motion were corrected using EDDY with corresponding rotation of the gradient vectors [30]. Brain masks derived from the mean b = 0 image using BET were visually inspected [31].

Diffusion tensors were fitted voxel wise using DTIFIT to generate fractional anisotropy (FA), mean diffusivity (MD), and radial diffusivity (RD) maps. Mean diffusion values were extracted from six prespecified Johns Hopkins University tract families with more than 1500 skeleton voxels: anterior thalamic radiation, corticospinal tract, forceps minor, inferior fronto-occipital fasciculus, inferior longitudinal fasciculus, and superior longitudinal fasciculus [32]. Bilateral values were averaged except for midline forceps minor. Global metric values were tract means weighted by voxel count. Further details can be found in the Supplementary Methods.

Diffusion-perivascular imaging markers

DTI-ALPS was calculated from tensor-derived directional diffusivities at the lateral ventricular body using the method introduced by Taoka and colleagues [17]. Bilateral 5-mm-diameter spherical regions of interest were placed in projection- and association-fibre regions adjacent to the medullary veins. The ALPS index was defined as the mean x-axis diffusivity in the projection and association regions divided by the mean of y-axis diffusivity in the projection region and z-axis diffusivity in the association region. Left and right indices were averaged. Region placement and calculation were performed without knowledge of POD status or other imaging results. DTI-ALPS was treated as an indirect perivascular diffusion marker rather than a direct clearance measure [18–20].

Basal ganglia enlarged perivascular spaces (BG-EPVS) were automatically segmented and quantified on T1-weighted structural images using the T1-PVS v4 deep learning model [33]. Brain-extracted T1-weighted images were spatially and intensity standardised before inference. Each image was passed through the five-fold ensemble, fold probabilities were averaged and thresholded at 0.5, and masks were mapped back to native T1 space. PVS count was defined as the number of three-dimensional connected components in the binary PVS mask within predefined basal ganglia and deep grey matter regions. Further details can be found in the Supplementary Methods.

Healthy-control-referenced burden scores

Healthy-control-referenced scores were constructed to place markers with different units and biological directions on a common scale. Separate internal reference regression models in the 32 healthy controls included age and sex for tract diffusion metrics, DTI-ALPS, and BG-EPVS count. Each participant's residual was divided by the healthy-control residual standard deviation, and the coefficients and residual standard deviations were then applied to all participants without refitting.

Directions were harmonised so that higher values indicated greater abnormality. FA and DTI-ALPS scores were sign-reversed because lower values were considered adverse; MD, RD, and BG-EPVS retained their original directions. DTI burden was the mean of metric-level FA, MD, and RD scores, each averaged across six tract families. Perivascular burden was the mean of sign-reversed DTI-ALPS and BG-EPVS scores. The combined diffusion-perivascular burden was the arithmetic mean of DTI and perivascular burdens, preventing the diffusion domain from receiving extra weight because it contained more component measures. Grey matter volume regions were excluded from the primary burden because they were selected from same-cohort voxel-wise results. Further details can be found in the Supplementary Methods.

Grey matter volume analysis

Preoperative T1-weighted images were processed using the Computational Anatomy Toolbox (CAT12) implemented in Statistical Parametric Mapping (SPM12) [34]. Processing comprised bias field correction, tissue classification into grey matter, white matter, and cerebrospinal fluid, spatial normalisation to Montreal Neurological Institute space, Jacobian modulation to preserve local tissue volume, and Gaussian smoothing. TIV was calculated as the sum of grey matter, white matter, and cerebrospinal fluid volumes.

Voxel-wise grey matter volume differences were examined using a general linear model with group as the factor of interest and age, sex, and TIV as covariates. An omnibus F contrast tested for any difference among healthy controls, non-POD patients, and POD patients. Three directional T contrasts then tested healthy controls > non-POD, healthy controls > POD, and non-POD > POD. Each statistical map was independently thresholded at voxel-level family-wise-error-corrected $P < .05$. Atlas-based characterisation of selected clusters used the Automated Anatomical Labelling Atlas 3 [35] and was descriptive rather than confirmatory. Further details can be found in the Supplementary Methods.

Statistical analysis

Continuous baseline variables were summarised as mean (standard deviation), except Hoehn and Yahr stage and POD outcomes, which were summarised as median [interquartile range]. Categorical variables were summarised as n (%). Baseline comparisons used Welch analysis of variance or t tests, Pearson chi-square tests, or Mann-Whitney U tests as appropriate and were descriptive.

The primary patient-only analysis tested the association between combined diffusion-perivascular burden and POD occurrence. Logistic regression used POD occurrence as the dependent variable and combined burden as the exposure of interest. The fully adjusted clinical model included age, sex, education, MMSE score, disease duration, and Movement Disorder Society Unified Parkinson Disease Rating Scale Part III score. Results were reported as odds ratios per 1-z-unit increase in burden with 95% confidence intervals. Because only 30 POD events were observed, Firth penalised logistic regression was implemented in Python using the Jeffreys-prior penalised likelihood [36], with profile-likelihood

confidence intervals for the combined burden coefficient. Continuous clinical covariates were z-standardised; combined burden was retained on its healthy-control-referenced z scale.

Secondary analyses compared DTI burden, perivascular burden, DTI-ALPS, and BG-EPVS among groups using age- and sex-adjusted heteroscedasticity-consistent linear models. Pairwise comparisons were corrected within predefined analysis families using Benjamini–Hochberg false-discovery-rate control [37]. Associations between burden scores and delirium duration or peak Nu-DESC score among affected patients were assessed using age- and sex-adjusted partial Spearman correlations with bootstrap confidence intervals.

Exploratory prediction analyses were limited to patients with PD. A clinical model including age, sex, disease duration, MMSE score, and motor severity was compared with the same model plus combined diffusion-perivascular burden. Elastic-net logistic regression [38] was evaluated using repeated nested stratified cross-validation, with imputation, standardisation, hyperparameter selection, and coefficient estimation restricted to training folds. Performance was summarised using area under the receiver-operating-characteristic curve (AUROC), average precision, Brier score, scaled Brier score, calibration slope and intercept, and decision-curve analysis. These analyses were treated as hypothesis-generating because no external validation cohort was available.

All statistical tests were two-sided except prespecified directional voxel-wise contrasts. P values and false-discovery-rate-adjusted q values were reported exactly when possible, with $P < .001$ reported for smaller values.

Data and code availability

Deidentified derived imaging features, clinical variables required to reproduce the reported analyses, and statistical code are available from the corresponding author on reasonable request subject to institutional approval. Raw MRI data are subject to institutional and consent restrictions.

Results

Participants

The final sample comprised 132 participants: 32 healthy controls and 100 patients with PD, of whom 70 did not develop POD and 30 developed POD. Age, sex distribution, and TIV did not differ across the three groups. Among patients with PD, the POD group had a lower MMSE score than the non-POD group (26.63 [SD 1.25] vs 27.21 [SD 1.17]) and numerically lower education (8.87 [SD 2.94] vs 10.03 [SD 2.26] years). Disease duration, levodopa equivalent daily dose, motor severity, mood symptoms, non-motor symptom burden, sleep scale score, and anaesthesia durations were similar between patient groups (Table 1 and Supplementary Table 1).

Table 1
Participant characteristics

Characteristic	HC	Non-POD	POD
Age, years	60.06 (6.28)	61.16 (6.04)	59.50 (6.39)
Male sex, n (%)	17/32 (53.1%)	42/70 (60.0%)	17/30 (56.7%)
Education, years		10.03 (2.26)	8.87 (2.94)
Disease duration, years		9.53 (2.35)	10.00 (2.73)
LEDD, mg/day		472.86 (118.84)	514.50 (211.49)
Hoehn and Yahr stage		3.5 [3.0,4.0]	3.0 [3.0,4.0]
MDS-UPDRS III		56.73 (11.05)	56.77 (12.22)
MMSE		27.21 (1.17)	26.63 (1.25)
HAMD		15.69 (3.19)	15.33 (3.45)
HAMA		15.31 (2.55)	15.47 (2.87)
NMSS		41.13 (12.52)	43.20 (16.44)
PDSS		104.23 (14.04)	101.33 (12.25)
Local anesthesia time, min		32.73 (5.45)	33.20 (6.28)
General anesthesia time, min		65.36 (8.01)	66.23 (7.97)
POD duration, days			2.0 [1.0,3.0]
Peak Nu-DESC score			4.0 [3.0,5.0]
Data are mean (SD), median [IQR], or n (%). HC, healthy control; POD, postoperative delirium.			

White matter diffusion findings

Global weighted diffusion summaries differed across groups (Fig. 1B). Compared with HC, non-POD patients had lower FA (adjusted difference, -0.0205; $q=.016$), higher MD (4.0×10^{-5} ; $q=.013$), and higher RD (4.7×10^{-5} ; $q=.002$). The corresponding differences were larger for POD versus HC: FA, -0.0399; MD, 1.21×10^{-4} ; and RD, 1.13×10^{-4} (all $q<.001$). In the direct POD versus non-POD contrast, POD was associated with lower FA (-0.0194; $q=.032$), higher MD (8.1×10^{-5} ; $q<.001$), and higher RD (6.6×10^{-5} ; $q=.001$). A five-tract MD sensitivity analysis excluding the superior longitudinal fasciculus yielded an almost identical direct difference (8.0×10^{-5}).

Tract-level analyses showed a metric-dependent pattern. In the POD versus non-POD contrast, no individual FA tract survived metric-wise FDR correction (oriented standardized effects, 0.421 to 0.537; $q=.072$ to .080; Fig. 1C,D). MD was higher in POD across all six tract families (oriented standardized effects, 0.745 to 0.926; $q=.002$ for the anterior thalamic radiation and $q<.001$ for the remaining tracts).

RD was also higher in POD across all six tract families (oriented standardized effects, 0.706 to 0.885; $q=.004$ for the anterior thalamic radiation, $q=.002$ for forceps minor, and $q<.001$ for the remaining tracts). Further details can be found in the Supplemental Table 2, 3.

Perivascular imaging findings

Adjusted mean DTI-ALPS was 1.172 (95% CI, 1.141 to 1.202) in HC, 1.067 (1.042 to 1.093) in non-POD, and 1.016 (0.976 to 1.056) in POD (Fig. 2B). DTI-ALPS was lower in non-POD than HC (adjusted difference, -0.104; $q<.001$), POD than HC (-0.156; $q<.001$), and POD than non-POD (-0.052; $q=.036$).

Adjusted mean BG-EPVS counts were 4.92 (95% CI, 4.57 to 5.27) in HC, 6.88 (6.56 to 7.20) in non-POD, and 8.13 (7.69 to 8.57) in POD (Fig. 2B). Counts were higher in non-POD than HC by 1.96, in POD than HC by 3.21, and in POD than non-POD by 1.25 (all $q<.001$).

Within the POD group, lower DTI-ALPS was associated with longer POD duration (partial $\rho=-0.679$; 95% CI, -0.890 to -0.352; $q<.001$) and higher peak Nu-DESC score ($\rho=-0.582$; 95% CI, -0.812 to -0.266; $q=.002$) (Fig. 2C). Higher BG-EPVS count was associated with longer POD duration ($\rho = 0.672$; 95% CI, 0.347 to 0.893; $q<.001$) and higher peak Nu-DESC score ($\rho = 0.521$; 95% CI, 0.186 to 0.802; $q=.004$) (Fig. 2D). Further details can be found in the Supplemental Table 4.

Component diffusion and perivascular findings

The primary burden was supported by both diffusion and perivascular components. DTI burden was higher in the POD group than in the non-POD group by 1.164 z units (95% CI, 0.496 to 1.832; $q=.001$), and perivascular burden was higher by 0.978 z units (95% CI, 0.495 to 1.460; $q<.001$). At the individual-marker level, DTI-ALPS was lower in patients with POD than in those without POD (adjusted difference, -0.052; 95% CI, -0.100 to -0.003; $q=.036$), whereas BG-EPVS count was higher (adjusted difference, 1.245; 95% CI, 0.691 to 1.799; $q<.001$; Fig. 3A).

Within patients with POD, perivascular burden was associated with longer delirium duration (partial $\rho = 0.808$; 95% CI, 0.593 to 0.931; $q<.001$) and higher peak Nu-DESC score ($\rho = 0.675$; 95% CI, 0.385 to 0.869; $q<.001$). DTI burden was also associated with delirium duration ($\rho = 0.483$; 95% CI, 0.118 to 0.754; $q=.011$) and peak score ($\rho = 0.419$; 95% CI, 0.014 to 0.712; $q=.027$). Combined burden was associated with duration ($\rho = 0.683$; 95% CI, 0.428 to 0.878; $q<.001$) and peak score ($\rho = 0.580$; 95% CI, 0.241 to 0.818; $q=.002$) (Fig. 3B). Further details can be found in the Supplemental Table 4.

Primary healthy-control-referenced diffusion-perivascular burden

The combined diffusion-perivascular burden showed a graded pattern across healthy controls, patients without POD, and patients with POD. In age- and sex-adjusted comparisons, combined burden was higher in patients with POD than in patients without POD by 1.091 z units (95% CI, 0.605 to 1.577; $q<.001$). The corresponding differences were 1.245 z units for non-POD patients versus healthy controls

(95% CI, 0.881 to 1.610; $q < .001$) and 2.336 z units for POD patients versus healthy controls (95% CI, 1.866 to 2.806; $q < .001$).

In patient-only logistic regression, combined diffusion-perivascular burden was associated with POD after adjustment for age and sex (odds ratio per 1-z-unit increase, 2.76; 95% CI, 1.64 to 4.62; $P < .001$). The association remained in the fully adjusted clinical model including age, sex, education, MMSE score, disease duration, and motor severity (odds ratio, 2.68; 95% CI, 1.54 to 4.68; $P < .001$). Bootstrap resampling gave a wider confidence interval for the fully adjusted model (95% CI, 1.71 to 7.59), consistent with limited precision from 30 delirium events.

Using Firth penalised logistic regression, higher combined diffusion-perivascular burden remained associated with POD after adjustment for age, sex, education, MMSE score, disease duration, and motor severity (odds ratio per 1-z-unit increase, 2.41; profile-likelihood 95% CI, 1.49 to 4.25; Wald $P < .001$; Fig. 3C and). Further details can be found in the Supplemental Table 5.

Exploratory grey matter volume localisation

The omnibus voxel wise analysis identified group differences in distributed frontal, insular, temporal, occipital, and subcortical regions (voxel-level FWE-corrected $P < 0.05$; Fig. 4A). Relative to HC, non-POD patients showed lower GMV predominantly in frontal and fronto-insular regions (Fig. 4B), whereas the HC > POD contrast extended to frontal, temporal, occipital, and subcortical regions (Fig. 4C). The difference in spatial extent between these separately thresholded HC referenced maps was descriptive. In the direct patient contrast, POD was associated with lower GMV in medial and lateral frontal, orbitofrontal, thalamic, and posterior cortical regions (voxel-level FWE-corrected $P < 0.05$; Fig. 4D).

Atlas based characterisation summarised participant level distributions for four representative regions (Fig. 4E). R-IFGoperc GMV was lower in non-POD than HC ($q = 0.007$), without evidence of further reduction in POD. R-MFG GMV was lower in non-POD than HC, POD than HC, and POD than non-POD (all $q < 0.001$). L-OFCmed GMV was lower in POD than HC ($q = 0.006$) and non-POD ($q = 0.044$). L-THA GMV was lower in POD than HC ($q = 0.035$) and non-POD ($q = 0.044$). Further details can be found in the Supplemental Table 6.

Exploratory internal prediction analysis

In exploratory repeated nested cross-validation among patients with PD, the clinical model yielded an AUROC of 0.538 (95% CI, 0.404 to 0.666), average precision of 0.329 (95% CI, 0.233 to 0.482), Brier score of 0.213, scaled Brier score of -0.012, calibration intercept of -0.992, and calibration slope of -0.165. Adding combined diffusion-perivascular burden yielded an AUROC of 0.701 (95% CI, 0.586 to 0.810), average precision of 0.517 (95% CI, 0.349 to 0.695), Brier score of 0.188, scaled Brier score of 0.103, calibration intercept of -0.035, and calibration slope of 0.946. Relative to the clinical model, the imaging-augmented model increased AUROC by 0.163 (95% CI, 0.038 to 0.281), average precision by 0.187 (95% CI, 0.060 to 0.314), and scaled Brier score by 0.115 (95% CI, 0.004 to 0.224; Fig. 5).

Discussion

In this single-centre cohort of patients with PD undergoing bilateral STN-DBS, POD was associated with higher preoperative healthy-control-referenced diffusion-perivascular MRI burden. The combined burden showed a graded pattern from healthy controls to patients without delirium and patients with delirium, differentiated the two patient groups, and was associated with delirium duration and peak Nu-DESC score among affected patients. The association remained in Firth penalised logistic regression after adjustment for clinical vulnerability variables. Exploratory internal resampling suggested that the burden provided incremental information beyond clinical variables.

These findings support a multidomain brain vulnerability framework for POD in PD. Delirium is unlikely to arise from a single anatomical lesion or isolated imaging abnormality. Instead, the results suggest that patients who develop delirium after STN-DBS may enter surgery with a greater burden of diffuse white matter and perivascular-related tissue abnormalities. This interpretation is consistent with the concept that an acute perioperative stressor can precipitate delirium when baseline neural reserve is limited.

The diffusion and perivascular components appeared complementary. Diffusion abnormalities were distributed across association, callosal, and projection pathways and were more consistent for MD and RD than for FA after tract-level correction. Because tensor metrics are influenced by fibre geometry, partial-volume effects, and tissue water properties, they should be interpreted as non-specific markers of microstructural vulnerability rather than direct histopathological measures [15]. The anatomical distribution provides a plausible functional context for these associations. The anterior thalamic radiation and forceps minor contribute to frontal–thalamic and interhemispheric frontal connectivity; lesions in strategically located portions of these pathways have been associated with executive and visuomotor performance [32, 39]. The superior longitudinal fasciculus participates in a lateralised parieto-frontal network supporting visuospatial attention [40]. The inferior fronto-occipital fasciculus links posterior association regions with frontal control systems, and tractography work has related its subdivisions to semantic access and visually guided behavioural control [41]. The inferior longitudinal fasciculus connects occipital and anterior temporal regions and has been implicated in visual recognition and higher-order visual cognition [42]. These associations provide anatomical hypotheses for how distributed white-matter vulnerability might contribute to disorientation or attentional dysregulation after surgery; they do not demonstrate disruption of any specific functional network.

The perivascular related findings provide a complementary dimension to the structural and diffusion findings. Recent PD studies have reported lower DTI-ALPS in patients with cognitive impairment or poorer longitudinal outcomes, including associations with subsequent cognitive decline and disease progression, while continuing to treat the index as an indirect imaging proxy rather than a direct clearance measure [21, 22]. MRI-visible EPVS, particularly BG-EPVS, have also been associated with worse cognitive phenotypes in PD; in addition, a longitudinal PD-MCI study found that lower DTI-ALPS and greater BG-PVS burden co-occurred in patients who later converted to dementia [23, 24]. The present study differs from much of that literature by measuring DTI-ALPS and BG-EPVS before a prospectively

ascertained postoperative outcome. DTI-ALPS decreased and BG-EPVS counts increased across HC, non-POD, and POD groups, and both markers remained different between the two PD groups after age and sex adjustment. Within POD, lower DTI-ALPS and higher BG-EPVS counts were associated with longer duration and greater peak Nu-DESC severity, indicating that these markers tracked variation in the delirium phenotype as well as case status. One hypothesis is that these markers capture a component of perivascular or vascular tissue homeostasis relevant to the brain's response to perioperative stress.

The healthy control referenced diffusion-perivascular burden provides a pragmatic way to summarise imaging abnormalities that differ in scale and direction. By expressing each marker as an age and sex adjusted residual z score relative to the internal healthy control distribution, and orienting scores so that higher values indicate greater abnormality, the approach allowed selected diffusion and perivascular measures to be examined within a common framework. The resulting burden scores showed a graded pattern across HC, non-POD, and POD groups and were higher in patients who developed POD, supporting the interpretation that POD after STN-DBS is associated with greater preoperative cerebral vulnerability. The associations between diffusion-perivascular burden and POD duration and peak Nu-DESC score further suggest that preoperative imaging vulnerability may relate not only to POD occurrence but also to its clinical expression among affected patients. However, these within POD analyses were exploratory and based on 30 events.

Exploratory grey matter volume results localised part of the vulnerability pattern to frontal, orbitofrontal, and thalamic regions. These regions are biologically plausible in relation to attention, cognitive control, and PD cognitive networks [3], but the atlas-based region analyses were selected from same-cohort voxel-wise findings and should not be treated as independent confirmatory biomarkers. Their main value is anatomical interpretation and generation of hypotheses for future prespecified region-of-interest studies.

The prediction analyses should be read as exploratory. Adding DTI burden, perivascular burden, or their diffusion-perivascular composite modestly improved internally cross-validated performance relative to the specified clinical model, with the composite model yielding the highest point estimates in the main framework (AUROC, 0.701), the clinical model yielded an AUROC of 0.538. These results indicate potential incremental information from preoperative imaging, but they do not establish an individual-level clinical prediction tool. The clinical comparator was limited, the number of POD events was small, bootstrap intervals were based on fixed aggregated out-of-fold predictions, and no external validation was available. The appropriate interpretation is hypothesis generation for future risk-stratification studies, not clinical deployment.

Several limitations define the scope of inference. First, the study was single-centre, and all validation was internal. Second, the number of POD events limited precision for clinical-adjusted modelling, within-delirium severity analyses, and exploratory prediction. Finally, the healthy-control reference included 32 participants, making the normative residual scale potentially sample-sensitive.

Conclusion

Preoperative healthy-control-referenced diffusion-perivascular MRI burden was associated with POD after STN-DBS in PD and was related to delirium duration and severity among affected patients. The findings support this burden as a candidate marker of perioperative neuropsychiatric vulnerability rather than a causal mechanism or a clinically deployable prediction tool. Multicentre external validation with prespecified imaging features, robust delirium adjudication, and transparent data and code sharing is required.

Abbreviations

AP

average precision

AUROC

area under the receiver-operating-characteristic curve

BG-EPVS

basal ganglia enlarged perivascular spaces

DBS

deep brain stimulation

DTI

diffusion tensor imaging

DTI-ALPS

diffusion tensor image analysis along the perivascular space

EPVS

enlarged perivascular spaces

FA

fractional anisotropy

FDR

false-discovery rate

GMV

grey matter volume

HC

healthy control

MD

mean diffusivity

MMSE

Mini-Mental State Examination

Nu-DESC

Nursing Delirium Screening Scale

PD

Parkinson disease

POD
postoperative delirium
RD
radial diffusivity
STN
subthalamic nucleus
TIV
total intracranial volume.

Declarations

Funding

The Excellent Research and Innovation Team Project of Anhui Province University, 2023AH010080; Fundamental Research Funds for the Central Universities, YD9110002037; China Postdoctoral Science Foundation, 2023M743393.

Competing interests

The authors declare no competing interests.

Ethics approval

This study was approved by the institutional review board of the First Affiliated Hospital of USTC (Ethics Review No. 2024KY-099).

Data availability

Data generated or analysed during the study are available from the corresponding author on reasonable request subject to institutional approval.

References

1. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet*. 2014;383:911–22.
2. Goldberg TE, Chen C, Wang Y, et al. Association of delirium with long-term cognitive decline: a meta-analysis. *JAMA Neurol*. 2020;77:1373–81.
3. Gratwicke J, Jahanshahi M, Foltynie T. Parkinson's disease dementia: a neural networks perspective. *Brain*. 2015;138:1454–76.

4. Deuschl G, Antonini A, Costa J, et al. European Academy of Neurology/Movement Disorder Society-European Section guideline on the treatment of Parkinson's disease: I. Invasive therapies. *Mov Disord*. 2022;37:1360–74.
5. Yousef O, Abouelmagd ME, Abbas A, et al. Delirium after deep brain stimulation for Parkinson's disease: a meta-analysis of incidence and risk factors. *Neurosurg Rev*. 2025;48:73.
6. Ozluk E, Ozturk G. Risk factors for delirium occurring after deep brain stimulation surgery in patients with Parkinson's disease. *Acta Neurochir (Wien)*. 2024;166:474.
7. Lu W, Chang X, Bo L, et al. Risk factors for delirium after deep brain stimulation surgery under total intravenous anesthesia in Parkinson's disease patients. *Brain Sci*. 2023;13:25.
8. Astalosch M, Mousavi M, Ribeiro LM, et al. Risk factors for postoperative delirium severity after deep brain stimulation surgery in Parkinson's disease. *J Parkinsons Dis*. 2024;14:1175–92.
9. Inouye SK, Charpentier PA. Precipitating factors for delirium in hospitalized elderly persons: predictive model and interrelationship with baseline vulnerability. *JAMA*. 1996;275:852–7.
10. Kant IMJ, de Bresser J, van Montfort SJT, et al. Preoperative brain MRI features and occurrence of postoperative delirium. *J Psychosom Res*. 2021;140:110301.
11. Fislage M, Feinkohl I, Pischon T, et al. Presurgical thalamus volume in postoperative delirium: a longitudinal observational cohort study in older patients. *Anesth Analg*. 2022;135:136–42.
12. Fislage M, Winzeck S, Stamatakis EA, et al. Presurgical diffusion metrics of the thalamus and thalamic nuclei in postoperative delirium: a prospective two-centre cohort study in older patients. *Neuroimage Clin*. 2022;36:103208.
13. Cavallari M, Dai W, Guttman CRG, et al. Neural substrates of vulnerability to postsurgical delirium as revealed by presurgical diffusion MRI. *Brain*. 2016;139:1282–94.
14. Fislage M, Winzeck S, Woodrow R, et al. Structural disconnectivity in postoperative delirium: a perioperative two-center cohort study in older patients. *Alzheimers Dement*. 2024;20:2861–72.
15. Jones DK, Knösche TR, Turner R. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *Neuroimage*. 2013;73:239–54.
16. Duering M, Biessels GJ, Brodtmann A, et al. Neuroimaging standards for research into small vessel disease—advances since 2013. *Lancet Neurol*. 2023;22:602–18.
17. Taoka T, Masutani Y, Kawai H, et al. Evaluation of glymphatic system activity with the diffusion MR technique: diffusion tensor image analysis along the perivascular space (DTI-ALPS) in Alzheimer disease cases. *Jpn J Radiol*. 2017;35:172–8.
18. Taoka T, Ito R, Nakamichi R, et al. Diffusion tensor image analysis along the perivascular space (DTI-ALPS): revisiting the meaning and significance of the method. *Magn Reson Med Sci*. 2024;23:268–90.
19. Georgiopoulos C, Werlin A, Lasic S, et al. Diffusion tensor imaging along the perivascular space: the bias from crossing fibres. *Brain Commun*. 2024;6:fcae421.

20. Mossige I, Valnes LM, Storås TH, et al. Comparing glymphatic function measures: diffusion tensor image analysis along perivascular spaces versus intrathecal contrast-enhanced MRI. *Radiology*. 2026;318:e252070.
21. Wood KH, Nenert R, Miften AM, et al. Diffusion tensor imaging-along the perivascular-space index is associated with disease progression in Parkinson's disease. *Mov Disord*. 2024;39:1504–13.
22. Zarkali A, Thomas G, Paterson R, et al. Impaired glymphatic clearance, measured using diffusion tensor image analysis along the perivascular space (DTI-ALPS), is linked to poor cognitive outcomes in Parkinson's disease. *Mov Disord*. 2025;40:2367–80.
23. Kim S, Na HK, Sun Y, et al. Regional burden of enlarged perivascular spaces and cognition and neuropsychiatric symptoms in drug-naive patients with Parkinson disease. *Neurology*. 2024;102:e209483.
24. Pang H, Wang J, Yu Z, et al. Glymphatic function from diffusion-tensor MRI to predict conversion from mild cognitive impairment to dementia in Parkinson disease. *J Neurol*. 2024;271:5598–609.
25. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370:1453–7.
26. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162:55–63.
27. Gaudreau JD, Gagnon P, Harel F, et al. Fast, systematic, and continuous delirium assessment in hospitalized patients: the Nursing Delirium Screening Scale. *J Pain Symptom Manage*. 2005;29:368–75.
28. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004;23(Suppl 1):S208–19.
29. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage*. 2003;20:870–88.
30. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage*. 2016;125:1063–78.
31. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002;17:143–55.
32. Mori S, Oishi K, Jiang H, et al. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage*. 2008;40:570–82.
33. Boutinaud P, Tsuchida A, Laurent A, et al. 3D segmentation of perivascular spaces on T1-weighted 3 Tesla MR images with a convolutional autoencoder and a U-shaped neural network. *Front Neuroinform*. 2021;15:641600.
34. Gaser C, Dahnke R, Thompson PM, et al. CAT: a computational anatomy toolbox for the analysis of structural MRI data. *Gigascience*. 2024;13:giae049.

35. Rolls ET, Huang CC, Lin CP, et al. Automated anatomical labelling atlas 3. *Neuroimage*. 2020;206:116189.
36. Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993;80:27–38.
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57:289–300.
38. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33:1–22.
39. Biesbroek JM, Weaver NA, Hilal S, et al. Impact of strategically located white matter hyperintensities on cognition in memory clinic patients with small vessel disease. *PLoS One*. 2016;11:e0166261.
40. de Schotten MT, Dell’Acqua F, Forkel SJ, Simmons A, Vergani F, Murphy DGM, et al. A lateralized brain network for visuospatial attention. *Nat Neurosci*. 2011;14:1245–6.
41. Gonzalez Alam TR, Arias JC, Jefferies E, Smallwood J, Leemans A, Davolos JM. Ventral and dorsal aspects of the inferior fronto-occipital fasciculus support verbal semantic access and visually-guided behavioural control. *Brain Struct Funct*. 2024;229:207–21.
42. Herbet G, Zemmoura I, Duffau H. Functional anatomy of the inferior longitudinal fasciculus: from historical reports to current hypotheses. *Front Neuroanat*. 2018;12:77.

Figures

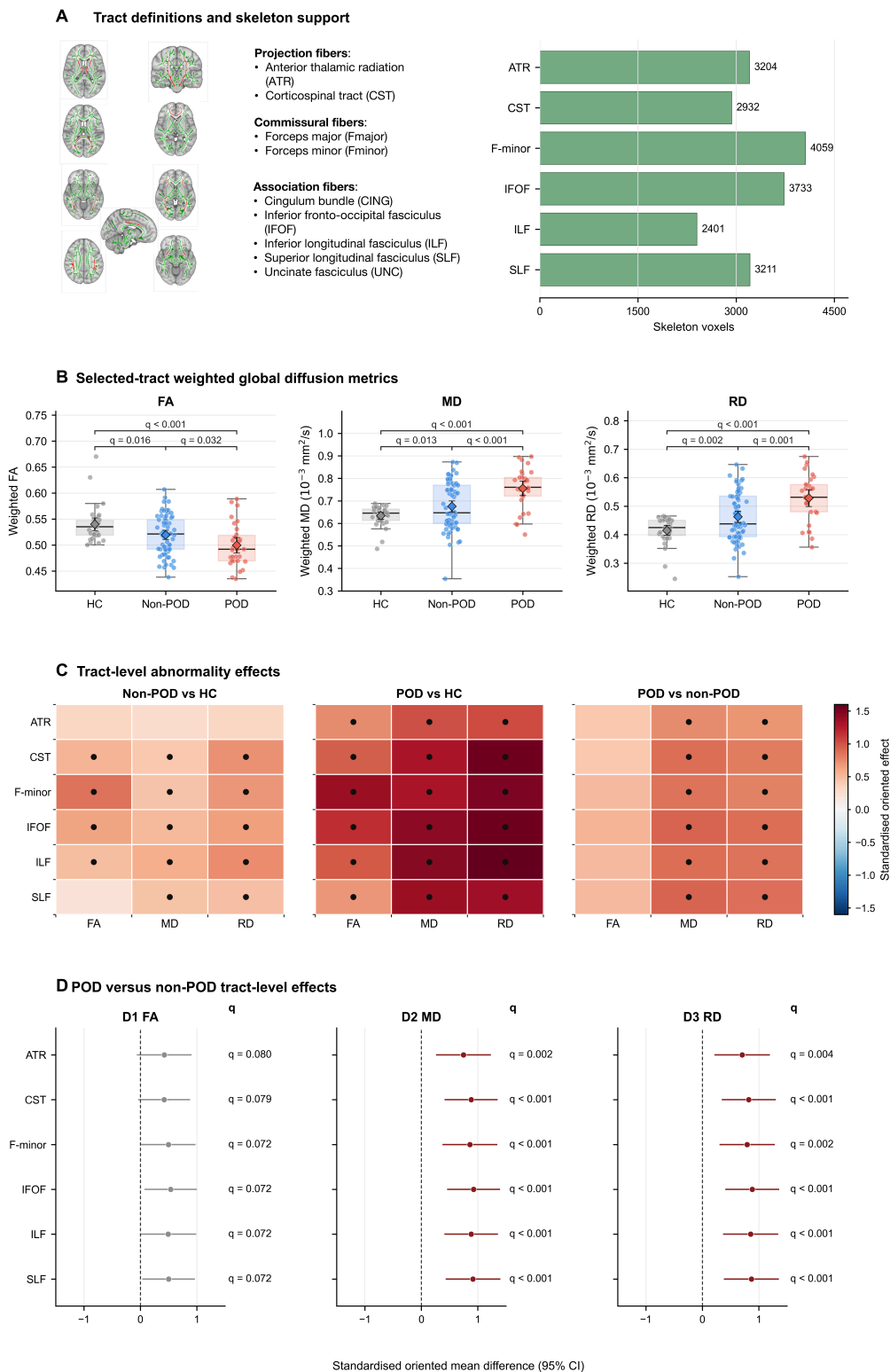


Figure 1

Diffusion measures across the selected white matter tracts.

A, Anatomical tract illustrations and skeleton voxel counts for the six tract families included in the quantitative analysis. B, Voxel-count-weighted FA, MD, and RD summaries across the selected tracts in healthy controls (n=32), non-POD patients (n=70), and POD patients (n=30). Points show participants,

boxes show medians and interquartile ranges, and diamonds with error bars indicate age- and sex-adjusted means and 95% confidence intervals. MD and RD are displayed in units of $10^{-3} \text{ mm}^2/\text{s}$. C, Oriented standardised group effects; positive values indicate lower FA or higher MD/RD in the first-named group. Dots mark comparisons meeting the displayed FDR criterion. D, POD versus non-POD effects with 95% confidence intervals, corrected separately by metric. No individual FA tract met $q < 0.05$. ATR, anterior thalamic radiation; CST, corticospinal tract; F-minor, forceps minor; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; SLF, superior longitudinal fasciculus.

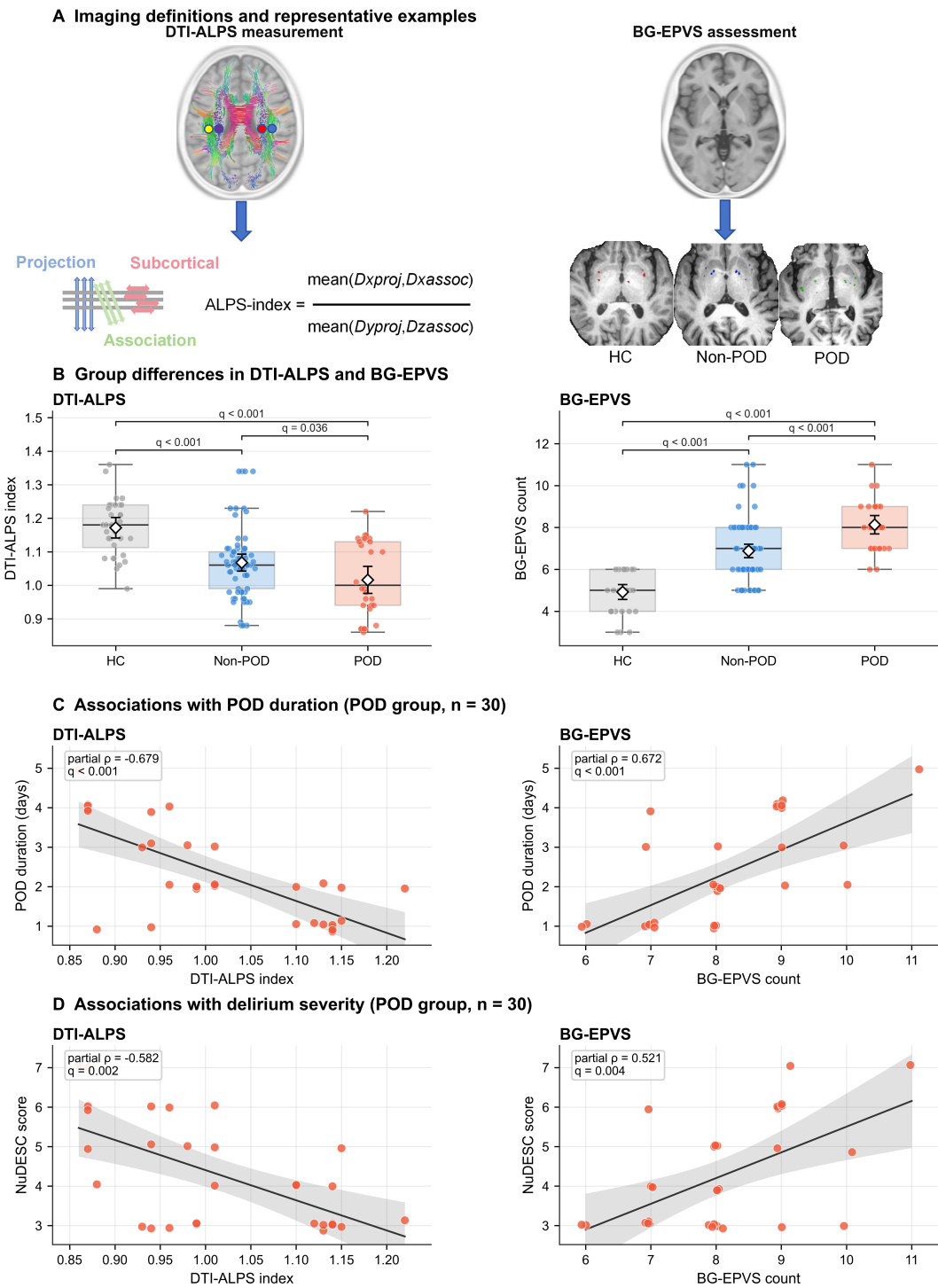


Figure 2

Perivascular imaging measures and delirium phenotype.

A, DTI-ALPS region placement and calculation, with representative images illustrating BG-EPVS assessment. B, DTI-ALPS and BG-EPVS distributions in healthy controls (n=32), non-POD patients (n=70), and POD patients (n=30). Points show participants, boxes show medians and interquartile ranges, and diamonds with error bars indicate age- and sex-adjusted means and 95% confidence intervals. C and D, Associations with POD duration and peak Nu-DESC score among patients with POD (n=30). Annotations report age- and sex-adjusted partial Spearman correlations. q denotes FDR-adjusted P.

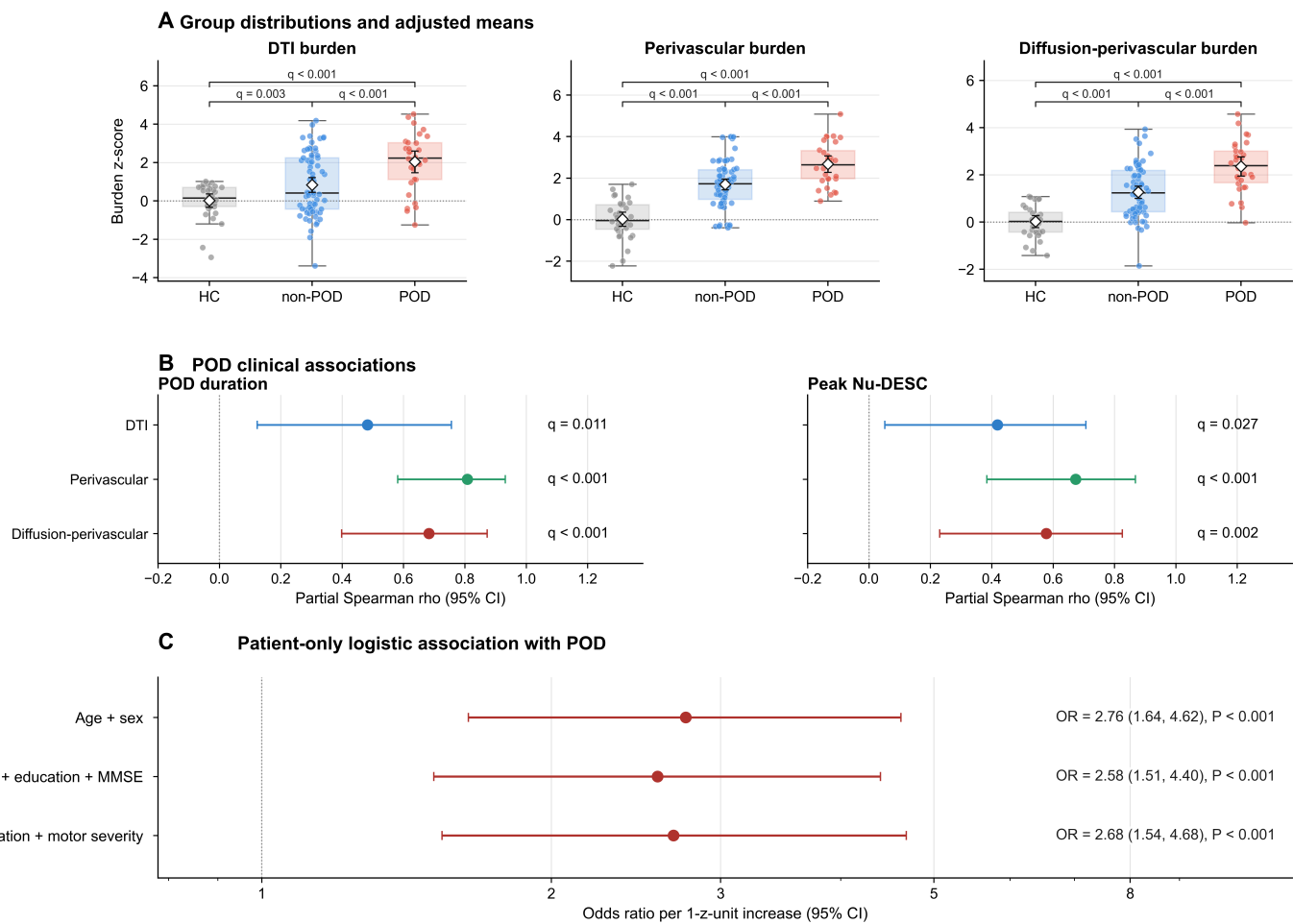


Figure 3

Combined diffusion-perivascular burden and clinical associations.

A, Participant distributions and age- and sex-adjusted means of DTI, perivascular, and combined burdens in healthy controls (n=32), non-POD patients (n=70), and POD patients (n=30). Boxes show medians and interquartile ranges; points represent participants. Diamonds and error bars show adjusted means and 95% confidence intervals. B, Age- and sex-adjusted partial Spearman correlations with POD duration and peak Nu-DESC score among patients with POD (n=30), with bootstrap 95% confidence intervals. C, Firth

penalised logistic regressionanalyses of combined burden and POD under increasing clinical adjustment.

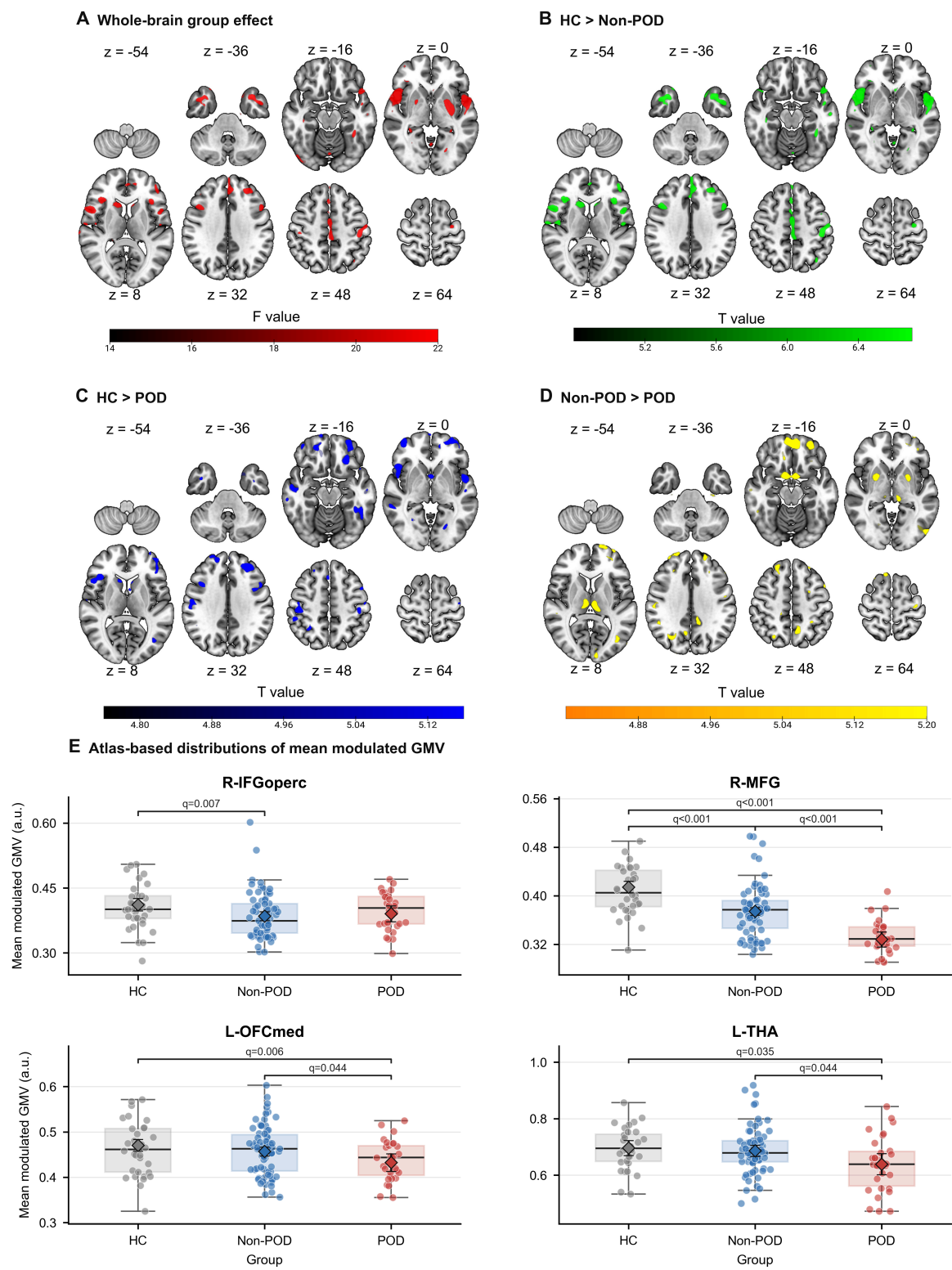


Figure 4

Exploratory whole-brain morphometry and atlas-based summaries.

A, Omnibus group F contrast. B–D, Directional contrasts for healthy controls > non-POD, healthy controls > POD, and non-POD > POD. Voxel-wise models include age, sex, and TIV; each map is separately thresholded at whole-brain voxel-level FWE-corrected $P < 0.05$. Colour bars show F or T statistics. E, Mean modulated GMV within the right inferior frontal pars opercularis, right middle frontal gyrus, left medial orbitofrontal cortex, and left thalamus. Points show participants, boxes show medians and interquartile ranges, and diamonds with error bars indicate adjusted means and 95% confidence intervals.

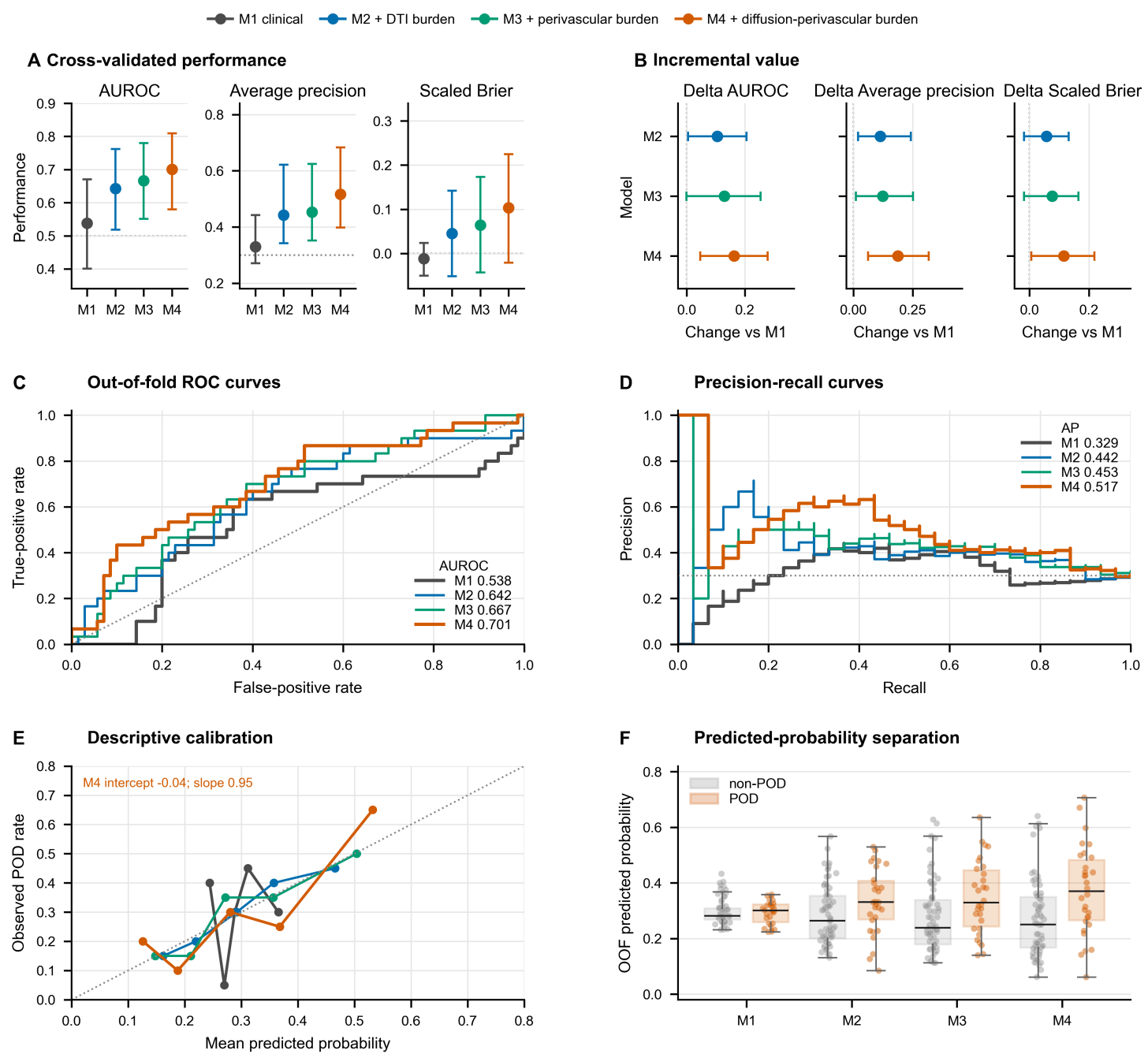


Figure 5

Exploratory internal prediction of POD.

A, Internally evaluated AUROC, average precision, and scaled Brier score, with resampling intervals. B, Changes relative to M1. C and D, ROC and precision–recall curves from aggregated out-of-fold predictions. The precision–recall reference line is the observed POD proportion of 0.30. E, Descriptive calibration curves; the diagonal marks agreement between predicted probabilities and observed frequencies. F, Distribution of out-of-fold predicted probabilities by observed POD status. M1 contains age, sex, disease duration, MMSE, and motor severity. M2, M3, and M4 each extend M1 by adding DTI burden, perivascular burden, or combined burden, respectively.

Supplementary Files

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

- [SupplementaryMethods.docx](#)
- [Supplementaryfigureandtable.docx](#)