Tract-based Lesion Metrics for Resolving the Clinico-Radiological Paradox in Multiple Sclerosis

Osama Abdullah[1,2], Puti Wen*[2,3], Abdelmalek W. Abdelrazeq*[,2], Mariya Matrosova*[5], Lev Brylev[4], Vasiliy Bryukhov[5], David Melcher[2,3], Bas Rokers[2,3]*

Author affiliations:

1 Core Technology Platforms, New York University Abu Dhabi, Abu Dhabi, UAE

2 Center for Brain and Health, New York University Abu Dhabi, Abu Dhabi, UAE

3 Department of Psychology, New York University Abu Dhabi, Abu Dhabi, UAE

4 Yas Clinic Khalifa City, Abu Dhabi, UAE

5 Russian Center of Neurology and Neurosciences, Moscow, Russia

* denotes equal contribution

Abstract

The clinico-radiological paradox in multiple sclerosis (MS) arises because conventional MRI measures, particularly total lesion volume, fail to capture the true burden of disability. These broad volumetric measures overlook the dual influence of where lesions occur and what is their microstructural composition. This study aimed to improve the clinical relevance of lesion analysis in MS by refining both how and where lesions are measured. While traditional approaches emphasize anatomical lesion burden, studies of structural disconnection in neurological disease have shown that functional deficits often arise from tract-specific interruptions within distributed brain networks. We hypothesized that filtering lesions through clusters of functionally meaningful white matter tracts, rather than broad anatomical compartments, would enable more accurate identification of clinically relevant damage. To test our hypothesis, we studied 132 participants, including 89 patients with MS (49

relapsing–remitting, 19 primary progressive, 21 secondary progressive; 61 women and 28 men; mean age 41.7 years, range 21–67) and 43 healthy controls (27 women and 16 men; mean age 40.2 years, range 23–65). All underwent standardized 3 Tesla MRI including fluid-attenuated inversion recovery (FLAIR), T1 mapping, magnetization transfer ratio (MTR), and diffusion tensor imaging (DTI) with fractional anisotropy (FA) and mean diffusivity (MD). Lesions were segmented from FLAIR, mapped to both classical anatomical regions and individualized white matter tracts, then enriched with quantitative MRI metrics.

Associations between imaging measures and disability were evaluated using Spearman or Pearson correlations, and predictive performance was assessed with ridge-penalized regression across Expanded Disability Status Scale (EDSS), motor and cognitive scores, as well as progression risk defined by the MSPro scale. Filtering lesions through functionally critical tracts revealed associations obscured by conventional approaches. Tract-based lesion T1 values correlated more strongly with disability than classical regional metrics (e.g. EDSS $r = 0.82$ in occipitoparietal tracts vs $r = 0.59$ in periventricular regions). Multimodal tract-specific metrics including T1, MTR, FA, and MD significantly predicted disability (odds ratios 2.2–7.7), with explanatory power up to $R^2 = 0.65$. Importantly, tract-based T1 stratified patients with relapsing–remitting MS into stable versus early-progressive groups with 92% accuracy (area under the curve = 0.90).

By integrating lesion location and microstructural composition, tract-based quantitative MRI resolves key aspects of the clinico-radiological paradox. This framework enhances disability prediction, enables early identification of patients at risk of progression, and provides a clinically interpretable biomarker to support individualized monitoring and therapeutic decision-making.

1. Introduction

Multiple sclerosis (MS) is a chronic, immune-mediated disease of the central nervous system characterized by inflammatory demyelination and neurodegeneration, most prominently in white matter, with post-mortem studies showing that lesions vary widely in demyelination, axonal loss, gliosis, and microglial activity depending on stage, location, and patient-specific factors [1–3]. This heterogeneity is mirrored clinically in the diverse disability profiles seen in patients, ranging

from motor and sensory impairment to visual dysfunction and cognitive decline, raising the challenge of how lesion characteristics can be meaningfully linked to functional outcomes[4–8]. Conventional MRI, particularly T2 and FLAIR, guides diagnosis and monitoring but depicts lesions as uniform bright spots in broad anatomical regions, so lesion counts and volumes often miss true disease burden. This shortcoming gives rise to the clinico-radiological paradox, where clinical disability and disease course often do not align with visible lesion load on routine anatomical MRI, underscoring the need for more advanced imaging approaches that can bridge this disconnect [9–12].

Quantitative MRI (qMRI) offers promising complementary measures that augment conventional MRI by probing tissue microstructure beyond visible lesions. Metrics such as magnetization transfer ratio (MTR) and T1 relaxation times are sensitive to myelin and axonal content[13,14]. Diffusion tensor imaging (DTI) probes tissue microstructure by characterizing water diffusion. Scalar metrics such as reduced fractional anisotropy (FA) and increased mean (MD) or radial diffusivity (RD) indicate axonal disorganization and demyelination, even in normal-appearing white matter[15]. Building on these principles, several studies have applied qMRI to characterize the pathological burden of MS. MTR and T1 metrics have been used to quantify demyelination within lesions, successfully stratifying them by disease severity and stage[16,17]. A particularly compelling finding came from a longitudinal study showing that reductions in MTR could be detected in normally-appearing white matter (NAWM) up to four months before new lesions became visible on conventional MRI [18]. In addition, by reconstructing white matter tracts from diffusion imaging, studies have moved beyond gross anatomical compartments, providing a more precise mapping of disease impact and revealing stronger correlations with clinical disability than region-based approaches[19–29]. Together, these findings highlight the potential of qMRI and tract-based approaches to capture clinically meaningful pathology that conventional lesion measures overlook.

Parallel to imaging advances, clinical outcomes are now assessed with multiple domain-specific scales: the Expanded Disability Status Scale (EDSS, motor deficit)[30], Symbol Digit Modalities Test (SDMT, processing speed)[31], Timed 25-Foot Walk (T25-FW, gait)[32], Nine-Hole Peg Test (9-HPT, manual dexterity)[33], and MSPro scale (progression risk)[34]. While these developments enrich both imaging and disability assessment, most previous studies relied on simple univariate

correlations, often using a single clinical score. Moreover, previous studies frequently used gross anatomical compartments to explain disability (for example, periventricular lesion-load; $r = 0.331$, $p = 0.005$)[35], an approach that, although informative, lacks functional specificity and reinforces oversimplified interpretations of "global disability." This reliance on univariate and regionally crude analyses introduces confounds, observer bias, and noise, thereby perpetuating the paradox.

In this study, we propose a tract-aware, multimodal framework that integrates qMRI metrics with the functional anatomy of white matter tracts. By quantifying both lesion location (*where*) and microstructural composition (*what*), and correlating these measures with multiple disability scores across progressively structured statistical models, we aim to establish stronger and more specific structure-function relationships. We hypothesize that this comprehensive approach will better account for lesion heterogeneity and the multidimensional nature of disability, enabling improved patient stratification and more accurate distinction of MS phenotypes compared to conventional lesion volume measures

2. Materials and Methods

2.1 Subject Selection, Clinical Assessment, and Data Management

A total of 132 participants— 89 MS patients and 43 healthy controls—were recruited between June 2022 and June 2024 at the Research Center of Neurology in Moscow, Russia. Consent was obtained according to the Declaration of Helsinki and that it has been approved by the ethical committee.

MS patients were clinically classified to relapsing remitting (RRMS), (primary progressive) PPMS, or secondary progressive (SPMS) subtypes based on established diagnostic criteria, with inclusion limited to individuals aged 18–70. Control subjects, aged 18–75, were included if they had no cognitive or neurological disorders, no severe somatic illness, and no structural brain abnormalities on MRI beyond age-related nonspecific white matter hyperintensities. General exclusion criteria for all participants included MRI contraindications (e.g., pacemaker), pregnancy, recent corticosteroid therapy (within 30 days for MS patients), structural brain pathology, or corrupted MRI data. All subjects provided written informed consent prior to

participation. Clinical disability was assessed using the EDSS, SDMT, T25FW, 9HPT (dominant and non-dominant hand), and MSPRO. Tables 1–2, Fig 1, and Fig 2 summarize the full demographic and clinical profiles.

Table 1 *Summary of demographic and clinical metrics across MS subtypes and healthy controls, reporting group-wise means and standard deviations. It highlights differences in age, disease duration, disability scores, and cognitive performance between subtypes. DMT: Disease modifying treatment*

Parameter	Healthy Controls	RRMS	PPMS	SPMS
N (% Female)	43 (65%)	49 (76%)	17 (59%)	23 (65%)
Age	35.60 ± 10.83	35.65 ± 10.25	47.88 ± 7.01	48.87 ± 12.50
Disease duration (years)	-	9.17 ± 7.48	7.53 ± 5.26	17.22 ± 8.37
EDSS	-	1.81 ± 0.95	3.94 ± 1.47	5.17 ± 1.14
DMT: Yes, n (%)	-	5 (12%)	9 (52,9%)	6 (26,1%)
Therapy: 1st line	-	29 (65.9%)	2 (22.2%)	11 (64.7%)
Therapy: 2nd line	-	15 (34.1%)	7 (77.8%)	6 (35.3%)
T25FW (s)	5.63 ± 0.52	6.23 ± 1.01	9.76 ± 3.57	13.95 ± 9.94
9HPT-D (s)	18.47 ± 2.31	19.81 ± 2.65	24.12 ± 4.51	28.61 ± 9.37
9HPT-ND (s)	20.00 ± 2.87	21.56 ± 3.89	27.39 ± 5.58	35.39 ± 14.16
SDMT Total	61.88 ± 10.24	56.49 ± 11.03	46.88 ± 8.78	38.43 ± 8.07
SDMT Mean	61.11 ± 10.37	55.88 ± 11.34	46.29 ± 9.02	37.50 ± 8.14
MSPro	-	1.22 ± 0.59	2.53 ± 0.62	2.55 ± 0.74

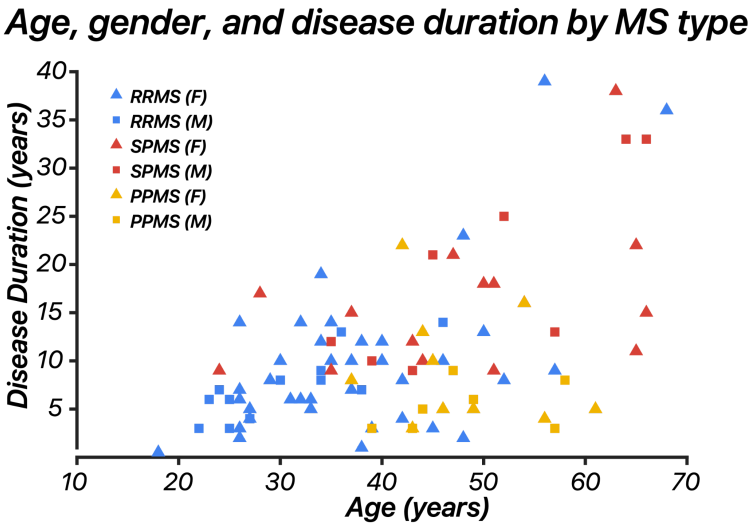


Figure 1 *Summarizes demographics (age, gender, and disease duration) across MS subtypes. For each subject age (x-axis), gender (male: square, female: triangle), disease duration in years (y-axis), and MS subtype (RRMS: blue, PPMS: yellow, SPMS: red) are represented.*

Table 2 *Comparative overview of the six disability assessment tools employed in this study: EDSS, SDMT, T25FW, 9HPT (dominant and non-dominant hands), and MSPro. Each scale is characterized along three key dimensions: (1) the functional domains it is designed to assess, (2) a concise description of the task(s) involved in administering the test, and (3) the specific scoring system used to quantify patient performance or impairment.*

Scale	Expanded Disability Status Scale (EDSS)	Symbol Digit Modalities Test (SDMT)	Timed 25-Foot Walk (T25FW)	Nine-Hole Peg Test (9HPT)	MSPRO
Functional Domains	**Pyramidal** (muscle weakness or difficulty moving limbs) **Cerebellar** (coordination, tremor, and balance) **Brainstem** (speech, swallowing, and eye movement) **Sensory** (numbness or loss of sensations) **Bowel** and bladder function **Visual** function **Cerebral** or mental function (cognition and mood) **Ambulation** (walking distance and use of aids)	**Visual scanning** and visual-spatial processing **Attention** and concentration **Working memory** and incidental learning **Psychomotor** speed **Visual-motor coordination** **Executive** functioning	**Ambulation** and walking speed **Lower limb** strength and coordination **Balance** and motor control Mobility-related **functional independence**	**Fine motor** coordination **Finger dexterity** **Visual-motor** integration **Bilateral upper limb** function (dominant and non-dominant hands) **Independence** in activities of daily living (ADLs)	**Disease activity** (relapse frequency, recovery quality) **Symptoms**: Motor, visual, coordination/balance, speech, sensory, bladder/bowel **Cognitive** and fatigue-related functions Impact on daily living activities and independence: **Mobility**, self-care, occupational functioning, hobbies, and social engagement
Type of Tasks	**Motor tasks**: Evaluating muscle strength and coordination **Sensory tests**: Assessing response to touch, pain, and vibration **Visual assessments**: Testing visual acuity and field **Cognitive evaluations**: Observing mental status and mood **Gait analysis**: Measuring walking distance and stability	In **writing** (written format), or **Orally** (oral format), depending on administration choice. The written form primarily assesses motor and visual processing, while the oral version helps distinguish whether deficits are motor-based or cognitive.	The individual is instructed to walk 25 feet as quickly and safely as possible. The time taken to complete the walk is recorded. The test is typically performed twice, and the average time is calculated.	Pick up nine pegs one at a time and place them into nine holes on a pegboard as quickly as possible. Once all pegs are placed, remove them one at a time and return them to the container. Each hand is tested separately, typically with two trials per hand. The task is completed under standardized conditions, with instructions to move "as quickly and accurately as possible."	Symptom History (last 6 months) Impact Assessment impact severity on areas of functioning.
Scoring System	0.0–3.5: No to moderate disability but is fully ambulatory 4.0–5.5: Increasing difficulty walking but still ambulatory without aid for 300+ meters 6.0–7.5: Ambulation severely restricted; requires assistance (e.g., cane, walker, or wheelchair) 8.0–9.5: Essentially restricted to bed or chair; retains some self-care functions 10.0: Death due to MS	The raw score is the number of correct substitutions completed within 90 seconds. Standardized scores are not consistently available in the test manual; instead, performance is typically interpreted based on standard deviations (SD) from the mean	Healthy Controls: Median T25FW time is approximately 3.7 seconds (range: 2.8–5.2 seconds) 20% or greater increase in T25FW time is considered a clinically meaningful change, indicating potential disease progression	Normative Data: Norms vary based on age and sex. Cut-off values: >18 seconds is generally considered abnormal for the dominant hand. <0.5 pegs/s may indicate high risk for activity limitations. 33.3 seconds (corresponding to ~0.27 pegs/s) separates mild from more severe impairment	Section A: Relapse and Recovery (max 14 points) Section B: Symptoms (max 91 points) Section C: Impacts (max 14 points) Total Score = Sum of all above sections Interpretation Output: 1: (<46.3): Unlikely progression 2: (46.3–57.8): Possible progression 3: (>57.8): Likely progression

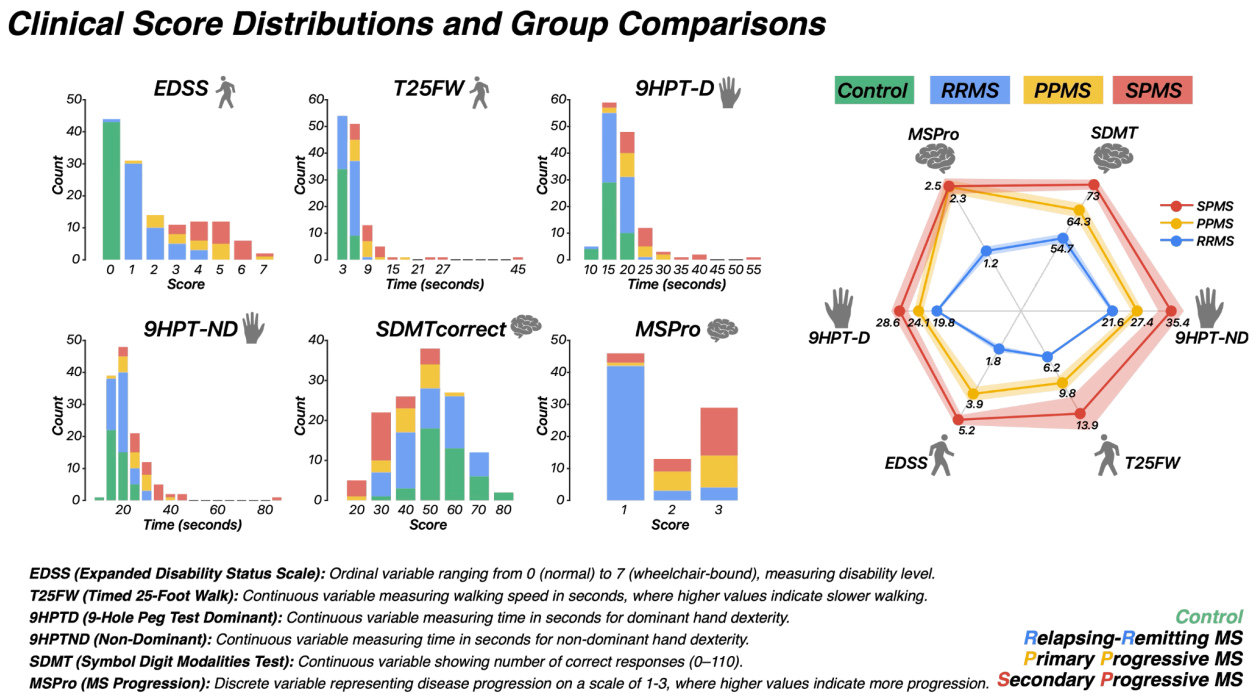


Figure 2 Represents the cohort's clinical score distribution subtypes comparisons. *The radar plot (right) displays the mean and standard deviation for key clinical scores - EDSS, SDMT, 9HPTD, 9HPTND, T25FW, and MSPRO--across RRMS, SPMS, and PPMS groups. Clear separation is observed between subtypes, particularly between RRMS and SPMS, reflecting increasing disability with disease severity. Bar plots (left) represent each individual clinical disability score and their accompanying numerical scales. Controls were included when appropriate.*

2.2 MRI Acquisition

A multimodal 3T MRI protocol was acquired on a Siemens Magnetom Prisma (Syngo MR VE11) at the Research Center of Neurology in Moscow, using a 64-channel Head/Neck coil. FLAIR imaging was acquired using a 3D T2-weighted SPACE sequence with the following parameters: TR = 7000 ms, TE = 390 ms, TI = 2200 ms, flip angle = 120°, voxel size = 0.47 × 0.47 × 0.60 mm³, matrix size = 512 × 512, FOV = 240 × 240 mm². T1-mapping was performed using a 3D MP2RAGE sequence with TR = 5000 ms, TE = 2.98 ms, TI1 = 700 ms, TI2 = 2500 ms, flip angles = 4° and 5° for the first and second inversions respectively, and isotropic voxel size = 1.00 × 1.00 × 1.00 mm³, matrix size = 256 × 240, FOV = 256 × 240 mm². Magnetization transfer imaging was acquired using a 3D gradient-echo sequence with TR = 34 ms, TE = 6 ms, flip angle = 10°, voxel size = 1.00 × 1.00 × 3.00 mm³, matrix size = 256 × 192, FOV = 256 × 192 mm². Both MT-on and MT-off images were collected for MTR calculation. Diffusion-weighted imaging was performed using a single-shot echo-planar imaging sequence

with TR = 5600 ms, TE = 82 ms, flip angle = 90°, matrix size = 122 × 122, multiband acceleration factor = 2. A multi-shell diffusion scheme was employed with b-values of 0, 1000, and 2500 s/mm², using 64 diffusion directions per shell. The total acquisition time was approximately one hour.

2.3 Image Preprocessing

Anatomical segmentation utilized the UNI image from the MP2RAGE acquisition after noise removal with the PreSurfer toolkit (https://github.com/srikash/presurfer). Cortical and subcortical structures were reconstructed using FreeSurfer (v7.4.1, recon-all)[36], and FLAIR images were registered to the structural space. Lesions were segmented with a custom-trained U-Net implemented in MONAI[37] and manually corrected in FSLeyes[38] by a physician (AA), with verification by a trained neurologist (LB). Quantitative parameter mapping was then performed using qMRLab[39]. T1 relaxation maps were estimated from the MP2RAGE sequence with the mp2rage model (TI1/TI2 = 700/2500 ms, flip angles = 4°/5°, TR = 5000 ms). Magnetization transfer ratio (MTR) maps were computed from MT-on and MT-off images using the formula: $MTR = [(S_0 - S_{MT})/S_0] \times 100\%$, where S_0 and S_{MT} represent signal without and with MT saturation, respectively. Diffusion preprocessing was carried out with TractoFlow (Nextflow pipeline) in the Scilus 1.5.0 Singularity container[40], including susceptibility distortion correction (TOPUP) and eddy current correction (EDDY) in FSL. DTI reconstruction yielded FA and MD maps. Local deterministic tractography was performed in DSI Studio with FA-based seeding (threshold = 0.1). Preprocessed DWI, b-values, and b-vectors were converted into diffusion source (.src) files, followed by generalized Q-sampling imaging (GQI; sampling length ratio = 1.25) aligned to the T1 space. Using the resulting fiber orientation distribution (.fib) files, 27 canonical white matter tracts were automatically reconstructed with DSI Studio's ICBM-152 template-based deterministic tracking[41]. All imaging metrics (FLAIR, T1 map, MTR, FA, MD) and tract masks were co-registered to the individual T1 structural space for consistency in subsequent analyses (Fig. 3).

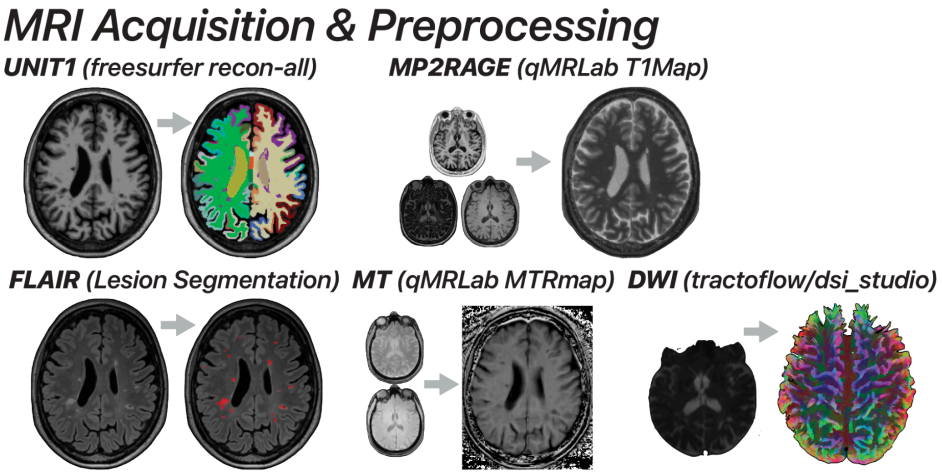


Figure 2 MRI acquisition and preprocessing. *UNI images from MP2RAGE were used for anatomical segmentation in FreeSurfer. FLAIR scans provided lesion masks generated by a U-Net model with manual correction. Quantitative maps included T1 relaxation and magnetization transfer ratio, and diffusion processing enabled tractography of major white-matter pathways.*

2.4 Lesion Load and Quantitative Metrics

Lesion quantification was performed using two complementary frameworks: a classical anatomical definition, familiar to neurologists (periventricular, juxtacortical, infratentorial, and deep white matter), and a tract-based definition grounded in functional white matter pathways. In the classical framework, lesion masks were first intersected with the four canonical anatomical compartments, and lesion load was recorded as (1) lesion number (LN, count of discrete lesions), (2) lesion volume (LV, cumulative lesion volume in mm^3), and (3) normalized lesion volume (Lnorm, LV expressed as a percentage of total intracranial volume). In the tract-based frameowork, 27 subject-specific white matter fiber pathways which were collapsed into four functional clusters (association, projection/brainstem, occipitoparietal, and cerebellar pathways). For each cluster, lesion load was likewise defined as LN, LV, and Lnorm (LV expressed as a percentage of tract volume). Specific tract assignments are listed in Table 3 with representative renderings in Fig. 4. This dual-scale (regional vs tract-based) framework enabled direct testing of whether pathway-specific lesion mapping offers greater clinical insight than conventional anatomical compartmentalization. Note that although MS lesions are often prominent in periventricular regions, callosal fibers (except Forceps Major) were excluded to avoid

double-counting in overlapping regions, despite their clear clinical relevance in periventricular MS.

For each participant, mean T1, MTR, FA, and MD values were calculated within lesions falling inside (1) the four classical compartments and (2) the four tract-based clusters. Because tract clusters form clearly defined fiber bundles, the same metrics were also sampled from NAWM by excluding lesion voxels from tract masks, yielding an internal reference free of visible pathology on 3D FLAIR. This generated, for each subject, a matrix of eight lesion-level measurements (four metrics × two location schemes) and a set of tract-specific NAWM values. NAWM sampling was not attempted for the classical compartments, whose diffuse anatomical borders preclude a reliable lesion-free reference. The combined lesion-based and tract-NAWM metrics constituted the lesion-composition feature set supplied to the statistical modeling stage

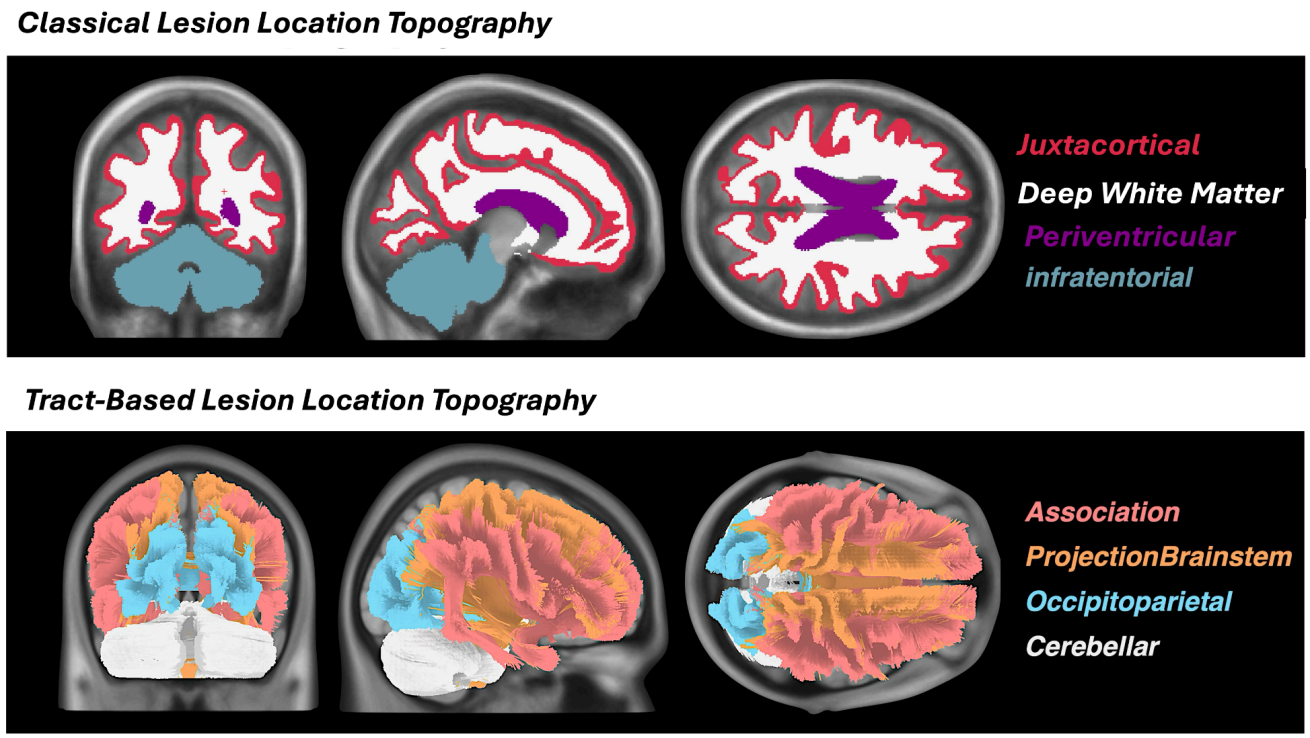


Figure 4 *Summarizes the classical lesion location and the tract based locations as they are annotated anatomically in the brain.*

Table 3 *Categorizes major white matter tracts into functional groups, including occipitoparietal, association, projection, brainstem, cerebellar, and commissural systems. Each group lists specific tracts (e.g., corticospinal tract under projection, corpus callosum under occipitoparietal) that were likely analyzed for lesion burden or structural integrity in the study.*

Tracts			
Occipitoparietal	**Association**	**Projection Brainstem**	**Cerebellar**
Projection Basal Ganglia: Optic Radiation	Arcuate Fasciculus	Corticospinal Tract	Inferior Cerebellar Peduncle
Vertical Occipital Fasciculus	Cingulum	Corticobulbar Tract	Vermis
Corpus Callosum: Forceps Major	Extreme Capsule	Corticopontine Tract	Middle Cerebellar Peduncle
	Frontal Aslant Tract	Medial Lemniscus	Superior Cerebellar Peduncle
	Hippocampus Alveus	Dentatorubro-thalmic Tract	
	Inferior Fronto Occipital Fasciculus	Non-Decussating Dentatorubro-thalamic Tract	
	Inferior Longitudinal Fasciculus	Reticular Tract	
	Middle Longitudinal Fasciculus	Medial Forebrain Bundle	
	Parietal Aslant Tract		
	Superior Longitudinal Fasciculus		
	Uncinate Fasciculus		
	Vertical Occipital Fasciculus		

2.5 Statistical Analysis

2.5.1 Group-wise Metric Comparison

We first quantified lesion location by extracting lesion load measures from both classical anatomical regions and tract-based clusters. For classical regions (periventricular, juxtacortical, infratentorial, and deep white matter) LN, LV, and Lnorm were calculated. In parallel, whole-brain tractography defined four bilateral tract-clusters, and the same lesion load metrics (LN, LV, Lnorm) were computed for each cluster, with left–right hemispheric measures retained when applicable.

Lesion composition was then assessed by qMRI metrics. Signal intensity on T1-maps, FA, MD, and MTR were extracted from lesions located in both classical regions and tract-based clusters. In addition, the same measures were obtained from NAWM for comparison across subtypes. For

both lesion location and lesion composition (including NAWM), group-wise mean values and standard errors of the mean (SEM) were calculated across the MS subtypes. No hypothesis testing was performed at this stage; the objective was exploratory profiling of subtype separation, visualized using polar (radar) plots.

2.5.2 Univariate Correlation Analysis with Clinical Outcomes

Correlations were obtained between the various lesion metrics and clinical disability scores (EDSS, T25FW, 9HPTD, 9HPTND). For ordinal clinical variables (EDSS), Spearman's rank correlation was used, while Pearson correlation was applied to continuous variables. To control for potential confounding effects, partial correlations were calculated using patient sex, age, and disease duration as covariates. For each pairwise combination of a lesion metric and clinical score, the correlation coefficient (r) and the corresponding p-values are reported.

2.5.3 Multivariate Predictive Modeling

The objective of multivariate predictive modeling was to evaluate the association between MRI metrics and clinical disability outcomes in MS, and to compare the predictive performance of the Tract-Based and the Classical Model. Clinical outcomes (EDSS, T25FW, 9HPTD, 9HPTND) were first transformed into three ordinal categories (Low, Medium, High). Thresholds were selected to achieve approximately equal group sizes. This transformation was motivated by several considerations (a) avoiding a restrictive linearity assumption and (b) focusing on functional rank ordering rather than arbitrary numeric distances. We implemented a two-stage modeling approach. Ridge regularization was employed in the first stage to address the potential multicollinearity among imaging features, to prevent overfitting, and to reduce the bias toward models with more predictors. In the first stage, imaging features were reduced to linear predictors using ridge-penalized regression (*lassoglm* with $\alpha = 0.05$, 10-fold cross-validation to select λ minimizing deviance). For each metric (Lnorm, T1, MTR, FA, MD), this yielded two linear predictors. The tract-based model combined 16 regressors from four bilateral tract groups (Association, Cerebellar, Occipitoparietal, Projection Brainstem) across lesion and NAWM tissue types. The classical model utilized 4 regressors from classic classical regions (periventricular, juxtacortical, infratentorial, deep white matter) with metrics extracted from within lesions only. In the second stage, we fitted a combined ordinal logistic regression model incorporating both linear predictors plus demographics (age, sex, and disease duration). In the combined model,

significance for either the tract-based or classical approach indicated that it captured disability-relevant variance not explained by the competing model. We reported standardized odds ratios (per 1-SD increase) with 95% confidence intervals, along with corresponding p-values for both approaches in the combined model. Significant demographic confounders (age, sex, and disease duration) were also identified.

To demonstrate a potential practical use case of our model, we used tract-based T1 mapping to predict progression risk as determined by the clinical measure MSPro. The intended clinical application is to identify at-risk RR patients prior to transition to the progressive stage. MSPro was incorporated to capture patient-reported disability and to assess the risk of conversion from relapsing to progressive MS subtypes. MSPro score was dichotomized (MSPro = 1, clinically stable; MSPro = 2 - 3, progression). After excluding controls and participants with missing data, the final sample included 88 patients (46 stable, 42 progressive). The imaging predictors comprised 16 mean T1 intensities, calculated separately for lesions and NAWM across four bilateral tract systems (association, cerebellar, occipitoparietal, projection brainstem). Age, sex, and disease duration were included as additional covariates. Classification relied on ridge-penalized logistic regression implemented in MATLAB R2024b (*fitclinear*, Learner = "*logistic*", $\alpha = 0.05$). The regularization constant λ was re-estimated within each training set by nested ten-fold cross-validation. Generalization was assessed with leave-one-out cross-validation: for every held-outpatient the model was refit on the remaining N – 1 cases, λ was re-selected, and the posterior probability of progression was stored. Overall accuracy, sensitivity, specificity, and the area under the receiver-operating-characteristic curve (AUC) were computed.

3. Results

3.1 Classical versus tract-based lesion load: Group-wise Comparison of Lesion Frameworks

3.1.1 Lesion Location

The overall lesion load in the Classical method represents the total lesion volume derived from whole-brain 3D FLAIR. Overall lesion load was lowest in RRMS, increased in PPMS, and more than doubled in SPMS as shown in Fig 4a. Periventricular lesions accounted for the greatest share of this volume, followed by juxtacortical and deep-white-matter compartments, with infratentorial regions contributing the least. The tract-based metric, which includes only lesions intersecting predefined white-matter fiber bundles (Table 3), captured 22% of the total lesion volume in RRMS, 37% in PPMS, and 27% in SPMS. Within these restricted tract-based masks, lesion burden rose progressively from RRMS through PPMS to SPMS, reproducing the phenotype-dependent escalation observed with the classical measure.

Tract-based lesion loads—especially the Lnorm—separate the three MS subtypes far more cleanly than classical region-based metrics, while simultaneously highlighting the functional pathways most relevant to disease evolution (radar plots in Fig 4b). Across all three tract-level measures—raw LN, LV and Lnorm—the distributions for RRMS, SPMS and PPMS are minimally or non-overlapping, mirroring the known escalation of disease burden. Lnorm, which adjusts for tract size and hemispheric side, produces the widest inter-quartile gaps. On the other hand, classical region-based metrics show inconsistent patterns, for instance the periventricular band, classical LN provides almost no separation between MS subtypes, whereas periventricular LV rises sharply, exaggerating group differences. The opposite pattern emerges in juxtacortical and deep-white-matter regions, where LN inflates differences and LV compresses them. This inconsistency underscores the sensitivity of classical metrics to region size and lesion confluence.

In the analyzed tracts, lesion load was highest in association fibers, followed by projection/brain-stem tracts, occipito-parietal bundles and cerebellar pathways. Because these tracts underpin vision, cognition, sensorimotor control and balance, their progressive

involvement aligns with the stepwise clinical worsening observed across disability scores in Fig 1—information lost when lesions are pooled into broad anatomical zones.

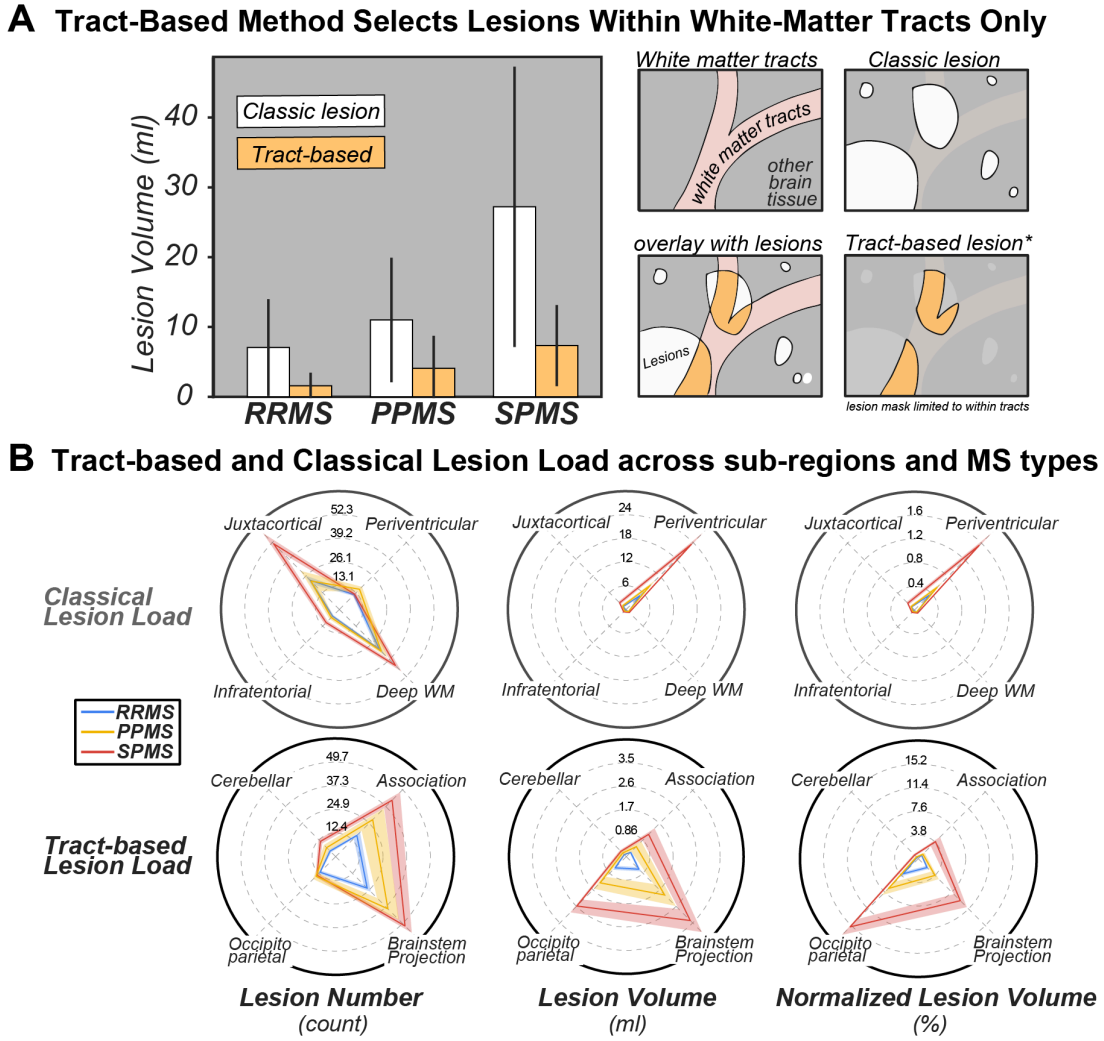


Figure 5 (a) *Left panel: Box plots compare absolute lesion volumes (in mL) across MS subtypes using classical (white) versus tract-based (yellow) definitions. Error bars represent the standard deviation of the lesion volumes across subjects within each MS subtype group. Classical lesion volume reflects total lesion burden irrespective of anatomical specificity, treating each lesion as a unified mask. In contrast, tract-based lesion volume quantifies only the lesion portions intersecting with defined white-matter tracts, thereby filtering out non-tract areas. Right panel: Schematic illustration of the filtering mechanism used in tract-based lesion localization. Only voxels within white matter tracts or tract intersections are included in the tract-based volume, whereas the classical approach includes entire lesions regardless of anatomical context.* ***(b)*** *Demonstrates radar plots of lesion location metrics in MS subtypes groups (SPMS: red, PPMS: yellow, RRMS: blue). The figure shows a comparison of each metric in classical lesion location versus tract-based lesion location. The right-column: Lesion Number, middle-column: Lesion Volume, and left-column: Normalized Lesion Volumes.*

3.1.2 Lesion Composition

Quantitative assessment of lesion composition using tract-based regions provided greater separation between MS subtypes than classical anatomical lesion mapping. As shown in Fig 5, T1, MTR, MD, and FA values measured in lesions located within classical regions showed minimal differences between RRMS, PPMS, and SPMS groups. In contrast, tract-based lesion mapping showed subtype-specific trends in several white matter systems. The occipitoparietal tract demonstrated the most consistent separation across all tissue metrics when comparing MS subtypes. From RRMS to PPMS, we observed a 10.0% increase in T1, a 5.4% decrease in MTR, a 7.8% increase in MD, and a 1.2% decrease in FA. Progressing further from PPMS to SPMS, the same tract showed a 7.7% increase in T1, an additional 3.9% decrease in MTR, a 6.0% increase in MD, and a more pronounced 9.0% decrease in FA. These directional changes across all four metrics reflect cumulative microstructural damage and provide a coherent tissue signature of progression within the occipitoparietal system. Other tract clusters demonstrated similar directionality but with varying magnitudes. In the cerebellar tract, FA and MTR values were greater in SPMS than PPMS, deviating from the trend seen in other regions.

Tract-based NAWM analysis showed a monotonic pattern across subtypes for all four metrics. T1 and MD values in occipitoparietal NAWM increased progressively from control to RRMS, PPMS, and SPMS. MTR and FA values showed a corresponding decrease across these groups. These directional changes were observed in most tract clusters, including the projection-brainstem and association tracts. In contrast, cerebellar NAWM again deviated from this pattern in FA, with SPMS showing greater values than controls.

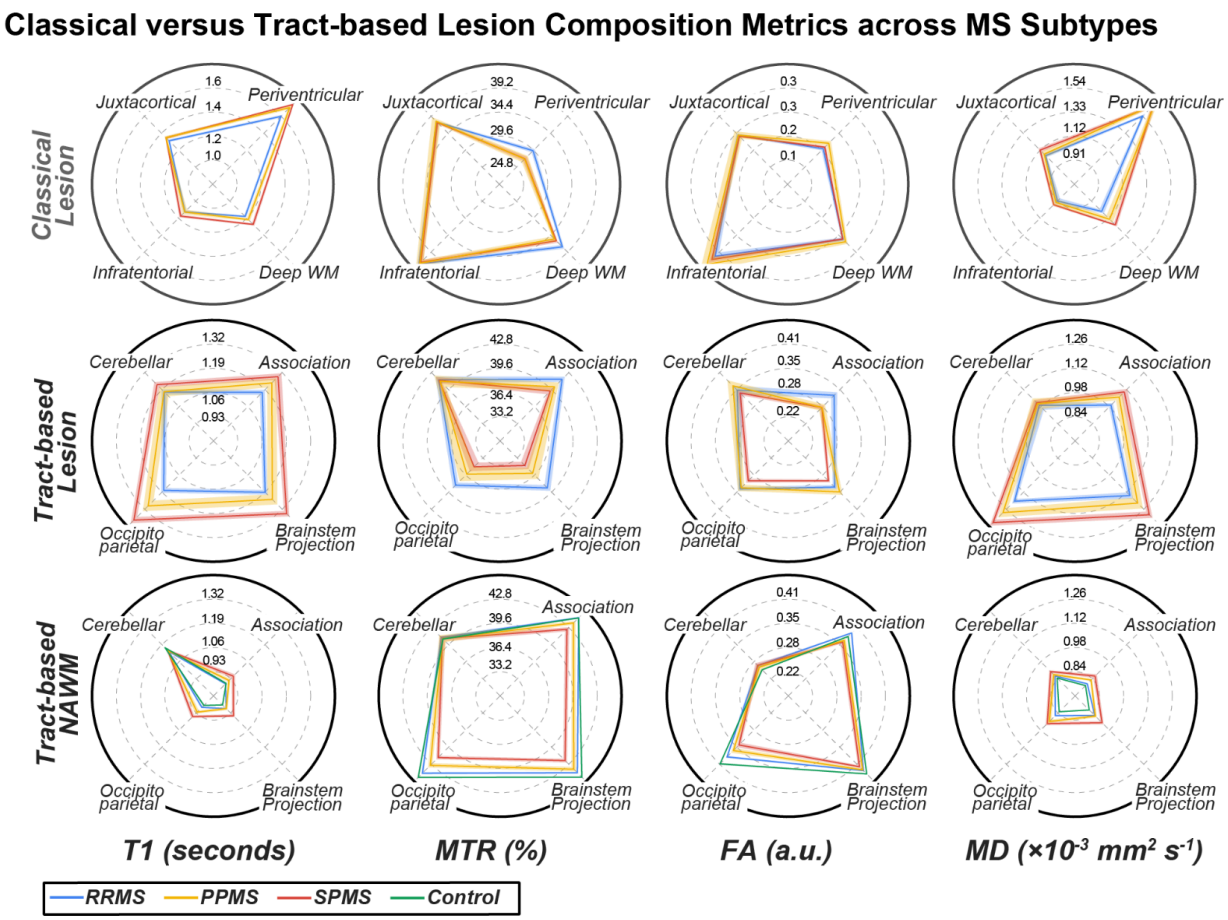


Figure 6 *Demonstrates radar plots of lesion composition metrics in MS subtypes groups (SPMS: red, PPMS: yellow, RRMS: blue, Control: green). The figure shows comparison of each lesion MRI metric in classical lesions, tract-based lesions, and tract-based NAWM. The first-row top (a) compares T1 values, second-row (b) MTR values, third-row (c) MD values, and bottom-row (d) FA values.*

3.2 Classical versus tract-based: Univariate correlation analysis with clinical outcomes

In the previous section, we reported that group-level comparisons (Fig 4b) showed that tract-based filtering of lesion location reproduced trends observed across clinical phenotypes, suggesting that tract-based clustering captures disease-relevant patterns more effectively than anatomical-based lesion clustering. To extend this analysis, we next evaluated the correlations between MRI biomarkers and disability measures, comparing tract-based versus anatomical

lesion clustering, and testing whether lesion location alone is sufficient to explain clinical outcomes or whether adding lesion composition provides additional predictive value.

3.2.1 Lesion Location and Disability Correlation

Spearman's correlation analysis between lesion number and disability scores showed generally low associations across both classical and tract-based regions (Fig S1). The highest correlation was observed in the infratentorial region with 9HPTND ($r = 0.51$) after adjustment for sex, age, and disease duration. Among tract-based measures, the cerebellar and projection-brainstem tracts showed the strongest correlation with EDSS ($r = 0.44$). Correlation values across other regions and disability scores remained low. These findings highlight that lesion location when taken alone poorly captures disability severity, reinforcing the disconnect between radiological lesion count and clinical outcomes frequently observed in MS.

Lesion volume demonstrated stronger correlations with disability than lesion number (Fig S2). In classical regions, the periventricular area showed the highest associations with EDSS and 9HPTND ($r = 0.58$ for both). Tract-based metrics showed the strongest correlations in the occipitoparietal (EDSS: $r = 0.59$; 9HPTND: $r = 0.59$) and projection-brainstem tracts (EDSS: $r = 0.56$; 9HPTND: $r = 0.55$). Other regions, including deep white matter and juxtacortical areas, had consistently low correlations (e.g., 9HPTND: $r = 0.25$, MSPro: $r = 0.08$). Across all regions, MSPro showed weak and statistically significant correlations, suggesting limited dependence on lesion volume for patient-reported outcomes.

Lnorm showed similar spatial patterns to absolute volume demonstrating moderate to low correlation strength across disability scales (Fig S3). In classical regions, the periventricular correlation with EDSS was $r = 0.50$ and with 9HPTND $r = 0.56$. Tract-based measures showed moderate associations, particularly in the occipitoparietal (EDSS: $r = 0.50$; 9HPTND: $r = 0.55$) and projection-brainstem tracts (EDSS: $r = 0.47$; 9HPTND: $r = 0.58$). Other disability scores and regions showed lower correlations, with MSPro correlations remaining the weakest overall.

3.2.2 Lesion Composition and Disability Correlation

Across all metrics (Figs S4-S7), tract-based lesion regions consistently showed stronger correlations with EDSS compared to classical lesion regions, with particularly high associations observed in the occipitoparietal and projection-brainstem tracts. NAWM in these tracts also

showed moderate associations with EDSS, while correlations with other disability measures (e.g., T25FW, 9HPT, MSPro) were generally lower and more variable.

T1 values in tract-based lesions correlated strongly with EDSS, especially in the occipitoparietal (r = 0.82), projection-brainstem (r = 0.79), association (r = 0.77), and cerebellar (r = 0.70) tracts. In contrast, classical lesion regions such as periventricular and juxtacortical showed weaker EDSS correlations (r=0.59 and r=0.26, respectively). T1 values in tract-based NAWM showed moderate EDSS associations, with the highest correlations in the occipitoparietal (r = 0.52) and projection-brainstem (r = 0.53) tracts. MSPro and other functional scores showed generally weak correlations across both lesion and NAWM regions (e.g., MSPro–occipitoparietal lesion T1: r = 0.27; NAWM: r = 0.26).

MTR values in tract-based lesions showed moderate EDSS correlations, led by cerebellar (r = 0.59), association (r = 0.58), occipitoparietal (r = 0.54), and projection-brainstem (r = 0.51) tracts. In classical regions, EDSS associations were weaker (juxtacortical: r = 0.24; periventricular: r = 0.43). In tract-based NAWM, MTR–EDSS correlations peaked in the occipitoparietal (r = 0.63) and projection-brainstem (r = 0.56) tracts, while other tracts showed weaker associations. Functional scores were weakly correlated across all regions (MSPro range: 0.00–0.37).

MD values in tract-based lesions yielded strong EDSS correlations: occipitoparietal and projection-brainstem tracts both r = 0.80, followed by association (r = 0.77) and cerebellar (r = 0.69). Functional scores showed moderate correlations (T25FW: 0.38–0.44; 9HPT: 0.37–0.39), while MSPro remained weakly associated (0.07–0.12). In classical lesions, MD–EDSS correlations were modest (deep white matter: r = 0.43; others < 0.33), with similarly low correlations for functional scores. Tract-based NAWM MD also showed moderate EDSS associations (projection-brainstem: r = 0.66; occipitoparietal: r = 0.59), while functional correlations remained lower (T25FW and 9HPTND: 0.11–0.24).

FA in tract-based lesions correlated moderately with EDSS, most strongly in cerebellar (r = 0.66) and occipitoparietal (r = 0.61) tracts, followed by association and projection-brainstem (both r = 0.59). Classical lesion regions showed weak-to-negligible EDSS associations (range: 0.01–0.14). FA values in tract-based NAWM were moderately associated with EDSS in occipitoparietal (r = 0.65) and projection-brainstem (r = 0.37) tracts, with weaker correlations in other tracts (r =

0.13). Functional scores showed modest associations (T25FW, 9HPT: 0.03–0.39), and MSPro correlations remained low (0.01–0.23).

In summary, univariate analyses showed that lesion location alone provided only modest associations with disability, with correlations generally weak across both classical and tract-based regions. Incorporating lesion composition strengthened these associations, particularly for tract-based metrics, where T1 and MD measures in occipitoparietal and projection–brainstem tracts correlated strongly with EDSS. However, correlations with other functional outcomes such as T25FW, 9HPT, and MSPro remained limited. These findings indicate that while tract-based filtering enhances the biological relevance of MRI metrics, univariate correlations alone are insufficient to fully capture the complexity of disability in MS.

3.3 Classical versus tract-based: Predictive Modeling & Risk Stratification

To address the limitations of univariate analysis, we next employed multivariate predictive modeling to test whether combining lesion location and composition across tracts can improve disability prediction and risk stratification, and to evaluate whether tract-based frameworks outperform classical region-based approaches when considered in aggregate.

3.3.1 Lesion Location: Predictive Multivariate Analysis

(Table 4) In the combined ordinal logistic regression model, normalized Lnorm did not reach statistical significance across most disability outcomes. Tract-based Lnorm was significantly associated with 9HPTD (OR = 4.37 [1.72, 11.07], p = 0.002), and classical Lnorm showed significance for 9HPTND (OR = 4.09 [1.08, 15.53], p = 0.038). Lnorm predictors were not significant for EDSS, T25FW, or MSPro in either anatomical model (all $p > 0.05$). Among demographic covariates, age was significantly associated with multiple disability outcomes.

3.3.2 Lesion Composition Multivariate Analysis

(Table 3) Lesion-composition predictors in the tract-based model show consistently stronger and more frequent associations with disability than the same metrics extracted from classical regions. In the combined ordinal-logistic framework, tract-based T1, MTR, FA, and MD reached statistical significance for four of the five disability scales, whereas their classical counterparts were rarely significant and, when they were, exhibited smaller effect sizes.

Table 4 *Demonstrates the combined ordinal logistic regression model incorporating both linear predictors plus demographics (confounds). Significant predictions are highlighted in cyan color.*

Ordinal logistic regression results (Odds Ratio)

		Tract-based (OR [CI], p)	Classical-based (OR [CI], p)	Confounds
EDSS	Lnorm	2.59 [0.90, 7.47], p = 0.078	1.75 [0.61, 4.97], p = 0.296	Age, p = 0.013
	T1	5.70 [2.28, 14.26], p = 0.000	1.30 [0.57, 2.95], p = 0.531	None
	MTR	0.08 [0.03, 0.22], p = 0.000	0.54 [0.27, 1.08], p = 0.082	None
	FA	0.12 [0.05, 0.26], p = 0.000	0.72 [0.40, 1.30], p = 0.280	None
	MD	4.78 [1.83, 12.48], p = 0.001	1.13 [0.42, 3.03], p = 0.802	None
T25FW	Lnorm	1.06 [0.35, 3.18], p = 0.916	2.22 [0.74, 6.63], p = 0.155	Age, p = 0.014
	T1	2.22 [1.01, 4.90], p = 0.048	1.58 [0.65, 3.81], p = 0.309	None
	MTR	0.25 [0.11, 0.57], p = 0.001	0.92 [0.45, 1.89], p = 0.821	None
	FA	0.17 [0.07, 0.41], p = 0.000	0.80 [0.34, 1.87], p = 0.603	None
	MD	2.88 [1.33, 6.28], p = 0.008	1.04 [0.44, 2.43], p = 0.936	None
9HPTD	Lnorm	4.37 [1.72, 11.07], p = 0.002	1.07 [0.45, 2.54], p = 0.879	None
	T1	1.84 [0.95, 3.58], p = 0.072	1.27 [0.63, 2.58], p = 0.503	None
	MTR	0.67 [0.37, 1.23], p = 0.195	0.50 [0.26, 0.96], p = 0.039	None
	FA	0.23 [0.11, 0.51], p = 0.000	0.55 [0.28, 1.09], p = 0.087	None
	MD	1.55 [0.81, 2.96], p = 0.184	2.57 [1.14, 5.79], p = 0.023	None
9HPTND	Lnorm	1.45 [0.47, 4.51], p = 0.516	4.09 [1.08, 15.53], p = 0.038	None
	T1	3.24 [1.59, 6.60], p = 0.001	1.19 [0.61, 2.31], p = 0.614	None
	MTR	0.22 [0.10, 0.47], p = 0.000	0.74 [0.40, 1.37], p = 0.337	None
	FA	0.13 [0.06, 0.31], p = 0.000	0.58 [0.31, 1.09], p = 0.090	None
	MD	7.72 [2.78, 21.44], p = 0.000	1.45 [0.66, 3.16], p = 0.354	None
MSPro	Lnorm	1.78 [0.53, 5.97], p = 0.349	1.36 [0.46, 4.02], p = 0.574	Age, p = 0.000
	T1	3.13 [1.22, 7.98], p = 0.017	1.13 [0.43, 2.99], p = 0.798	Age, p = 0.000
	MTR	0.40 [0.19, 0.81], p = 0.012	0.79 [0.43, 1.44], p = 0.439	Age, p = 0.000
	FA	0.30 [0.15, 0.61], p = 0.001	0.45 [0.22, 0.91], p = 0.025	Age, p = 0.000
	MD	2.92 [1.05, 8.13], p = 0.040	1.21 [0.49, 2.99], p = 0.685	Age, p = 0.000

3.4 Tract-based T1 signals stratify RRMS patients into MSPro stable and MSpro progressive risk groups.

Leveraging the MSPro score, we dichotomized the relapsing–remitting cohort into clinically stable (score = 1) and progressive risk (score ≥ 2) categories and then applied ridge-penalized logistic regression to the mean T1 intensity of 16 bilateral tract compartments. In the entire sample of 88 patients (46 stable, 42 progressive) the model correctly classified 75 cases, yielding an accuracy of 85%, sensitivity of 86 %, specificity of 85 %, and an AUC of 0.90.

(Fig 6) A follow-up subset analysis, limited to only RRMS patients (42 stable, 7 progressive according to MSPro score), showed even stronger discrimination: accuracy rose to 92%, with sensitivity 86% and specificity 93%. These results indicate that a tract-based T1 signature, supplemented only by age, sex and disease duration, can meaningfully flag RRMS individuals at heightened short-term risk of progression.

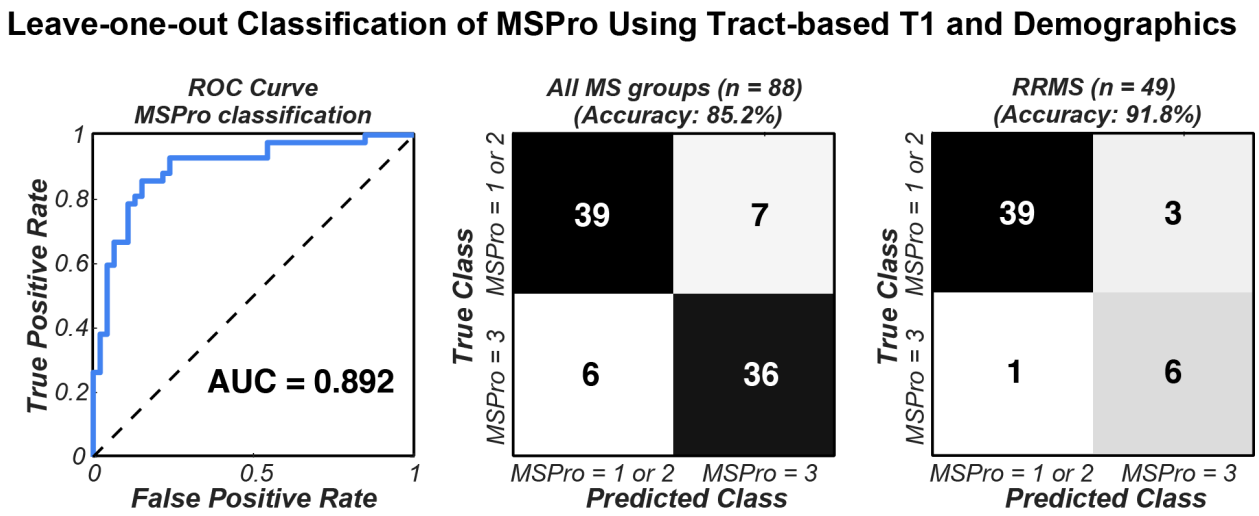


Figure 7 *Demonstrates that tract-based T1 signal analysis, combined with clinical variables, (a) represents the ROC curve, (b) confusion matrix for the full cohort, and (c) confusion matrix for the RR group.*

4. Discussion

4.1 Study recap

This study aimed to improve the clinical relevance of lesion analysis in multiple sclerosis by refining how and where lesions are measured. While traditional approaches focus on anatomical lesion burden, "disconnection" studies have demonstrated that functional deficits in neurological disease often arise from tract-specific interruptions in distributed brain networks [20,22,29].

We hypothesized that filtering lesions through clusters of functionally meaningful white matter tracts rather than broad anatomical compartments would allow for more accurate identification of clinically relevant damage. The premise is intuitive: lesions that disrupt specific tracts or their intersections likely carry greater functional weight than those in other regions. Such filtering provides a biologically grounded and reproducible framework for lesion inclusion and exclusion, enabling more meaningful analysis of tissue properties within these tracts.

The clinico-radiological paradox in MS persists because not all lesions are the same. While lesion location has traditionally been used to explain disability [42], our findings confirm that lesion position alone, even when filtered through functionally relevant white matter tracts, offers limited explanatory power. This is expected: lesion number, volume, or normalized volume may reflect spatial patterns of disease but cannot differentiate between biologically distinct lesion types.

4.2 Tract-Based Filtering and The Value of Lesion Location

MS lesions are biologically heterogeneous both across and within patients. Plaques in RRMS differ markedly from those in SPMS in chronicity and in the degree and nature of inflammation, demyelination, axonal loss, and gliosis grey [11,43]. T1 mapping and MTR captured those differences and aligned closely with postmortem histology, reinforcing their pathological validity [44]. Yet even within a single patient lesions vary in their microstructural composition and functional relevance, a variability documented by numerous studies demonstrating that lesions in distinct anatomical locations exert differential impacts on neurological function and correlate variably with disease severity [45]

Hence, capturing MS disease severity and progression with quantitative MRI hinges on two complementary dimensions: lesion topography (precise spatial distribution) and lesion microstructure (intrinsic pathology such as demyelination, axonal transection, and gliosis). Nonetheless, anatomically defined ROIs, like the periventricular compartment, treat every lesion that contacts the ependymal surface as equivalent, regardless of its microstructural damage or functional connectivity. Because the periventricular mask can extend far from the ventricular lining whenever plaques merely "touch" the ROI, confluent lesions in SPMS disproportionately inflate total burden and obscure stage-specific changes. Likewise, classical anatomical masks based solely on McDonald criteria, potentially including lesions irrelevant to key neural circuits while excluding strategically disruptive foci.

By restricting analysis to lesions within well-characterized, functionally important white-matter tracts, tract-based filtering dramatically reduces this noise and sharpens specificity. As illustrated in Fig 4, our approach isolates the spatial zones where lesion microstructure mirrors the connections/disconnections inside an ROI. Moreover, this framework now enables analysis of NAWM within restricted dimensions defined by each tract, thereby reflecting underlying biological and molecular heterogeneity.

4.3 Integrating Tissue Composition Enhances Clinical Associations

Building on tract-based spatial filtering, we incorporated quantitative MRI modalities (T1 and MTR) to probe underlying lesion pathology. While previous studies reported modest correlations between whole-brain T1 values and EDSS [16] or between periventricular MTR and EDSS [46], these approaches treated all lesions equally. In contrast, tract-based analysis minimizes contributions from silent regions and amplifies signals from strategically damaging lesions. T1 values from tract-based lesion masks showed strong associations with EDSS ($r_{occipitoparietal} = 0.82$, $r_{projection\text{-}brainstem} = 0.79$, $r_{association} = 0.77$, $r_{cerebellar} = 0.70$). These substantially outperformed classical regions such as periventricular ($r = 0.59$) and juxtacortical ($r = 0.26$). A similar pattern was observed with MTR that tract-based lesion MTR values yielded stronger correlations compared to the classical compartments. Importantly, these high correlations emerged only when location and composition were considered jointly; either factor alone offered diminished explanatory power. In NAWM, T1 and MTR values in the occipitoparietal and projection-brainstem tracts correlated with EDSS up to $r = 0.63$, supporting the role of diffuse microstructural injury in disability progression. This

layered analysis also revealed that not all tracts contribute equally, while averaged tract clusters captured moderate correlations with other disability scores. Individual tracts, particularly those involved in sensorimotor integration and balance, showed sharper association.

4.4 Diffusion Metrics Anchor Structure to Function

Diffusion-based MRI methods reconstruct white matter architecture in vivo, enabling interrogation of structural integrity along biologically meaningful pathways. While classical lesion compartments (e.g., periventricular or juxtacortical regions) group tissue based on anatomical proximity, their clinical relevance has often been diluted by arbitrary anatomical ROIs that ignore white matter architecture. By anchoring diffusion metrics to specific tracts, we observed markedly stronger and more interpretable correlations with disability. This is not surprising: FA and MD reflect fiber coherence and water diffusivity along directional axes, and thus their biological relevance emerges most clearly when measured along coherent, functionally relevant fiber systems. Tracts such as the occipitoparietal and association bundles consistently showed strong correlations with EDSS across both lesion and NAWM compartments, suggesting that damage in these pathways integral to visuospatial coordination, attention, and sensorimotor integration [47] may be particularly impactful on global disability. On the other hand, the cerebellar tracts showed weaker associations, likely reflecting well-known challenges in tractography in this region, including signal dropout, complex fiber geometry, and distortion artifacts [48]. This again highlights that the quality of the spatial model directly affects the clinical signal: even biologically important structures may fail to correlate if their spatial definition is noisy.

4.5 Multivariate Comparison of Predictors

To practically assess how different lesion‑derived metrics behave in predicting ordered disability outcomes, we employed an ordinal logistic regression with ridge regularization, not as a direct “tract vs. classical” showdown, but as a unified framework for comparing model performance. We extracted reduced feature sets from both classical anatomical ROIs and tract-based regions, each capturing lesion and NAWM burden plus microstructural properties, and entered them together with demographic covariates into a cumulative logit model. In this context, a predictor’s significance indicates its independent ability to shift the probability distribution across disability categories after adjusting for all other variables. Tract-informed predictors consistently yielded

higher odds ratios, demonstrating that filtering imaging metrics through functionally defined tracts enhances both biological relevance and predictive power. In contrast, classical ROI metrics reached significance infrequently or showed attenuated effect sizes, reflecting their coarser granularity.

4.6 Translating Tract-Based Lesion Metrics to Risk Stratification

A period of diagnostic uncertainty regarding the transition from RRMS to SPMS is present in many MS patients, with a mean time of around 3 years [49, 50]. MSPro tool classified SPMS cases with a sensitivity of 82% and specificity of 84%; it also distinguished relapsing–remitting disease with 83% sensitivity and 82% specificity [51].

The classification results using MSPro support the central premise of this study: that combining lesion location with quantitative assessment of tissue composition captures clinically relevant variance that is not explained by lesion burden alone. MSPro was chosen as an outcome measure because it reflects subtle, patient-reported functional decline that may precede overt progression on conventional scales such as the EDSS. This makes it particularly suitable for evaluating early risk in RRMS patients who may appear clinically stable but are exhibiting underlying disease activity.

The ridge-penalized logistic regression model trained on tract-based T1 metrics reliably identified individuals with elevated MSPro scores, both across the full cohort and within the RRMS subgroup. This indicates that microstructural abnormalities within functionally relevant tracts can signal emerging disability before overt clinical worsening. These findings demonstrate that tract-informed T1 metrics might identify early subclinical deterioration with functional relevance, particularly in patients not yet classified as progressive by conventional clinical criteria.

4.7 Methodological Considerations and Future Directions

While our findings highlight the clinical relevance of tract-informed lesion and NAWM metrics, several methodological limitations merit consideration. One concern is the generalizability of MP2RAGE-based T1 mapping. Although this sequence provides improved tissue contrast and quantitative stability, it is not yet widely implemented in routine clinical protocols. MP2RAGE

typically requires longer acquisition times compared to conventional sequences, making it less practical in time-constrained clinical workflows[52].

More broadly, while our tract-based framework enables refined tissue sampling, future work should further tailor tract filtering to specific clinical endpoints rather than applying universal lesion mask across all outcomes. The inclusion of spinal cord atrophy measures and advanced imaging modalities such as quantitative susceptibility mapping (QSM) could improve sensitivity to neurodegeneration, particularly in progressive MS[53]. Longitudinal studies will also be critical to validate whether these tract-specific markers predict clinical decline over time. Follow-up imaging will allow us to assess whether early T1 abnormalities evolve in parallel with symptom progression, helping define actionable thresholds. Finally, methodological advances in cerebellar tract segmentation and motion correction will likely enhance correlation detection in regions currently affected by signal instability or tracking artifacts[48].

5. Conclusion

This study demonstrated that combining lesion location with microstructural composition through tract-based qMRI mapping offers a clinically meaningful advancement in MS progression prediction. The integration of spatial and biological lesion attributes translates into clinically interpretable risk estimation beyond statistical measures. As quantitative MRI becomes standardized and computational modeling more accessible, this framework enables imaging-driven tools for identifying at-risk patients and supporting tailored therapeutic decisions in routine practice. Ultimately, we believe this approach transforms lesion analysis from a descriptive exercise to a functional biomarker, bridging radiological findings with clinical outcomes.

Acknowledgements

This study was supported by the Core Technology Platforms (CTP), Biomedical Imaging Core at New York University Abu Dhabi (NYUAD). We also acknowledge the Computational Research Center at NYUAD for providing supercomputing resources.

Funding

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under the NYU Abu Dhabi Research Institute grant CG012. OA and AWA were supported by a grant from the National Multiple Sclerosis Society, UAE, under project LAMINATE.

Competing interests

The authors report no competing interests.

References

1. Lassmann H. The pathology of multiple sclerosis and its evolution. *Philos Trans R Soc Lond B Biol Sci*. 1999;354(1390):1635-1640.

2. Allen IV, McKeown SR. A histological, histochemical and biochemical study of the macroscopically normal white matter in multiple sclerosis. *J Neurol Sci*. 1979;41(1):81-91.

3. Bitsch A, Schuchardt J, Bunkowski S, Kuhlmann T, Brück W. Acute axonal injury in multiple sclerosis. Correlation with demyelination and inflammation. *Brain*. 2000;123 (Pt 6):1174-1183.

4. Cameron MH, Nilsagard Y. Balance, gait, and falls in multiple sclerosis. *Handb Clin Neurol*. 2018;159:237-250.

5. Younger DS. Multiple sclerosis: Motor dysfunction. *Handb Clin Neurol*. 2023;196:119-147.

6. Benedict RHB, DeLuca J, Enzinger C, Geurts JJG, Krupp LB, Rao SM. Neuropsychology of multiple sclerosis: Looking back and moving forward. *J Int Neuropsychol Soc*. 2017;23(9-10):832-842.

7. Glanz BI, Holland CM, Gauthier SA, et al. Cognitive dysfunction in patients with clinically isolated syndromes or newly diagnosed multiple sclerosis. *Mult Scler*. 2007;13(8):1004-1010.

8. Rae-Grant AD, Eckert NJ, Bartz S, Reed JF. Sensory symptoms of multiple sclerosis: a hidden reservoir of morbidity. *Mult Scler*. 1999;5(3):179-183.

9. Filippi M, Rocca MA, Comi G. use quantitative magnetic-resonance- based techniques monitor evolution multiple sclerosis. *Lancet Neurol*. 2003;2:337-346.

10. Barkhof F. The clinico-radiological paradox in multiple sclerosis revisited. *Curr Opin Neurol*. 2002;15(3):239-245.

11. Geurts JJG, Barkhof F. Grey matter pathology in multiple sclerosis. *Lancet Neurol*. 2008;7(9):841-851.

12. Solomon AJ, Pettigrew R, Naismith RT, Chahin S, Krieger S, Weinshenker B. Challenges in multiple sclerosis diagnosis: Misunderstanding and misapplication of the McDonald criteria. *Mult Scler*. 2021;27(2):250-258.

13. Moccia M, van de Pavert S, Eshaghi A, et al. Pathologic correlates of the magnetization transfer ratio in multiple sclerosis. *Neurology*. 2020;95(22):e2965-e2976.

14. Rahmanzadeh R, Weigel M, Lu PJ, et al. A comparative assessment of myelin-sensitive measures in multiple sclerosis patients and healthy subjects. *NeuroImage Clin*. 2022;36(103177):103177.

15. Zollinger LV, Kim TH, Hill K, Jeong EK, Rose JW. Using diffusion tensor imaging and immunofluorescent assay to evaluate the pathology of multiple sclerosis. *J Magn Reson Imaging*. 2011;33(3):557-564.

16. Thaler C, Faizy TD, Sedlacik J, et al. T1 recovery is predominantly found in black holes and is associated with clinical improvement in patients with multiple sclerosis. *AJNR Am J Neuroradiol*. 2017;38(2):264-269.

17. Martínez-Heras E, Solana E, Prados F, et al. Characterization of multiple sclerosis lesions with distinct clinical correlates through quantitative diffusion MRI. *NeuroImage Clin*. 2020;28:102411.

18. Gloor M, Andelova M, Gaetano L, et al. Longitudinal analysis of new multiple sclerosis lesions with magnetization transfer and diffusion tensor imaging. *Eur Radiol*. 2024;34(3):1680-1691.

19. Wilson M, Tench C, Morgan P, Blumhardt L. Pyramidal tract mapping by diffusion tensor magnetic resonance imaging in multiple sclerosis: improving correlations with disability. *Journal of Neurology, Neurosurgery, and Psychiatry*. 2003;74:203-207.

20. Rimkus C de M, Junqueira T de F, Callegaro D, Otaduy MCG, Leite C da C. Segmented corpus callosum diffusivity correlates with the Expanded Disability Status Scale score in the early stages of relapsing-remitting multiple sclerosis. *Clinics (Sao Paulo)*. 2013;68(8):1115-1120.

21. Sechi E, Keegan BM, Kaufmann TJ, Kantarci OH, Weinshenker BG, Flanagan EP. Unilateral motor progression in MS: Association with a critical corticospinal tract lesion: Association with a critical corticospinal tract lesion. *Neurology*. 2019;93(7):e628-e634.

22. Chylińska M, Karaszewski B, Komendziński J, et al. The association between white matter tract structural connectivity and information processing speed in relapsing-remitting multiple sclerosis. *Neurol Sci*. 2023;44(9):3221-3232.

23. Madsen MAJ, Wiggermann V, Marques MFM, et al. Linking lesions in sensorimotor cortex to contralateral hand function in multiple sclerosis: a 7 T MRI study. *Brain*. 2022;145(10):3522-3535.

24. Kim M, Seo JW, Kim MS, Lee KH, Kim M. White matter tract density index is associated with disability in multiple sclerosis. *Neurobiol Dis*. 2024;198(106548):106548.

25. Winter M, Tallantyre EC, Brice TAW, Robertson NP, Jones DK, Chamberland M. Tract-specific MRI measures explain learning and recall differences in multiple sclerosis. *Brain Commun*. 2021;3(2):fcab065.

26. Ravano V, Piredda GF, Krasensky J, et al. Tract-wise microstructural analysis informs on current and future disability in early multiple sclerosis. *J Neurol*. 2024;271(2):631-641.

27. Harrison DM, Shiee N, Bazin PL, et al. Tract-specific quantitative MRI better correlates with disability than conventional MRI in multiple sclerosis. *J Neurol*. 2013;260(2):397-406.

28. Storelli L, Pagani E, Meani A, Preziosa P, Filippi M, Rocca MA. Advanced diffusion-weighted imaging models better characterize white matter neurodegeneration and clinical outcomes in multiple sclerosis. *J Neurol*. 2022;269(9):4729-4741.

29. Ravano V, Andelova M, Fartaria MJ, et al. Validating atlas-based lesion disconnectomics in multiple sclerosis: A retrospective multi-centric study. *NeuroImage Clin*. 2021;32(102817):102817.

30. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983;33(11):1444-1452.

31. Smith A. Symbol Digit Modalities Test. *PsycTESTS Dataset*. Published online June 13, 2016. doi:10.1037/t27513-000

32. Motl RW, Cohen JA, Benedict R, et al. Validity of the timed 25-foot walk as an ambulatory performance outcome measure for multiple sclerosis. *Mult Scler*. 2017;23(5):704-710.

33. Feys P, Lamers I, Francis G, et al. The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. *Mult Scler*. 2017;23(5):711-720.

34. Ziemssen T, Vandercappellen J, Jordan Mondragon V, Giovannoni G. MSProDiscuss™ clinical decision support tool for identifying multiple sclerosis progression. *J Clin Med*. 2022;11(15):4401.

35. Pirpamer L, Kincses B, Kincses ZT, et al. Periventricular magnetisation transfer abnormalities in early multiple sclerosis. *NeuroImage Clin*. 2022;34(103012):103012.

36. Fischl B. FreeSurfer. *Neuroimage*. 2012;62(2):774-781.

37. Cardoso MJ, Li W, Brown R, et al. MONAI: An open-source framework for deep learning in healthcare. *arXiv [csLG]*. Published online November 4, 2022. http://arxiv.org/abs/2211.02701

38. Woolrich MW, Jbabdi S, Patenaude B, et al. Bayesian analysis of neuroimaging data in FSL. *Neuroimage*. 2009;45(1 Suppl):S173-86.

39. Karakuzu A, Boudreau M, Duval T, et al. qMRLab: Quantitative MRI analysis, under one umbrella. *J Open Source Softw*. 2020;5(53):2343.

40. Mitra-Behura S, Fiolka RP, Daetwyler S. Singularity containers improve reproducibility and ease of use in computational image analysis workflows. *Front Bioinform*. 2021;1:757291.

41. Hsu YC, Lo YC, Chen YJ, Wedeen VJ, Isaac Tseng WY. NTU-DSI-122: A diffusion spectrum imaging template with high anatomical matching to the ICBM-152 space. *Hum Brain Mapp*. 2015;36(9):3528-3541.

42. Sanderson-Cimino M, Panizzon MS, Elman JA, et al. Periventricular and deep abnormal white matter differ in associations with cognitive performance at midlife. *Neuropsychology*. 2021;35(3):252-264.

43. Dutta R, Trapp BD. Relapsing and progressive forms of multiple sclerosis: insights from pathology. *Curr Opin Neurol*. 2014;27(3):271-278.

44. Righart R, Biberacher V, Jonkman LE, et al. Cortical pathology in multiple sclerosis detected by the T1/T2-weighted ratio from routine magnetic resonance imaging. *Ann Neurol*. 2017;82(4):519-529.

45. Bø L, Vedeler CA, Nyland H, Trapp BD, Mørk SJ. Intracortical multiple sclerosis lesions are not associated with increased lymphocyte infiltration. *Mult Scler*. 2003;9(4):323-331.

46. Amann M, Papadopoulou A, Andelova M, et al. Magnetization transfer ratio in lesions rather than normal-appearing brain relates to disability in patients with multiple sclerosis. *J Neurol*. 2015;262(8):1909-1917.

47. De Benedictis A, Duffau H, Paradiso B, et al. Anatomo-functional study of the temporo-parieto-occipital region: dissection, tractographic and brain mapping evidence from a neurosurgical perspective. *J Anat*. 2014;225(2):132-151.

48. Habas C, Manto M. Probing the neuroanatomy of the cerebellum using tractography. *Handb Clin Neurol*. 2018;154:235-249.

49. Katz Sand I, Krieger S, Farrell C, Miller AE. Diagnostic uncertainty during the transition to secondary progressive multiple sclerosis. *Mult Scler*. 2014;20(12):1654-1657.

50. Rojas JI, Patrucco L, Alonso R, et al. Diagnostic uncertainty during the transition to secondary progressive multiple sclerosis: Multicenter study in Argentina. *Mult Scler*. 2021;27(4):579-584.

51. Ziemssen T, Piani-Meier D, Bennett B, et al. A physician-completed digital tool for evaluating disease progression (multiple sclerosis progression discussion tool): Validation study. *J Med Internet Res*. 2020;22(2):e16932.

52. Marques JP, Kober T, Krueger G, van der Zwaag W, Van de Moortele PF, Gruetter R. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *Neuroimage*. 2010;49(2):1271-1281.

53. Zivadinov R, Weinstock-Guttman B, Pirko I. Iron deposition and inflammation in multiple sclerosis. Which one comes first? *BMC Neurosci*. 2011;12(1):60.

Supplementary

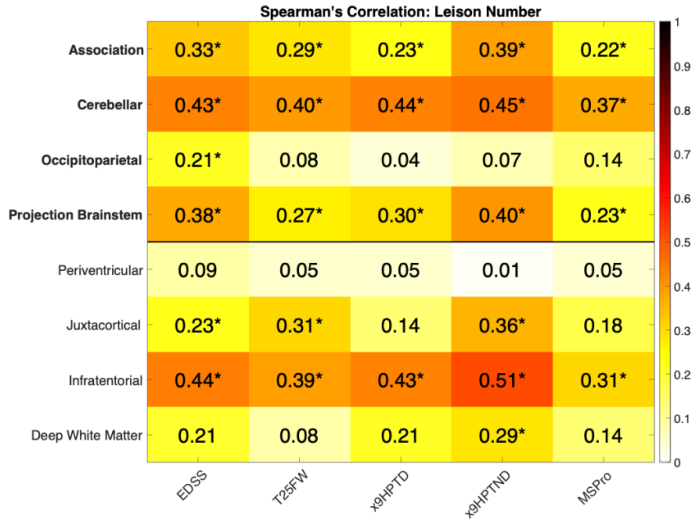


Figure S1 *Demonstrates Spearman's correlation between lesion number in different location (Tract-based vs Classical-based) and disability scores. In the heatmaps, darker shades indicate stronger absolute correlation values (range 0–1), with statistically significant associations annotated.*

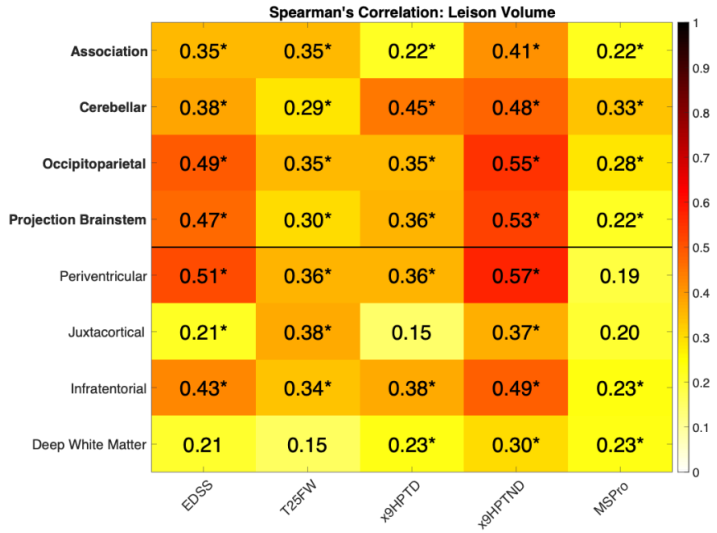


Figure S2 *Demonstrates Spearman's correlation between lesion volume in different location (Tract-based vs Classical-based) and disability scores.*

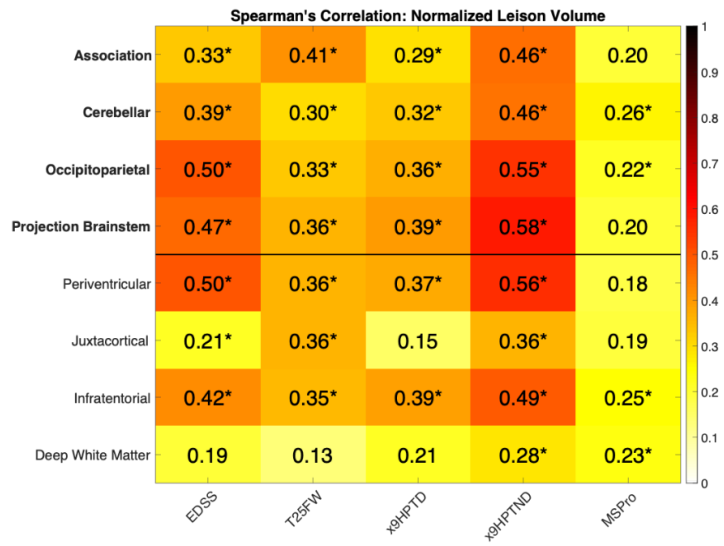


Figure S3 *Demonstrates Spearman's correlation between normalized lesion volume in different location (Tract-based vs Classical-based) and disability scores.*

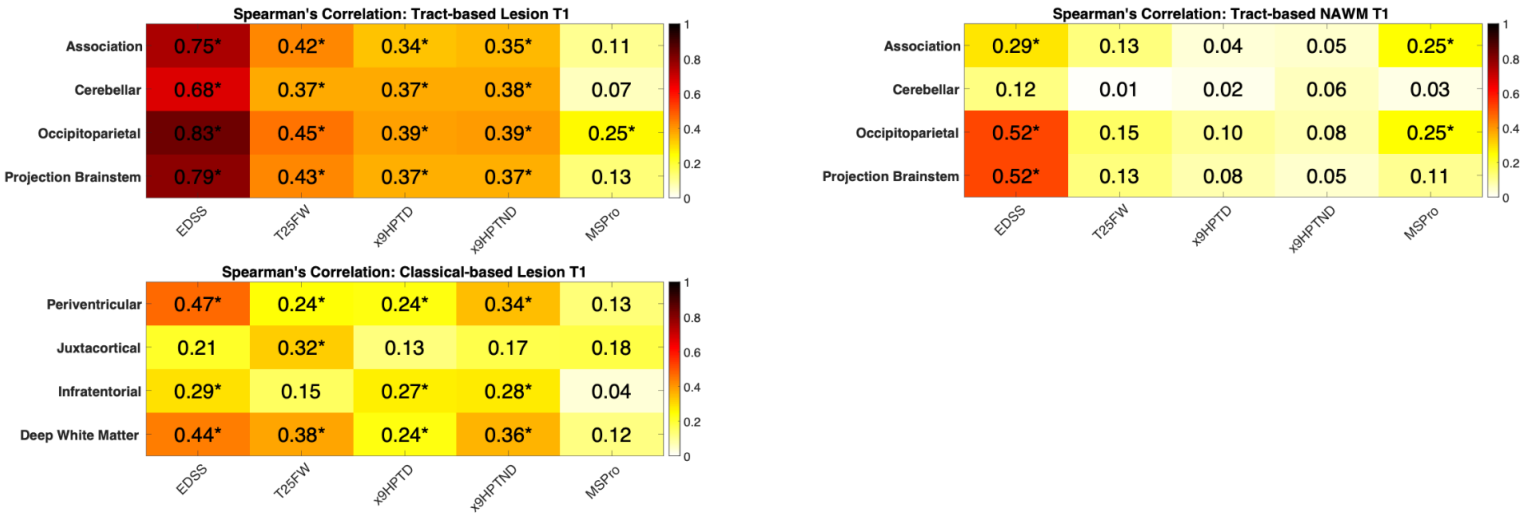


Figure S4 *Demonstrates Spearman's correlation between T1 values of lesions (right) and NAWM (left), Tract-based vs Classical-based, and disability scores.*

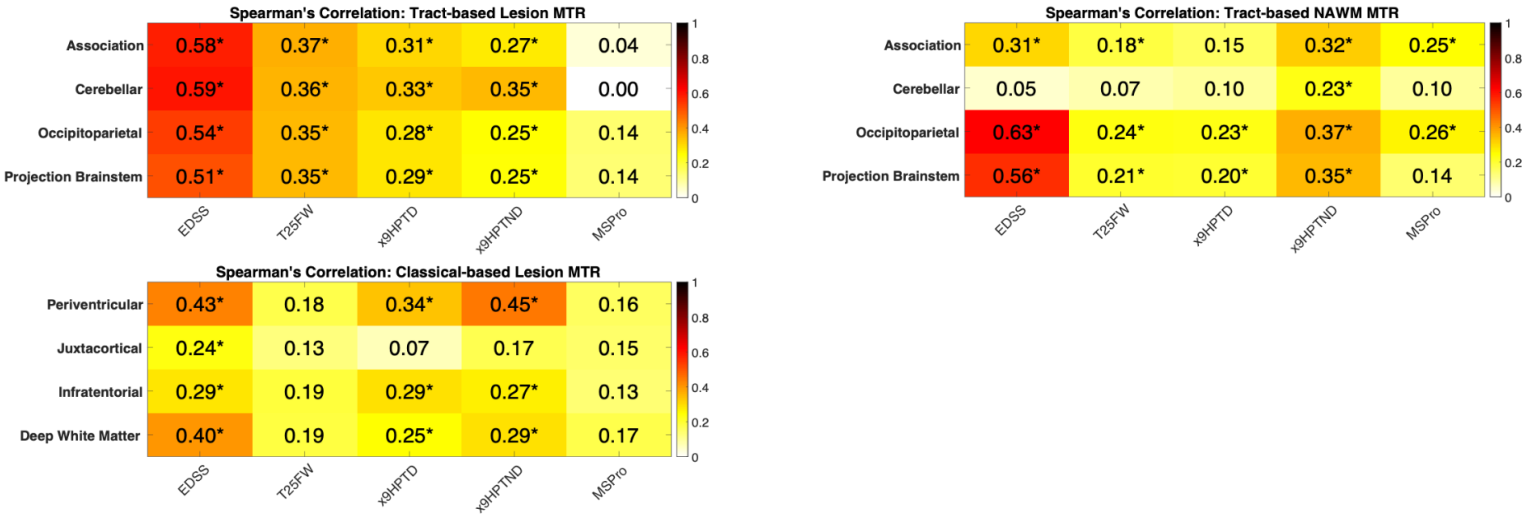


Figure S5 *Demonstrates Spearman's correlation between MTR values of lesions (right) and NAWM (left), Tract-based vs Classical-based, and disability scores.*

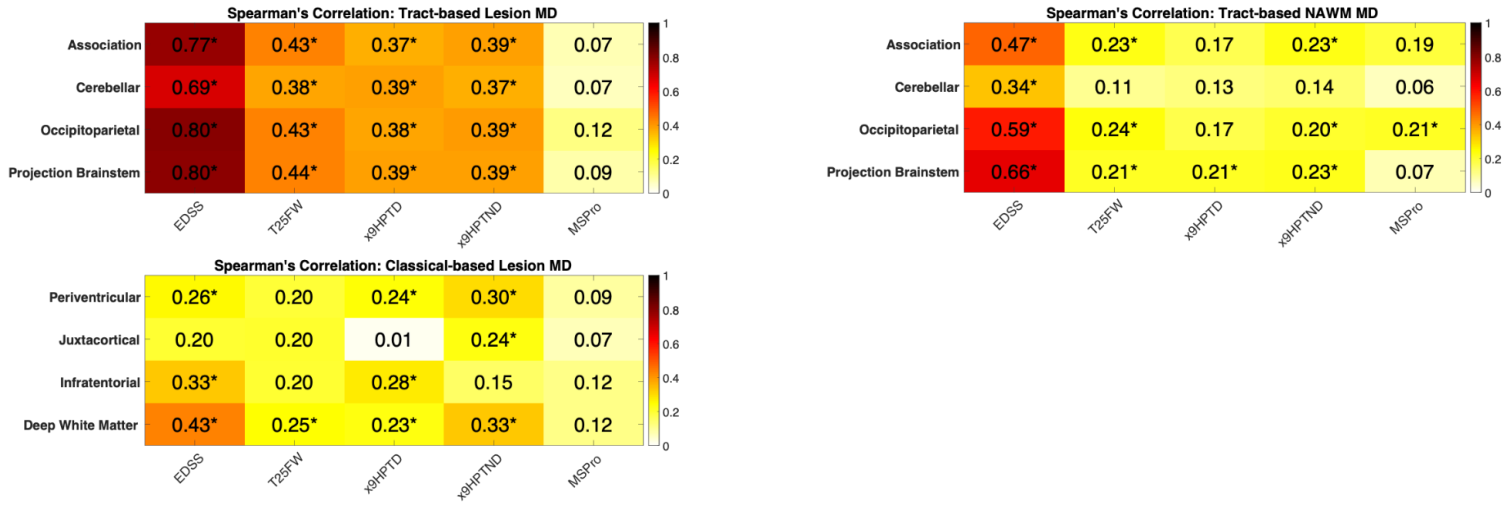


Figure S6 *Demonstrates Spearman's correlation between MD values of lesions (right) and NAWM (left), Tract-based vs Classical-based, and disability scores.*

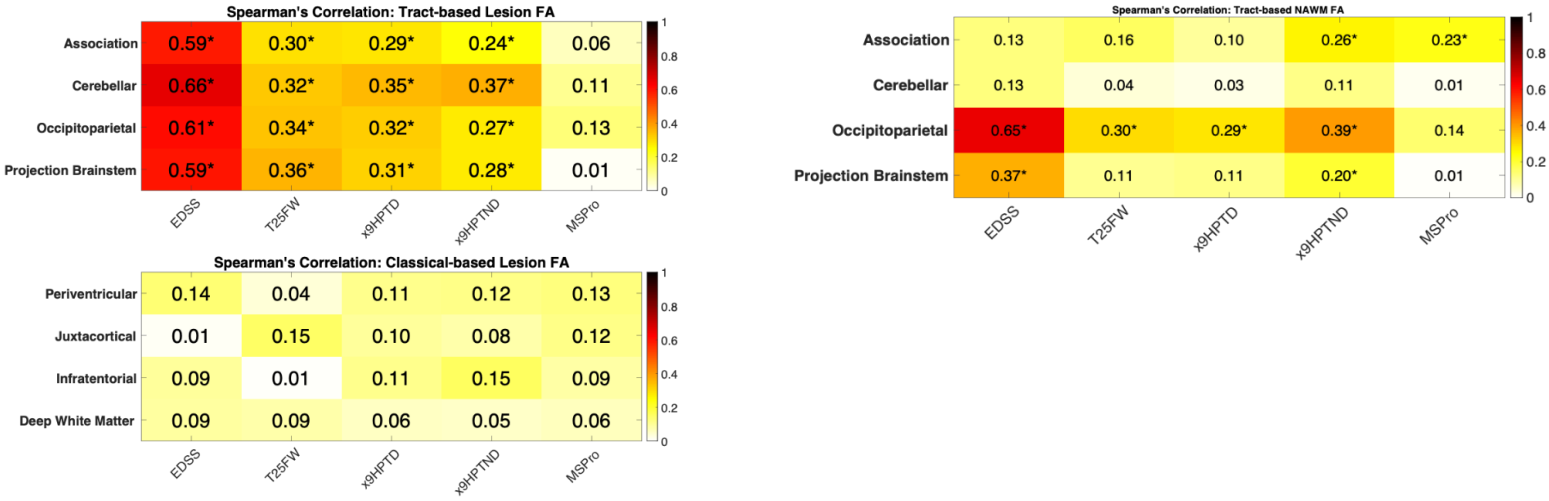
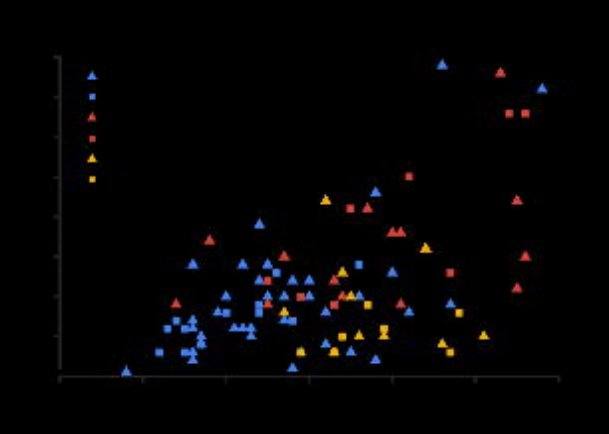
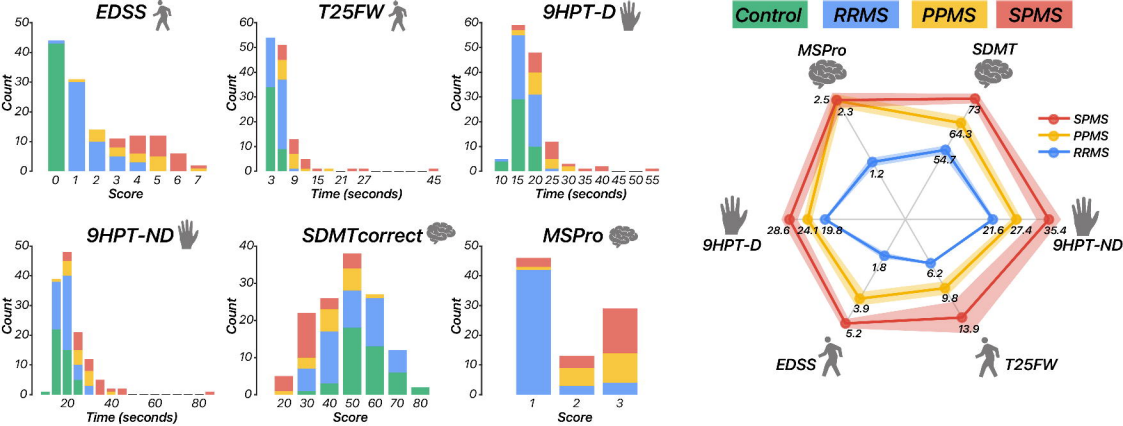


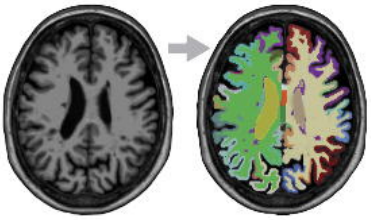
Figure S7 *Demonstrates Spearman's correlation between FA values of lesions (right) and NAWM (left), Tract-based vs Classical-based, and disability scores.*




EDSS
Count
Score
T25FW
Time (seconds)
9HPT-D
9HPT-ND
SDMTcorrect
MSPro
Control
RRMS
PPMS
SPMS
2.5
2.3
73
64.3
54.7
1.2
28.6
24.1
19.8
21.6
27.4
35.4
1.8
6.2
9.8
3.9
5.2
13.9

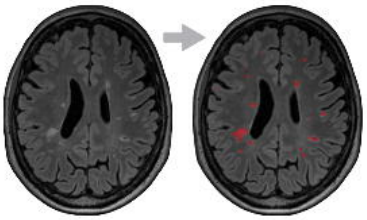
MRI Acquisition & Preprocessing

UNIT1 *(freesurfer recon-all)*

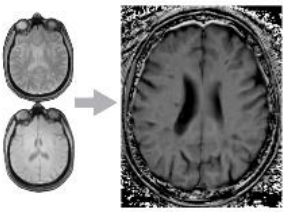


MP2RAGE *(qMRLab T1Map)*

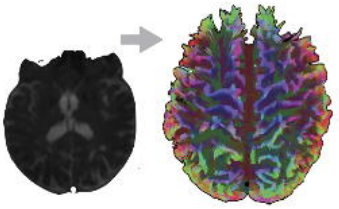
FLAIR *(LST-AI lesion)*



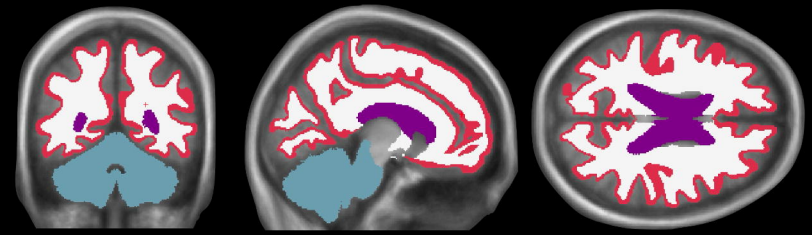
MT *(qMRLab MTRmap)*



DWI *(tractoflow/dsi_studio)*



Classical Lesion Location Topography

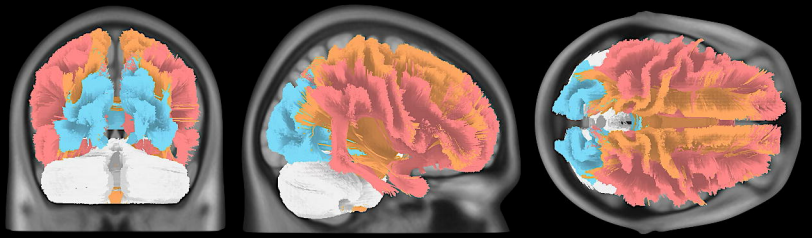


Juxtacortical

Deep White Mattrer

Periventricular

Infratentorial

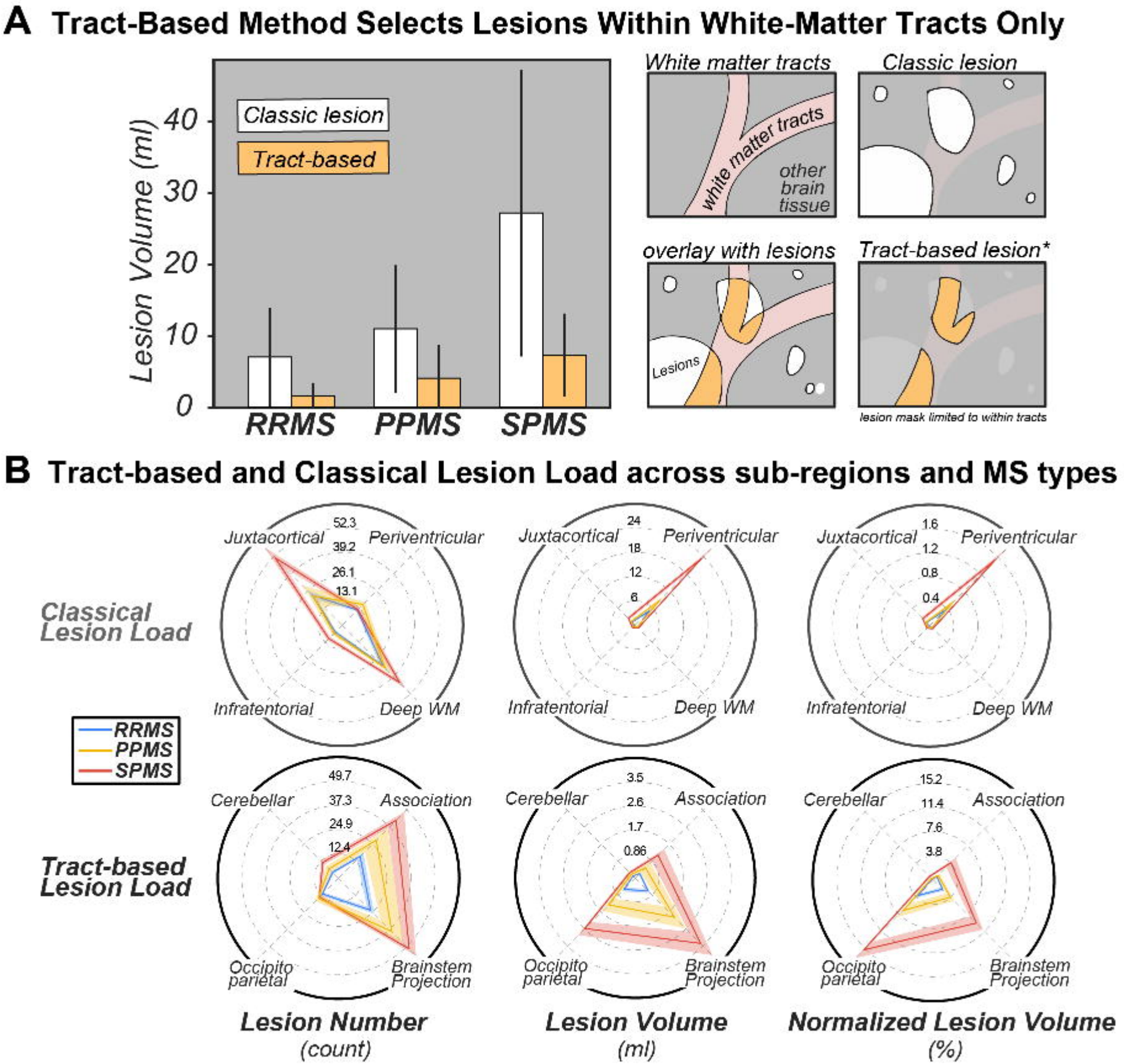


Association

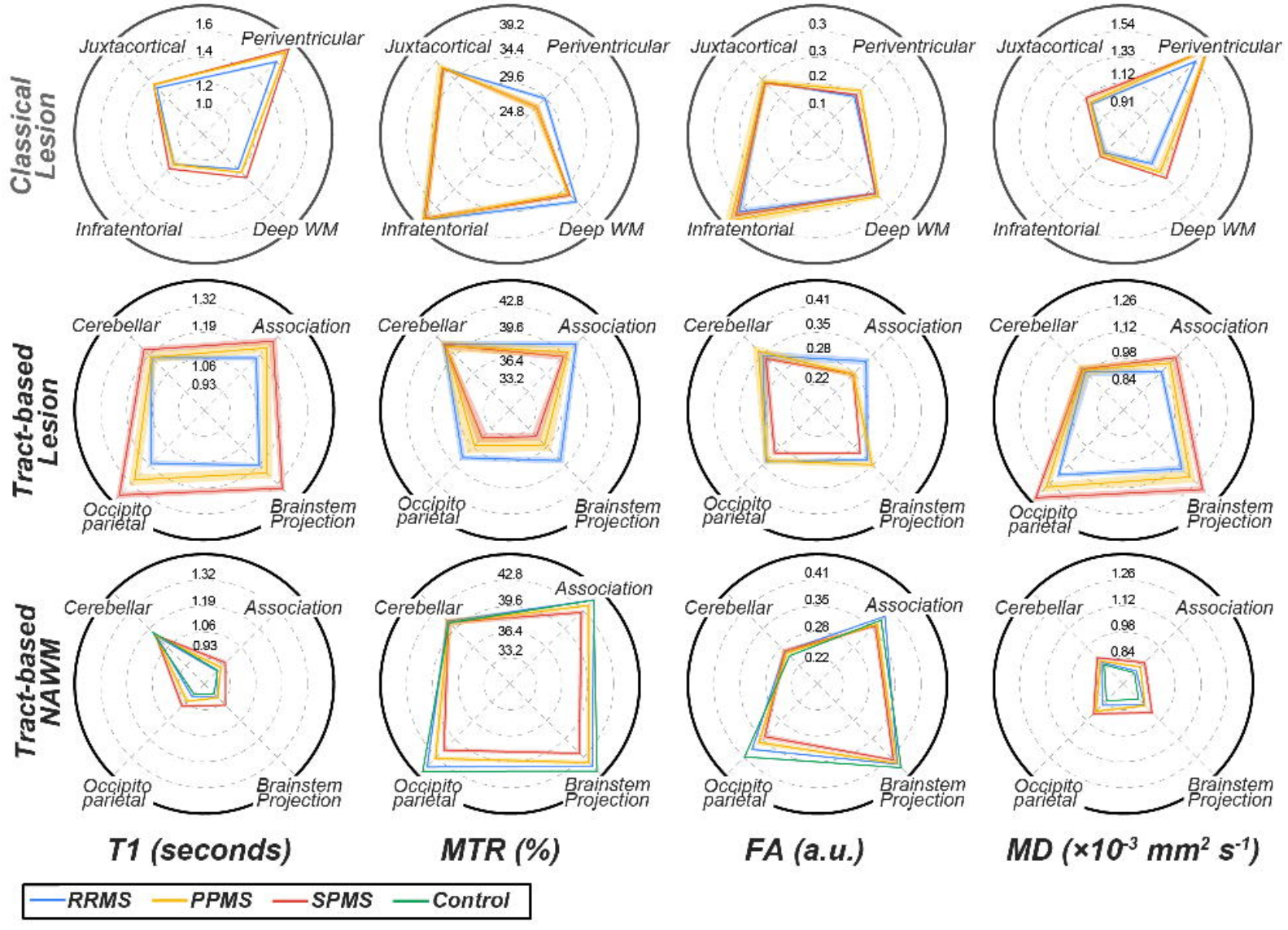
ProjectionBrainstem

Occipitoparietal

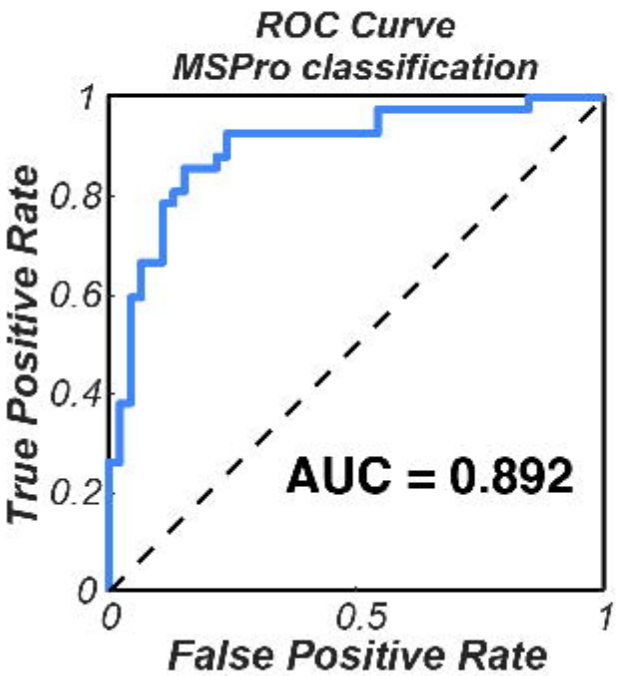
Cerebellar

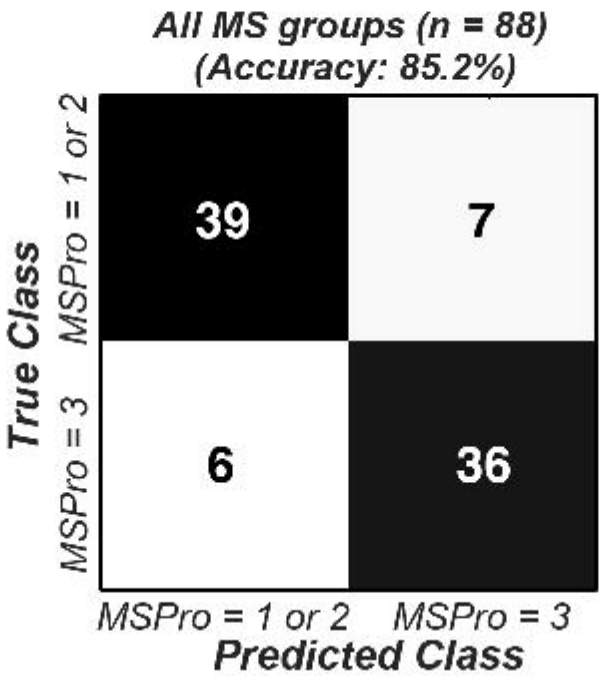

A Tract-Based Method Selects Lesions Within White-Matter Tracts Only
Lesion Volume (ml)
40
30
20
10
0
Classic lesion
Tract-based
RRMS
PPMS
SPMS
White matter tracts
white matter tracts
other brain tissue
Classic lesion
overlay with lesions
Lesions
Tract-based lesion*
lesion mask limited to within tracts
B Tract-based and Classical Lesion Load across sub-regions and MS types
Classical Lesion Load
Tract-based Lesion Load
RRMS
PPMS
SPMS
Juxtacortical
Periventricular
Infratentorial
Deep WM
Cerebellar
Association
Occipito parietal
Brainstem Projection
52.3
39.2
26.1
13.1
24
18
12
6
1.6
1.2
0.8
0.4
49.7
37.3
24.9
12.4
3.5
2.6
1.7
0.86
15.2
11.4
7.6
3.8
Lesion Number
(count)
Lesion Volume
(ml)
Normalized Lesion Volume
(%)

Classical versus Tract-based Lesion Composition Metrics across MS Subtypes



Leave-one-out Classification of MSPro Using Tract-based T1 and Demographics





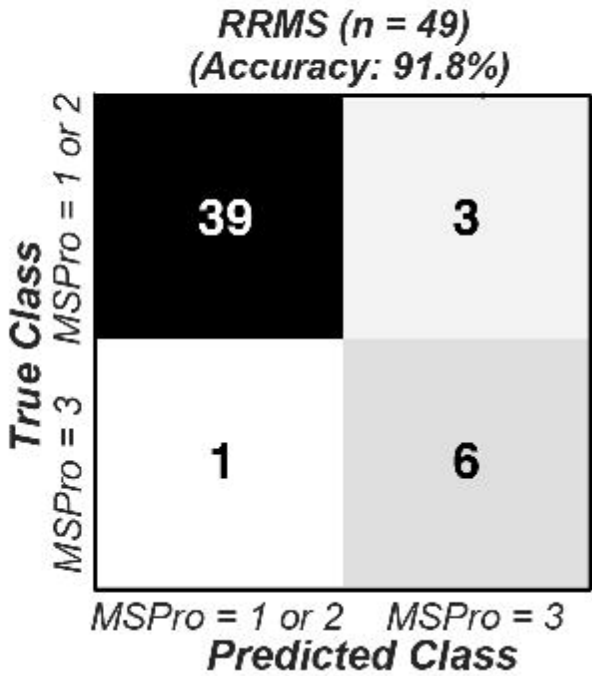


Table 1 Summary of demographic and clinical metrics across MS subtypes and healthy controls

Parameter	Healthy Controls	RRMS	PPMS	SPMS
N (% Female)	43 (65%)	49 (76%)	17 (59%)	23 (65%)
Age	35.60 ± 10.83	35.65 ± 10.25	47.88 ± 7.01	48.87 ± 12.50
Disease duration (years)	-	9.17 ± 7.48	7.53 ± 5.26	17.22 ± 8.37
EDSS	-	1.81 ± 0.95	3.94 ± 1.47	5.17 ± 1.14
DMT: Yes, n (%)	-	5 (12%)	9 (52,9%)	6 (26,1%)
Therapy: 1st line	-	29 (65.9%)	2 (22.2%)	11 (64.7%)
Therapy: 2nd line	-	15 (34.1%)	7 (77.8%)	6 (35.3%)
T25FW (s)	5.63 ± 0.52	6.23 ± 1.01	9.76 ± 3.57	13.95 ± 9.94
9HPT-D (s)	18.47 ± 2.31	19.81 ± 2.65	24.12 ± 4.51	28.61 ± 9.37
9HPT-ND (s)	20.00 ± 2.87	21.56 ± 3.89	27.39 ± 5.58	35.39 ± 14.16
SDMT Total	61.88 ± 10.24	56.49 ± 11.03	46.88 ± 8.78	38.43 ± 8.07
SDMT Mean	61.11 ± 10.37	55.88 ± 11.34	46.29 ± 9.02	37.50 ± 8.14
MSPro	-	1.22 ± 0.59	2.53 ± 0.62	2.55 ± 0.74

Table 2 Categorizes major white matter tracts into functional groups, including occipitoparietal, association, projection, brainstem, cerebellar, and commissural systems.

Occipitoparietal	Association	Projection Brainstem	Cerebellar
Projection Basal Ganglia: Optic Radiation	Arcuate Fasciculus	Corticospinal Tract	Inferior Cerebellar Peduncle
Vertical Occipital Fasciculus	Cingulum	Corticobulbar Tract	Vermis
Corpus Callosum: Forceps Major	Extreme Capsule	Corticopontine Tract	Middle Cerebellar Peduncle
	Frontal Aslant Tract	Medial Lemniscus	Superior Cerebellar Peduncle
	Hippocampus Alveus	Dentatorubro-thalmic Tract	
	Inferior Fronto Occipital Fasciculus	Non-Decussating Dentatorubro-thalamic Tract	
	Inferior Longitudinal Fasciculus	Reticular Tract	
	Middle Longitudinal Fasciculus	Medial Forebrain Bundle	
	Parietal Aslant Tract		
	Superior Longitudinal Fasciculus		
	Uncinate Fasciculus		
	Vertical Occipital Fasciculus		

Table 3 Summarizes the ordinal logistic regression results (Odds Ratio)

Variable	Tract-based	Non-tract-base	Confounds
EDSS: Lnorm	2.59 [0.90, 7.47], p = 0.078	1.75 [0.61, 4.97], p = 0.296	**Age, *p = 0.013**
EDSS: T1	**5.70 [2.28, 14.26], * p = 0.000**	1.30 [0.57, 2.95], p = 0.531	None
EDSS: MTR	**0.08 [0.03, 0.22], *p = 0.000**	0.54 [0.27, 1.08], p = 0.082	None
EDSS: FA	**0.12 [0.05, 0.26], *p = 0.000**	0.72 [0.40, 1.30], p = 0.280	None
EDSS: MD	**4.78 [1.83, 12.48], *p = 0.001**	1.13 [0.42, 3.03], p = 0.802	None
T25FW: Lnorm	1.06 [0.35, 3.18], p = 0.916	2.22 [0.74, 6.63], p = 0.155	**Age, *p = 0.014**
T25FW: T1	**2.22 [1.01, 4.90], *p = 0.048**	1.58 [0.65, 3.81], p = 0.309	None
T25FW: MTR	**0.25 [0.11, 0.57], *p = 0.001**	0.92 [0.45, 1.89], p = 0.821	None
T25FW: FA	**0.17 [0.07, 0.41], *p = 0.000**	0.80 [0.34, 1.87], p = 0.603	None
T25FW: MD	**2.88 [1.33, 6.28], *p = 0.008**	1.04 [0.44, 2.43], p = 0.936	None
9HPTD: Lnorm	**4.37 [1.72, 11.07], * p = 0.002**	1.07 [0.45, 2.54], p = 0.879	None
9HPTD: T1	1.84 [0.95, 3.58], p = 0.072	1.27 [0.63, 2.58], p = 0.503	None
9HPTD: MTR	0.67 [0.37, 1.23], p = 0.195	**0.50 [0.26, 0.96], *p = 0.039**	None
9HPTD: FA	**0.23 [0.11, 0.51], *p = 0.000**	0.55 [0.28, 1.09], p = 0.087	None
9HPTD: MD	1.55 [0.81, 2.96], p = 0.184	**2.57 [1.14, 5.79], *p = 0.023**	None
9HPTND: Lnorm	1.45 [0.47, 4.51], p = 0.516	**4.09 [1.08, 15.53], *p = 0.038**	None
9HPTND: T1	**3.24 [1.59, 6.60], *p = 0.001**	1.19 [0.61, 2.31], p = 0.614	None
9HPTND: MTR	**0.22 [0.10, 0.47], *p = 0.000**	0.74 [0.40, 1.37], p = 0.337	None
9HPTND: FA	**0.13 [0.06, 0.31], *p = 0.000**	0.58 [0.31, 1.09], p = 0.090	None
9HPTND: MD	**7.72 [2.78, 21.44], *p = 0.000**	1.45 [0.66, 3.16], p = 0.354	None
MSPro: Lnorm	1.78 [0.53, 5.97], p = 0.349	1.36 [0.46, 4.02], p = 0.574	**Age, *p = 0.000**
MSPro: T1	**3.13 [1.22, 7.98], *p = 0.017**	1.13 [0.43, 2.99], p = 0.798	**Age, *p = 0.000**
MSPro: MTR	**0.40 [0.19, 0.81], *p = 0.012**	0.79 [0.43, 1.44], p = 0.439	**Age, *p = 0.000**
MSPro: FA	**0.30 [0.15, 0.61], *p = 0.001**	0.45 [0.22, 0.91], p = 0.025	**Age, *p = 0.000**
MSPro: MD	**2.92 [1.05, 8.13], *p = 0.040**	1.21 [0.49, 2.99], p = 0.685	**Age, *p = 0.000**