Received: 26 April 2024 | Revised: 19 August 2024 | Accepted: 21 August 2024

DOI: 10.1002/alz.14371

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Microstructural mapping of neural pathways in Alzheimer's disease using macrostructure-informed normative tractometry

Yixue Feng[1] | Bramsh Q. Chandio[1] | Julio E. Villalon-Reina[1] | Sophia I. Thomopoulos[1] | Talia M. Nir[1] | Sebastian Benavidez[1] | Emily Laltoo[1] | Tamoghna Chattopadhyay[1] | Himanshu Joshi[2] | Ganesan Venkatasubramanian[3] | John P. John[2] | Neda Jahanshad[1] | Robert I. Reid[4,5] | Clifford R. Jack[5] | Michael W. Weiner[6] | Paul M. Thompson[1] | for the Alzheimer's Disease Neuroimaging Initiative

[1]Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Marina del Rey, California, USA

[2]Multimodal Brain Image Analysis Laboratory, National Institute of Mental Health and Neuro Sciences (NIMHANS), Bangalore, India

[3]Translational Psychiatry Laboratory, National Institute of Mental Health and Neuro Sciences (NIMHANS), Bangalore, India

[4]Department of Information Technology, Mayo Clinic and Foundation, Rochester, Minnesota, USA

[5]Department of Radiology, Mayo Clinic and Foundation, Rochester, Minnesota, USA

[6]Department of Radiology and Biomedical Imaging, UCSF School of Medicine, San Francisco, California, USA

Correspondence
Paul Thompson and Yixue Feng, Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Marina del Rey, CA 90292, USA.
Email: pthomp@usc.edu and yixuefen@usc.edu

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data, but most of them did not participate in the analysis or writing of this report. A complete list of ADNI investigators may be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Abstract

INTRODUCTION: Diffusion-weighted magnetic resonance imaging (dMRI) is sensitive to the microstructural properties of brain tissues and shows great promise in detecting the effects of degenerative diseases. However, many approaches analyze single measures averaged over regions of interest without considering the underlying fiber geometry.

METHODS: We propose a novel macrostructure-informed normative tractometry (MINT) framework to investigate how white matter (WM) microstructure and macrostructure are jointly altered in mild cognitive impairment (MCI) and dementia. We compared MINT-derived metrics with univariate diffusion tensor imaging (DTI) metrics to examine how fiber geometry may impact the interpretation of microstructure.

RESULTS: In two multisite cohorts from North America and India, we find consistent patterns of microstructural and macrostructural anomalies implicated in MCI and dementia; we also rank diffusion metrics' sensitivity to dementia.

Funding information
NIA, Grant/Award Numbers: R01 AG057892, U19 AG024904; US Alzheimer's Association, Grant/Award Number: AARG-23-1149996; NIH, Grant/Award Number: R01 AG060610; Department of Science and Technology, Government of India, Grant/Award Numbers: DSTSR/CSI/73/2011 (G), DST-SR/CSI/70/2011 (G), DST/CSRI/2017/249 (G)

DISCUSSION: We show that MINT, by jointly modeling tract shape and microstructure, has the potential to disentangle and better interpret the effects of degenerative disease on the brain's neural pathways.

KEYWORDS
Alzheimer's disease, anomaly detection, deep generative models, diffusion magnetic resonance imaging, normative modeling, tractometry, transfer learning

Highlights

- Changes in diffusion tensor imaging metrics may be due to macroscopic changes.
- Normative models encode normal variability of diffusion metrics in healthy controls.
- Variational autoencoder applied on tractography can learn patterns of fiber geometry.
- WM microstructure and macrostructure are modeled with multivariate methods.
- Transfer learning uses pretraining and fine-tuning for increased efficiency.

1 | BACKGROUND

Diffusion-weighted magnetic resonance imaging (dMRI) is often used to investigate the microstructure and geometry of the brain's neural pathways.[1,2] As water diffusion is constrained along these pathways, tractography can reconstruct the 3D geometry of the brain's major fiber bundles from dMRI signals. Pathological processes in degenerative diseases, such as the loss of neurons and myelin and inflammation, affect tissue diffusion properties, resulting in changes to tissue microstructure and pathway geometry. Because of this, dMRI is sensitive to pathological processes that are not detectable with standard anatomical MRI.

Anisotropy and diffusivity measures computed from diffusion tensor imaging (DTI) are widely used to characterize white matter (WM) microstructural properties. As many types of molecular pathology influence the dMRI signal, including amyloid and tau protein accumulation in the brain, numerous studies have focused on charting WM anomalies in neurodegenerative disorders such as Alzheimer's disease (AD),[3–5] Parkinson's disease,[6] and other dementias. DTI metrics have been found to be associated with age and dementia severity,[4,7] particularly in the cingulum, fornix, and corpus callosum. A follow-up study[5] found that cortical dMRI metrics mediated the relationship between cerebrospinal fluid (CSF) markers of AD and delayed logical memory performance, which is typically impaired in early AD. Lower CSF amyloid beta 1-42 (Aβ142) and higher phosphorylated tau 181 (pTau181) were associated with cortical dMRI measures reflecting less restricted diffusion; this apparent link between AD pathology and diffusion metrics has bolstered interest in using dMRI to study AD. Even so, standard analysis methods typically reduce microstructural indices to summaries over relatively large regions of interest. These limitations have stimulated efforts to map disease effects on brain microstructure along WM tracts at a finer anatomical scale.[5,8,9]

DTI-based metrics such as fractional anisotropy (FA), radial diffusivity (RD), and axial diffusivity (AxD) are contaminated by the presence of multiple fiber populations in a single voxel[10,11] and not specific to any individual fiber population. While voxel-based, fixel-based,[12] and tractometry approaches* have been proposed to help resolve crossing fibers, microstructural measures are still calculated at the voxel level. Current tractometry methods often compute statistics for each bundle separately using a univariate approach without accounting for the complex pattern of intersecting fibers in the brain. Tractography data can also be used to study the macrostructural or "shape" properties of WM bundles. In a large multisite cohort of elderly adults, Schilling et al.[13] found heterogeneous patterns of age-related WM macrostructural changes compared to the more homogeneous microstructural changes across bundles. A recent study[14] found that early tau-related WM changes in AD may be due to fiber macrostructure using fixel-based analysis. To our knowledge, no work has examined how WM micro- and macrostructure are jointly altered in neurodegenerative conditions such as AD using tractometry methods, which we address in the current study.

This study proposes macrostructure-informed normative tractometry (MINT)† to jointly model microstructural measures and co-occurring variations in bundle geometry with a variational autoencoder (VAE). When used as a normative model, VAE can encode normal variability for diffusion metrics in healthy controls. This multivariate model integrates multiple complementary microstructural features and accounts for the statistical covariance among different dMRI metrics as well as along-tract spatial correlations. We compare MINT-derived measures with traditional univariate measures derived from DTI to investigate characteristic patterns of WM abnormalities in mild cognitive impairment (MCI) and dementia in a large multisite sample. We also examine how WM abnormalities relate to clinical measures of dementia severity. As there is interest in identifying the microstructural metrics useful for detecting and tracking dementia, we also rank

DTI metrics by evaluating their sensitivity to dementia. After visualizing WM microstructural abnormalities in dementia and MCI in two cohorts of different ancestry and demographics, we examine how they relate to overall tract geometry, noting ambiguities of interpretation that can be resolved by joint statistical modeling of microstructure and shape.

RESEARCH IN CONTEXT

1. **Systematic review**: dMRI can be used to characterize the microstructure of the brain's neural pathways or WM tracts. Metrics calculated from dMRI are sensitive to AD and clinically defined dementia. Advanced tractography methods, which delineate WM geometry, can help chart WM abnormalities in greater anatomical detail, further enhanced by deep learning methods.
2. **Interpretation**: In two cohorts of people with AD, MCI, and healthy controls from North America and India, we found consistent abnormalities in tissue microstructure and WM geometry in MCI and dementia. Joint modeling of microstructural and macrostructural measures facilitated the interpretation of disease effects.
3. **Future directions**: More advanced dMRI measures will be assessed to evaluate their relationship to additional AD biomarkers, including cerebral amyloid load and tau.

2 | METHODS

2.1 | Participant demographics

In total, we analyzed diffusion data of 1032 subjects from two cohorts – 730 participants from the Alzheimer's Disease Neuroimaging Initiative (ADNI) Phase 3 collected using seven acquisition protocols and an independent sample of 302 Indian-ancestry participants, evaluated and scanned at two scanners at the National Institute of Mental Health and Neurosciences (NIMHANS) in Bangalore, India. The subset of the ADNI cohort that was scanned with dMRI consisted of 447 cognitively normal (CN) participants, 214 diagnosed with MCI, and 69 with dementia. For the current study, we analyzed their cross-sectional scans only, although most of these participants also have longitudinal dMRI data. The NIMHANS cohort consisted of dMRI data from 123 CN participants, 89 people with MCI, and 90 with dementia. The demographic information for all subjects categorized by protocols and scanners is detailed in Table 1.

2.2 | Data preprocessing

2.2.1 | Alzheimer's Disease Neuroimaging Initiative

One of the two datasets analyzed in this article was obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of MCI and early AD.

We analyzed dMRI data from Phase 3 of ADNI, which recently concluded, from 730 participants, collected using seven acquisition protocols – S31, S55, S127, P33, P36, GE36, and GE54 (Table 1)[‡] – consisting of three scanner vendors and varying numbers of gradient directions. All protocols were single-shell except S127, and all DTI scans were derived from b0 and b1000 shells. Data from all protocols were preprocessed with the same pipeline detailed elsewhere.[4,15]

Clinical assessments used in the study for the ADNI cohort included the Clinical Dementia Rating Sum of Boxes (CDRsb, 0.97 ± 1.9),[16] the Alzheimer's Disease Assessment Scale (ADAS13, 16.28 ± 8.26),[17] and Mini-Mental State Examination (MMSE, 27.99 ± 2.91).[17,18] The CDRsb and ADAS measures are widely used as a primary outcome measure in clinical trials of anti-amyloid treatments and other interventions.[19,20] The availability of these measures listed by protocol is in Section S2.

2.2.2 | National Institute of Mental Health and Neurosciences

For the first scanner used at NIMHANS – a Philips 3T Ingenia scanner – dMRI was acquired using a single-shot, diffusion-weighted echo-planar imaging sequence (repetition time [TR] = 7441 ms, time to echo [TE] = 85 ms, total acquisition time [TA] = 630 s, voxel size: 2×2×2 mm^3, 64 slices, flip angle = 90°, field of view [FOV] 224 mm). Detailed descriptions of the acquisitions and quality assurance can be found in Parekh et al.[21] For the Siemens 3T Skyra scanner, dMRI data were also collected using a single-shot, diffusion-weighted echo-planar imaging sequence (TR = 8400 ms, TE = 91 ms, TA = 546 s, FOV 240 mm).[21–23] For both scanners, transverse sections 2 mm thick were acquired parallel to the anterior commissure–posterior commissure (AC–PC) line, and diffusion weighting was encoded along 64 independent spherical directions using a *b*-value of 1000 s/mm^2 with 1 b0 s/mm^2 for the Siemens scanner and 2 b0 s/mm^2 for the Philips scanner. Preprocessing steps for all dMRI volumes in the NIMHANS cohort closely followed the ADNI preprocessing pipeline[4] and included denoising using local principal component analysis (PCA)[24] implemented in DiPy,[25] Gibbs ringing removal[26,27] implemented in DiPy,[25] susceptibility-induced distortion, eddy current, and motion correction[28] using FSL's *topup*[29] and *eddy_cuda*[30] and bias field inhomogeneity correction[31] using MRtrix's *dwibiascorrect*. While NIMHANS Philips b0s were acquired with reverse phase-encoded blips, the Siemens protocol did not acquire dMRI scans with blip-up blip-down acquisitions; Synb0-DisCo[32] was therefore used to generate synthetic non-distorted b0 images, to estimate the distortions with *topup*. After processing dMRI, diffusion tensors were fitted at each voxel using the non-linear least-squares method to produce FA, MD, RD, and AxD scalar maps in the dMRI native space. Fiber orientations were reconstructed using Robust and

TABLE 1 Demographics for ADNI and NIMHANS data used in this study.

Protocol	Total *N*	*N* (CN)	*N* (MCI)	*N* (Dementia)	Age (Mean ± SD)	Male (%)
ADNI S31	96	59	27	10	72.03 ± 8.51	39 (40.62%)
ADNI S55	270	175	71	24	74.41 ± 7.93	120 (44.44%)
ADNI S127	112	68	36	8	73.66 ± 7.58	53 (47.32%)
ADNI P33	56	37	14	5	75.74 ± 7.68	34 (60.71%)
ADNI P36	46	20	23	3	73.14 ± 6.93	24 (52.17%)
ADNI GE36	42	21	17	4	72.24 ± 7.15	24 (57.14%)
ADNI GE45	108	67	26	16	76.11 ± 8.06	55 (50.93%)
ADNI Total	730	447	214	69	74.13 ± 7.93	349 (47.81%)
NIMHANS Siemens	192	91	45	56	67.02 ± 8.26	111 (57.81%)
NIMHANS Philips	110	32	44	34	67.65 ± 7.13	66 (60.0%)
NIMHANS Total	302	123	89	90	67.25 ± 7.86	177 (58.61%)
Total	1032	570	570	159	72.12 ± 8.5	526 (50.97%)

Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; CN, cognitively normal; MCI, mild cognitive impairment; NIMHANS, National Institute of Mental Health and Neurosciences.

Unbiased Model-Based Spherical Deconvolution (RUMBA-SD) along with contextual enhancement.[33,34] Particle filtering tracking[35] was used to generate whole-brain tractograms, with eight seeds per voxel generated from the WM mask, a step size of 0.2 mm, angular threshold of 30°, and the continuous map stopping criterion.

The NIMHANS Neuropsychological Battery for the Elderly (NNB-E)[36] was developed for older Indian adults with considerations for cultural background and language. To assess participants' episodic memory, NNB-E measures used in this study included word list delayed recall (WLDR: 3.61 ± 3.12) for verbal memory and story memory immediate recall (SMIR: 7.3 ± 3.59) and delayed recall (SMDR: 5.12 ± 4.3) for logical memory. The availability of these three measures listed by protocol is in Section S2.

Comparing the two cohorts, ADNI (74.13 ± 7.93) was significantly older than the NIMHANS cohort (67.25 ± 7.86); $t(1030) = 12.72$, $p = 1.5 \times 10^{-34}$. The proportions of diagnosis were significantly different across the two cohorts: $X^2(2, N = 1030) = 73.68$, $p = 1{\times}10^{-16}$. For the ADNI cohort, diagnosis is determined based on clinical assessments, including CDR and MMSE. For the NIMHANS cohort, diagnosis was determined according to the National Institute on Aging and Alzheimer's Association (NIA-AA) criteria and CDR, where CDR = 0.5 is MCI and CDR = 1 denotes dementia. Future work will assess congruence of these historical definitions of AD with recent biological definitions of AD,[37] which are rapidly evolving.

2.2.3 | TractoInferno

We used a deep learning method to encode the mathematical properties of normal statistical variation in the brain's WM bundles, assisted by auxiliary information from another high-quality dataset. Such an auxiliary dataset is referred to as a pretraining dataset, which can help robustly encode and estimate normal variation in DTI measures and tract geometry in healthy subjects. In our study, this pretraining dataset consisted of 198 single-shell dMRI volumes already preprocessed from the publicly available TractoInferno training dataset,[38] acquired on 3T scanners with *b*-values around 1000 s/mm^2 from six studies.[§] Constrained spherical deconvolution[39] and deterministic tracking[25] were used to generate whole-brain tractograms.

2.2.4 | Bundle segmentation

For all ADNI, NIMHANS, and the TractoInferno datasets, we used the bundle analytics (BUAN) tractometry framework[8] for bundle segmentation, subsequent bundle profiling, harmonization, and statistics (detailed in Sections 2.3.3 and 2.4). DiPy's[25] autocalibrated RecoBundles method[8,40] was used to segment 30 WM bundles using the population-averaged HCP-842 atlas[41] in both the native and Montreal Neurological Institute (MNI) space (Figure 1 and Section S3). These 30 bundles have been successfully reconstructed in single-shell dMRI data from participants with Parkinson's disease and bipolar disorder.[8,42] The autocalibration step[8] performs additional pruning and can help reduce false positives and extract bundles of various shapes and sizes without manually tuning parameters. Bundles with fewer than 96 streamlines were removed.

2.3 | Macrostructure-informed normative tractometry

2.3.1 | Macrostructure modeling

In our macrostructure-informed normative tractometry (MINT) framework, a key component is the VAE model, which learns patterns of macrostructural (tract shape and geometry) variation in a healthy population based on the 3D streamline coordinates provided by tractography in addition to the microstructural features along those coordinates.

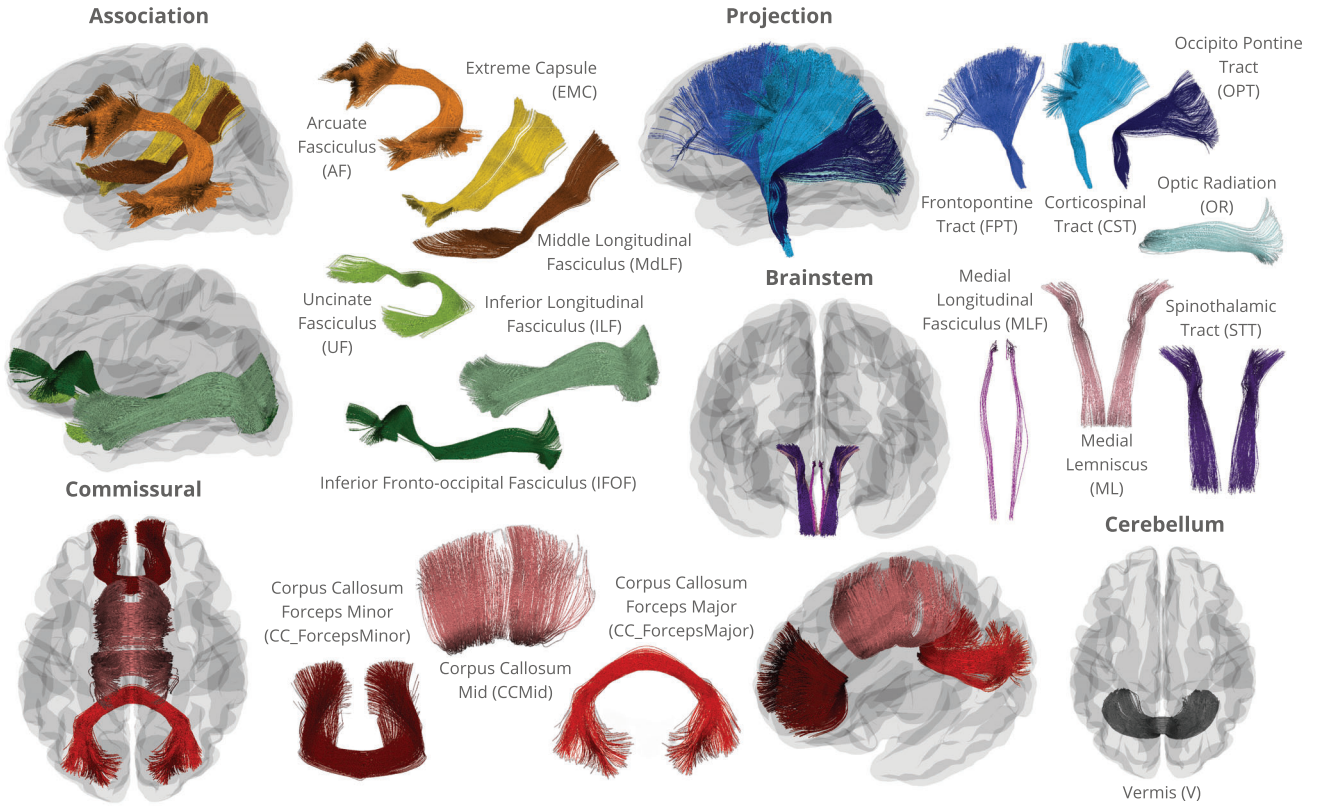

FIGURE 1 WM bundles from HCP-842 atlas[15] used in RecoBundles[16] in this study. Bundles visualized individually outside of the glass brain are not to scale. While bilateral association and projection bundles were used in this study, only those in the left hemisphere are shown.

VAE[37] is considered to be one of many non-linear extensions of dimensionality reduction methods, such as PCA, where the variation in a multidimensional dataset is often encoded using a small number of latent factors. These underlying factors represent a probability distribution over the input data and enable statistical inference in a fully generative model.

In MINT, we use 1D convolutional layers to model streamline-level information. This method was proposed to enforce sequential dependency of 3D points along a streamline.[43] Illustrated in Figure 2A, a 1D convolution learns from local neighborhood points along a streamline using a "sliding window," and multiple convolutional layers can be stacked to learn information from points farther away. This approach has been successfully applied to problems such as streamline filtering, clustering,[43,44] and generative sampling of streamlines for bundle segmentation.[45]

MINT can learn the patterns of statistical variation in whole-brain macrostructure by training it on streamlines (3D curves) from multiple bundles. While prior studies used the term *macrostructure* to describe bundle-level shape information, here we use this term to describe the overall, whole-brain fiber geometry derived from tractography (including information from multiple bundles). We illustrate this in Figure 2B showing examples of projection fibers (1, 2, and 4) and commissural fibers (3). If each streamline were modeled independently, with rotational invariance enforced, 1D convolution would be able to distinguish streamlines 1, 2, and 4 from streamline 3 due to their different lengths and curvatures. Still, it may have difficulty distinguishing between streamlines 2 and 4. On the other hand, if streamlines from multiple bundles are normalized with respect to the whole brain, the model can learn that streamlines 1 and 2 are more likely to belong to the same bundle and share similar microstructural properties than streamlines 1 and 4, because of their global coordinates. In the case of a fiber crossing, illustrated here for streamlines 3 and 4, the model can encode information on their spatial proximity. Through joint statistical modeling of micro- and macrostructure, variance introduced by the underlying fiber geometry – which may also influence microstructural measures such as FA – could be attributed to a difference in macrostructure instead. We previously showed that using this approach, a VAE with 1D convolutional layers (ConvVAE) preserves streamline and bundle distances in the latent space by training on only 3D coordinates.[46] Through unsupervised training, the model can learn that streamlines with similar shapes and in similar positions likely belong to the same bundle.

As shown in Figure 2C, MINT is trained on whole-brain macrostructure as well as the four standard DTI microstructural measures. By applying whole-brain normalization, MINT leverages the flexibility offered by streamline-based deep neural networks during training and inference. The model can produce a reconstruction for any collection of streamlines during inference, and the deviation score derived from the reconstruction error is calculated for every point on each streamline. The reconstruction error describes how well a new sample fits

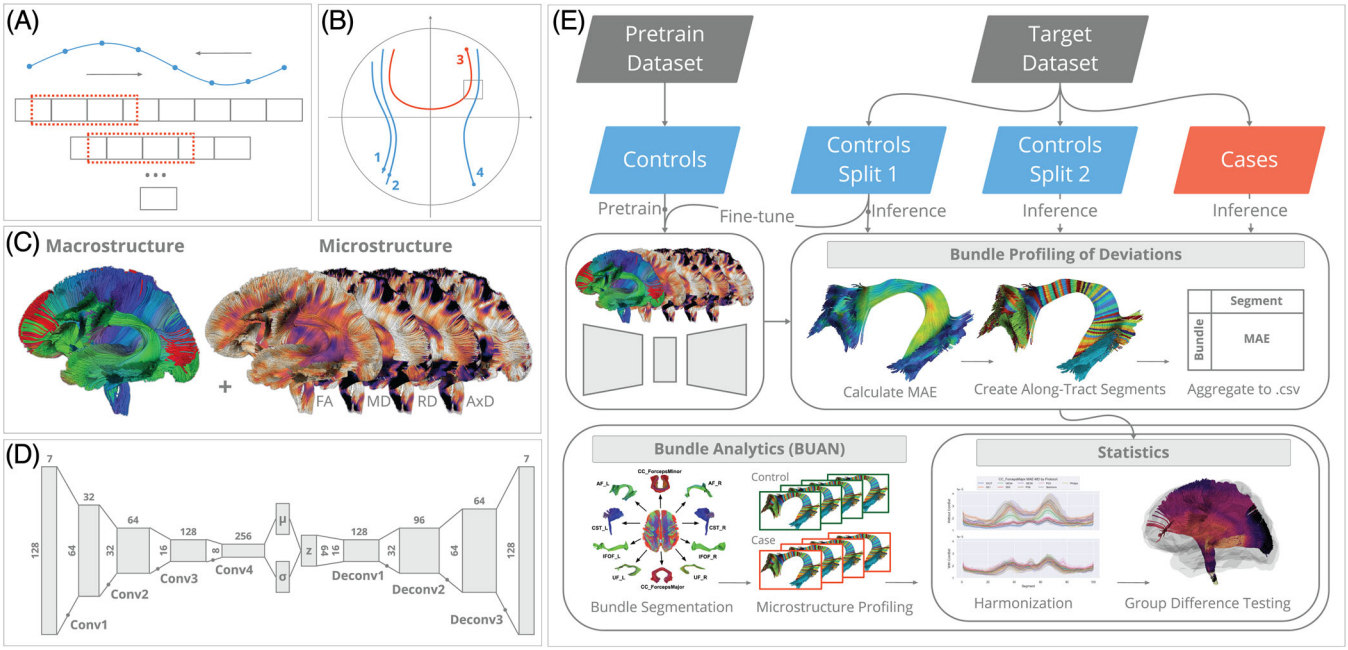

FIGURE 2 Modeling whole-brain macro- and microstructure in MINT. (A) MINT uses 1D convolutional layers in a VAE to model the sequential dependency along each streamline. (B) MINT models whole-brain fiber geometry by training the VAE on streamlines from multiple fiber bundles. Proximity and similarity between streamlines are encoded using global coordinates. (C) Whole-brain macrostructure is represented using tractography data, and whole-brain microstructure is represented by projecting scalar maps of DTI metrics onto every point on each streamline. Both types of data are combined to train the VAE model. (D) The VAE is a generative model, trained via backpropagation, to reconstruct its input data as accurately as possible while passing all the intermediate data through a narrow bottleneck layer (z in middle). VAE enforces the data to be optimally compressed into a latent set of parameters, and the means μ and covariance σ of these parameters follow a multivariate Gaussian distribution. This enables the rich complexity of variations in fiber microstructure and 3D geometry to be modeled using a compact representation. (E) MINT framework. A VAE model is first trained on healthy controls from a pretraining dataset (TractoInferno). For each target dataset (ADNI, NIMHANS), the controls are split into two sets (Split 1 and Split 2), and Split 1 is used to fine-tune the VAE model. We then perform model inference on all data from the target dataset to obtain their reconstruction for subsequent bundle profiling. For each bundle, the point-wise deviation score (MAE) is indexed using 100 along-tract segments created to align data across subjects. MAE for all points within each segment is averaged to create a bundle profile of deviation scores. The BUAN statistical framework is used for harmonization and group difference testing. In the harmonization step, bundle profiles created from data in Split 1 are used to train ComBat,[35,36] whose parameters are applied to bundle profiles from healthy controls in Split 2 and the cases (participants with MCI and AD). Harmonized bundle profiles are then used for group difference testing using linear regression. AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; MCI, mild cognitive impairment; NIMHANS, National Institute of Mental Health and Neurosciences; BUAN, bundle analytics; DTI, diffusion tensor imaging; MAE, mean absolute error; MINT, macrostructure-informed normative tractometry; VAE, variational autoencoder.

the distribution of the training data and can serve as a normative baseline when calculated on healthy controls. We expected the reconstruction error to be higher for data with abnormalities. This tactic of using reconstruction error from a VAE, or a related generative model, has been widely employed for anomaly detection in both biomedical and industrial contexts.[46–48] Next, we created along-tract segments aligned across subjects to localize tract-specific deviations. The methods to create along-tract profiles for these deviations are detailed in Section 2.3.3.

2.3.2 | Data preparation and training

Streamline coordinates used for training were derived from 30 bundles extracted in the MNI space coregistered across subjects. DTI metrics, FA, MD, RD, and AxD, were projected onto the corresponding bundle in the native space. Each streamline was resampled to 128 equidistant points – so that the resulting set of points could be used with 1D convolutional layers – with seven input features – x, y, z, FA, MD, RD, and AxD. To preserve the spatial relationships between streamlines within each subject, unit sphere normalization was applied to the 3D coordinates for all streamlines in each subject's bundles, using a centroid and radius calculated from the HCP-842 atlas tractogram.[41] Min-max scaling was applied to each of the four DTI metrics using preset ranges, and all seven input features were scaled between −1 and 1. As streamlines are also "flip-independent" – that is, one streamline can be traversed in either direction – we performed random flipping as an augmentation step[45] to further enforce this property, although 1D convolution can learn information from neighboring points in both directions.

As the VAE model was used here in a normative model setting, we need to be careful not to overparameterize the model to avoid reconstructing anomalies at inference time. One-dimensional

depth-wise separable convolution (SepConv)[49] was used in the encoder and decoder to accommodate a deeper feature space with efficient use of model parameters. Overall, the VAE model consists of 464,408 parameters: the encoder consists of four 1D SepConv layers with kernel sizes of 15, and 32, 64, 128, and 256 channels; the decoder consists of three 1D SepConv layers with the same kernel sizes of 15, and 128, 96, and 64 channels (Figure 2D). This asymmetric architecture, where the decoder is smaller than the encoder, was adopted to encourage the model to learn more robust representations[50] while reducing the number of parameters. The model was trained with the evidence-based lower bound (ELBO) loss.[51] We used cyclical annealing on the Kullback—Leibler (KL) divergence term[52] to prevent posterior collapse, a common issue during VAE training. The pretraining data, consisting of streamlines from 30 bundles from 198 subjects in the TractoInferno dataset, contains over 100 million streamlines. To improve computational efficiency during data loading, the VAE model sampled a batch of 512 streamlines, with replacement, for each parameter update and was pretrained for 24,000 steps.

For both the ADNI and NIMHANS datasets, data from the CN participants were split into independent subsets of size 35%/15%/50% for training/validation/testing. Data from participants with MCI or dementia were concatenated with the 50% CN subjects to create the test set. The test set contained a large percentage of CN subjects – 226 subjects from ADNI and 62 from NIMHANS – to better estimate group difference statistics from the resulting deviation scores. The validation set was used to adjust the model's fine-tuning parameters, such as batch size and number of training steps, and inspect the reconstruction quality. The pretrained VAE model was fine-tuned on 30 bundles from the subjects in the training set, in each cohort, for 5000 steps. We did not pool the training data in ADNI and NIMHANS together in this step, to simulate the scenario where sites cannot share raw data, but can still share bundle profiles computed from their own fine-tuned model. A complete diagram of the MINT framework – including training, inference, bundle profiling, and group difference testing – is shown in Figure 2E.

2.3.3 | Bundle profiling

At inference time (ie, when the model is applied to new datasets without updating its parameters), each bundle X from the test set of participants in both ADNI and NIMHANS was passed through its respective fine-tuned model, to obtain their reconstruction X' of the same shape. We then used the bundle profiling approach from BUAN[8,15] to create 100 along-tract segments, aligned across subjects. Every point in X was assigned a segment index (1 to 100), and the same index was applied to the reconstructed bundle. We then calculated the deviation using mean absolute error (MAE) for each segment i and feature k as follows:

$$\mathrm{MAE}_{ik} = \frac{\sum_{j=1}^{N_i} \left| X'_{ijk} - X_{ijk} \right|}{N_i}.$$

Here, N_i is the number of points assigned to segment i in X, and the features include shape (calculated from 3D coordinates), FA, MD, RD, and AxD. To compare the MAE scores with the original DTI microstructural measures, we also calculated the mean along-tract profile M as follows:

$$M_{ik} = \frac{\sum_{j=1}^{N_i} X_{ijk}}{N_i},$$

where the features contain four DTI measures. For each subject's bundle, X, we created along-tract profiles for a total of nine features: MAE scores were calculated for five features at 100 along-tract segments, which we will refer to as MAE-Shape, MAE-FA, MAE-MD, MAE-RD, and MAE-AxD, to distinguish them from the mean along-tract DTI profile, calculated for four features at the same 100 along-tract segments – DTI-FA, DTI-MD, DTI-RD, and DTI-AxD. The same bundle profiling procedures were repeated for the training set in both cohorts for ComBat harmonization (detailed next).

Along-tract bundle profiles for all nine features were harmonized using ComBat using the bundle-wise approach described in Chandio et al.[15] ComBat is a "location and scale" model, where we assume that each measurement is affected by the additive and multiplicative effects of the scanner and protocol,[15,53] estimated using an empirical Bayes method. Covariates can also be specified in ComBat to preserve the known biological variability across subjects (eg, age and sex). In this study, ComBat was run for each of the 30 bundles and nine features, by pooling information across 100 segments of each bundle, with age and sex modeled as covariates. We used the Python implementation of ComBat (https://github.com/Jfortin1/neuroCombat). To prevent data leakage,[54,55] we estimated ComBat parameters on bundle profiles created from the training set for both cohorts previously used in fine-tuning. The training set contained only CN subjects, so the diagnosis was not included in the ComBat model during training. The estimated ComBat parameters were then applied to bundle profiles from the test set, which included CN and subjects with MCI and dementia, for subsequent statistical analysis.

2.4 | Statistical analysis of diagnostic group differences

To evaluate the effects of MCI and dementia on MINT-derived along-tract MAE and DTI measures and how they differed between the ADNI and NIMHANS cohorts, we used the BUAN statistics framework[8] to investigate the association of nine diffusion measures calculated in Section 2.3.4 with diagnosis. A linear regression model was fitted for each of the 100 segments in 30 bundles with harmonized diffusion metrics modeled as the dependent variable, and diagnosis (DX), age, and sex as independent variables ($MAE \sim Age + Sex + DX$, $M \sim Age + Sex + DX$). The CN group was used as the reference group against which we compared MCI and dementia. Linear regression was fitted separately for each cohort to compare MCI and dementia effects between ADNI and NIMHANS. We corrected for multiple comparisons using the

false discovery rate (FDR) method[19] applied to 3000 tests across the whole brain. We then repeated this procedure on bundle profiles for the dementia group with MCI as the reference. We also conducted an additional analysis on diffusion measures from the merged cohorts. Although DTI and MAE measures were harmonized with ComBat across all nine protocols, regional differences in demographics and cognitive assessments across cohorts might have affected diagnosis. Instead of a linear regression, we fitted a linear mixed model for the merged analysis where cohort was modeled as a random effect, and age, sex, and DX were modeled as fixed effects. Only bundle profiles from the test set were used for statistical analysis in this step and Section 2.5.

There is considerable interest in determining which dMRI-derived metrics are most effective in picking up subtle and distributed effects of the disease in the brain. As the statistics are defined at multiple locations along a tract, we then used a quantile-quantile (Q-Q) plot to gauge how much of the tract shows significance and at what level. These plots are cumulative histograms of the p values of association between dementia diagnosis and diffusion metrics for all segments in a tract from the linear regression model. If a curve rises sharply from the origin in the Q-Q plot, it means that a large proportion of the tract has very low p values. They can be used to rank and summarize the extent of significant associations for different brain biomarkers.

2.5 | Associations with clinical metrics

To assess how MAE and DTI measures were associated with standardized clinical measures, we computed partial correlations between cognitive assessment scores and nine bundle profile features, removing the effect of age and sex. For the ADNI cohort, we used CDRsb, ADAS13, and MMSE; for the NIMHANS cohort, we used three measures from NNB-E: SMIR and SMDR tests and WLDR. Each bundle profile from the test set was averaged to obtain one scalar value per bundle per feature, and participants with missing measures were removed from the corresponding analysis. Correlation was calculated using the Spearman rank-order correlation, and we report ρ correlation coefficient and its 95% confidence interval. We used the partial correlation implementation from the Pingouin Python package version 0.5.4.[56]

3 | RESULTS

3.1 | Visualizing MCI and dementia effects on WM tracts

In Figures 3–5, we visualize along-tract group differences (i.e., the regression coefficient for the effect of diagnosis on microstructure, β(DX) of MD and FA measures), for both MCI versus CN and dementia versus CN comparisons in the ADNI and NIMHANS cohorts. The along-tract β(DX) for the remaining diffusion measures and corresponding $-\log_{10}P$ values after FDR correction are available in Section S4.1. Results for dementia versus MCI comparisons by cohort are in Section S4.2, and results from the merged cohort analysis are in Section S4.3. Quantitative results, including FDR-corrected p values, β(DX), and standard errors per bundle segment are available in Supplementary Tables. All visualizations in these figures were created using the FURY package.[25,57,58] As expected from the intensification and spread of disease pathology of MCI and dementia, we see greater effects of dementia than MCI in almost all WM regions for all nine measures compared to CN. Overall, mean diffusivity (MD), which is commonly reported as increased in dementia, was higher in MCI and higher still in dementia. Consistent with the temporal lobe atrophy, that is, typical of both MCI and dementia, the abnormal elevation in MD showed the strongest effects in the temporal lobe pathways (inferior longitudinal fasciculus (ILF), temporal uncinate fasciculus (UF), and arcuate fasciculus (AF) bundles), and in the corpus callosum at the midline.

For both DTI and MAE diffusivity measures in general, higher MD (Figure 3), RD, and AxD (Section S4.1) were found in dementia and MCI compared to control participants. Lower DTI-FA (Figure 4A), higher MAE-FA (Figure 4B), and higher MAE-Shape (denoting a bundle's geometrical distortion compared to its normal configuration learned from the VAE model; Figure 5) are found in dementia and MCI compared to CN in most brain regions. For all MAE metrics, these findings are consistent with our expectation that statistical deviation from the norm is greater in the MCI and dementia groups than in the CN group. Trends and directions of deviation for DTI metrics are also consistent with prior reports on regional DTI measures in AD.[7,59–62] Patterns of dementia and MCI effects were also mostly consistent across the two cohorts we examined, with stronger statistical significance in ADNI compared to NIMHANS.

For MAE-MD and DTI-MD, we see widespread effects of dementia across the whole brain, with the largest effects in the posterior and temporal regions, including the OPT and OR bundles in the projection pathways, the ILF, posterior regions of the MdLF and EMC and temporal regions of the UF and AF bundles in the association pathways, and the forceps major in the commissural pathways. The patterns are largely similar for both DTI-MD and MAE-MD and across the two cohorts. We see smaller effects in NIMHANS than ADNI, except for the brainstem bundles – where effects could be inflated due to the ComBat harmonization (Section S5).

Without considering crossing fibers, the overall diffusivity can be decomposed into a diffusivity along the axons (AxD) and a component transverse to axons (RD). In the context of neurodegeneration, the latter is increased when cell membranes are damaged, neurons are lost, or myelin degrades. MCI and dementia effects on MAE-RD were similar to those found for MAE-MD in all regions, whereas their effects on MAE-AxD and MAE-MD are more similar in some bundles in the association, brainstem, and projection pathways. The ADNI cohort showed greater MCI and dementia effects with stronger significance for the three MAE diffusivity measures in most regions relative to the NIMHANS cohort. The difference across cohorts is most pronounced for MAE-AxD. Compared to ADNI, in the NIMHANS cohort, we found a smaller effect of dementia on MAE-AxD in the temporal association pathways; this cross-cohort difference is larger compared to MAE-MD and MAE-RD

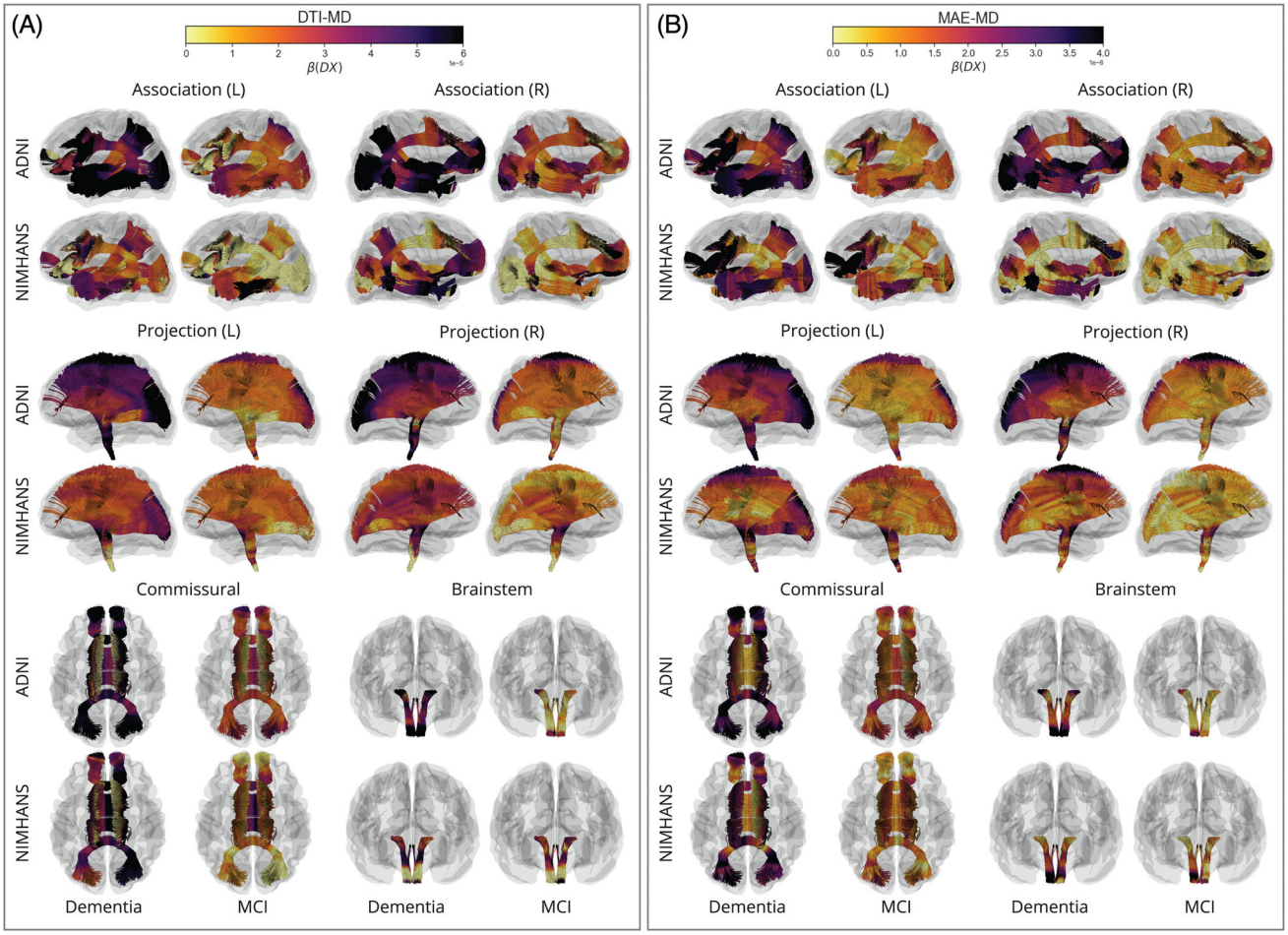

FIGURE 3 Effects of dementia and MCI on MD. Here we show along-tract effects of diagnostic group by color-coding the regression coefficients associated with diagnostic group, β(DX), for DTI-MD and MAE-MD, for dementia versus CN and MCI versus CN in the ADNI and NIMHANS cohorts, categorized by WM pathways. The colors are mapped using different scales for MAE and DTI values, as MAE measures – calculated from reconstruction error – generally are smaller in magnitude than the DTI measures. The plots show patterns of effects but are not intended for direct comparison of effect sizes across measures, except across three corresponding diffusivity measures where the color map range is the same. ADNI, Alzheimer's Disease Neuroimaging Initiative; CN, cognitively normal; DTI-MD, diffusion tensor imaging-mean diffusivity; MAE, mean absolute error; MCI, mild cognitive impairment; MD, mean diffusivity; NIMHANS, National Institute of Mental Health and Neuro Sciences; WM, white matter.

measures in the same regions. In the commissural pathways, however, the NIMHANS cohort shows a higher effect of dementia on DTI-AxD compared to ADNI in the forceps major and the middle section of the CCMid bundle.

For the univariate DTI measures, all three diffusivity measures were increased in both dementia and MCI across the two cohorts, with similar patterns in the association, commissural, and brainstem pathways across measures. Projection pathways showed a widespread elevation in DTI-MD in both dementia and MCI, but the patterns of increased DTI-RD and DTI-AxD differ along the lengths of the tracts. The sharp increase of DTI-RD from the proximal parts closer to the deep nuclei and brainstem to the distal parts closer to the cortex – and the increase of DTI-AxD in the opposite direction – result in a fan-shaped pattern of the projection pathways. This pattern is most pronounced for DTI-FA (Figure 4A), where FA is lower in the distal parts in dementia and higher in the proximal parts for both dementia and MCI.

This fan-shaped pattern is more prominent for the DTI metrics (FA, RD, and AxD) than their counterpart MAE metrics. The primary difference between DTI and MAE metrics is that they are modeled jointly along with macrostructure in MINT. Notably, patterns of dementia effects on DTI-FA are more similar to dementia effects on MAE-Shape than MAE-FA. The fan-shaped pattern in the projection pathways is seen for both MAE-Shape and DTI-FA, where the proximal parts show lower MAE-Shape and higher DTI-FA and distal parts show higher MAE-Shape and lower DTI-FA in both dementia and MCI. Aside from the projection pathways, the similarity between MAE-Shape and DTI-FA is also seen in the posterior and temporal regions of the association pathways and the forceps major in the commissural pathways. After accounting for the effects of whole-brain macrostructure using MINT, patterns of dementia and MCI differences in MAE-FA show the greatest deviation from their DTI counterpart compared to the three diffusivity measures. Most bundle segments that showed significant

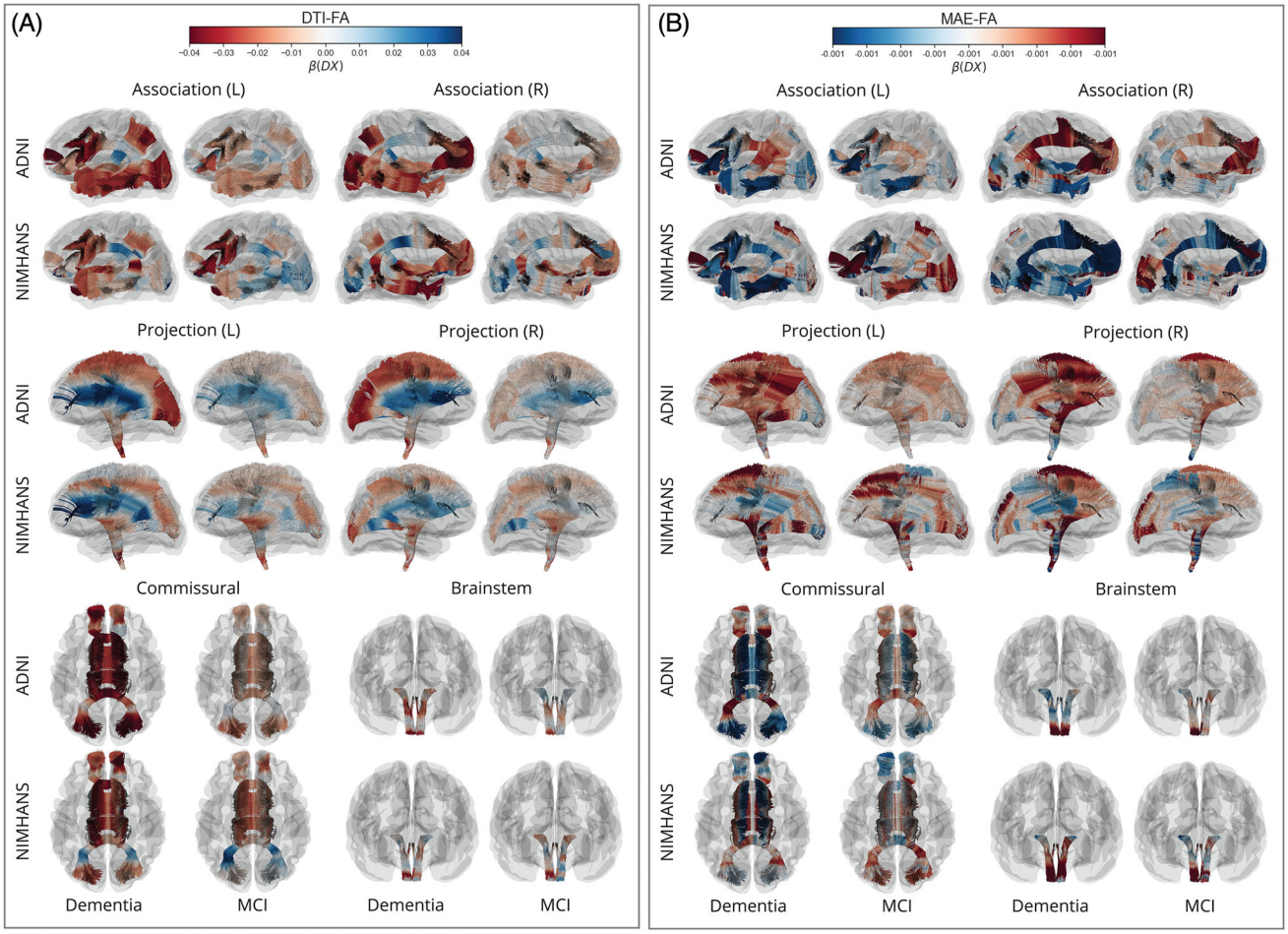

FIGURE 4 Effects of dementia and MCI on FA. Along-tract β(DX) for DTI-FA and MAE-FA, for dementia versus CN and MCI versus CN in the ADNI and NIMHANS cohorts, categorized by WM pathways. The color map is flipped for DTI-FA compared to MAE-FA and MAE-Shape in Figure 5, to better illustrate disease effects on these measures (red). ADNI, Alzheimer's Disease Neuroimaging Initiative; CN, cognitively normal; DTI-FA, diffusion tensor imaging-fractional anisotropy; FA, fractional anisotropy; MAE-FA, mean absolute error-fractional anisotropy; MCI, mild cognitive impairment; NIMHANS, National Institute of Mental Health and Neurosciences; WM, white matter.

group differences in DTI-FA no longer pass the significance threshold after FDR correction for MAE-FA; the remaining significant differences are primarily in the projection and commissural pathways.

The fan-shaped pattern in the projection pathways is seen in both ADNI and NIMHANS cohorts for DTI-RD, DTI-AxD, and DTI-FA, with highly similar patterns across cohorts. For MAE measures, however, MCI and dementia participants from the ADNI cohort show a more homogeneous along-tract increase in MAE-RD, MAE-AxD, and MAE-FA, whereas the along-tract patterns are not consistent in the NIMHANS cohort. The NIMHANS cohort shows larger dementia and MCI effects on MAE-Shape in the distal parts of the projection pathways compared to ADNI, but smaller effects in the association and commissural pathways.

3.2 | Association of diffusion metrics with clinical metrics

We plotted the partial Spearman rank-order correlation coefficient ρ and its 95% confidence interval for each bundle to assess the association of nine diffusion-derived measures (described in Section 2.3.3) averaged over 100 segments per bundle with CDRsb measures of dementia severity for the ADNI cohort and WLDR for the NIMHANS cohort (Figure 6). We compared the four DTI-derived measures with their corresponding MINT-derived measures, as well as MAE-Shape. For the ADNI cohort, both the DTI and MAE diffusivity measures were very similar in their associations with CDRsb (Figure 6A), MMSE, and ADAS13 (Section S6.1), where each metric showed a significant positive correlation in most bundles, except IFOF_L and V. MAE diffusivity measures aligned most with DTI measures for MD, followed by RD and AxD, where the association with RD measures was slightly higher for MAE in the association and projection pathways, and the association with AxD measures was slightly lower for MAE in the projection pathways.

For the NIMHANS cohort, however, MAE diffusivity measures showed a significant negative correlation with WLDR in more bundles than DTI measures (Figure 6B). The differences between the corresponding DTI and MAE diffusivity measures were more prominent than those for CDRsb in the ADNI cohort. Significant negative correlations with WLDR for MAE-MD and MAE-RD were widespread in the

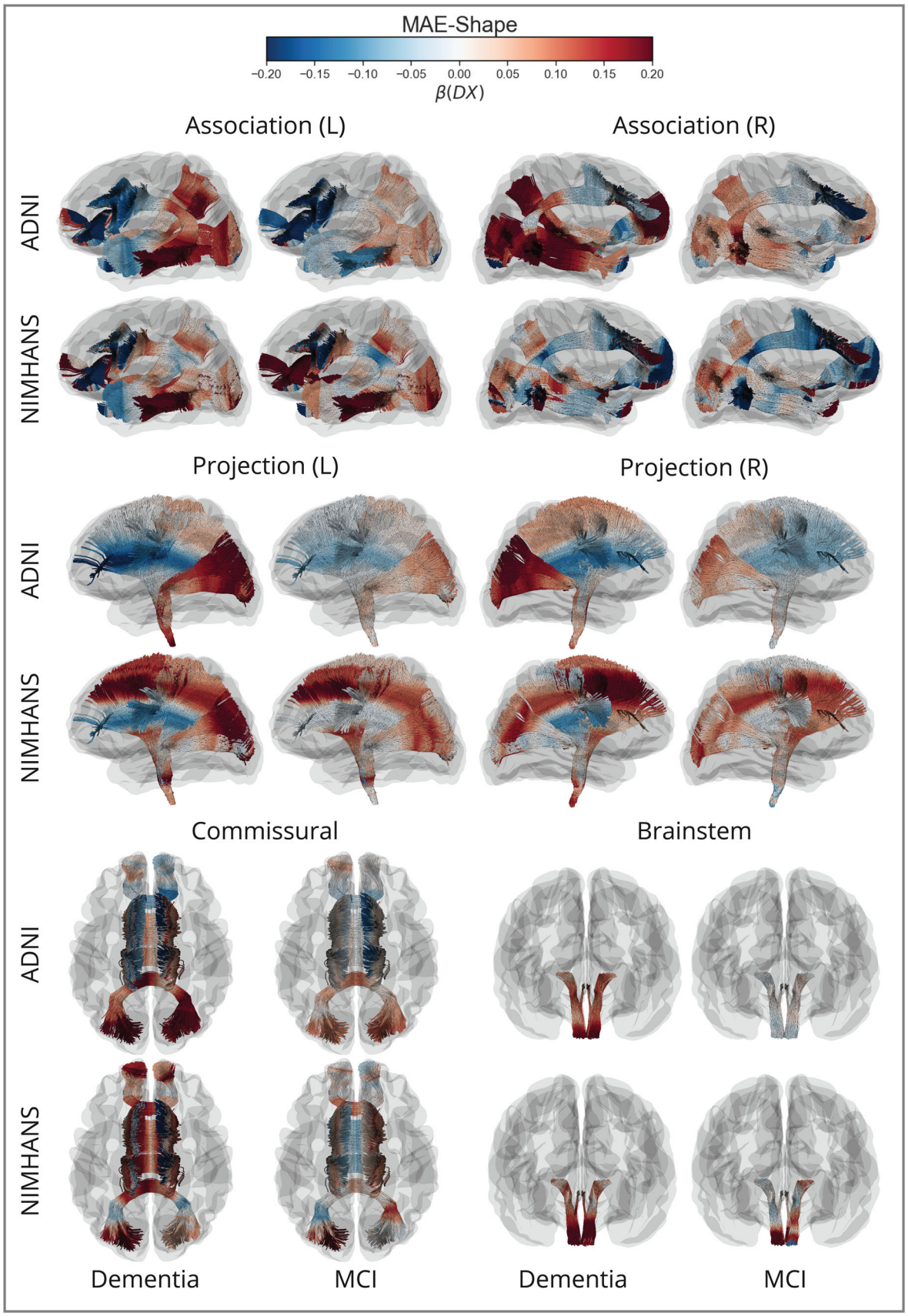

FIGURE 5 Effects of dementia and MCI on MAE-Shape. Along-tract β(DX) for MAE-Shape, for dementia versus CN and MCI versus CN in the ADNI and NIMHANS cohorts, categorized by WM pathways. ADNI, Alzheimer's Disease Neuroimaging Initiative; CN, cognitively normal; MAE, mean absolute error; MCI, mild cognitive impairment; NIMHANS, National Institute of Mental Health and Neurosciences; WM, white matter.

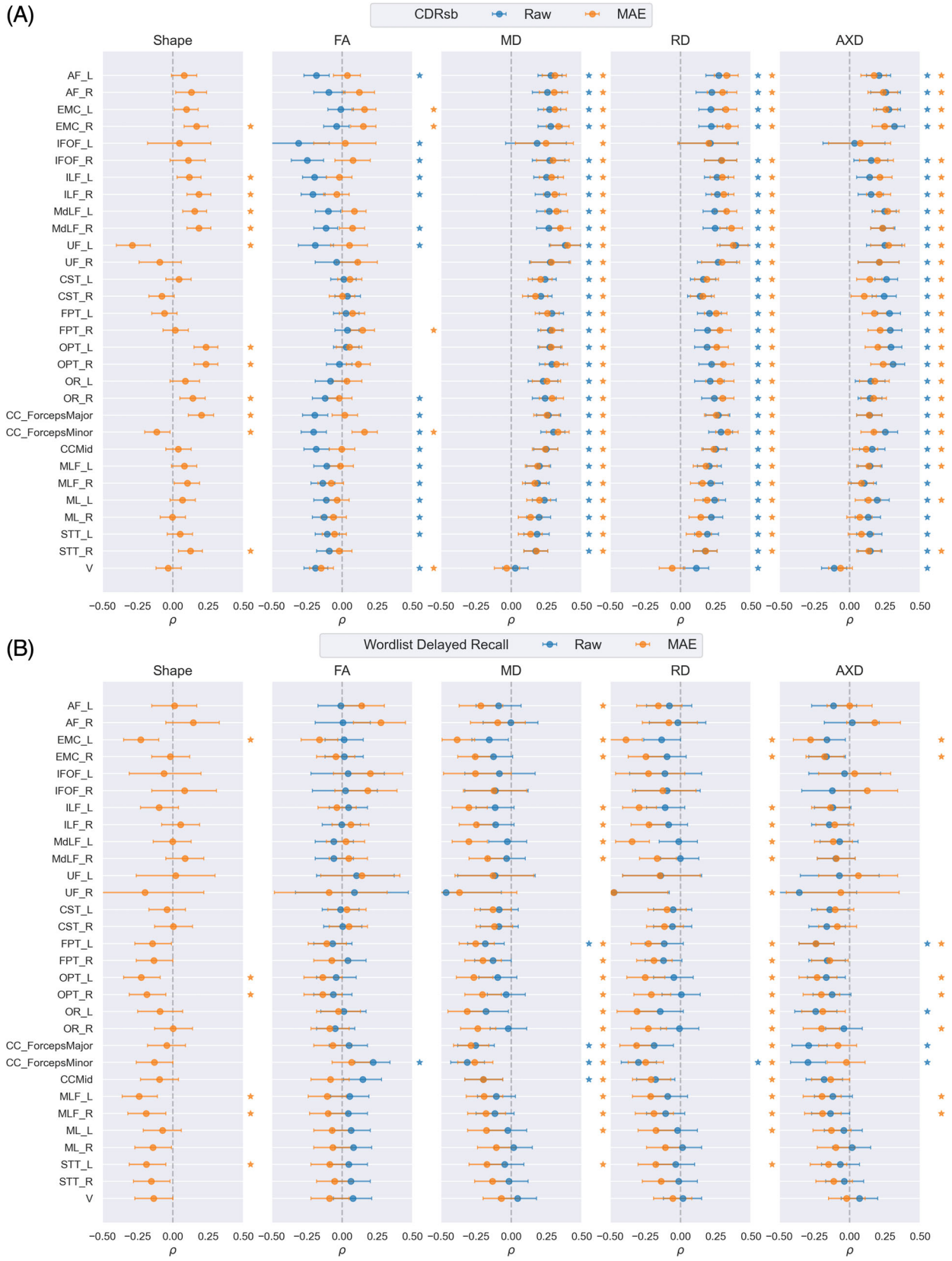

FIGURE 6 Partial Spearman rank-order correlation coefficient (ρ) and its 95% confidence interval for all 30 bundles between nine diffusion measures and CDRsb (A) and WLDR (B), calculated in Section 2.3.3. The corresponding DTI and MAE measures for each anisotropy and diffusivity metric are shown in the same panel, and MAE-Shape is shown separately. Correlation coefficients that are statistically significant after FDR correction are marked with a star to the right of each plot with the corresponding color. CDRsb, Clinical Dementia Rating Sum of Boxes; DTI, diffusion tensor imaging; FDR, false discovery rate; MAE, mean absolute error; WLDR, word list delayed recall.

posterior and temporal association pathways (EMC, ILF, and MdLF), posterior projection pathways (FPT, OPT, and OR), and corpus callosum. Significant associations with SMDR (Section S6.2) – although less widespread compared to WLDR – were seen in more bundles for MAE-MD and MAE-RD compared to the corresponding DTI measures.

Since we expected DTI-FA to be lower and MAE-FA to be higher in participants with greater cognitive impairment, correlations for these two measures had opposite signs for many bundles. There was a weak correlation between MAE-FA and all clinical metrics, lower than that for DTI-FA in most bundles, except for EMC and FPT, where associations with CDRsb and ADAS13 were comparable or higher. No bundle, except for the forceps minor, exhibited significant correlation between FA measures with WLDR and SMIR. Notably, MAE-Shape showed a significant positive correlation with CDRsb, MMSE, and ADAS13 in the association and projection pathways and a significant negative correlation with WLDR in EMC_L, OPT, MLF, and STT_L. Overall, EMC showed statistically significant associations with more clinical metrics in more diffusion measures than any other bundles, followed by OPT and forceps minor.

Overall, MAE measures were comparable to DTI measures in their associations with clinical metrics for the ADNI cohort. MAE-MD and MAE-RD outperformed the corresponding DTI measures in their association with episodic memory for the NIMHANS cohort. For metrics from both cohorts, MAE-Shape was comparable to DTI-FA in their association with clinical metrics, indicating that macrostructural changes may be associated with neurodegeneration pathologies.

3.3 | Comparing DTI- and MINT-derived diffusion measures

To evaluate how DTI- and MINT-derived measures differed in detecting group differences, we display a Q-Q plot of $-\log_{10}(P)$ for the dementia versus CN comparison in the ADNI cohort in Figure 7. We selected ILF_L, OPT_L, and CC_ForcepsMajor – three WM bundles from each association, projection, and commissural pathways – which showed significant dementia effects in most along-tract diffusion measures (Section 3.1). For both DTI- and MINT-derived measures, MD and RD (shown in orange and green colors) associated most strongly with dementia compared to AxD and FA for all three bundles. This is consistent with prior work examining dMRI metrics in regions of interest, where MD and RD typically showed the strongest association with dementia.[7] The ranking of DTI-AxD and DTI-FA differed by bundle, with DTI-AxD being more sensitive to dementia effects than DTI-FA in OPT_L and less in CC_ForcepsMajor. For MAE measures, however, MAE-FA was the weakest at detecting dementia effects in all bundles, followed by MAE-Shape and MAE-AxD. MAE-Shape outperformed both MAE-FA and MAE-AxD in OPT_L and CC_ForcepsMajor. These findings might indicate that metrics describing diffusion along the principal direction of axons (FA and AxD) may be more prone to crossing fibers and macrostructural changes (characterized by MAE-Shape) than MD and RD. We also observed that the difference in statistical significance between corresponding MAE and DTI measures varied across bundles, where the p values changed most, and for most measures, for CC_ForcepsMajor compared to ILF_L and OPT_L. These differences may reflect a bundle's sensitivity to macrostructural changes, which is accounted for in MINT, as some variations in DTI metrics may be explained away by MAE-Shape. While MAE-MD and MAE-RD remained most sensitive to dementia effects compared to other MINT-derived measures, they were less sensitive than their DTI counterparts for CC_ForcepsMajor and OPT_L. If macrostructural changes were disentangled from such compound measures, it could be expected that MAE measures might have a lower sensitivity to dementia for some bundles. In ILF_L, on the other hand, multivariate modeling of DTI and whole-brain fiber geometry showed comparable performance in detecting dementia effects on MD and RD with univariate modeling and could boost sensitivity for AxD measures.

4 | DISCUSSION

4.1 | Microstructural brain abnormalities in dementia and MCI

In this study, we proposed the MINT framework for mapping micro- and macrostructural WM abnormalities using a joint multivariate model. Our novel method contrasts with the univariate approach used in existing tractometry methods, where metrics defined on each tract are considered without reference to other related metrics in the rest of the brain.[||] We investigated group differences in dMRI-derived metrics in subjects with dementia and MCI, relative to matched healthy controls, for both North American and Indian cohorts. We computed bundle profiles of the four widely used DTI measures and five MINT-derived measures quantifying deviations from the norm, corresponding to each DTI metric and an additional shape measure. Overall, diffusivity increased in MCI and further in dementia, with stronger abnormalities in the temporal lobe tracts and the corpus callosum. We detected widespread and largely similar patterns of abnormalities across two cohorts with stronger statistical significance for disease effects in ADNI than NIMHANS. Of all DTI-derived metrics, MD was the most sensitive to dementia, followed by RD, AxD, and FA. Patterns of dementia and MCI effects on MD were also consistent across univariate and multivariate approaches. This aligns with our understanding that MD is less susceptible to crossing fibers, and macrostructural changes are linked to variations in FA more than MD.

4.2 | Paradoxical patterns of FA

Notably, patterns of dementia and MCI effects on RD and AxD were similar between the corresponding DTI- and MINT- derived features except in the projection pathways, for which they exhibited a fan-shaped pattern where the corona radiata and the corpus callosum intersect. This pattern was prominent for MAE-Shape and DTI-FA, where the proximal parts had higher DTI-FA in dementia compared to controls, contrary to what is commonly expected in WM degeneration.

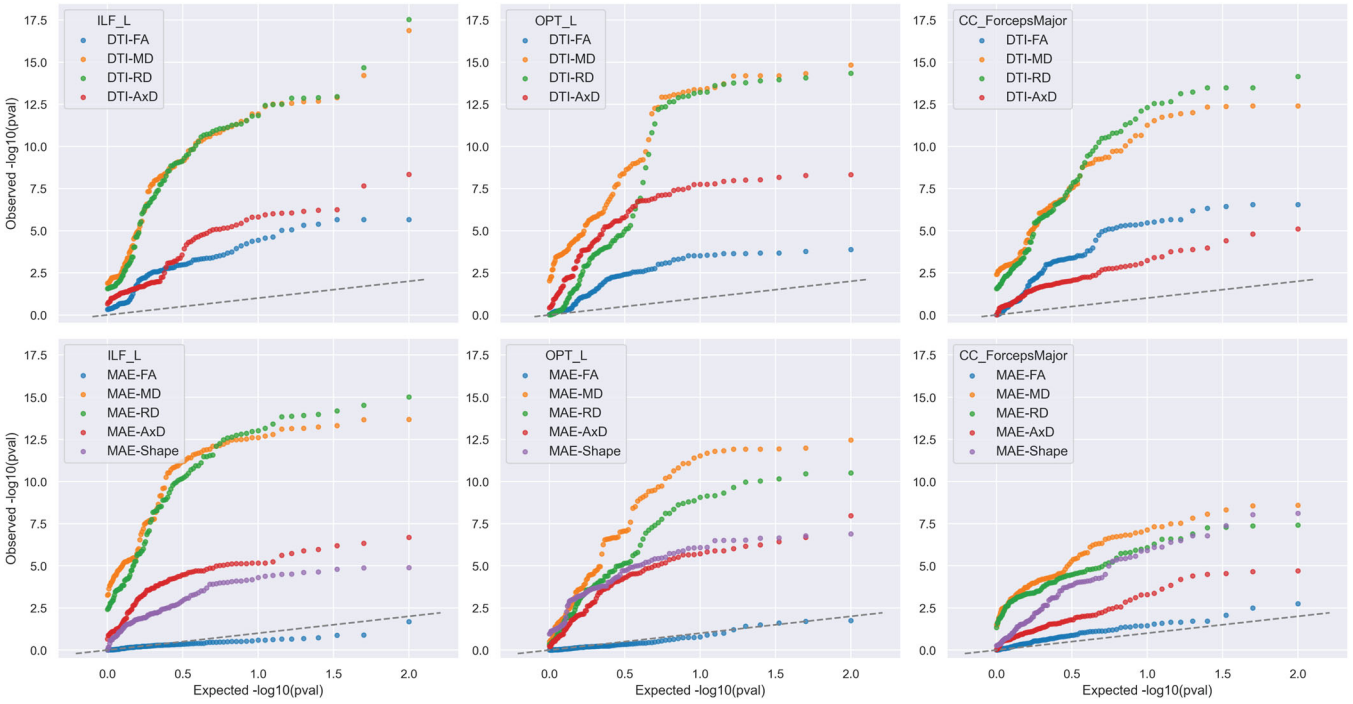

FIGURE 7 Ranking diffusion metrics for their sensitivity to dementia. Here we show a Q-Q plot of $-\log_{10}(P)$ after FDR correction for dementia versus CN group comparison in ADNI cohort, for four DTI- and five MINT-derived diffusion measures in ILF_L, OPT_L, and CC_ForcepsMajor. Microstructural metrics that are associated most strongly with dementia are typically MD and RD, shown in orange and green colors. The dashed line in these plots shows the pattern of *p* values that would be expected for null measures, that is, brain metric that shows no association with dementia. For some tracts, the MAE-FA is not much better than a null measure (ie, shows no detectable association with dementia). ADNI = Alzheimer's Disease Neuroimaging Initiative. CN, cognitively normal; DTI, diffusion tensor imaging; FDR, false discovery rate; MAE-FA, mean absolute error-fractional anisotropy; MINT, macrostructure-informed normative tractometry; QQ, quantile-quantile; RD, radial diffusivity.

Areas showing higher DTI-FA in dementia and MCI largely corresponded to areas with lower macrostructural anomalies characterized by MAE-Shape, except for the bilateral OPT tracts. This paradoxical pattern of FA was previously reported in dementia and MCI by Nir et al.[7] and Douaud et al.[63] in regional DTI analyses. In healthy aging, FA has also been found to have no detectable association – or even a paradoxical, positive association with age – in the projection fibers.[64–66] Using fixel-based analysis, Han et al.[67] reported fiber-specific age associations in healthy adults and found evidence of selective WM degeneration. This fan-shaped pattern in the projection pathways in our findings was likely due to the fiber crossings near the corona radiata and the centrum semiovale. The complex fiber configurations in the commissural and U-fibers, compared to the more homogeneous fiber organization in the deep proximal regions, may contribute to this spatially varying pattern in the long-range projection pathways. While such fiber crossings would presumably exist in healthy controls, we found that this pattern was enhanced in dementia and MCI in both cohorts after accounting for age and sex effects. In a voxel-wise DTI analysis, Racine et al.[68] reported higher FA in the cingulum and fornix in preclinical AD subjects who were amyloid positive. FA is a composite measure and is not a reliable metric for interpreting fiber integrity.[11] For similar reasons, RD and AxD are also susceptible to fiber crossings, where changes in one measure can lead to spurious changes in the other.[69] When studying neurodegenerative diseases such as AD, where the underlying pathology may affect the biophysical properties of WM, the effects of fiber crossing can overshadow (and be confounded with) disease-related microstructural changes in FA, AxD, and RD, when considered independently. Our novel MINT framework accounts for whole-brain macrostructure using tractography in a multivariate model. This may help to disentangle intrinsic changes in DTI measures from signals affected by crossing fibers, to better reflect the underlying pathology. While FA is the metric that is most correlated with alterations in WM macrostructure – and few regions pass the statistical significance threshold compared to univariate DTI measures – MINT still identified widespread abnormalities in RD and AxD in dementia and MCI and higher correlation with cognitive measures compared to DTI values in some WM regions. Although MD is less prone to the effect of crossing fibers than other DTI metrics, we note that it is susceptible to partial volume effect. In WM regions close to the ventricles, such as the forceps major, CSF may contribute to changes in MD that could be attributed to atrophy instead of WM microstructural changes. We will investigate multicompartment diffusion metrics such as free water[70] in the MINT framework in our future work to better account for various sources of signal in the multivariate model.

4.3 | Relative strengths and limitations of regional DTI metrics versus tractometry

Although FA, RD, and AxD have major limitations in characterizing WM microstructure, they still have their place in dMRI analysis, as do simpler, region-of-interest approaches such as TBSS – as tractography often requires high *b*-values (dMRI gradient strength) and high angular resolutions and not all data can be analyzed with tractometry. Even in acquisitions with sufficient angular resolution for tractography, widely used tractography and segmentation pipelines systematically produce false positive streamlines.[71] For this study, we selected 30 bundles from a total of 80 bundles from the HCP-842 atlas.[8,41] State-of-the-art fiber orientation reconstruction and tractography methods used in this study[33,35] can sufficiently reconstruct fibers from these bundles in single-shell dMRI acquisitions with low *b*-values.[72] However, some key bundles, such as the fornix and the cingulum, are difficult to reconstruct using automated methods due to their unique shape and location,[72,73] although the fornix is a key structure showing early changes in aging.[74] Even among the 30 bundles we used in this study, UF and IFOF bundles were not well reconstructed in all subjects across all protocols, which might have affected downstream harmonization and statistical analysis (Section S5). Skeleton-based methods such as TBSS, compared to tractometry, examine a relatively small subset of the brain's anatomy. These methods can be more computationally and memory efficient, facilitating efficient large-scale international analyses such as those by the ENIGMA-DTI working group.[74,75] FA abnormalities in neurodegenerative diseases are robustly reported in numerous studies, as are age effects across the human lifespan in normative models of vast international samples.[76,77] While FA, RD, and AxD are not specific indices of microstructural integrity per se, they are still useful for group comparisons and for detecting anomalies in acquisitions where the WM tracts of interest are difficult to reconstruct using tractography.

4.4 | Comparing findings across cohorts

Although multisite data can be harmonized across scanners and protocols, we analyzed ADNI and NIMHANS data separately because participants in these two cohorts were recruited from different geographic regions and had very different demographics. Cognitive assessments conducted in different languages were not harmonized prior to diagnosis. In our analysis of dementia and MCI effects on dMRI metrics, findings were similar between the ADNI and NIMHANS cohorts. The difference in the apparent effect of diagnosis on brain microstructure, β(DX), and its statistical significance were likely due to the difference in sample sizes and age, as the NIMHANS sample was younger than the ADNI one. In ADNI, where the sample size was larger, MAE measures for RD and AxD were largely homogeneous along-tract compared to the smaller NIMHANS cohort. Compared to ADNI, fewer subjects in NIMHANS were used for model fine-tuning and statistical analysis. Assuming all other parameters remain the same, a larger sample size for either the reference group (CN) or the test group (MCI and dementia) would result in smaller standard errors and, therefore, smaller p values. In our cohort-specific analysis, the standard error from the linear regression model was higher in the NIMHANS cohort than ADNI for almost all bundles and diffusion. MINT may more effectively encode patterns of whole-brain fiber geometry when a larger sample of healthy controls is used for fine-tuning. However, as shown in Figure 5, the NIMHANS cohort showed stronger dementia and MCI effects on macrostructural alterations characterized by MAE-Shape than the ADNI cohort in projection pathway regions close to the cortex. These anomalies were also concentrated in the temporal regions of the left hemisphere association pathways compared to ADNI, where more widespread alterations were detected bilaterally in the posterior and temporal association pathways. While this may be an effect of sample size, or even overfitting, we cannot discount the possibility that the impact of dementia and MCI on WM may differ across Indian and North American populations. Confounders such as education and lifestyle factors are not currently available for the NIMHANS cohort, although they might be associated with diffusion metrics analyzed in this study. Future work will assess these factors and how dMRI measures relate to specific types of neuropsychological decline (eg, language, memory, executive function), as well as amyloid and tau measures. One additional limitation of this study is that the dMRI data from 730 ADNI subjects were all baseline scans, although many ADNI participants underwent repeated scans. Data from repeated visits and new subjects from ADNI3 and ADNI4 will be included in future longitudinal studies.

5 | CONCLUSION

In this study, we introduced MINT for along-tract mapping of WM microstructure and macrostructure. MINT employs a multivariate deep learning model to capture the covariance of four DTI measures, whole-brain fiber organization from tractography, and their spatial dependency. Trained on a large sample of healthy controls, MINT learns the normative distribution and can score individual deviations. In a sample of 1032 participants from two independent cohorts from North America and India, we identified widespread diffusivity abnormalities in MCI and dementia compared to healthy controls using multivariate MINT measures and the BUAN statistical framework. Diffusivity patterns identified by MINT aligned with univariate DTI measures, showing large effects in the posterior-temporal region and corpus callosum, and had a stronger association with clinical metrics of episodic memory. While FA is the least robust DTI measure in its association with MCI and dementia and most susceptible to fiber crossings, MINT may help distinguish microstructural changes from macrostructural alterations. With the growing number of diffusion models and larger international multicenter studies, examining the covariance or mutual information among diffusion metrics and brain regions using multivariate normative tractometry can yield new insights into how brain WM pathways are altered in neurodegenerative diseases.

ACKNOWLEDGMENTS

We thank the research study participants and their families, whose contributions made this work possible. We also thank the DIPY, FURY, and pytorch-lightning contributors and community, whose tools facilitated the analyses in this study. Algorithm development for this project was supported by the National Institute on Aging (NIA) under Grant R01 AG057892. This work was also supported by the US Alzheimer's Association (AARG-23-1149996). The NIMHANS data collection was supported by National Institutes of Health (NIH) Grant R01 AG060610 and the Department of Science and Technology, Government of India, Grants DSTSR/CSI/73/2011 (G), DST-SR/CSI/70/2011 (G), and DST/CSRI/2017/249 (G). Data collection and sharing for ADNI is funded by the NIA (NIH Grant U19 AG024904). The grantee organization is the Northern California Institute for Research and Education. In the past, ADNI also received funding from the National Institute of Biomedical Imaging and Bioengineering, the Canadian Institutes of Health Research, and private-sector contributions through the Foundation for the National Institutes of Health (FNIH), including generous contributions from the following: AbbVie; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interests. Author disclosures are available in the Supporting Information.

CONSENT STATEMENT

All participants in the studies used in this manuscript provided written informed consent.

ORCID

Yixue Feng https://orcid.org/0000-0003-1015-4209

ENDNOTES

* For a summary of related work on methods for dMRI analysis, please see Section S1.

† Code and the pretrained model are available at https://github.com/wendyfyx/MINT.

‡ S, P, and GE denote Siemens, Philips, and General Electric MRI scanners, and the number in the abbreviation denotes the number of diffusion-weighted gradients in each protocol; if other factors are constant, higher numbers of gradients tend to give better signal-to-noise ratio in dMRI.

§ All participants in the TractoInferno are healthy adults. Demographics including age and sex are available elsewhere.[38]

‖ A detailed discussion on additional considerations of univariate and multivariate normative model for tractometry are available in Section S7.

REFERENCES

1. Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*. 1986;161(2):401–407.
2. Basser PJ, Mattiello J, LeBihan D. MR diffusion tensor spectroscopy and imaging. *Biophys J*. 1994;66(1):259–267.
3. Zavaliangos-Petropulu A, Nir TM, Thomopoulos SI, et al. Diffusion MRI indices and their relation to cognitive impairment in brain aging: the updated multi-protocol approach in ADNI3. *Front Neuroinform*. 2019;13:2.
4. Thomopoulos SI, Nir TM, Villalon-Reina JE, et al. Diffusion MRI metrics and their relation to dementia severity: effects of harmonization approaches. In: Walker A, Rittner L, Romero Castro E, Lepore N, Brieva J, Linguraru MG, eds. *17th International Symposium on Medical Information Processing and Analysis*. SPIE; 2021:79.
5. Nir TM, Villalón-Reina JE, Salminen LE, et al. Cortical microstructural associations with CSF amyloid and pTau. *Mol Psychiatry*. 2024;29;257–268. doi:10.1038/s41380-023-02321-7
6. Owens-Walton C, Nir TM, Al-Bachari S, et al. A worldwide study of white matter microstructural alterations in people living with Parkinson's disease. *NPJ Parkinsons Dis*. 2024;10(1):151. doi:10.1038/s41531-024-00758-3
7. Nir TM, Jahanshad N, Villalon-Reina JE, et al. Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging. *NeuroImage Clinical*. 2013;3:180–195.
8. Chandio BQ, Risacher SL, Pestilli F, et al. Bundle analytics, a computational framework for investigating the shapes and profiles of brain pathways across populations. *Sci Rep*. 2020;10(1):17149.
9. Chandio BQ. *Advancing White Matter Tractometry of the Brain Using Diffusion MRI and Machine Learning*. Indiana University; 2022.
10. Jeurissen B, Leemans A, Tournier JD, Jones DK, Sijbers J. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Hum Brain Mapp*. 2013;34(11):2747–2766.
11. Figley CR, Uddin MN, Wong K, Kornelsen J, Puig J, Figley TD. Potential pitfalls of using fractional anisotropy, axial diffusivity, and radial diffusivity as biomarkers of cerebral white matter microstructure. *Front Neurosci*. 2022;15:799576.
12. Raffelt DA, Smith RE, Ridgway GR, et al. Connectivity-based fixel enhancement: whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres. *Neuroimage*. 2015;117:40–55.
13. Schilling KG, Archer D, Yeh FC, et al. Aging and white matter microstructure and macrostructure: a longitudinal multi-site diffusion MRI study of 1218 participants. *Brain Struct Funct*. 2022;227(6):2111-2125.
14. Ahmadi K, Pereira JB, Van Westen D, et al. Fixel-based analysis reveals tau-related white matter changes in early stages of Alzheimer's disease. *J Neurosci*. 2024;44:e0538232024.
15. Chandio BQ, Villalon-Reina JE, Nir TM, et al. Bundle analytics based data harmonization for multi-site diffusion MRI tractometry. *bioRxiv*. Published online February 4, 2024. doi:10.1101/2024.02.03.578764
16. Berg L. Clinical dementia rating (CDR). *Psychopharmacol Bull*. 1988;24(4):637–639.
17. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141(11):1356–1364.
18. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189–198.
19. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57(1):289–300.
20. Van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in early Alzheimer's disease. *N Engl J Med*. 2023;388(1):9–21.

21. Parekh P, Bhalerao GV, Rao R, et al. Protocol for magnetic resonance imaging acquisition, quality assurance, and quality check for the accelerator program for discovery in brain disorders using stem cells. *Int J Methods Psychiatr Res*. 2021;30(3):e1871. doi:10.1002/mpr.1871
22. Bharath S, Joshi H, John JP, et al. A multimodal structural and functional neuroimaging study of amnestic mild cognitive impairment. *Am J Geriatr Psychiatry*. 2017;25(2):158–169.
23. Joshi H, Bharath S, Balachandar R, et al. Differentiation of early Alzheimer's disease, mild cognitive impairment, and cognitively healthy elderly samples using multimodal neuroimaging indices. *Brain Connect*. 2019;9(9):730–741.
24. Manjón JV, Coupé P, Concha L, Buades A, Collins DL, Robles M. Diffusion weighted image denoising using overcomplete local PCA. *PLoS One*. 2013;8(9):e73021.
25. Garyfallidis E, Brett M, Amirbekian B, et al. Dipy, a library for the analysis of diffusion MRI data. *Front Neuroinform*. 2014;8:8. doi:10.3389/fninf.2014.00008
26. Neto Henriques R. Advanced Methods for Diffusion MRI Data Analysis and their Application to the Healthy Ageing Brain. Published online September 2017. doi:10.17863/CAM.29356
27. Kellner E, Dhital B, Kiselev VG, Reisert M. Gibbs-ringing artifact removal based on local subvoxel-shifts: Gibbs-ringing artifact removal. *Magn Reson Med*. 2016;76(5):1574–1581.
28. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage*. 2016;125:1063–1078.
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(2):870–888.
30. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *Neuroimage*. 2012;62(2):782–790.
31. Tournier JD, Smith R, Raffelt D, et al. MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage*. 2019;202:116137.
32. Schilling KG, Blaber J, Huo Y, et al. Synthesized b0 for diffusion distortion correction (Synb0-DisCo). *Magn Reson Imaging*. 2019;64:62–70.
33. Canales-Rodríguez EJ, Daducci A, Sotiropoulos SN, et al. Spherical deconvolution of multichannel diffusion MRI data with non-Gaussian noise models and spatial regularization. *PLoS One*. 2015;10(10):e0138910.
34. Portegies JM, Fick RHJ, Sanguinetti GR, Meesters SPL, Girard G, Duits R. Improving fiber alignment in HARDI by combining contextual pde flow with constrained spherical deconvolution. *PLoS One*. 2015;10(10):e0138122.
35. Girard G, Whittingstall K, Deriche R, Descoteaux M. Towards quantitative connectivity analysis: reducing tractography biases. *Neuroimage*. 2014;98:266–278.
36. Tripathi R, Kumar JK, Bharath S, Marimuthu P, Varghese M. Clinical validity of NIMHANS neuropsychological battery for elderly: a preliminary report. *Indian J Psychiatry*. 2013;55(3):279–282.
37. Jack CR Jr, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimers Dement*. 2024;20(8):5143-5169. doi:10.1002/alz.13859
38. Poulin P, Theaud G, Rheault F, et al. TractoInferno—A large-scale, open-source, multi-site database for machine learning dMRI tractography. *Scientific Data*. 2022;9(1):725.
39. Tournier JD, Calamante F, Connelly A. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage*. 2007;35(4):1459–1472.
40. Garyfallidis E, Côté MA, Rheault F, et al. Recognition of white matter bundles using local and global streamline-based registration and clustering. *Neuroimage*. 2018;170:283–295.
41. Yeh FC, Panesar S, Fernandes D, et al. Population-averaged atlas of the macroscale human structural connectome and its network topology. *Neuroimage*. 2018;178:57–68.
42. Nabulsi L, Chandio BQ, McPhilemy G, et al. Multi-site statistical mapping of along-tract microstructural abnormalities in bipolar disorder with diffusion MRI tractometry. In: *2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM)*. IEEE; 2023:1–5.
43. Legarreta JH, Petit L, Rheault F, et al. Filtering in tractography using autoencoders (FINTA). *Med Image Anal*. 2021;72:102126.
44. Legarreta JH, Petit L, Jodoin PM, Descoteaux M. Clustering in tractography using autoencoders (CINTA). In: Cetin-Karayumak S, Christiaens D, Figini M, eds. *Computational Diffusion MRI*. Springer Nature Switzerland; 2022:125–136.
45. Dumais F, Legarreta JH, Lemaire C, et al. FIESTA: autoencoders for accurate fiber segmentation in tractography. *Neuroimage*. 2023;279:120288.
46. Feng Y, Chandio BQ, Chattopadhyay T, et al. Learning optimal white matter tract representations from tractography using a deep generative model for population analyses. In: *18th International Symposium on Medical Information Processing and Analysis*. SPIE; 2023:48.
47. Baur C, Wiestler B, Albarqouni S, Navab N. Deep Autoencoding models for unsupervised anomaly segmentation in brain MR images. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, Van Walsum T, eds. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Springer International Publishing; 2019:161–169.
48. Simon C, Barr J. *Deep Learning and XAI Techniques for Anomaly Detection: Integrate The Theory and Practice of Deep Anomaly Explainability*. Packt Publishing; 2023.
49. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. Published online April 2017. Accessed October 16, 2023. http://arxiv.org/abs/1610.02357
50. Chen X, Kingma DP, Salimans T, et al. Variational Lossy Autoencoder. Published online March 2017. Accessed April 1, 2024. http://arxiv.org/abs/1611.02731
51. Kingma DP, Welling M, Auto-Encoding Variational Bayes. arXiv:1312 6114 [cs, stat]. Published online May 2014. Accessed October 8, 2021. http://arxiv.org/abs/1312.6114
52. Fu H, Li C, Liu X, Gao J, Celikyilmaz A, Carin L, Cyclical Annealing Schedule: A Simple Approach to Mitigating KL Vanishing. Published online June 2019. Accessed November 22, 2022. http://arxiv.org/abs/1903.10145
53. Fortin JP, Parker D, Tunç B, et al. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage*. 2017;161:149–170.
54. Pomponio R, Erus G, Habes M, et al. Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *Neuroimage*. 2020;208:116450.
55. Rosenblatt M, Tejavibulya L, Jiang R, Noble S, Scheinost D. Data leakage inflates prediction performance in connectome-based machine learning models. *Nat Commun*. 2024;15(1):1829.
56. Vallat R. Pingouin: statistics in python. *J Open Source Softw*. 2018;3(31):1026.
57. Garyfallidis E, Koudoro S, Guaje J, et al. FURY: advanced scientific visualization. *J Open Source Softw*. 2021;6(64):3384.
58. Guaje J, Koudoro S, Garyfallidis E. Horizon, closing the gap between cinematic visualization and medical imaging. *medRxiv*. Published online July 2023:2023.07.11.23292519.
59. Bozzali M. White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. *J Neurol Neurosurg Psychiatry*. 2002;72(6):742–746.
60. DeSimone JC, Wang WE, Loewenstein DA, et al. Diffusion MRI relates to plasma Aβ42/40 in PET negative participants without dementia. *Alzheimers Dement*. 2024;20(4):2830–2842.
61. Yang Y, Schilling K, Shashikumar N, et al. White matter microstructural metrics are sensitively associated with clinical staging in Alzheimer's disease. *Alzheimers Dement*. 2023;15(2):e12425.

62. Dalboni da Rocha JL, Bramati I, Coutinho G, Tovar Moll F, Sitaram R. Fractional anisotropy changes in parahippocampal cingulum due to Alzheimer's disease. *Sci Rep*. 2020;10(1):2660.
63. Douaud G, Jbabdi S, Behrens TEJ, et al. DTI measures in crossing-fibre areas: increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *Neuroimage*. 2011;55(3):880–890.
64. Madden DJ, Bennett IJ, Burzynska A, Potter GG, Chen NK, Song AW. Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochimica et Biophysica Acta*. 2012;1822(3):386–400.
65. Chad JA, Pasternak O, Chen JJ. Orthogonal moment diffusion tensor decomposition reveals age-related degeneration patterns in complex fiber architecture. *Neurobiol Aging*. 2021;101:150–159.
66. Schilling KG, Chad JA, Chamberland M, et al. White matter tract microstructure, macrostructure, and associated cortical gray matter morphology across the lifespan. *Imaging Neuroscience*. 2023;1:1–24.
67. Han A, Dhollander T, Sun YL, Chad JA, Chen JJ. Fiber-specific age-related differences in the white matter of healthy adults uncovered by fixel-based analysis. *Neurobiol Aging*. 2023;130:22–29.
68. Racine AM, Adluru N, Alexander AL, et al. Associations between white matter microstructure and amyloid burden in preclinical Alzheimer's disease: a multimodal imaging investigation. *NeuroImage Clinical*. 2014;4:604–614.
69. Wheeler-Kingshott CAM, Cercignani M. About "axial" and "radial" diffusivities. *Magn Reson Med*. 2009;61(5):1255–1260.
70. Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*. 2012;61(4):1000–1016.
71. Maier-Hein KH, Neher PF, Houde JC, et al. The challenge of mapping the human connectome based on diffusion tractography. *Nat Commun*. 2017;8(1):1349.
72. Jones DK. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study†. *Magn Reson Med*. 2004;51(4):807–815.
73. Rheault F, Roy M, Cunnane S, Descoteaux M. Bundle-specific fornix reconstruction for dual-tracer PET-tractometry. *bioRxiv*. Published online September 2018. doi:10.1101/423459
74. Chen DQ, Strauss I, Hayes DJ, Davis KD, Hodaie M. Age-related changes in diffusion tensor imaging metrics of fornix subregions in healthy humans. *Stereotact Funct Neurosurg*. 2015;93(3):151–159.
75. Kochunov P, Hong LE, Dennis EL, et al. ENIGMA-DTI: translating reproducible white matter deficits into personalized vulnerability metrics in cross-diagnostic psychiatric research. *Hum Brain Mapp*. 2022;43(1):194–206.
76. Villalón-Reina JE, Zhu AH, Nir TM, et al. Large-scale Normative Modeling of Brain Microstructure. In: *2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM)*. IEEE; 2023:1–5.
77. Zhu AH, Nir TM, Javid S, et al. Reference curves for harmonizing multi-site regional diffusion MRI metrics across the lifespan. *bioRxiv*. Published online February 2024. doi:10.1101/2024.02.22.581646

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Feng Y, Chandio BQ, Villalon-Reina JE, et al. Microstructural mapping of neural pathways in Alzheimer's disease using macrostructure-informed normative tractometry. *Alzheimer's Dement*. 2024;1-18. https://doi.org/10.1002/alz.14371