



OPEN Evaluating the impact of denoising diffusion MRI data on tractometry metrics of optic tract abnormalities in glaucoma

Daiki Taguma^{1,2,✉}, Shumpei Ogawa³ & Hiromasa Takemura^{1,2,4,5}

Diffusion MRI (dMRI)-based tractometry is a non-invasive neuroimaging method for evaluating white matter tracts in living humans, capable of detecting abnormalities caused by disorders. However, measurement noise in dMRI data often compromises the signal quality. Several denoising methods for dMRI have been proposed, but the extent to which denoising affects tractometry metrics of white matter tissue properties associated with disorders remains unclear. We evaluated how denoising affects tractometry along the optic tract (OT) in patients with glaucoma. Because glaucoma damages retinal ganglion cells, the OT in patients with glaucoma is likely to exhibit tissue abnormalities. Therefore, we examined dMRI data from patients with glaucoma to evaluate how two widely used denoising methods (MPPCA and Patch2Self) affect tractometry metrics regarding the expected tissue changes in the OT. We found that denoising affected the appearance of diffusion-weighted images, increased the estimated signal-to-noise ratio, and reduced residuals in voxelwise model fitting. However, denoising had a limited impact on the differences in tractometry metrics of the OT between patients with glaucoma and controls. Moreover, we found no evidence that denoising improved the reproducibility of tractometry. These findings suggest that the current denoising methods have a limited impact when used together with a tractometry framework.

Keywords Diffusion MRI, White matter, Glaucoma, Denoising, Optic tract, Tractometry

Vision is an essential sensory modality for humans as it provides vital information for avoiding collisions, finding food, and communicating with others. Vision impairment significantly reduces the quality of life¹ and increases the risk of death². Glaucoma is a leading cause of visual loss in modern aging societies because its prevalence rate is higher in elderly populations³. Therefore, an accurate understanding of the impact of glaucoma on the nervous system, including white matter pathways that carry visual signals, is important⁴.

Diffusion-weighted MRI (dMRI)-based tractometry is a non-invasive neuroimaging method for quantifying the degree of tissue abnormalities along a specific white matter tract in living humans^{5,6} (Fig. 1). Although this approach has proven useful for identifying tissue changes caused by glaucoma⁷, its clinical application is limited by measurement noise, such as thermal noise during dMRI acquisition⁸. Researchers have developed denoising methods for dMRI data by using either principal component analysis (PCA)-based^{9,10} or machine learning approaches^{11–13}. These methods identify “noise” in dMRI data based on expected characteristics of noise or random fluctuations, and subsequently remove identified noise from original images to generate “denoised” dMRI data.

Nonetheless, previous neuroimaging research has highlighted the potential risk that denoising methods may remove true signals or introduce unwanted biases¹⁴. This poses a serious issue when applied to understanding clinical disorders. Therefore, establishing the extent to which different denoising techniques alter the data is critical for both basic and clinical science. However, assessing their impact on in vivo human dMRI is challenging, given that the ground truth of true signals versus noise is unknown. Prior studies have evaluated dMRI denoising by analyzing datasets with simulated noise¹⁰, visual inspection of the image quality¹⁰, quantifying residual

¹Division of Sensory and Cognitive Brain Mapping, Department of System Neuroscience, National Institute for Physiological Sciences, 38 Nishigonaka Myodaiji, Okazaki 444-8585, Aichi, Japan. ²The Graduate Institute of Advanced Studies, SOKENDAI, Hayama, Japan. ³Department of Ophthalmology, The Jikei University School of Medicine, Tokyo, Japan. ⁴Core for Spin Life Sciences, Okazaki Collaborative Platform, National Institutes of Natural Sciences, Okazaki, Japan. ⁵Center for Information and Neural Networks (CiNet), Advanced ICT Research Institute, National Institute of Information and Communications Technology, Suita, Japan. ✉email: dtagama@nips.ac.jp

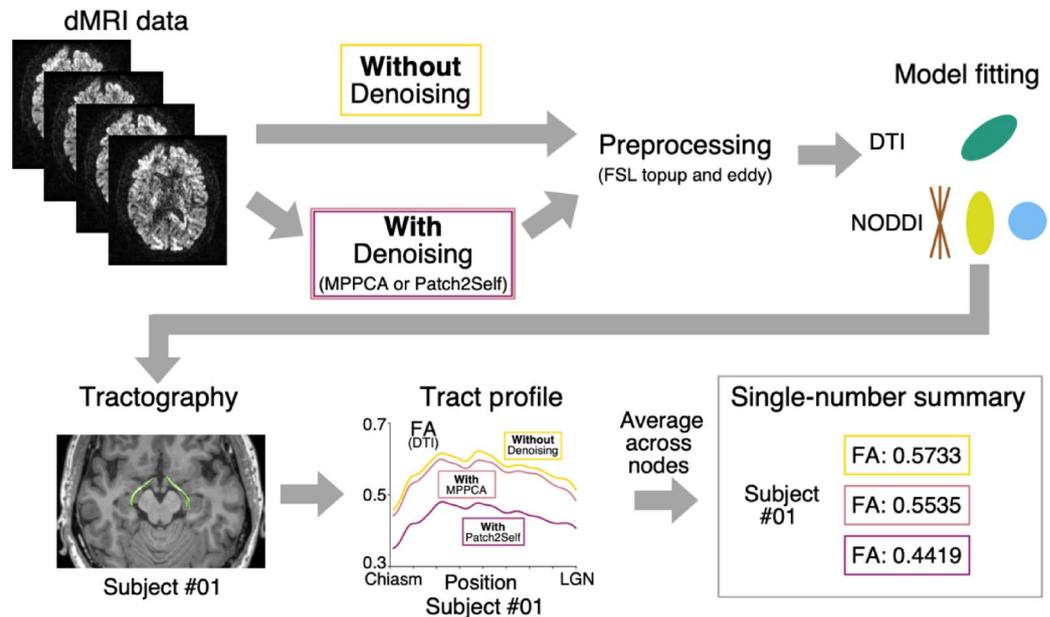


Fig. 1. Schematic diagram of the data processing pipeline for dMRI-based tractometry approaches. In common practice, after acquiring dMRI data (left top panel), researchers may apply denoising algorithms on the dMRI dataset. We compared the analysis results with (purple) and without denoising (yellow) while keeping the subsequent processing procedure the same. dMRI data are then typically preprocessed to correct for susceptibility- and eddy-current distortions^{19,20}. After preprocessing, researchers fit voxelwise diffusion models (diffusion tensor imaging, DTI; neurite orientation and dispersion imaging, NODDI) to dMRI data in each voxel to quantify white matter microstructural properties. Tractography is used to identify a white matter tract of interest (in this study, the optic tract; green in the left bottom panel). Researchers can then calculate a tract profile^{5,6}, which is a summary of voxelwise measurements along the tract (bottom middle figures). Finally, these tract profiles were averaged along the spatial position along the tract to obtain a single-number summary per subject, for each metric and tract. We compared these metrics per subject between data with and without denoising.

in voxelwise model fitting¹², and the geometry of tractography in a single tract¹². While these approaches contribute to understanding the impact of denoising, they did not individually address the extent to which denoising affects tractometry metrics for detecting neurobiological tissue changes caused by disorders, such as glaucoma in empirical dMRI data.

In this study, we assess how denoising applied to dMRI data affects the ability of tractometry to detect optic tract (OT) damage in patients with glaucoma. As glaucoma is a disorder that damages retinal ganglion cells and the optic nerve⁴, it likely affects OT tissues, which comprise axons from these cells. Multiple research groups have reported that dMRI and tractometry can detect tissue changes along the OT in patients with glaucoma^{15–18}. Therefore, if denoising improves the ability of tractometry to identify white matter tissue abnormalities, it would be easier to distinguish data acquired from patients with glaucoma and controls after denoising. We re-analyzed dMRI data acquired from patients with glaucoma and controls¹⁸ and compared the tractometry results on this dataset with and without the application of two widely used denoising methods (Fig. 1; MPPCA¹⁰, Patch2Self¹²). We also quantified how denoising affected voxelwise model fitting, the estimated signal-to-noise ratio (SNR) of the image, and the scan-rescan reliability of tractometry on the OT.

Methods

Dataset

We analyzed a dataset presented in a previous study¹⁸. Below, we briefly describe the details of the dataset.

Subjects

We analyzed T1-weighted structural images and dMRI data acquired from 17 patients with glaucoma (mean age = 56.6 years, age range = 24–72, 8 females) and 30 healthy controls with normal visual function (mean age = 51.4 years, age range = 36–71, 14 females). All subjects were recruited from the Tokyo area, Japan. Data acquisition and analysis were approved by the ethics committees of the Jikei University School of Medicine, Tamagawa University, and the National Institute for Physiological Sciences. All subjects provided written informed consent. All methods were performed in accordance with relevant guidelines and regulations. All patients were diagnosed with primary or secondary open-angle glaucoma by experienced ophthalmologists at the Department of Ophthalmology, Jikei University School of Medicine. Detailed information on the clinical features is available in a previous publication¹⁸. Information regarding comorbidities, pharmacological profiles,

and types of treatments was unavailable in the original dataset; therefore, we could not consider these factors in this study.

Visual field test

All patients with glaucoma underwent visual field testing using the Humphrey Field Analyzer (HFA) with the 24–2 or 30–2 Swedish Interactive Thresholding Algorithm (SITA) Standard (Carl Zeiss Meditec, Dublin, CA, USA). The HFA results of all patients were quantified in units of mean deviation from healthy populations.

MRI data acquisition methods

MRI data were acquired using a 3T MAGNETOM Trio, Tim System with a 32-channel head coil (Siemens, Erlangen, Germany) at the Tamagawa University Brain Science Institute in Machida, Japan.

T1-weighted structural images were acquired from all subjects using magnetization and prepared rapid acquisition with gradient echo (MPRAGE) sequence with 1 mm isotropic voxels (repetition time [TR], 2000 ms; echo time [TE], 1.98 ms; in-plane acceleration [generalized autocalibrating partially parallel acquisitions; GRAPPA²¹, 2; see Supplementary Table S1 for T1-weighted structural data acquisition parameter details).

dMRI images were acquired from all subjects using single-shot spin-echo, echo planar imaging with multiple b-values²² (EPI; 32 directions with $b = 700 \text{ s/mm}^2$; 64 directions with $b = 2000 \text{ s/mm}^2$; 1.7 mm isotropic voxels; TR, 4500 ms; TE, 94 ms; in-plane acceleration [GRAPPA], 2; multiband factor, 3; phase partial Fourier, 6/8; diffusion scheme, monopolar; see Supplementary Table S2 for dMRI acquisition parameter details) implemented in a multiband-accelerated EPI pulse sequence provided by the Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota (<https://www.cmrr.umn.edu/multiband/>)²³. In addition to diffusion-weighted images, 12 low b-value ($b = 0 \text{ s/mm}^2$) images were obtained. Two dMRI image sets were acquired with reversed phase-encoding directions (anterior-posterior and posterior-anterior) to correct susceptibility-induced distortion during preprocessing. These two image sets were also used to quantify the scan-rescan reliability of tractometry.

Data analysis

dMRI data denoising

We employed two widely used denoising approaches. One was the MPPCA method¹⁰ implemented in the “dwidenoise” command of the MRTrix3²⁴ (<https://mrtrix.readthedocs.io/en/dev/reference/commands/dwidenoise.html>). This method uses the Marchenko-Pastur (MP) distribution to define noise in dMRI dataset and performs PCA-based denoising. The other was Patch2Self¹², which is distributed as part of the DIPY toolbox²⁵ (https://docs.dipy.org/stable/examples_built/preprocessing/denoise_patch2self.html). Patch2Self is a self-supervised learning method that does not assume that signal characteristics correspond to noise; instead, it aims to learn random fluctuations in four-dimensional dMRI data and then builds a regressor to remove these fluctuations in each three-dimensional volume.

We applied one of these methods to the dMRI dataset before preprocessing, using the default parameters of each command. Subsequently, the dMRI data with and without denoising were preprocessed using identical procedures.

dMRI data preprocessing

The dMRI data were preprocessed using the TOPUP and EDDY tools in FSL to correct for susceptibility-induced distortions, eddy-current induced distortions, and subject’s motion^{19,20}. Subsequently, the dMRI data were co-registered with T1-weighted structural MRI data acquired from the same subject. Further details on the preprocessing methods are described in the original study¹⁸.

Fitting voxel-wise models

Diffusion tensor imaging (DTI) We fitted the diffusion tensor model to the preprocessed dMRI data to calculate the fractional anisotropy (FA) and mean diffusivity (MD)^{26–28} using iterated weighted least-squares algorithms implemented in MRTrix3²⁹.

Neurite orientation dispersion and density imaging (NODDI) We also fitted NODDI³⁰ to the dMRI data to estimate the intracellular volume fraction (ICVF) and orientation dispersion index (ODI) using the NODDI MATLAB toolbox (<http://mig.cs.ucl.ac.uk/index.php?n=Tutorial.NODDI matlab>).

Tractography

The OT and optic radiation in each subject were identified in a previous study¹⁸ using probabilistic tractography (ConTrack³¹) on the dMRI data without denoising. In brief, we identified the optic chiasm, lateral geniculate nucleus, and primary visual cortex from structural T1-weighted images and used these areas as seed regions for tractography. We then generated streamlines connecting these regions and selected streamlines based on the scoring process implemented in ConTrack. Additionally, we applied an outlier streamline rejection procedure proposed in a previous study⁶. The detailed procedures for tractography and their anatomical validity have been described previously^{18,32} (see Supplementary Table S3 for tractography parameters).

We used the OT and optic radiation identified from the dMRI data without denoising to perform a subsequent tract profile analysis on data with denoising, ensuring consistent voxel selection across all cases.

Calculating tract profiles

We used the Automated Fiber Quantification (AFQ) MATLAB toolbox⁶ (<https://github.com/yeatmanlab/AFQ>) to calculate tract profiles representing tissue properties along the OT and optic radiation. In brief, after

resampling the streamlines into 100 nodes, the tissue properties (FA, MD, ICVF, and ODI) were calculated at each node. These metrics were summarized using a weighted average based on the distance between the tract core and each voxel. We excluded the first and last 10 nodes from each tract to minimize the impact of crossing with superficial white matter and partial volume effect with gray matter. The data from the remaining 80 nodes were averaged to produce a single-number summary of each metric for each subject. Additionally, the data from the left and right hemispheres were averaged. Finally, the data from the two dMRI image sets were averaged for analyses, except for the scan-rescan reliability analysis, which aimed to quantify the consistency between the sets.

Assessment of image appearance and estimated SNR

We assessed the impact of denoising (MPPCA and Patch2Self) on dMRI image quality using two strategies. First, we visually inspected diffusion-weighted images with and without denoising (Fig. 2). We also calculated the difference map by subtracting the data with denoising from the data without denoising, allowing for visual characterization of the impact of each denoising method.

Second, we investigated the impact of denoising on the estimated SNR. Accurately estimating the SNR for diffusion-weighted images is challenging because diffusion signals are represented as decreased image intensity and vary when a motion probing gradient is applied with different orientations. In this dataset, diffusion-weighted images were not acquired twice with identical gradient directions and acquisition parameters, making it difficult to calculate the SNR based on the variability across multiple scans. Therefore, we estimated the SNR using low b-value images acquired multiple times in a single run. We applied the following formula to calculate the SNR^{33–35}:

$$\text{SNR} = \sqrt{2} \times \text{SI (first image)} / \text{SD (subtracted image)} \quad (1)$$

where SI (first image) is the mean signal intensity of the first low b-value image within the OT, and SD (subtracted image) is the standard deviation of the signal intensity differences between the first and second low b-value images within the OT. Using this method, we estimated the SNR of the low b-value images in the OT of each subject by pooling the OT voxels from the left and right hemispheres. Finally, we compared the SNR of the OT between the data with and without MPPCA. We did not evaluate Patch2Self because it does not apply denoising to low b-value images, indicating that no SNR impact was expected. Data visualization was performed using a function that generates violin plots in MATLAB³⁶ (<https://zenodo.org/records/4559847>).

Quantification of the impact of denoising on voxelwise model fitting

We quantified the effect of denoising on the fitting of voxelwise models (DTI and NODDI) on diffusion signals in the OT.

Based on previous studies^{37,38}, we used the root mean squared error (RMSE) for the DTI:

$$\text{RMSE} = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}} \quad (2)$$

where \hat{y}_i represents the signals estimated by fitting the DTI in the voxel i , y_i represents the measured signals, and n represents the number of voxels in the OT for each subject.

For NODDI, we used the Rician log-likelihood ($\log L$):

$$\log L = \sum_{i=1}^n \left[\log \frac{x_i}{\sigma^2} - \frac{x_i^2 + s^2}{2\sigma^2} + \log I_0 \left(\frac{x_i s}{\sigma^2} \right) \right] \quad (3)$$

where x_i represents the measured signal in voxel i , s represents the signal estimated by NODDI, σ represents the estimated standard deviation of the noise, and I_0 represents the modified Bessel function. This metric was chosen because it does not assume Gaussian noise, which may not directly apply to NODDI, which uses non-Gaussian biophysical models.

For both metrics, we averaged the RMSE and Rician log-likelihood in the OT across the hemispheres and two datasets to obtain a single summary value for each subject. We then compared them between data with and without denoising (MPPCA and Patch2Self) to quantify the impact of denoising on the voxelwise model fitting. To statistically assess this, we performed a two-tailed paired t-test (for calculating P-values) using the MATLAB Statistics and Machine Learning Toolbox and a Bayesian paired t-test (for calculating the Bayes factor, BF_{10}) using the bayesFactor MATLAB package³⁹ (<https://github.com/klabhub/bayesFactor>).

Investigation of the impact of denoising on differences between glaucoma and control data

We investigated the impact of denoising (MPPCA and Patch2Self) on diffusivity-based metrics (FA, MD, ICVF, and ODI) along the OT, which were single-number summary metrics obtained by averaging along nodes, data from two hemispheres, and data from two runs (see “Calculating tract profiles” above). Specifically, we assessed the effect of denoising on the differences in these metrics between patients with glaucoma and controls. First, we calculated the degree of deviation from the control mean for each glaucoma patient in units of the standard deviation of the control subjects for each metric and dataset (without denoising, with MPPCA, and with Patch2Self). We visualized the distribution of the degree of deviation from the controls using a violin plot.

Then, we calculated the statistical differences in the deviations between the data with and without denoising (MPPCA and Patch2Self). To determine the strength of the effect of denoising on the differences between patients

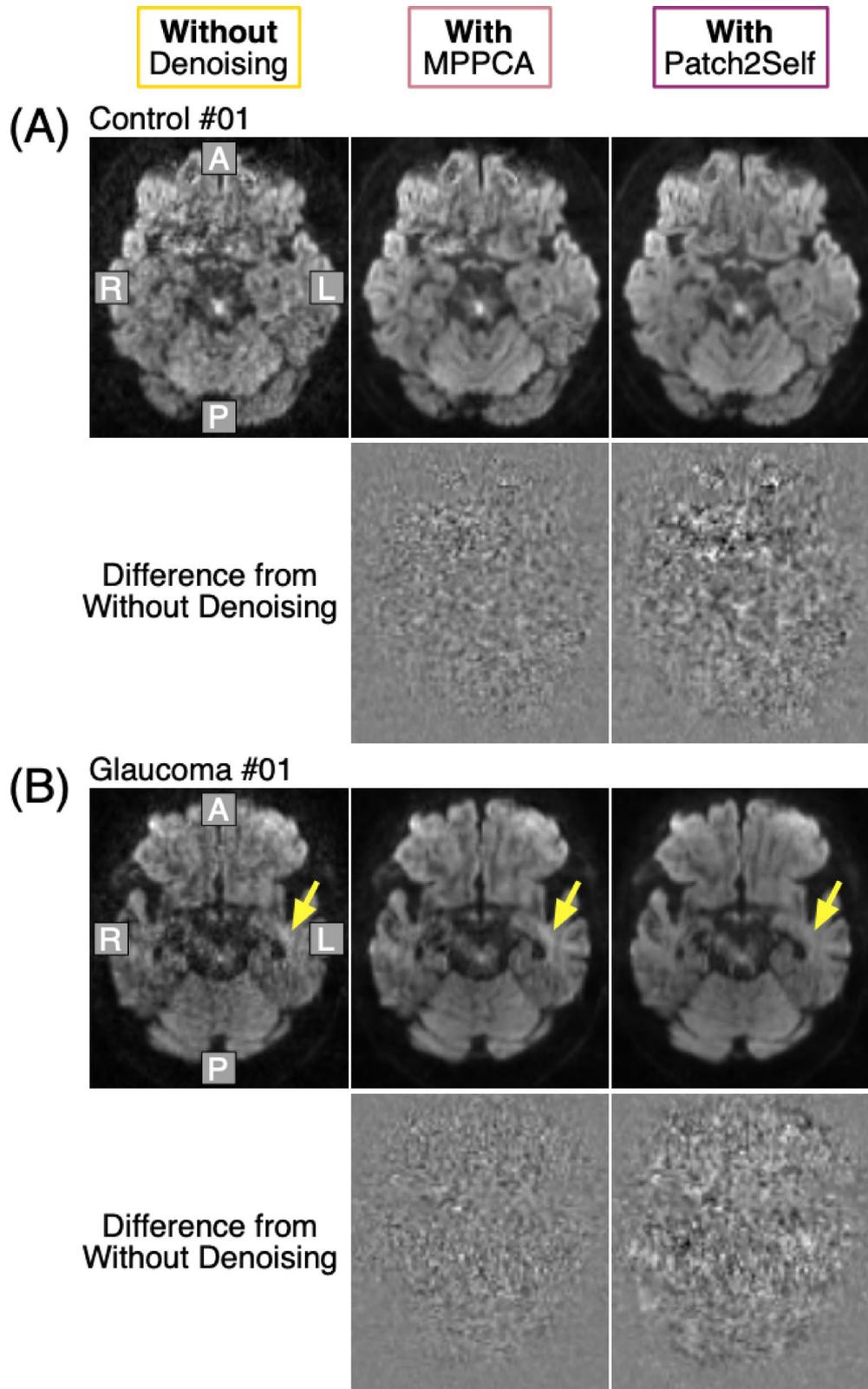


Fig. 2. Diffusion-weighted images with and without denoising in representative subjects (A a healthy control, B a patient with glaucoma). The top panel depicts diffusion-weighted images in an axial section (left, image without denoising; middle, image with MPPCA; right, image with Patch2Self). The bottom panel depicts difference maps between data without denoising and data with one of the denoising methods (MPPCA and Patch2Self). Letters in the image denote image orientation (A: anterior, P: posterior, L: left, R: right). Yellow arrows in panel B depict the location of white matter regions with high signal intensity in a diffusion-weighted image without denoising; the signal intensity is reduced after applying Patch2Self.

with glaucoma and controls, we calculated the effect size (Cohen's d). Additionally, we determined the statistical differences between datasets by performing a two-tailed paired t -test and a Bayesian paired t -test to calculate the P -values and BF_{10} , respectively. P -values greater than 0.025 (equivalent to applying a Bonferroni correction for two comparisons, given $P=0.05$) were considered insignificant because we conducted two comparisons on the dataset without denoising (comparison between data without denoising vs. data with MPPCA; data without denoising vs. data with Patch2Self).

In addition to the univariate analysis for each metric, we performed a linear discriminant analysis to assess the impact of denoising on the discriminability between patients with glaucoma and controls when multiple tractometry metrics were used for classification. This analysis was performed using the `fitcdiscr` function in the MATLAB Statistics and Machine Learning Toolbox. The misclassification probability was used as a measure of classification performance, and differences between data with and without denoising were statistically evaluated using McNemar's test (<https://github.com/dnafinder/mcnemar>;⁴⁰).

The same analyses were performed on data along the optic radiation to assess the specificity of the results in the OT.

Calculation of correlation between visual field test and OT tissue properties

We performed an analysis to calculate the correlation between visual field test scores and OT tissue properties, following the approach used in Ogawa et al. (2022)¹⁸. In brief, we normalized the diffusivity-based metrics (FA, MD, ICVF, and ODI) of patients with glaucoma by calculating their deviation from the control mean, using units of standard deviation in the control group. The visual field test scores for each patient (mean deviation of HFA) were averaged across both eyes. We then calculated the Pearson correlation between diffusivity measurements along the OT and the visual field test results in patients with glaucoma. This analysis was performed on data without denoising, data with MPPCA denoising, and data with Patch2Self denoising.

Quantification of scan-rescan reliability

We quantified the scan-rescan reliability of tractometry for each metric (FA, MD, ICVF, and ODI), dataset (data without denoising, with MPPCA, and with Patch2Self), and tract (OT and optic radiation) by comparing data from two dMRI runs acquired from all subjects (healthy controls and patients with glaucoma). Because these two runs were collected with different phase-encoding directions (anterior-posterior and posterior-anterior), we expected systematic differences in the tractometry data between them. Nonetheless, a higher inter-subject correlation would indicate that tractometry provides reproducible measurements for identifying the properties of white matter tracts in individual subjects. We calculated the Pearson correlation coefficient (R) to assess the scan-rescan reliability. Additionally, we performed linear regression on the scan-rescan data to estimate the regression line using the least squares method in the MATLAB Statistics and Machine Learning Toolbox.

Results

In this study, we determined the impact of two different types of denoising methods (MPPCA and Patch2Self) on dMRI data along the OT of patients with glaucoma and healthy controls to assess the extent to which denoising affects tractometry metrics and reliability.

Qualitative comparisons of image appearance

First, we qualitatively compared the appearance of the dMRI data with and without denoising. Figure 2 shows the axial diffusion-weighted images acquired from representative subjects (one healthy control and one patient with glaucoma). Compared to the original images, the denoised images exhibited clearer tissue contrast between the gray and white matter. Additionally, the cerebellum exhibited improved visibility of macrostructures, such as folds, after denoising. Furthermore, the difference maps between the images with and without denoising did not reveal obvious tissue borders, indicating that denoising may not remove information that contributes to the tissue contrast. While both methods enhanced the visual appearance of the tissue, we observed that Patch2Self excluded more high-contrast information than MPPCA based on visual comparisons of the difference maps (see white matter regions highlighted by yellow arrows in Fig. 2B).

We then evaluated how denoising affected the appearance of tissue parameter maps obtained by fitting voxelwise models. Figure 3 shows a comparison of the FA and ICVF maps derived from DTI and NODDI, respectively. The overall appearance of these maps did not appear to differ significantly between the data with and without denoising in terms of tissue contrast. However, we found that the FA values were generally lower after Patch2Self was applied compared to the data without denoising. In addition, the ODI values increased after Patch2Self (see Supplementary Fig. S1 for comparisons of each diffusivity metric with and without Patch2Self in the OT). In the following sections, we quantitatively evaluate the effect of denoising on these metrics along the OT.

Impact of denoising on the estimated SNR of images

To quantitatively assess the impact of denoising on the dMRI data along the OT, we first compared how denoising affected the estimated SNR of the dMRI measurements along the OT. We evaluated the SNR based on comparisons of multiple low b -value images^{33–35}. This analysis was not performed on data with Patch2Self because this method, by default, does not apply to low b -value images¹².

Figure 4 shows a comparison of the estimated SNR of the measurements along the OT between the data without denoising and the data with MPPCA. The data with MPPCA exhibited a significantly higher estimated SNR than those without denoising (Fig. 4B; $d=1.87$; $BF_{10}>100$; two-tailed paired t -test, $t_{46}=12.80$, $P<0.001$). This suggests that MPPCA may improve the SNR of measurements along the OT. Furthermore, the difference

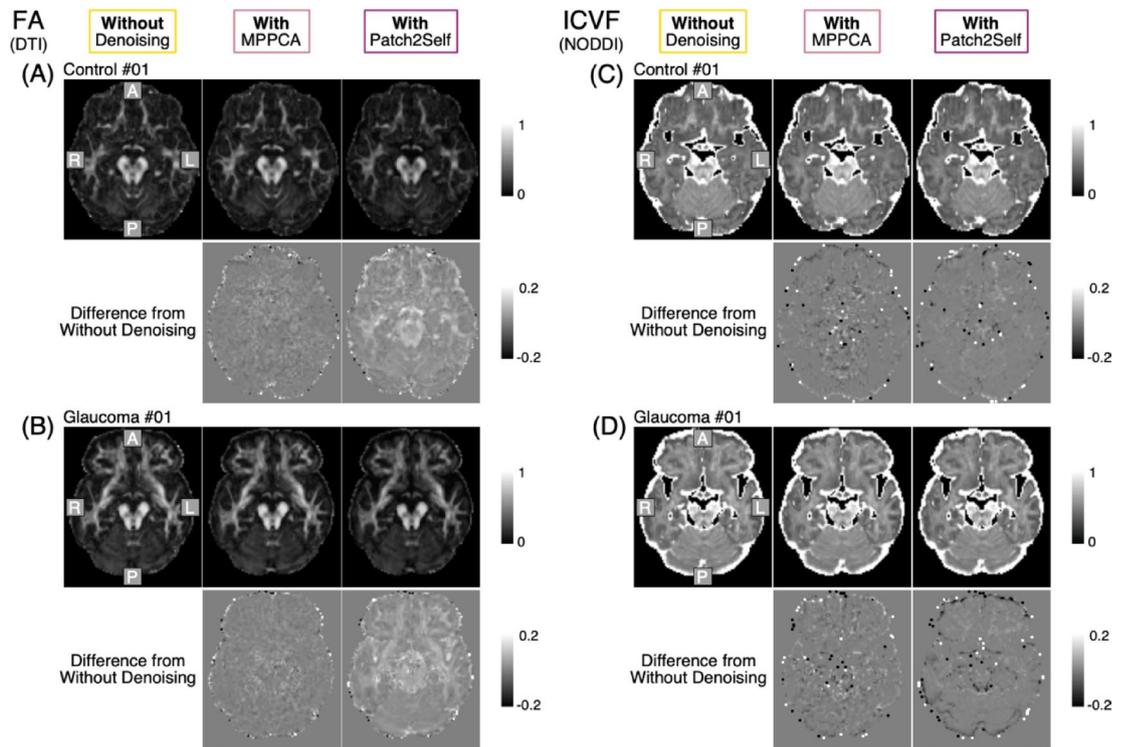


Fig. 3. Impact of denoising on parameter maps estimated using DTI and NODDI in representative subjects (A, C a healthy control; B, D a patient with glaucoma). The top panels depict parameter maps calculated by DTI (A, B) and NODDI (C, D) in an axial section (left, image without denoising; middle, image with MPPCA; right, image with Patch2Self). This axial section is identical to that in Fig. 2. The bottom panels depict difference maps between data without denoising and data with one of the denoising methods (MPPCA and Patch2Self). Letters in the top left image of each panel denote image orientation (A: anterior, P: posterior, L: left, R: right).

in the estimated SNR between the data without denoising and the data with MPPCA was significant in the optic radiation (Supplementary Fig. S2; $d = 4.10$; $BF_{10} > 100$, $t_{46} = 28.08$, $P < 0.001$).

Impact of denoising on voxelwise model fitting

We evaluated the impact of denoising on the fitting of voxelwise models (DTI and NODDI) for dMRI data in the OT. To assess this, we calculated metrics describing model fitting errors in each voxel along the OT (RMSE for DTI, Rician log-likelihood for NODDI) and compared these metrics between the data with (MPPCA and Patch2Self) and without denoising.

Figure 5A shows a comparison of the RMSE in the OT when DTI was used as a voxelwise model for diffusion signals. We found that the data with MPPCA exhibited a significantly lower RMSE compared to the data without denoising ($d = -5.42$; $BF_{10} > 100$; two-tailed paired t-test, $t_{46} = 37.18$, $P < 0.001$). Similarly, the data with Patch2Self also showed a significantly lower RMSE compared to the data without denoising ($d = -6.82$; $BF_{10} > 100$; $t_{46} = 46.74$, $P < 0.001$). Additionally, the data with Patch2Self demonstrated a significantly lower RMSE than the data with MPPCA ($d = -6.16$; $BF_{10} > 100$; $t_{46} = 42.25$, $P < 0.001$). These results suggest that the model prediction using DTI was closer to the measured diffusion signals in the OT after denoising. Notably, this analysis included multiple comparisons across three conditions, but the conclusions would remain unchanged even if Bonferroni correction was applied.

Figure 5B shows a comparison of the NODDI model fitting in the OT using Rician log-likelihood (higher values indicate smaller errors). Consistent with the DTI results, we found that denoising significantly reduced the NODDI model fitting error in the OT voxels (comparison between without denoising and with MPPCA: $d = 3.28$; $BF_{10} > 100$; $t_{46} = 22.50$, $P < 0.001$; comparison between without denoising and with Patch2Self: $d = 9.16$; $BF_{10} > 100$; $t_{46} = 62.76$, $P < 0.001$). Additionally, the data with Patch2Self showed a reduced NODDI fitting error compared to the data with MPPCA ($d = 1.59$; $BF_{10} > 100$; $t_{46} = 10.89$, $P < 0.001$).

We also found similar effects in the voxelwise model fitting for the optic radiation (Supplementary Fig. S3), indicating that denoising substantially reduced errors between the voxelwise model prediction and the measured diffusion signals in both the OT and optic radiation.

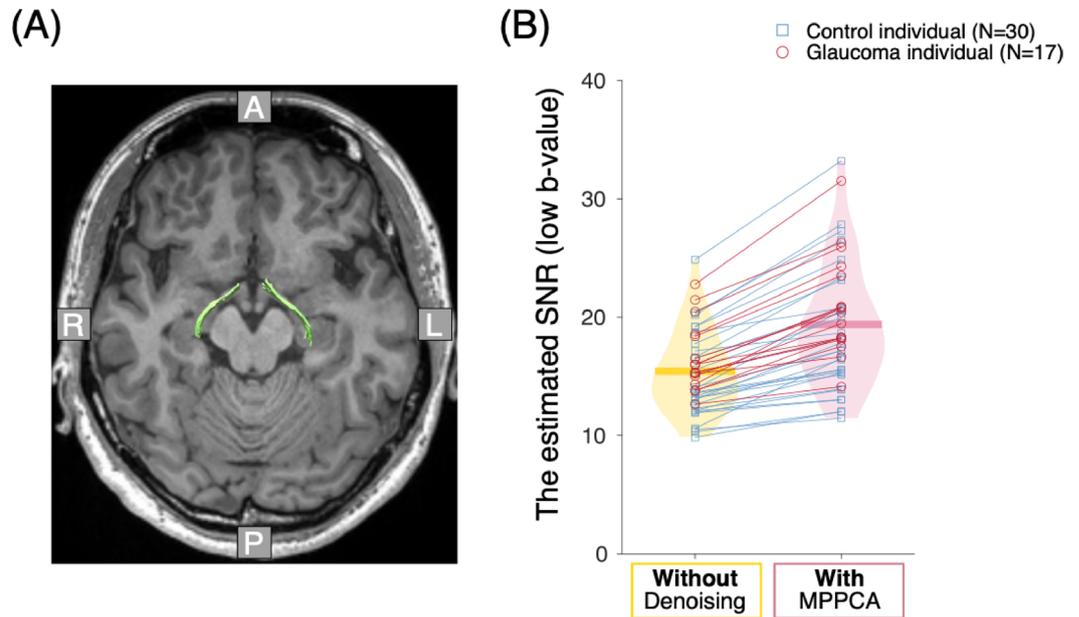


Fig. 4. Comparison of the estimated SNR on low b-value images along the OT between data with and without denoising. **(A)** The OT (green) of a representative subject (Control #01) overlaid on an axial section of the T1-weighted image. Letters in the image denote image orientation (A: anterior, P: posterior, L: left, R: right). **(B)** Comparison of the estimated SNR on low b-value data between data without denoising (yellow) and data with MPPCA. The vertical axis represents the estimated SNR. Blue squares depict data from individual control subjects whereas purple dots depict data from individual patients with glaucoma. Data points connected by lines were acquired from identical subjects. Thick horizontal lines in the violin plot represent the mean across subjects, whereas the widths of the shadowed areas represent the approximate frequency of data points.

Impact of denoising on tractometry metrics to identify tissue changes caused by glaucoma

In the previous sections, we demonstrated that denoising dMRI data increased the estimated SNR and reduced the voxelwise model fitting error in the OT. Here, we address how denoising affects tractometry analysis for identifying differences in OT tissue properties between patients with glaucoma and controls.

Figure 6 shows a comparison of the tractometry data along the OT with and without denoising for each diffusivity metric (FA, MD, ICVF, and ODI). Individual glaucoma patient data were plotted based on each patient's deviation from the control mean using the standard deviation of the controls as the unit. Notably, as shown in the original study¹⁸, patients with glaucoma exhibited lower FA, lower ICVF, and higher ODI values than controls.

In DTI-based metrics (FA and MD), the data with MPPCA did not show any significant differences from that without denoising (FA, $d = 0.54$, $BF_{10} = 1.74$, $t_{16} = 2.23$, $P = 0.04$; MD, $d = 0.09$, $BF_{10} = 0.26$, $t_{16} = 0.37$, $P = 0.72$). Although the data with Patch2Self showed statistically significant differences from that without denoising (FA, $d = 0.72$, $BF_{10} = 5.87$, $t_{16} = 2.97$, $P = 0.009$; MD, $d = 2.64$, $BF_{10} > 100$, $t_{16} = 10.89$, $P < 0.001$), the differences between the controls and patients with glaucoma decreased after Patch2Self (Fig. 6).

The lack of evidence for a significant effect of denoising can be explained by the simplicity of the diffusion tensor model, which may not capture the signal characteristics improved by denoising. This prompted us to test NODDI, a more complex multi-compartment model of diffusion signals. While the data with MPPCA showed no significant difference in ICVF compared to that without denoising ($d = 0.17$, $BF_{10} = 0.31$, $t_{16} = 0.72$, $P = 0.48$), the data with Patch2Self showed a statistically significant difference from that without denoising ($d = -1.43$, $BF_{10} > 100$, $t_{16} = 5.91$, $P < 0.001$), demonstrating that patient data after Patch2Self deviated more from controls. However, the differences in ODI between controls and patients with glaucoma decreased after both MPPCA ($d = -0.53$, $BF_{10} = 1.61$, $t_{16} = 2.18$, $P = 0.05$) and Patch2Self ($d = -0.58$, $BF_{10} = 2.27$, $t_{16} = 2.40$, $P = 0.03$).

Because Patch2Self had opposing impacts on ICVF and ODI for distinguishing patients with glaucoma from controls, we conducted a supplementary analysis to evaluate how these changes affected classification when both metrics (ICVF and ODI) were used in a linear discriminant analysis. The classification performance of the data with Patch2Self was only slightly better than that of the data without denoising (Supplementary Fig. S4; misclassification probabilities, 0.13 for data without denoising, 0.11 for data with Patch2Self). The difference in the misclassification probabilities between the data with and without denoising was not statistically significant (McNemar's test; $P > 0.5$).

We also analyzed the impact of denoising on dMRI measurements of the optic radiation to identify tissue changes caused by glaucoma (Supplementary Fig. S5). Similar to the OT results, we did not find strong evidence that MPPCA improved the ability of dMRI to identify glaucoma-related white matter tissue changes. Patch2Self affected diffusivity metrics in the optic radiation, leading to significantly increased differences between patients

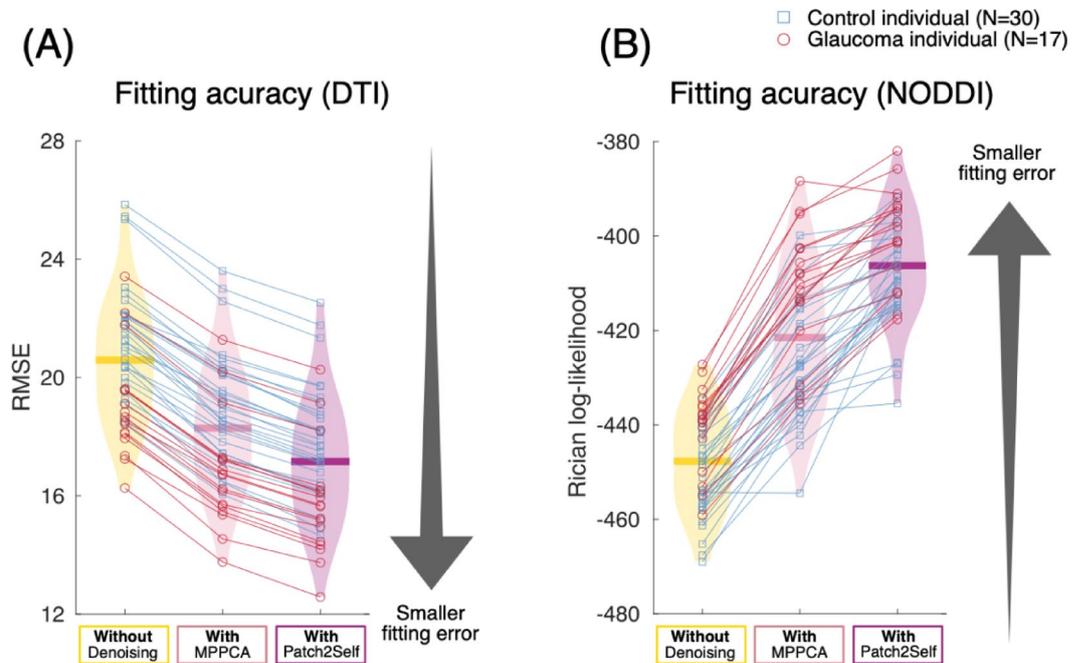


Fig. 5. Comparison of voxelwise model fitting in the OT between dMRI data with and without denoising. **(A)** Fitting of the diffusion tensor model (diffusion tensor imaging, DTI). The vertical axis represents the fitting error of the DTI quantified by the root mean square error (RMSE) for dMRI data with and without denoising (MPPCA and Patch2Self) in the OT, where a lower RMSE corresponds to a smaller error. Open squares/circles depict the data of individual subjects (blue square, controls; red circle, patients with glaucoma). Data points connected among different conditions (without denoising, with MPPCA, and with Patch2Self) by lines are data acquired from identical subjects. Thick horizontal lines in the violin plot represent the mean across subjects, whereas the widths of the violin plot represent the approximate frequency of data points in each condition and RMSE. **(B)** Fitting of the neurite orientation dispersion and density imaging (NODDI). The vertical axis depicts the fitting error of the NODDI quantified by the Rician log-likelihood for dMRI data with and without denoising in the OT. A higher Rician log-likelihood indicates smaller error. The other conventions are the same as those used in panel A.

with glaucoma and controls in MD and ICVF, but not in other metrics. Linear discriminant analysis revealed no statistically significant improvement in discriminability after Patch2Self (Supplementary Fig. S6).

Correlation between visual field test and OT tissue properties

We calculated the correlation between visual field test scores and OT tissue properties in data with and without denoising, to evaluate how denoising affects the correlation between diffusivity measurements and disease severity. In data without denoising, as already reported in Ogawa et al. (2022)¹⁸, ODI showed a significant negative correlation with the visual field test (Supplementary Fig. S7A; $R = -0.61$, $P = 0.009$). In addition, FA showed a significant positive correlation with visual field test (Supplementary Fig. S7A; $R = 0.69$, $P = 0.002$), while MD and ICVF did not (MD: $R = 0.38$, $P = 0.13$; ICVF: $R = 0.30$, $P = 0.24$). We found that data with MPPCA also showed significant correlations between OT tissue properties (ODI and FA) and visual field tests (Supplementary Fig. S7B; ODI: $R = -0.62$, $P = 0.008$; FA: $R = 0.70$, $P = 0.002$), and effect sizes of correlation was almost identical to those of data without denoising. These correlations were also significant in data with Patch2Self denoising, but slightly smaller than those of data without denoising (Supplementary Fig. S7C; ODI: $R = -0.54$, $P = 0.03$; FA: $R = 0.63$, $P = 0.007$).

Impact of denoising on scan-rescan reliability of tractometry

Finally, we evaluated the effect of denoising on the scan-rescan reliability of tractometry by calculating the correlation between datasets from two runs. Figure 7 shows the scan-rescan reliability of the FA and ICVF along the OT in the data with (MPPCA and Patch2Self) and without denoising. Because data from the two runs were acquired with reversed phase-encoding directions, a systematic difference in the ICVF measurements was observed between the two runs (Fig. 7, bottom panels). This likely reflects susceptibility-induced distortions in the OT⁴¹, with each dataset showing systematic differences owing to differences in the distortion correction procedure.

Nevertheless, the scan-rescan reliability of FA and ICVF measurements along the OT remained high in the dataset without denoising (FA, $R = 0.93$; ICVF, $R = 0.87$), demonstrating a high reproducibility of measurements despite the systematic differences found in ICVF. Importantly, the scan-rescan reliability in the data with denoising was not higher but sometimes slightly lower than that in the data without denoising (MPPCA, FA:

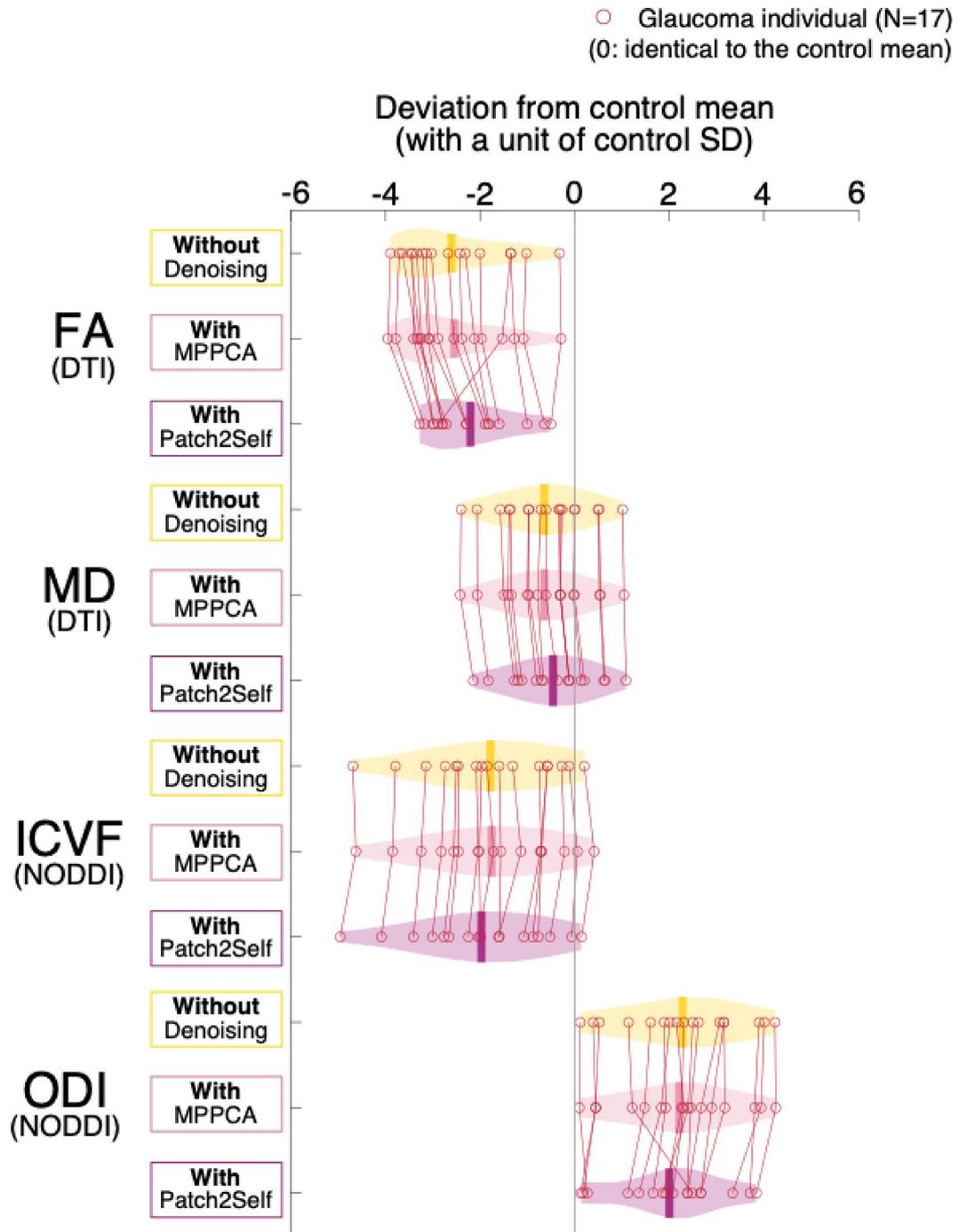


Fig. 6. Comparison of dMRI measurements to identify tissue property differences between patients with glaucoma and controls in the OT among data with and without denoising (MPPCA and Patch2Self). The horizontal axis represents the data of each dMRI-based metric (FA, MD, ICVF, and ODI) in patients with glaucoma normalized to the control mean. The unit of the horizontal axis indicates how much the data of patients with glaucoma deviated from the control mean (0) with a unit of the control standard deviation. The individual dots represent data of each patient with glaucoma, and dots connected by lines indicate data of identical subjects. Thick horizontal lines in the violin plot represent the mean among patients with glaucoma, whereas the widths of the violin plot represent the approximate frequency of data points in each metric and dataset (yellow, data without denoising; magenta, data with MPPCA; purple, data with Patch2Self).

$R=0.92$, ICVF: $R=0.85$; Patch2Self, FA: $R=0.87$, ICVF: $R=0.87$). Denoising did not improve the scan-rescan reliability of other metrics along the OT (MD and ODI, Supplementary Fig. S8).

Furthermore, we performed the same scan-rescan reliability analysis on the data along the optic radiation (Supplementary Fig. S9). Unlike the results in the OT, no strong systematic differences were observed between

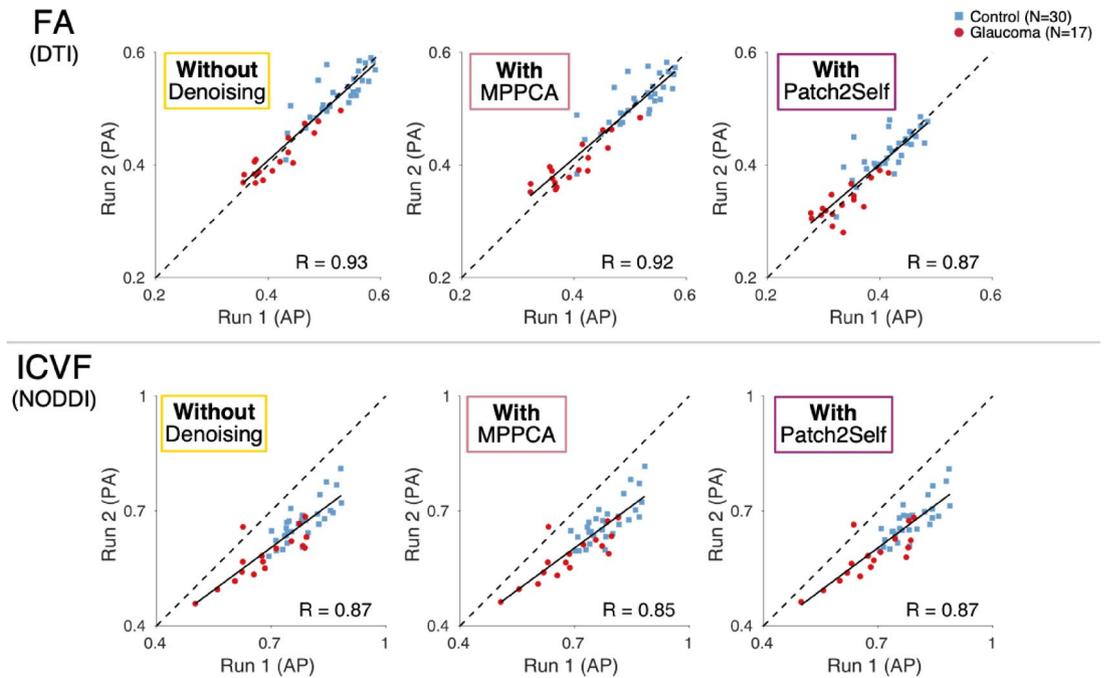


Fig. 7. Scan-rescan reliability of tractometry along the OT using two metrics (top, FA; bottom, ICVF) in data without and with denoising (left: data without denoising; middle, data with MPPCA; right, data with Patch2Self). The horizontal axis shows measurements in Run 1 (acquired with anterior-to-posterior [AP] phase encoding direction), whereas the vertical axis represents measurements in Run 2 (acquired with posterior-to-anterior [PA] phase encoding direction). Each dot represents data from individual subjects (blue squares, controls; red circles, patients with glaucoma). Solid lines indicate regression lines, while dotted lines represent identity lines.

the two datasets with reversed phase encoding directions in the optic radiation, most likely because this pathway is less affected by susceptibility-induced distortion⁴¹. Similar to the results in the OT, denoising did not improve the scan-rescan reliability of tractometry measurements in the optic radiation (Supplementary Fig. S9).

Additionally, we calculated the scan-rescan reliability separately for the control and glaucoma groups to evaluate whether the results could be generalized across both groups. We found no evidence that denoising improved the reproducibility of tractometry in either group (Supplementary Tables S4 and S5).

In summary, we found no evidence that denoising dMRI data improves tractometry scan-rescan reliability in the OT of healthy controls and patients with glaucoma.

Discussion

We aimed to evaluate how denoising affects tractometry on dMRI data acquired from patients with glaucoma to understand how much denoising may benefit the identification of neurobiological changes occurring in the white matter associated with the disease. Because glaucoma is a disorder that damages retinal ganglion cells and the optic nerve, it likely affects the diffusivity of the OT, which is part of the optic nerve. Therefore, we were primarily interested in how denoising can impact the ability of tractometry to identify differences between patients with glaucoma and controls, as well as the scan-rescan reliability in the OT. Our findings demonstrate that denoising (MPPCA and Patch2Self) altered the appearance of diffusion-weighted images (Fig. 2), increased the estimated SNR along the OT (Fig. 4), and reduced the fitting errors of the voxelwise diffusion models (DTI and NODDI, Fig. 5). In contrast, denoising had little or no impact on the ability of tractometry to distinguish between patients with glaucoma and controls (Fig. 6). Furthermore, there was no evidence supporting an improvement in scan-rescan reliability in the tractometry analysis of the OT (Fig. 7). Taken together, while we found that denoising improves dMRI data in some aspects, there is little evidence that it improves the detection and reproducibility of tractometry analysis for tissue changes caused by glaucoma.

Potential interpretations of the lack of evidence on the impact of denoising for tractometry

While we did not find strong evidence that denoising improves tractometry, we do not aim to underestimate the utility of denoising in dMRI. Since denoising clarifies images, it may benefit researchers interested in using dMRI for image diagnostics, although such practical advantages have not been tested within the scope of this study. We found that denoising had minimal impact on tractometry when used to analyze OT tissue properties, but we did not find any unwarranted effects¹⁴. Therefore, we consider that the utility of denoising may depend on the research goals and analysis strategies undertaken by researchers.

Although speculative, two possibilities may explain why denoising did not improve the tractometry results when evaluating the impact of glaucoma on the OT. The first hypothesis is that, while denoising alters the

appearance of images, it does not enhance neurobiologically meaningful information in the dMRI dataset. The second hypothesis is that denoising plays a role in improving dMRI data similar to that of the approach typically included in the tractometry pipeline: voxelwise diffusion model fitting. Because dMRI data contain measurement noise, a common practice is to fit a model, such as DTI or NODDI, to describe the diffusion signals in individual voxels. One purpose of model fitting is to characterize signals, and another is to reduce the impact of noise by fitting models that accurately describe expected signal characteristics³⁷. For this purpose, researchers have proposed model-fitting procedures that are robust against noise in the data⁴². Therefore, denoising and model fitting may have overlapping roles when used together for tractometry; in fact, we found that denoising reduced residuals in model fitting (Fig. 5), but this did not significantly impact subsequent tractometry analyses (Figs. 6 and 7). Therefore, we hypothesize that denoising may remove noise in dMRI data that would normally be excluded by voxelwise model fitting; when both strategies are used together, denoising does not significantly alter the results. Importantly, this possibility should be explored in future investigations to better understand the mechanisms underlying the observed effects.

This raises the question of whether denoising can improve tractometry if the voxelwise characterization of diffusion signals does not rely on fitting relatively simple models. For example, previous studies have proposed model-free methods that do not rely on voxelwise model fitting^{43–45} or more complex voxelwise diffusion models than DTI and NODDI^{37,46–51}. Denoising may have greater benefits when these methods are used to characterize diffusion signals. However, we note that our approach to use voxelwise model fitting using DTI and NODDI is a widely accepted approach in tractometry studies^{5,6,52,53}, as it is applicable to dMRI datasets with a moderate number of directions and b-values, such as clinical neuroimaging data. In fact, a recent systematic review showed that tractometry studies on patients with glaucoma adopted this approach, rather than model-free methods or more complex voxelwise models⁵⁴. Therefore, while the utility of denoising for model-free methods and more complex voxelwise models should be tested in future research on data that satisfy the prerequisites of each approach, such testing falls outside the scope of this study, which aims to test the impact of denoising on tractometry analysis for patients with glaucoma.

One might argue that we did not observe an improvement in the scan-rescan reliability owing to a ceiling effect, as it was already high in the original dataset. This is consistent with a previous study indicating that the reproducibility of tractometry analysis is high⁵². However, we note that the dataset we analyzed was not of a particularly high quality compared with that of some publicly available datasets (e.g., the Human Connectome Project Dataset⁵⁵), as it was acquired using a relatively old MRI scanner model (Siemens Trio Tim) with moderate gradient strength (40 mT/m). In addition, it was acquired with a modest number of b-values and directions within a relatively short acquisition time. Furthermore, we also found that denoising reduces voxelwise model fitting error (Fig. 5), suggesting that a considerable amount of noise was present in our dMRI datasets. While we cannot exclude the possibility that denoising may improve tractometry if applied to data of much lower quality, the quality of this dataset is comparable to those analyzed in many clinical neuroimaging studies. Therefore, we believe that there is still room for improvement in the ability to detect differences and in the reliability of tractometry in our dataset.

In parallel to the main findings, we observed that Patch2Self causes systematic effects on diffusivity metrics, such as a decrease in FA and an increase in ODI (Supplementary Fig. S1). This effect did not significantly affect group discriminability between patients with glaucoma and controls, most likely because diffusion metrics with and without Patch2Self are highly correlated across subjects (Figure S1). However, since there is no ground truth in *in vivo* dMRI data, we can only speculate on the underlying cause of this systematic effect. One hypothesis is that this effect may be associated with an increased SNR after denoising. However, this interpretation must be approached with caution, as a numerical simulation study suggested that the relationship between the diffusion anisotropy and SNR can be complex; a low SNR may lead to either underestimation or overestimation of the anisotropy⁵⁶. Another hypothesis is that denoising reduces the large signals in images when dMRI data are acquired in specific gradient orientations, thus causing a reduction in FA, which may be analogous to the over-smoothing effect on structural images discussed in a previous paper¹⁴. Visual inspection revealed that Patch2Self removed the high-contrast information visible in the original diffusion-weighted images (Fig. 2B), consistent with this hypothesis. Nevertheless, given the absence of a ground-truth, this study alone cannot provide a definitive interpretation of why this effect occurs and whether it is beneficial by reducing noise or disadvantageous by removing true diffusion signals and introducing potential biases, requiring future investigations.

Related studies

Evaluating the impact of denoising on neurobiological information measured by dMRI is generally challenging, given the absence of ground truth in *in vivo* dMRI data. One approach is to test dMRI data acquired from phantom⁵⁷ or biological specimens with simple fiber structure, such as plants⁵⁸. However, this approach has limitations since the configuration of axons in the brain is hard to reproduce in a phantom and not similar to those in plants, resulting in a method that works best for phantom or plant data does not always work the best for *in vivo* human data⁵⁹. The other approach was to add simulated noise to dMRI data and evaluate how much denoising could remove it^{60,61}. Although this approach has its merits, the evaluation of empirical datasets remains necessary because of the general challenge of ensuring that simulated noise follows the same characteristics as noise in empirical datasets.

For these reasons, other lines of previous studies have evaluated the impact of denoising on empirical datasets using various approaches available for *in vivo* dMRI data. For example, previous studies have reported improvements in image appearance^{10,12,50} and estimated SNR⁶² after denoising. Our observations are consistent with these findings (Figs. 2 and 4). Moreover, both a previous study¹² and our results (Fig. 5) indicate that voxelwise model fitting errors were reduced after denoising, highlighting certain advantages of denoising dMRI data.

However, Schilling et al.⁶¹ noted that after voxelwise model fitting, the resulting diffusivity measurements (such as FA and MD) were not significantly different between data with and without denoising in spinal cord voxels. Additionally, Sagawa et al.⁶² also pointed out that denoising only has a significant impact on FA in deep gray matter, but not in white matter. These findings are consistent with our results, showing that denoising with MPPCA had a small impact on the diffusivity metrics obtained by voxelwise model fitting in the OT (Fig. 2; but see Supplementary Fig. S1 for the impact of Patch2Self on diffusivity metrics). Our contribution is to demonstrate the limited impact of denoising on the ability to identify tissue changes caused by glaucoma and on the scan-rescan reliability, thereby providing further insights into the neurobiological significance of denoising when combined with current tractometry approaches.

While denoising had a small impact on the voxelwise diffusivity metrics obtained by model fitting, previous studies have also noted that it reduces the inter-voxel variability of diffusivity metrics^{61,62}. We speculate that although denoising affects inter-voxel variability, it may have a limited impact on tractometry, which is an approach that summarizes diffusivity metrics at each node of the tract based on a weighted sum across multiple voxels.

Scopes and Limitations of this study

We did not focus on the impact of denoising on tractography in this study because tractography results are unlikely to differ substantially if performed separately on data with and without denoising, as we focused on the OT, a straight white matter tract without fiber crossing⁶³. We also identified seed regions based on established anatomical definitions to identify these tracts. Therefore, we used a consistent definition of the OT and optic radiation across datasets with and without denoising. However, it is possible that denoising may improve the tractography of white matter tracts for which prior anatomical information is not established, which makes it challenging to precisely determine the tract identification protocol based on seed regions. While addressing this question is not within the scope of this study, the impact of denoising on tractography needs to be considered in future studies of other white matter tracts and disorders.

This study has several limitations. First, our evaluation did not include the NOise reduction with the DIstribution Corrected (NORDIC) PCA approach, which has been developed for denoising dMRI datasets with complex images and additional noise scans⁶⁴. Unfortunately, we could not test this approach because only magnitude images were available for the current datasets. The advantages of denoising using NORDIC or other methods requiring complex data⁶⁵ should be assessed in future investigations. Second, it is unclear whether our findings can be generalized to other visual disorders such as optic neuritis⁶⁶, because their underlying mechanisms causing white matter tissue damage may differ from those of glaucoma. Third, because we focused on glaucoma and early visual pathways (OT and optic radiation), we could not test the extent to which the findings can be generalized to other white matter tracts with different tissue properties, such as more complex fiber crossings (like acoustic radiation⁶⁷). These limitations pose challenges in broadly applying the findings of this study to different anatomical regions. Fourth, the dataset analyzed in this study may not have had sufficient statistical power to identify effects of denoising on tractometry if the effect size is very small. In addition, the sample size of patients with glaucoma in the dataset was insufficient to compare glaucoma subgroups and determine how denoising may affect subgroup differences, as dividing patients' data into multiple subgroups significantly limits statistical power. Future studies with larger cohorts may facilitate assessments considering variabilities in comorbidities, pharmacological profiles, and types of treatment as additional factors that may affect white matter tissue properties. The generalization of the conclusions to other brain regions, populations, and types of disorders remains a topic for future research.

Conclusion

We showed that denoising (MPPCA and Patch2Self) altered the overall image appearance, increased the estimated SNR, and reduced the voxelwise model fitting error in dMRI data along the OT. However, these two denoising methods had limited impact on the ability of tractometry analysis to identify white matter tissue changes caused by glaucoma. In addition, they did not have a major impact on the reproducibility of tractometry. Taken together, our results suggest that denoising has a limited impact on dMRI-based measurements of white matter tissue properties when combined with tractometry.

Data availability

We made the dataset and codes for replicating figures and statistical analyses publicly available on GitHub (<https://github.com/OkazakiTakemuraLab/GlaucomadMRIdenoising>; <https://doi.org/10.5281/zenodo.15015932>).

Received: 6 February 2025; Accepted: 7 July 2025

Published online: 16 July 2025

References

- Ramrattan, R. S. et al. Prevalence and causes of visual field loss in the elderly and associations with impairment in daily functioning: the Rotterdam Study. *Arch. Ophthalmol.* **119**, 1788–1794 (2001).
- McCarty, C. A., Nanjan, M. B. & Taylor, H. R. Vision impairment predicts 5 year mortality. *Br. J. Ophthalmol.* **85**, 322–326 (2001).
- Iwase, A., Suzuki, Y., Araie, M. & Tajimi Study Group. Characteristics of undiagnosed primary open-angle glaucoma: the Tajimi Study. *Ophthalmic Epidemiol.* **21**, 39–44 (2014).
- Gupta, N. & Yücel, Y. H. What changes can we expect in the brain of glaucoma patients? *Surv. Ophthalmol.* **52** (Suppl 2), S122–S126 (2007).
- Jones, D. K., Travis, A. R., Eden, G., Pierpaoli, C. & Basser, P. J. PASTA: Pointwise assessment of streamline tractography attributes. *Magn. Reson. Med.* **53**, 1462–1467 (2005).

6. Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A. & Feldman, H. M. Tract profiles of white matter properties: automating fiber-tract quantification. *PLoS One*. **7**, e49790 (2012).
7. Li, K. et al. Alteration of fractional anisotropy and mean diffusivity in glaucoma: novel results of a meta-analysis of diffusion tensor imaging studies. *PLoS One*. **9**, e97445 (2014).
8. Tax, C. M. W., Bastiani, M., Veraart, J., Garyfallidis, E. & Irfanoglu, M. O. What's new and what's next in diffusion MRI preprocessing. *Neuroimage* **249**, 118830 (2022).
9. Manjón, J. V. et al. Diffusion weighted image denoising using overcomplete local PCA. *PLoS One*. **8**, e73021 (2013).
10. Veraart, J. et al. Denoising of diffusion MRI using random matrix theory. *Neuroimage* **142**, 394–406 (2016).
11. Cheng, H. et al. Denoising diffusion weighted imaging data using convolutional neural networks. *PLoS One*. **17**, e0274396 (2022).
12. Fadnavis, S., Batson, J. & Garyfallidis, E. Patch2Self: denoising diffusion MRI with self-supervised learning. *ArXiv [cs LG]*. <https://doi.org/10.48550/ARXIV.2011.01355> (2020).
13. Muckley, M. J. et al. Training a neural network for Gibbs and noise removal in diffusion MRI. *Magn. Reson. Med.* **85**, 413–428 (2021).
14. Kay, K. The risk of bias in denoising methods: Examples from neuroimaging. *PLoS One*. **17**, e0270895 (2022).
15. Chen, Z. et al. Diffusion tensor magnetic resonance imaging reveals visual pathway damage that correlates with clinical severity in glaucoma. *Clin. Exp. Ophthalmol.* **41**, 43–49 (2013).
16. Hanekamp, S. et al. White matter alterations in glaucoma and monocular blindness differ outside the visual system. *Sci. Rep.* **11**, 6866 (2021).
17. Haykal, S., Curcic-Blake, B., Jansonius, N. M. & Cornelissen, F. W. Fixel-based analysis of visual pathway white matter in primary open-angle glaucoma. *Invest. Ophthalmol. Vis. Sci.* **60**, 3803–3812 (2019).
18. Ogawa, S. et al. Multi-contrast magnetic resonance imaging of visual white matter pathways in patients with glaucoma. *Invest. Ophthalmol. Vis. Sci.* **63**, 29 (2022).
19. Andersson, J. L. R., Skare, S. & Ashburner, J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* **20**, 870–888 (2003).
20. Andersson, J. L. R. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* **125**, 1063–1078 (2016).
21. Griswold, M. A. et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* **47**, 1202–1210 (2002).
22. Mattiello, J., Basser, P. J. & LeBihan, D. Analytical expressions for the b matrix in NMR diffusion imaging and spectroscopy. *J. Magn. Reson. A.* **108**, 131–141 (1994).
23. Setsompop, K. et al. Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *Neuroimage* **63**, 569–580 (2012).
24. Tournier, J. D. et al. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage* **202**, 116137 (2019).
25. Garyfallidis, E. et al. Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinform.* **8**, 8 (2014).
26. Basser, P. J., Mattiello, J. & LeBihan, D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson. B.* **103**, 247–254 (1994).
27. Basser, P. J. & Pierpaoli, C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J. Magn. Reson. B.* **111**, 209–219 (1996).
28. Pierpaoli, C. & Basser, P. J. Toward a quantitative assessment of diffusion anisotropy. *Magn. Reson. Med.* **36**, 893–906 (1996).
29. Veraart, J., Sijbers, J., Sunaert, S., Leemans, A. & Jeurissen, B. Weighted linear least squares Estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. *Neuroimage* **81**, 335–346 (2013).
30. Zhang, H., Schneider, T., Wheeler-Kingshott, C. A. & Alexander, D. C. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage* **61**, 1000–1016 (2012).
31. Sherbondy, A. J., Dougherty, R. F., Ben-Shachar, M., Napel, S. & Wandell, B. A. ConTrack: finding the most likely pathways between brain regions using diffusion tractography. *J. Vis.* **8**, 15.1–16 (2008).
32. Takemura, H. et al. Diffusivity and quantitative T1 profile of human visual white matter tracts after retinal ganglion cell damage. *NeuroImage Clin.* **23**, 101826 (2019).
33. Kida, I. et al. Comparison of diffusion-weighted imaging in the human brain using readout-segmented EPI and PROPELLER turbo spin echo with single-shot EPI at 7 T MRI. *Invest. Radiol.* **51**, 435–439 (2016).
34. Miyata, T., Benson, N. C., Winawer, J. & Takemura, H. Structural covariance and heritability of the optic tract and primary visual cortex in living human brains. *J. Neurosci.* **42**, 6761–6769 (2022).
35. Reeder, S. B. et al. Practical approaches to the evaluation of signal-to-noise ratio performance with parallel imaging: application with cardiac imaging and a 32-channel cardiac coil. *Magn. Reson. Med.* **54**, 748–754 (2005).
36. Bechtold, B., Fletcher, P., seamusholden & Gorur-Shandilya, S. *Bastibe/Violinplot-Matlab: A Good Starting Point* (Zenodo, 2021). <https://doi.org/10.5281/ZENODO.4559847>
37. Rokem, A. et al. Evaluating the accuracy of diffusion MRI models in white matter. *PLoS One*. **10**, e0123272 (2015).
38. Tuch, D. S. et al. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Reson. Med.* **48**, 577–582 (2002).
39. Krekelberg, B. *BayesFactor: A Matlab Package for Bayes Factor Statistical Analysis* (Github, 2024).
40. Cardillo, G. *McNemar: Permorm McNemar's Chi Square on a 2x2 Matrix* (Github, 2024).
41. Takemura, H., Liu, W., Kuribayashi, H., Miyata, T. & Kida, I. Evaluation of simultaneous multi-slice readout-segmented diffusion-weighted MRI acquisition in human optic nerve measurements. *Magn. Reson. Imaging.* **102**, 103–114 (2023).
42. Chang, L. C., Jones, D. K. & Pierpaoli, C. RESTORE: robust Estimation of tensors by outlier rejection. *Magn. Reson. Med.* **53**, 1088–1095 (2005).
43. Tuch, D. S. Q-ball imaging. *Magn. Reson. Med.* **52**, 1358–1372 (2004).
44. Van, A. T., Granziera, C. & Bammer, R. An introduction to model-independent diffusion magnetic resonance imaging. *Top. Magn. Reson. Imaging.* **21**, 339–354 (2010).
45. Wedeen, V. J., Hagmann, P., Tseng, W. Y. I., Reese, T. G. & Weisskoff, R. M. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn. Reson. Med.* **54**, 1377–1386 (2005).
46. Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y. & Basser, P. J. Axcaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magn. Reson. Med.* **59**, 1347–1354 (2008).
47. Assaf, Y., Freidlin, R. Z., Rohde, G. K. & Basser, P. J. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magn. Reson. Med.* **52**, 965–978 (2004).
48. Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* **34**, 144–155 (2007).
49. Palombo, M. et al. SANDI: A compartment-based model for non-invasive apparent soma and neurite imaging by diffusion MRI. *Neuroimage* **215**, 116835 (2020).
50. Schiavi, S. et al. Mapping tissue microstructure across the human brain on a clinical scanner with Soma and neurite density image metrics. *Hum. Brain Mapp.* **44**, 4792–4811 (2023).
51. Tournier, J. D., Calamante, F. & Connelly, A. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical Deconvolution. *Neuroimage* **35**, 1459–1472 (2007).

52. Kruper, J. et al. Evaluating the reliability of human brain white matter tractometry. *Apert Neuro* **1**, 1–25 (2021).
53. Yendiki, A. et al. Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Front. Neuroinform.* **5**, 23 (2011).
54. Kasa, L. W. et al. Application of advanced diffusion MRI based tractometry of the visual pathway in glaucoma: a systematic review. *Front. Neurosci.* **19**, 1577991 (2025).
55. Van Essen, D. C. et al. The Human Connectome Project: a data acquisition perspective. *Neuroimage* **62**, 2222–2231 (2012).
56. Jones, D. K. & Basser, P. J. “Squashing peanuts and smashing pumpkins”: how noise distorts diffusion-weighted MR data. *Magn. Reson. Med.* **52**, 979–993 (2004).
57. Côté, M. A. et al. Tractometer: towards validation of tractography pipelines. *Med. Image Anal.* **17**, 844–857 (2013).
58. Boujraf, S., Luytjaert, R., Eisendrath, H. & Osteaux, M. Echo planar magnetic resonance imaging of anisotropic diffusion in asparagus stems. *Magn. Reson. Mater. Phys., Biol. Med.* **13**, 82–90 (2001).
59. Jbadi, S. & Johansen-Berg, H. Tractography: where do we go from here? *Brain Connect.* **1**, 169–183 (2011).
60. Ades-Aron, B. et al. Evaluation of the accuracy and precision of the diffusion parameter Estimation with Gibbs and Noise removal pipeline. *Neuroimage* **183**, 532–543 (2018).
61. Schilling, K. G. et al. Denoising of diffusion MRI in the cervical spinal cord - effects of denoising strategy and acquisition on intra-cord contrast, signal modeling, and feature conspicuity. *Neuroimage* **266**, 119826 (2023).
62. Sagawa, H. et al. Deep learning-based noise reduction for fast volume diffusion tensor imaging: assessing the noise reduction effect and reliability of diffusion metrics. *Magn. Reson. Med. Sci.* **20**, 450–456 (2021).
63. Takemura, H., Kruper, J. A., Miyata, T. & Rokem, A. Tractometry of human visual white matter pathways in health and disease. *Magn. Reson. Med. Sci.* **23**, 316–340 (2024).
64. Moeller, S. et al. NOise reduction with distribution corrected (NORDIC) PCA in dMRI with complex-valued parameter-free locally low-rank processing. *Neuroimage* **226**, 117539 (2021).
65. Manzano Patron, J. P. et al. Denoising diffusion MRI: considerations and implications for analysis. *Imaging Neurosci.* **2**, 1–29 (2024).
66. Backner, Y. et al. Anatomical wiring and functional networking changes in the visual system following optic neuritis. *JAMA Neurol.* **75**, 287 (2018).
67. Maffei, C., Sarubbo, S. & Jovicich, J. Diffusion-based tractography atlas of the human acoustic radiation. *Sci. Rep.* **9**, 4046 (2019).

Acknowledgements

We thank Ariel Rokem for his suggestions for an earlier version of the manuscript, Isao Yokoi for preparing the computing environment, Junxiang Luo for assisting with visualization, and Gary H. Zhang for his suggestions on analyses using NODDI. We also thank Editage [<http://www.editage.com>] for editing and reviewing this manuscript for the English language.

Author contributions

Conceptualization: DT SO HT, Data curation: SO HT, Formal analysis: DT HT, Funding acquisition: DT SO HT, Investigation: DT, Project administration: HT, Resources: SO HT, Software: HT, Supervision: HT, Validation: DT HT, Visualization: DT, Writing—Original draft preparation: DT HT, Writing—Review & editing: DT SO HT.

Funding

This study was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (JP25KJ1324 to D.T.; JP20K18396 to S.O.; JP22K09841 and JP24K03240 to H.T.), the Joint Research Program (23NIPS141, 24NIPS253, and 25NIPS202) of the National Institute for Physiological Sciences, and the MEXT Promotion of Development of a Joint Usage/Research System Project: Coalition of Universities of Research Excellence Program (CURE; Grant Number: JPMXP1323015488; Spin-L Program Number: spin24XN018 and spin25XN011).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-10947-6>.

Correspondence and requests for materials should be addressed to D.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025