JMI Journal of Medical Imaging

RESEARCH PAPER

Characterizing patterns of diffusion tensor imaging variance in aging brains

Chenyu Gao,[a,*] Qi Yang,[b] Michael E. Kim,[b] Nazirah Mohd Khairi,[a] Leon Y. Cai,[c] Nancy R. Newlin,[b] Praitayini Kanakaraj,[b] Lucas W. Remedios,[b] Aravind R. Krishnan,[a] Xin Yu,[b] Tianyuan Yao,[b] Panpan Zhang,[d] Kurt G. Schilling,[e,f] Daniel Moyer,[b] Derek B. Archer,[g,h] Susan M. Resnick,[i] Bennett A. Landman,[a,b,c,e,f] for the Alzheimer's Disease Neuroimaging Initiative[†] and the BIOCARD Study team[‡]

[a]Vanderbilt University, Department of Electrical and Computer Engineering, Nashville, Tennessee, United States
[b]Vanderbilt University, Department of Computer Science, Nashville, Tennessee, United States
[c]Vanderbilt University, Department of Biomedical Engineering, Nashville, Tennessee, United States
[d]Vanderbilt University Medical Center, Department of Biostatistics, Nashville, Tennessee, United States
[e]Vanderbilt University Medical Center, Department of Radiology and Radiological Sciences, Nashville, Tennessee, United States
[f]Vanderbilt University, Vanderbilt University Institute of Imaging Science, Nashville, Tennessee, United States
[g]Vanderbilt Memory and Alzheimer's Center, Vanderbilt University School of Medicine, Nashville, Tennessee, United States
[h]Vanderbilt University Medical Center, Vanderbilt Genetics Institute, Nashville, Tennessee, United States
[i]National Institute on Aging, Laboratory of Behavioral Neuroscience, Baltimore, Maryland, United States

ABSTRACT. **Purpose:** As large analyses merge data across sites, a deeper understanding of variance in statistical assessment across the sources of data becomes critical for valid analyses. Diffusion tensor imaging (DTI) exhibits spatially varying and correlated noise, so care must be taken with distributional assumptions. Here, we characterize the role of physiology, subject compliance, and the interaction of the subject with the scanner in the understanding of DTI variability, as modeled in the spatial variance of derived metrics in homogeneous regions.

Approach: We analyze DTI data from 1035 subjects in the Baltimore Longitudinal Study of Aging, with ages ranging from 22.4 to 103 years old. For each subject, up to 12 longitudinal sessions were conducted. We assess the variance of DTI scalars within regions of interest (ROIs) defined by four segmentation methods and investigate the relationships between the variance and covariates, including baseline age, time from the baseline (referred to as "interval"), motion, sex, and whether it is the first scan or the second scan in the session.

Results: Covariate effects are heterogeneous and bilaterally symmetric across ROIs. Inter-session interval is positively related ($p \ll 0.001$) to FA variance in the cuneus and occipital gyrus, but negatively ($p \ll 0.001$) in the caudate nucleus. Males show significantly ($p \ll 0.001$) higher FA variance in the right putamen,

*Address all correspondence to Chenyu Gao, chenyu.gao@vanderbilt.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 did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

[‡]Data used in preparation of this article were derived from BIOCARD study data, supported by grant U19 –AG033655 from the National Institute on Aging. The BIOCARD study team did not participate in the analysis or writing of this report; however, they contributed to the design and implementation of the study. A listing of BIOCARD investigators may be accessed at https://www.biocard-se.org/public/Core%20Groups.html

thalamus, body of the corpus callosum, and cingulate gyrus. In 62 out of 176 ROIs defined by the Eve type-1 atlas, an increase in motion is associated ($p < 0.05$) with a decrease in FA variance. Head motion increases during the rescan of DTI ($\Delta\mu = 0.045$ mm per volume).

Conclusions: The effects of each covariate on DTI variance and their relationships across ROIs are complex. Ultimately, we encourage researchers to include estimates of variance when sharing data and consider models of heteroscedasticity in analysis. This work provides a foundation for study planning to account for regional variations in metric variance.

 [DOI: 10.1117/1.JMI.11.4.044007]

Keywords: brain; aging; DTI; variance; motion

Paper 24075GR received Mar. 11, 2024; revised Jul. 26, 2024; accepted Jul. 30, 2024; published Aug. 24, 2024.

1 Introduction

Large datasets enable the exploration of questions that would be impractical with smaller- or moderate-sized datasets while giving rise to the development and application of deep learning models, which can assimilate complex data. One prevalent challenge is that large datasets often comprise samples aggregated from distinct sources at different time points using diverse technologies, causing data heterogeneity, experimental variations, and statistical biases if the analysis is not executed appropriately.[1] In such scenarios, understanding the variance and variability of data is of great importance. The general linear model,[2] a structured and widely used framework for relationship modeling, allows us to illustrate the importance. The general linear model is assessed through regression. A linear regression can be expressed by $Y = X\beta + \varepsilon$, where the response variable Y, the covariate matrix X, and the regression coefficients β are conventionally represented in matrix forms given by

$$Y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \\ \vdots \\ y^{(M)} \end{bmatrix} \quad X = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ x^{(3)} \\ \vdots \\ x^{(M)} \end{bmatrix} = \begin{bmatrix} 1 & x_1^{(1)} & \cdots & x_N^{(1)} \\ 1 & x_1^{(2)} & \cdots & x_N^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^{(M)} & \cdots & x_N^{(M)} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix}, \tag{1}$$

where we use M to denote the number of samples and N to denote the number of independent variables. The error term ε is given by

$$\varepsilon \sim \mathcal{N}(0, \Sigma), \tag{2}$$

where Σ represents the covariance matrix. If we assume the errors are uncorrelated, Σ is simplified to a diagonal matrix. We can simplify the estimation of Eq. (1) with a whitening matrix,[3] W

$$WY = WX\beta + W\varepsilon, \quad \text{where } W = W^T = \Sigma^{-1/2}. \tag{3}$$

Note $W\varepsilon \sim \mathcal{N}(0, I_M)$, where I_M denotes the identity matrix of dimension $M \times M$. We illustrate the practical importance of understanding the variance structure for reducing statistical errors (Fig. 1).

Diffusion tensor imaging (DTI)[4–6] is a modeling approach used in diffusion-weighted imaging,[7–9] a variant of conventional magnetic resonance imaging (MRI) based on the tissue water diffusion rate.[10] DTI allows for visualization and measurement of the degree of anisotropy and structural orientation of fibers in the brain and has been widely used in studies.[11–13] DTI is inherently subject to low signal-to-noise ratio (SNR), and the noise structure exhibits spatial variability and correlation, primarily attributed to fast imaging and noise suppression techniques.[14,15] Understanding the statistical nature of DTI variance or noise has been proven to be beneficial for diffusion tensor estimation,[15,16] outlier detection,[17] and reproducibility assessment.[18–20]

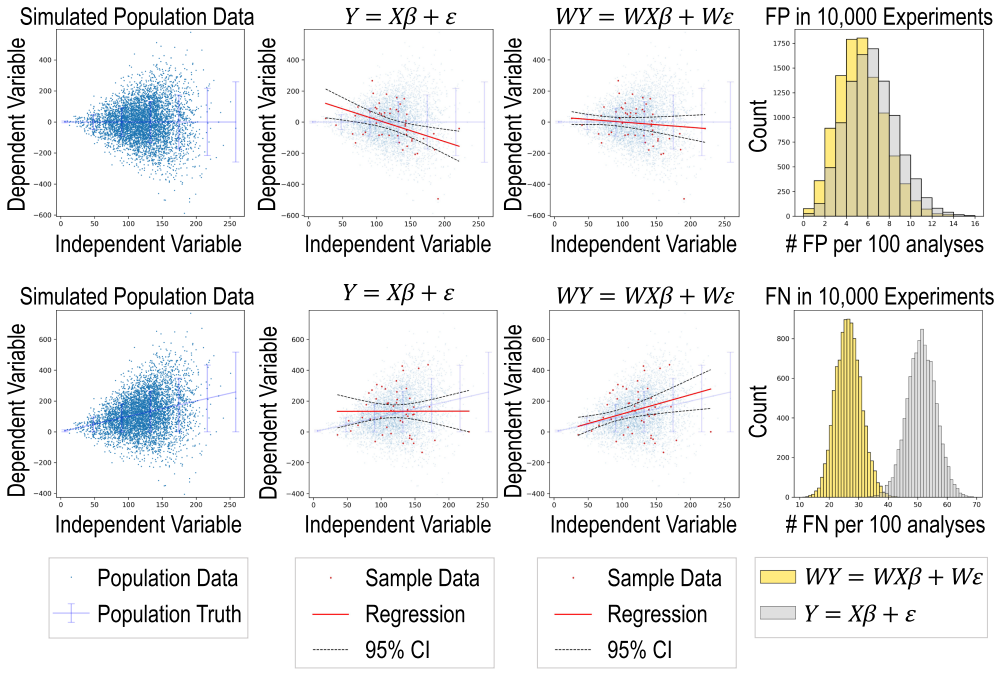

Fig. 1 Simulation shows that applying the whitening matrix to the standard linear regression equation reduces the number of false positives (FP) and false negatives (FN) under heteroskedasticity. In the top row, the population truth has zero slope. In data sampled from the synthetic population data, ordinary least square (OLS) regression using the standard equation generates FP, while the solution with whitening, W, does not falsely reject the null hypothesis (the horizontal line). After 10,000 experiments, the FP frequency is lower with whitening, centering at 5 per 100. In the second row, the population truth has a positive slope. In data sampled from the synthetic population data, OLS regression using the standard equation generates FN, while the solution with whitening, W, does not. After 10,000 experiments, the FN frequency with whitening is half that of the one without whitening.

Methods have been proposed for DTI variance (or noise) estimation, among which we recognize four types. The first type requires multiple repeated acquisitions (therefore, we refer to it as the "multiple acquisition method"). After the repeated acquisitions, we could take the standard deviation of the measurements, or perform data resampling, such as bootstrap or jackknife,[21] to quantify uncertainties of DTI parameters. The second type involves two repeated acquisitions, which we call the "scan-rescan method." We compute the difference between the images from each acquisition and then calculate the standard deviation of the difference across the space.[22,23] Note that the standard deviation must be divided by the square root of 2 to account for the combination of two random variables (noise in each image). The third and fourth types are used when we have only one acquisition. For the third type, we select a homogeneous region of interest (ROI) and compute the standard deviation of the measurements within this ROI ("ROI-based method").[24] The fourth type, often referred to as "model-based resampling," involves fitting a model (e.g., diffusion tensor) locally to the observed data. The residuals from the fitted model, along with the original observed data, are then used by data resampling techniques such as wild bootstrap to generate random subsets.[21,25] From each subset, we obtain an estimate of a specific parameter. Across all subsets, we get the distribution of the estimates and thus quantify the uncertainty of the DTI parameter. The first type (multiple acquisition method) makes no assumptions about the noise properties at the cost of requiring multiple acquisitions (for each diffusion gradient in the DTI scenario).[21] The second type (scan-rescan method) assumes that the noise is constant across space, or across the region from which the standard deviation is computed.[23] The third method (ROI-based method) assumes that both the noise and the signal are constant across the ROI.[24] The fourth method (model-based resampling) assumes that the non-constant variance of measured signals can be captured by the chosen model.[21] In this study, we choose the ROI-based method for estimating noise across brain regions in DTI. This is because it does not

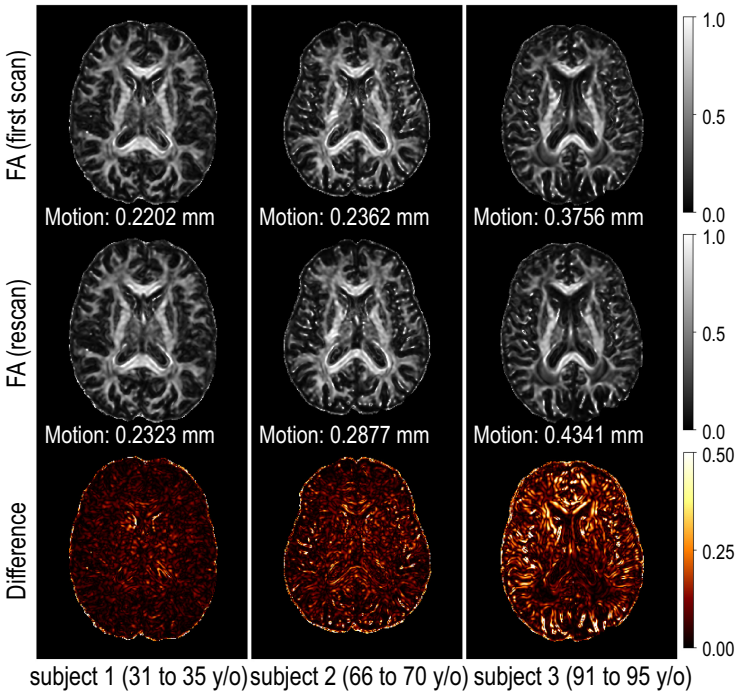

Fig. 2 We observe that the noise (approximated by the difference between the scan and rescan acquired within the same imaging session) in DTI scalar images, such as FA images, generally increases with age. (Subjects' ages are grouped into 5-year bins to respect privacy.) But motion is also considered to increase with age.[26,27] We would like to know the following: Which factor is associated with DTI variance? Where and how does this association manifest?

require repeated acquisitions, and one advantage is that we can compute noise from each scan within an imaging session, enabling inter-scan comparisons. First, this method can be applied to datasets that do not have scan-rescan data available, thus enabling validation of our findings using other datasets. Second, by using an individual scan for noise estimation, we can mitigate errors caused by motion and inter-scan misalignment of the brain, which could be problematic when using the multiple acquisition method or the scan-rescan method. Third, we want to avoid using the assumption of the fourth type (model-based resampling).[21]

Up to this point, we have been using the terms "variance" and "noise" interchangeably. In the following text, we use "variance" when referring to the measure of the dispersion of data and "noise" when referring to imaging noise such as MRI noise—primarily caused by thermal fluctuations and electrical noise—to avoid misinterpretation.

To gain a better understanding of DTI variance or noise, it is important to characterize the role of physiology, subject compliance, and the interaction between the subject and the scanner. Our approach is driven by two fundamental questions (Fig. 2): Which factors are associated with DTI variance? Where and how does this association manifest? We assess the variance of DTI scalars, including fractional anisotropy (FA), axial diffusivity (AD), mean diffusivity (MD), and radial diffusivity (RD), within ROIs, and investigate the associations between the variance and covariates, including baseline age, time from the baseline (referred to as "interval"), motion, sex, and whether it is the first or the second scan within the session, using linear mixed-effects models.[28]

2 Methods

We use the PreQual[29] pipeline for preprocessing and quality assurance of the DTI data. PreQual is an end-to-end pipeline that applies denoising, inter-scan normalization, susceptibility-induced distortion correction, eddy current-induced distortion correction, inter-volume motion correction, slice-wise signal dropout imputation, and more. PreQual also provides a summary of the data and preprocessing in a portable document format (PDF) report for more efficient quality assurance.

	Female	Male
# Subjects	562	473
Age at baseline		
Range	24.2 to 94.4	22.4 to 92.4
Mean (SD)	69.8 (12.5)	71.5 (13.2)
# Sessions	1497	1254
w/ rescan DTI	1471	1207
w/o rescan DTI	26	47
# Subjects by the total number of sessions		
1	205	172
2	112	94
3	86	77
4	69	60
5	51	33
6	17	20
7	11	10
8	6	1
9	2	3
10	1	0
11	0	3
12	2	0

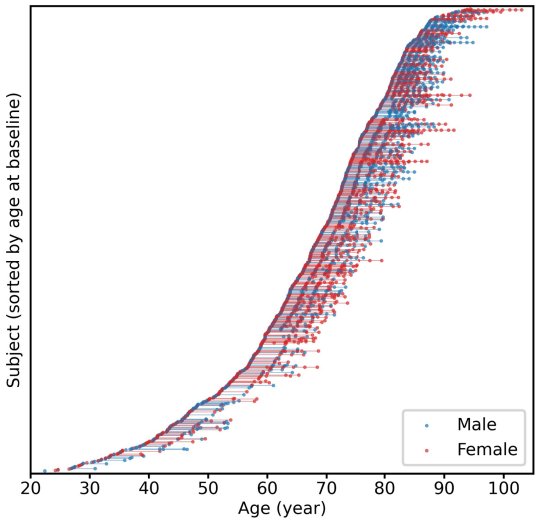

Fig. 3 The BLSA dataset we use has a slight imbalance between the number of females and males, but it is well-matched and appropriate for our research objectives in other aspects: (i) the age ranges of females and males align closely; (ii) rescan DTI data were acquired in most sessions, enabling inter-scan comparisons; and (iii) the distributions of sessions of females and males align closely.

We use data acquired from the neuroimaging substudy of the Baltimore Longitudinal Study of Aging (BLSA).[30,31] The BLSA is an extensive, ongoing research project that began in 1958, enrolling healthy volunteers aged 20 years and older to study normal aging through a longitudinal approach by following participants for their entire lives. We consider all subjects with at least one session comprising both T1-weighted (T1w) magnetization-prepared rapid gradient-recalled echo (MPRAGE) MRI data and DTI data. We exclude 49 DTI images exhibiting one or more of the following characteristics according to their potential impact on subsequent analyses:

(1) The presence of extreme susceptibility-induced distortion, motion artifacts, or eddy currents that resist correction.
(2) The failure of the preprocessed data to be fitted by the tensor model.
(3) An exceptionally low signal-to-noise ratio in the FA and MD images.

The exclusion of these cases results in the dataset depicted in Fig. 3. We identify 1035 subjects (562 F/473 M, 22.4 to 94.4 years old at baseline) with 2751 sessions (1497 F/1254 M). Detailed demographic information can be found in Table S1 and Fig. S1 in the Supplementary Material 1. The 2751 sessions of MRI data were acquired by four different scanners, including a 1.5 Tesla Philips Intera scanner (scanner A, 83 sessions) at the Kennedy Krieger Institute (KKI), two 3 Tesla Philips Achieva scanners using the same platform and protocol (scanners B and C, 16 sessions and 59 sessions, respectively) at the KKI, and a 3 Tesla Philips Achieva scanner (scanner D, 2593 sessions) at the National Institute on Aging. Detailed scanner information and protocol are provided in Table 1, which was previously reported.[32] Among the 1035 subjects, 59 switched to a different scanner in follow-up scans. During subsequent sessions, four female subjects and 10 male subjects were diagnosed with Alzheimer's disease, while the remaining subjects remained cognitively normal throughout all sessions.

2.1 ROI-Based DTI Variance Estimation

We use a registration-based approach for brain segmentation in the b0 (minimally weighted) volume (Fig. 4). We initiate the process with brain segmentations for the T1w images obtained through manual parcellations provided by the JHU-MNI-ss atlas ("Eve atlas")[33,34] and automated whole-brain segmentation by spatially localized atlas network tiles (SLANT).[35] For the Eve atlas, there are three types of parcellations available, each with a different regional focus.[33] For SLANT segmentation, labels for 132 regions covering the whole brain are provided.[35] We use the method

Table 1 Protocol for the T1w MPRAGE and DTI scans.

Parameter	Scanner A	Scanners B/C	Scanner D
MPRAGE			
Head coil	Philips 8-ch	Philips 8-ch	Philips 8-ch
Scan time (min:s)	3:58	10:52	10:52
Slice thickness (mm)	1.5	1.2	1.2
Number of slices	124	170	170
Flip angle (deg)	8	8	8
TR/TE (ms)	6.6/3.3	6.8/3.1	6.5/3.1
Field of view (mm)	240 × 240	256 × 240	256 × 240
Acquisition matrix	208 × 208	256 × 240	256 × 240
Reconstruction matrix	256 × 256	256 × 256	256 × 256
Reconstructed voxel size (mm)	0.94 × 0.94	1.00 × 1.00	1.00 × 1.00
DTI			
Head coil	Philips 8-ch	Philips 8-ch	Philips 8-ch
Scan time (min:s)	3:56	3:58	4:20
Number of gradients	30	32	32
Number of b0 images	1	1	1
Max b-factor (s/mm^2)	700	700	700
Number of signal averages (NSA)	1	1	1
Diffusion gradient timing DELTA/delta (ms)	39.2/15.1	36.3/16	36.3/13.5
Slice thickness (mm)	2.5	2.2	2.2
Number of slices	50	65	70
Flip angle (deg)	90	90	90
TR/TE (ms)	6210/80	6801/75	7454/75
Field of view (mm)	240 × 240	212 × 212	260 × 260
Acquisition matrix	96 × 96	96 × 95	116 × 115
Reconstruction matrix	256 × 256	256 × 256	320 × 320
Reconstructed voxel size (mm)	0.94 × 0.94	0.83 × 0.83	0.81 × 0.81

by Hansen et al.[36] to transfer these labels from T1w to b0 space. After label transferring, we manually review the segmentation to see if the labels align with the anatomical regions.

2.2 Linear Mixed-Effects Model

We use linear mixed-effects models[28] to analyze the association between DTI scalar standard deviation and covariates (*R* program, version 4.2.2;[37] Ubuntu 20.04.5 LTS; *R* package lme4, version 1.1.31;[38] *R* package lmerTest, version 3.1.3.[39]).

We study linear mixed-effects models of the form

$$\sigma_{ijkl} \sim \mathrm{Age}_{\mathrm{baseline},i} + \mathrm{Age}_{\mathrm{interval},ij} + \mathrm{Motion}_{\mathrm{ijl}} + \mathrm{Sex}_i + \mathrm{Rescan}_{\mathrm{ijl}} + r_{1,i} + r_{2,k} + \varepsilon_{ijkl}, \quad (4)$$

where σ_{ijkl} represents the standard deviation of a DTI scalar (FA, AD, MD, or RD) in a specific brain region of subject i at session j via scanner k in acquisition l, $\mathrm{Age}_{\mathrm{baseline},i}$ (hereafter referred

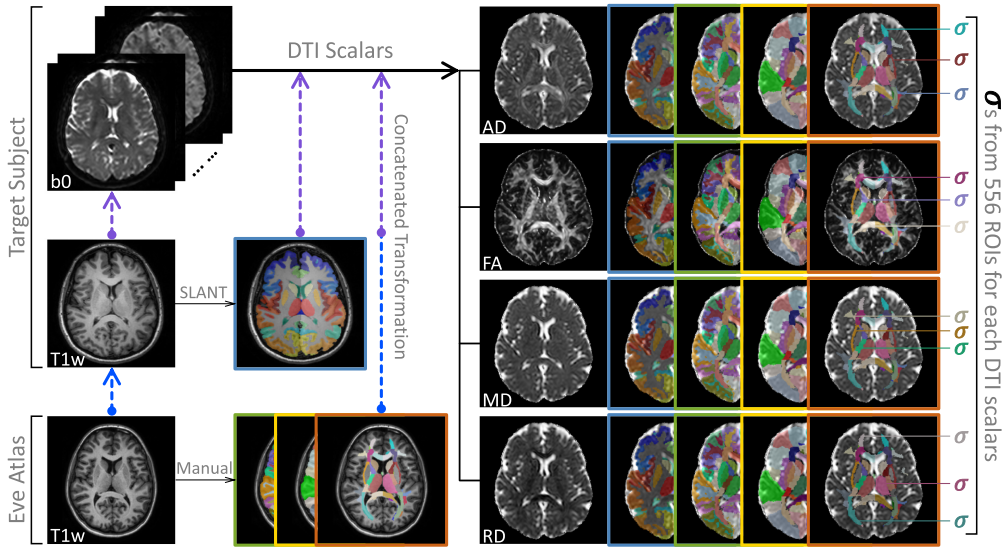

Fig. 4 Brain segmentation labels are obtained using the SLANT segmentation of the target subject's T1w image and using three types of manual parcellations provided by the Eve atlas. To generate transformation matrices for transferring these labels to DTI scalar images, intra- and inter-modality registrations are performed. Standard deviations of DTI scalars within each ROI are computed.

to as "baseline") is the age of subject i at baseline session (unit: decade), and $\text{Age}_{\text{interval},ij}$ (hereafter referred to as "interval") is the time between the current session, j, and the baseline session (unit: decade). Motion_{ijl} is a scalar value reflecting the degree of head movement of subject i at session j during acquisition l (calculated based on eddy movement, unit: millimeters),[40] Sex_i is the gender of subject i (0 for female and 1 for male), and Rescan_{ijl} is a binary variable indicating if the acquisition l is the first scan (coded 0) or the rescan (coded 1) of session j. We consider the subject and scanner as two random intercepts denoted by $r_{1,i}$ and $r_{2,k}$, respectively. Prior to fitting the models, we standardize the dependent variable σ. The results for the models are based on the standardized σ.

We have a total of 2224 models, derived from the four DTI scalars (FA, AD, MD, or RD), across varying ROIs defined by Eve type 1 (176 ROIs), Eve type 2 (130 ROIs), Eve type 3 (118 ROIs),[33,34] and SLANT (132 ROIs).[35] Each model starts with a full model, with all fixed effects and random effects, followed by an implementation of backward model selection.[39] The p-values for the fixed-effect terms are calculated based on the associated F tests.[39] To account for multiple comparisons, we adjust the p-values across the pairs of DTI scalar and ROI for a false discovery rate (FDR) of 0.05 using the Benjamini-Hochberg method.[41]

3 Results

The magnitude and direction of the effects of each covariate on DTI variance exhibit heterogeneous patterns across ROIs (Fig. 5). Specifically, interval is positively related to FA variance in ROIs such as the cuneus, middle occipital gyrus, superior occipital gyrus, medulla, and precuneus white matter, but negatively related in ROIs such as the caudate nucleus, posterior thalamic radiation, and superior fronto-occipital fasciculus (Table 2). Males have higher FA variance in the right putamen, thalamus, body of corpus callosum, and cingulum (cingulate gyrus), but lower FA variance in the middle frontal gyrus (Table 3). In the right inferior temporal gyrus, an increase of 1 mm in motion is associated with an increase of 2.211 standard deviations in the z-scored standard deviation (σ) of FA values ($\beta = 2.211$, $p \ll 0.001$). Interestingly and counter-intuitively, in several ROIs, including the medulla, middle occipital white matter, and cingulum (cingulate gyrus), an increase in motion is linked with a decrease in FA variance (Table 2).

In the lateral fronto-orbital gyrus, left insular, gyrus rectus, and inferior occipital gyrus, both motion and interval exhibit a positive association with FA variance (Table 2). In the left caudate

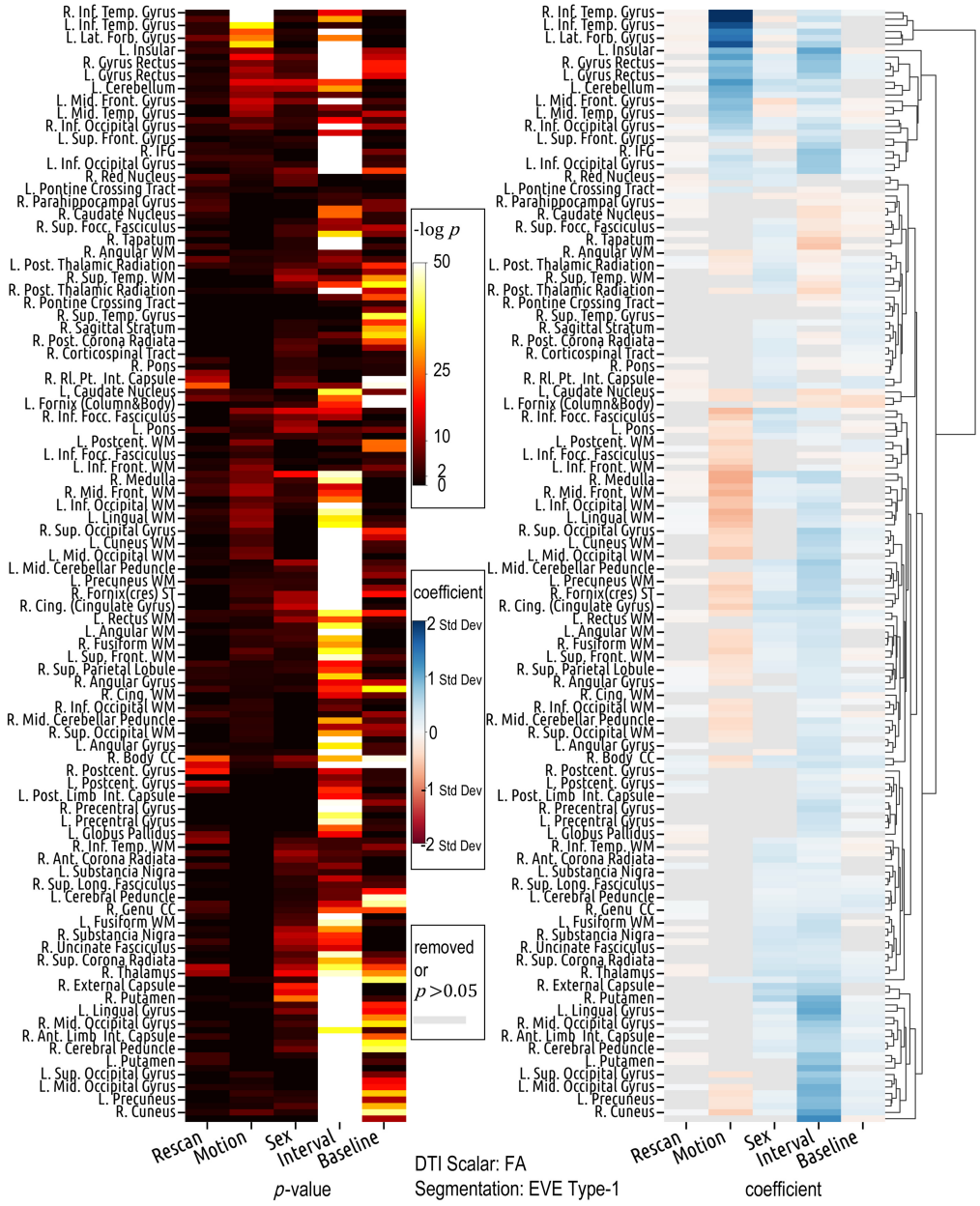

Fig. 5 Covariate effects on FA standard deviation (standardized) are region-specific. Motion and interval exhibit opposite effect directions in many ROIs. Gender differences exist in multiple ROIs. Counterintuitively, motion is negatively related to FA standard deviation in many ROIs. The lookup table for the abbreviation of the ROI name is in the Supplementary Material 2.

nucleus and right posterior thalamic radiation, they both show a negative association with FA variance (Table 2). In many other ROIs, such as the cuneus, lingual white matter, and middle occipital white matter, motion is negatively related to FA variance, while interval is positively related (Table 2). Results from the left ROI closely align with those from the corresponding right ROI (Table 2). There are some ROIs where the interval is significantly ($p \ll 0.001$) associated with FA variance, while motion either gets removed during the model selection or shows weak associations ($p \geq 0.05$) (Fig. 5).

On data extracted from ROIs defined by SLANT segmentation, which has a different regional focus and delineation than Eve type-1 segmentation, the aforementioned patterns of effects can also be observed (Figs. 6 and 7). For instance, in the left cerebellum exterior, both motion and interval are positively associated with FA variance, with motion's coefficient ($\beta = 0.960$, $p \ll 0.001$) higher than that of interval ($\beta = 0.485$, $p \ll 0.001$). This parallels the

Table 2 Covariate effects on FA standard deviation (standardized) in selected ROIs from Eve type-1 atlas.

Covariate	$Age_{interval}$				Motion			
ROI	Left		Right		Left		Right	
	β	p-value	β	p-value	β	p-value	β	p-value
Caudate nucleus	−0.383	4.0×10^{-41}	−0.317	2.4×10^{-26}	−0.337	2.1×10^{-03}	NA	NA
Posterior thalamic radiation	−0.091	1.1×10^{-07}	−0.414	7.7×10^{-94}	−0.186	5.0×10^{-03}	−0.188	1.8×10^{-02}
Superior fronto-occipital fasciculus	−0.358	2.5×10^{-35}	−0.220	6.4×10^{-11}	NA	NA	NA	NA
Middle occipital gyrus	0.957	7.2×10^{-218}	0.726	1.1×10^{-173}	−0.266	2.3×10^{-02}	NA	NA
Lateral fronto-orbital gyrus	0.427	5.5×10^{-28}	0.457	1.4×10^{-31}	1.481	3.6×10^{-28}	2.011	2.3×10^{-51}
Insular	1.055	1.7×10^{-132}	1.276	1.5×10^{-208}	0.935	5.8×10^{-10}	NA	NA
Gyrus rectus	0.636	4.6×10^{-59}	0.787	1.2×10^{-88}	0.887	2.1×10^{-10}	0.818	5.2×10^{-09}
Inferior occipital gyrus	0.746	1.6×10^{-109}	0.505	2.6×10^{-62}	0.422	9.3×10^{-4}	0.597	3.4×10^{-08}
Body of corpus callosum	0.439	7.6×10^{-60}	0.336	2.6×10^{-32}	−0.242	2.4×10^{-02}	−0.357	1.1×10^{-03}
Inferior cerebellar peduncle	0.560	2.9×10^{-55}	0.605	1.7×10^{-61}	−0.469	2.6×10^{-04}	−0.346	1.1×10^{-02}
Cingulum (cingulate gyrus)	0.247	1.9×10^{-17}	0.498	3.0×10^{-57}	−0.630	2.0×10^{-10}	−0.476	1.2×10^{-05}
Cuneus	0.951	1.1×10^{-288}	1.061	$\leq 2.3 \times 10^{-308}$	−0.318	3.2×10^{-04}	−0.471	5.4×10^{-07}
Cuneus WM	0.555	2.0×10^{-96}	0.585	1.4×10^{-95}	−0.383	4.0×10^{-05}	−0.393	7.9×10^{-05}
Fornix(cres) stria terminalis	0.340	5.4×10^{-41}	0.509	9.0×10^{-60}	−0.313	1.4×10^{-03}	−0.433	9.9×10^{-05}
Middle frontal WM	0.203	1.2×10^{-15}	0.275	5.4×10^{-22}	−0.623	8.0×10^{-12}	−0.716	2.5×10^{-12}
Superior frontal WM	0.341	2.4×10^{-48}	0.338	3.5×10^{-39}	−0.316	1.6×10^{-04}	−0.440	7.3×10^{-07}
Lateral fronto-orbital WM	0.372	7.3×10^{-28}	0.410	1.0×10^{-33}	−0.566	1.7×10^{-06}	−0.376	2.5×10^{-03}
Lingual WM	0.366	2.3×10^{-36}	0.426	6.1×10^{-45}	−0.640	6.4×10^{-10}	−0.695	1.5×10^{-10}
Medulla	0.537	1.9×10^{-48}	0.519	3.9×10^{-45}	−0.744	1.6×10^{-08}	−0.750	1.6×10^{-08}
Superior occipital gyrus	0.734	4.7×10^{-163}	0.644	1.3×10^{-134}	−0.311	1.1×10^{-03}	−0.445	1.7×10^{-06}
Middle occipital WM	0.544	3.5×10^{-102}	0.308	2.3×10^{-38}	−0.501	2.2×10^{-08}	−0.564	3.5×10^{-11}
Precuneus WM	0.549	3.0×10^{-96}	0.499	1.7×10^{-76}	−0.341	6.2×10^{-04}	−0.508	1.4×10^{-07}

Table 3 Effects of sex on FA standard deviation (standardized) in selected ROIs from Eve type-1 atlas.

Covariate	Sex			
ROI	Left		Right	
	β	p-value	β	p-value
Putamen	—	—	0.602	2.6×10^{-27}
Thalamus	0.388	5.9×10^{-13}	0.450	4.3×10^{-18}
Body of corpus callosum	0.287	4.1×10^{-07}	0.353	1.3×10^{-09}
Cingulum (cingulate gyrus)	0.509	1.7×10^{-16}	0.471	1.4×10^{-14}
Middle frontal gyrus	−0.370	1.8×10^{-09}	−0.148	1.9×10^{-02}

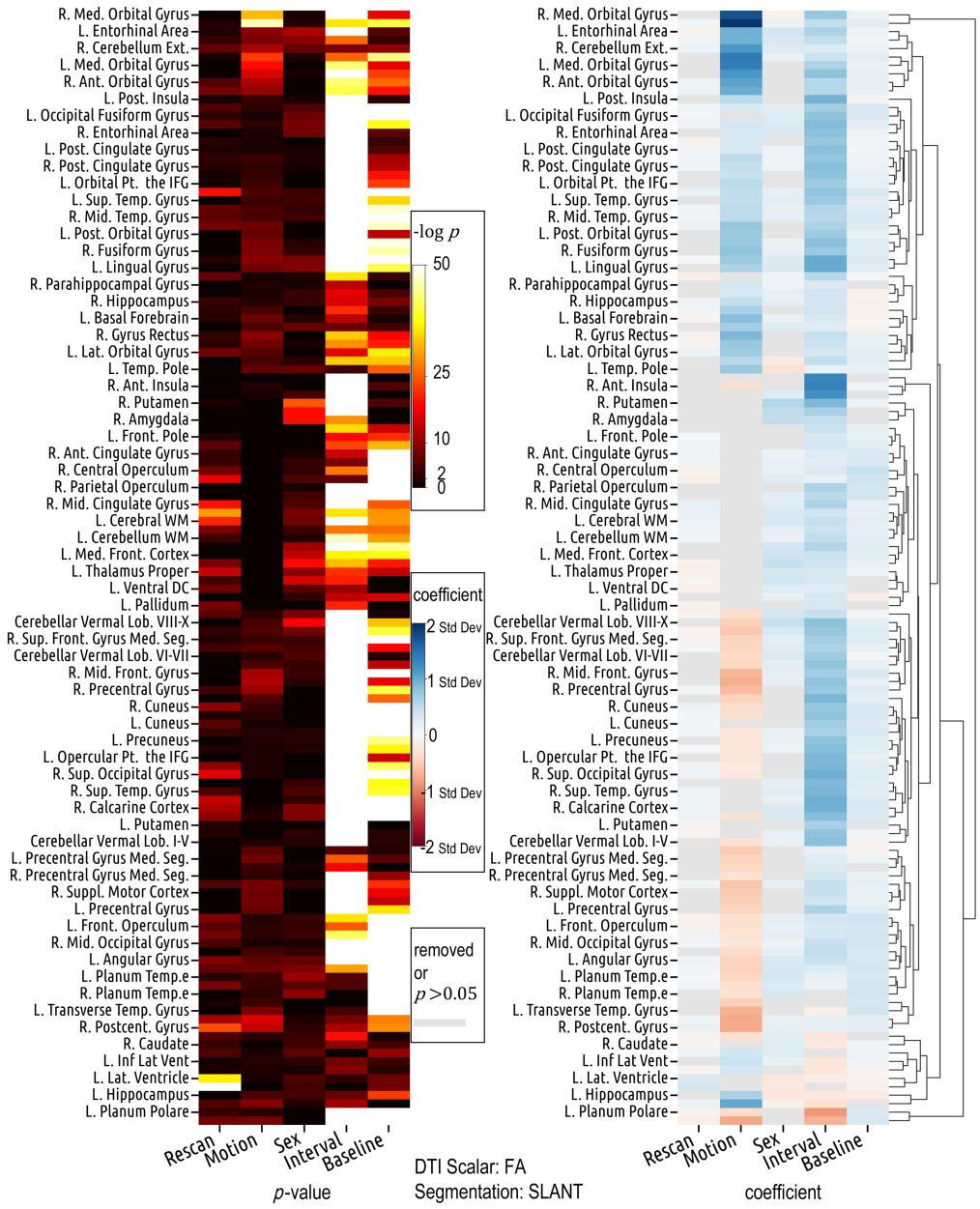

Fig. 6 The region-specific and bidirectional patterns of covariate effects are similarly observed in the results derived from SLANT segmentation, despite its differing definitions and delineations of ROIs compared to Eve type-1 segmentation (Fig. 5).

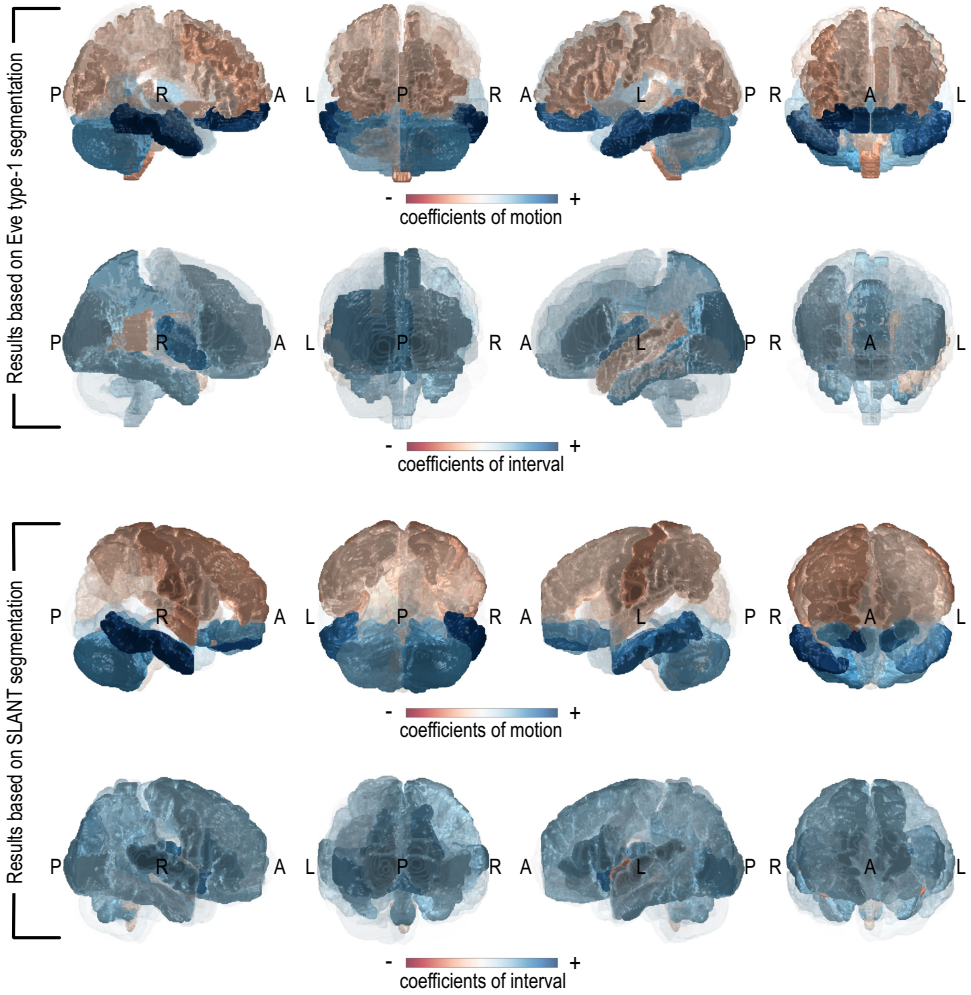

Fig. 7 Despite the different definitions and delineations of ROIs between Eve type-1 and SLANT segmentations, results based on the two segmentation methods are largely similar (comparable regions are colored similarly) and both show that the effects of motion and interval on FA variance vary across ROIs.

relationship observed between the motion and interval coefficients in the left cerebellum defined by the Eve type-1 segmentation ($\beta = 0.993$, $p \ll 0.001$ for motion; $\beta = 0.471$, $p \ll 0.001$ for interval). Similarly, in ROIs such as the right cuneus, left precuneus, right superior occipital gyrus, and right middle occipital gyrus, motion shows a negative association with FA variance, while interval shows a positive association. The exact coefficients and p-values are in the Supplementary Material 3, where we also provide the results from other pairs of DTI scalar and segmentation methods.

In the model selection process, we observe that the rescan and the motion terms appear mutually exclusive, with only one preserved post-selection in most models. This pattern is echoed in the heatmaps of coefficients (Figs. 5 and 6), where the cell in either the motion or the rescan column is colored gray. This hints at a correlation between rescan and motion. Supporting this observation, we detect an increase in head motion in the rescan of DTI acquired right after the first scan of DTI in the same session (mean shift $\Delta\mu = 0.045$ mm per volume, relative mean shift $\Delta\mu/\mu = 17.0\%$, coefficient of determination $R^2 = 0.065$).

To examine the differences between scanners, we measure the SNR of the FA images based on ROIs. There are differences between scanners regarding SNR across ROIs (Fig. 8). To validate our findings, we include two additional datasets, Alzheimer's Disease Neuroimaging Initiative (ADNI)[42] and Biomarkers of Cognitive Decline Among Normal Individuals (BIOCARD).[43] We exclude data points from subjects with cognitive impairment and use the remaining 1808 subjects

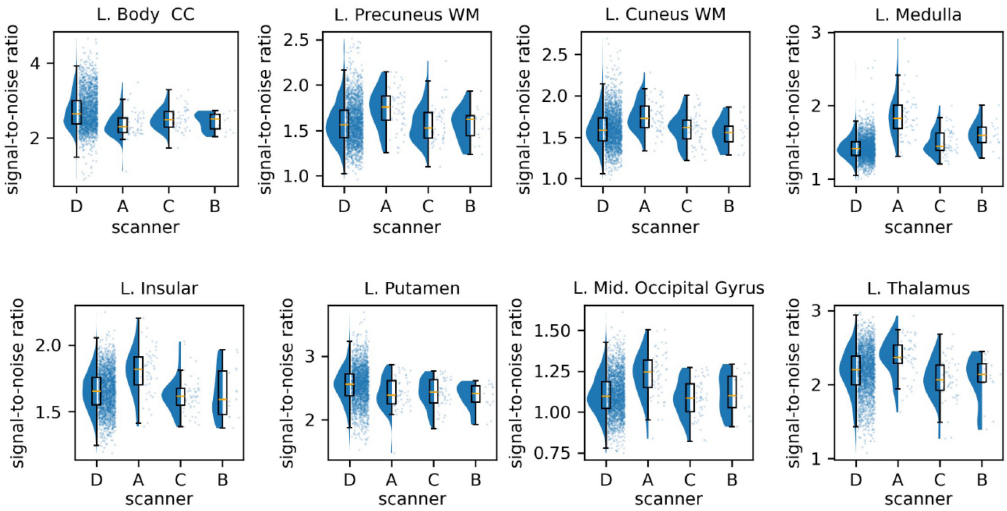

Fig. 8 SNR of the FA images across eight Eve type-1 atlas-defined ROIs, including white matter (WM) regions (body of the corpus callosum, precuneus WM, cuneus WM), gray matter regions (insular, putamen, middle occipital gyrus, thalamus), and mixed regions (medulla), in four different scanners of BLSA (where scanner A is the 1.5 Tesla Philips Intera scanner and scanners B/C/D are the 3 Tesla Philips Achieva scanners). "L." stands for the left hemisphere of the brain.

from the three datasets for the experiment. Detailed information about the data is included in Table S2 and Fig. S2 in the Supplementary Material 1. We reproduce the experiments of the linear mixed-effects models, except that the rescan term in Eq. (4) is omitted because the definitions of rescan in ADNI and BIOCARD differ from the definition of rescan in BLSA. The coefficients and p-values, visualized in Fig. 9, show a similar pattern to those derived from the BLSA, although the effect sizes differ, and the hierarchical clustering differs partly due to the omission of one covariate.

4 Discussion

While many studies have estimated and shown the spatial variability of DTI variance (or noise),[14,15,44] we characterize how DTI variance is associated with physiological and behavioral factors across brain regions. We answer the following questions: Which factors are associated with DTI variance? Where and how does this association manifest? We found region-specific and bidirectional effects of covariates—including interval (which captures the within-individual longitudinal change over time), motion, and sex—on FA variance across brain regions. For instance, FA variance is positively associated with an interval in the cuneus, but negatively associated with caudate nucleus. Long-standing research has demonstrated that there is a decline in white matter microstructure with aging,[45–50] with the consensus being that frontal and parietal areas are particularly vulnerable, and the occipital and motor areas are mostly preserved. The frontal lobe exhibits the most pronounced decline, with FA declining by ≈3% per decade starting at ∼35 years of age.[51] Although our study focuses on the standard deviation of FA, our results converge with these prior research studies as we have shown high sensitivity to aging in the frontal, parietal, and temporal areas. While it is unclear what mechanisms are driving the changes in these areas, potential culprits include the change in uniformity of fiber orientations and fiber density.[52–54]

Previous studies[55–57] have shown differences in FA between genders across brain regions. Oh et al.[55] found that males have significantly higher FA values in global corpus callosum structure areas, while they exhibit lower FA values than females in the partial areas of the rostrum, genu, and splenium. Menzler et al.[57] found that males show higher FA values in the thalamus, corpus callosum, and cingulum. Most of these regions previously identified in the literature also show significant ($p \ll 0.001$) associations between FA variance and sex in our study. While previous studies have reported changes in mean FA values, we offer a different perspective by depicting the variance of FA values.

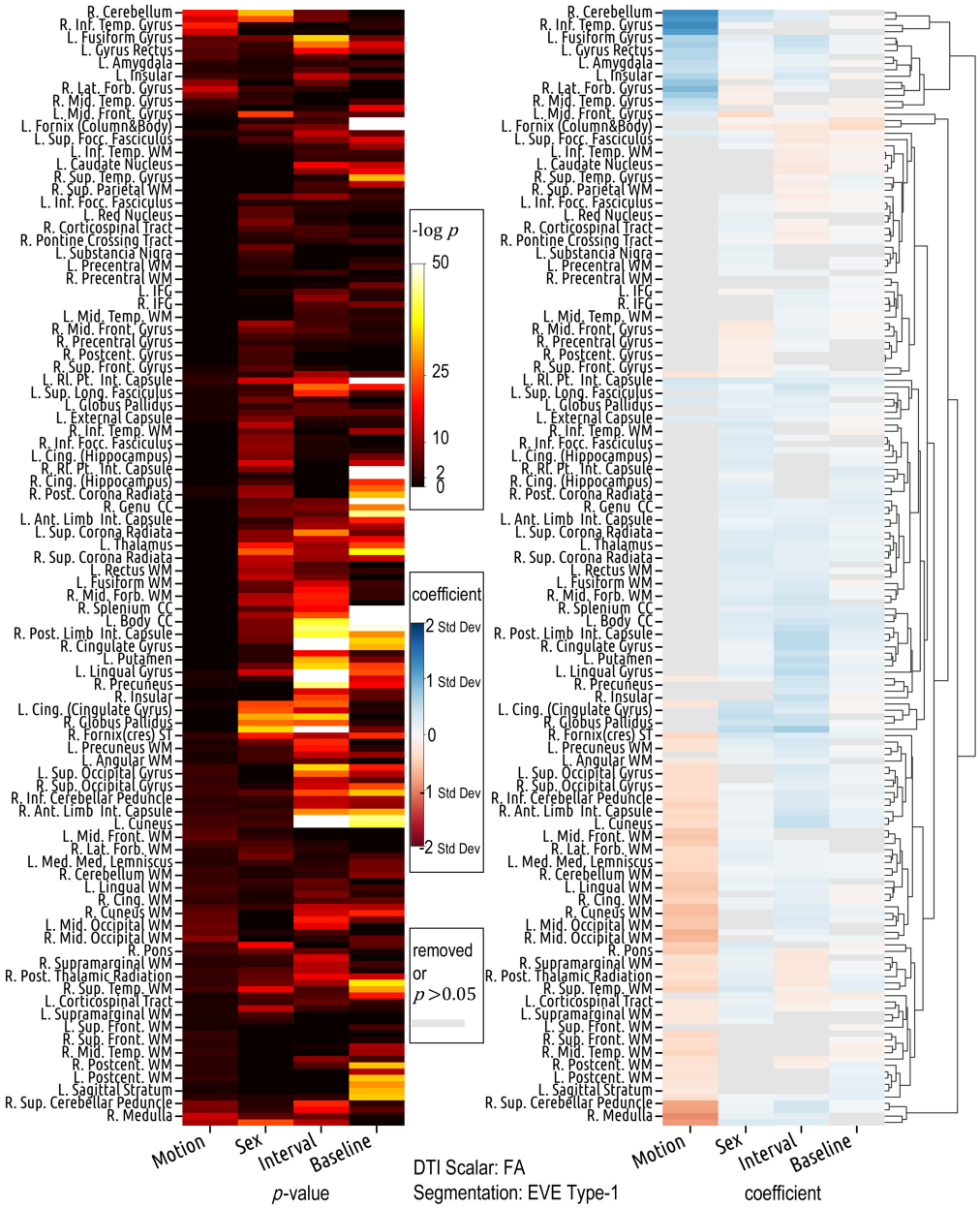

Fig. 9 To assess the generalizability of our findings, we include two additional datasets, ADNI and BIOCARD, and fit the linear mixed-effects models. The coefficients and *p*-values show similar patterns to those from BLSA alone, despite that the effect sizes and the hierarchical clustering are different, partly due to the omission of the rescan covariate.

The negative association observed between motion and FA variance in multiple regions, while counterintuitive, is not unreasonable. One might naturally expect that as motion increases, the uncertainty (reflected as variance) in the image should increase, given that motion leads to lower image quality, signal-to-noise ratio, and artifacts that can mislead image interpretation.[58,59] However, the images we use for analysis have undergone motion correction during preprocessing. Although in practice, motion artifacts cannot be fully eliminated from the image, the recorded motion value does not reflect the motion's impact on the image after preprocessing. Instead, it reflects the subject's motion during image acquisition. A higher motion value does not necessarily correspond to a noisier image post-preprocessing. Furthermore, Zeng et al.[60] found that head motion during brain imaging is not merely a technical artifact but a reflection of a neurobiological trait. Specifically, individuals with stronger distant connectivity in the default network could consistently refrain from moving and such "head motion tendency" remains

consistent within individuals. These points, taken together, provide explanations from image processing and biological perspectives, respectively, for why FA variance can decrease as motion increases.

This study underscores the significance of heteroscedasticity in diffusion-weighted MRI mega-analyses and provides a relatively straightforward approach to addressing this issue. Despite the longstanding recognition of heteroscedasticity in statistical analyses,[61–64] its application to diffusion-weighted MRI is still in its nascent stages. Recent advancements have started to bridge this gap, with emerging studies illustrating the importance of accounting for heteroscedasticity in MRI data.[65–67] These pioneering efforts are pivotal, yet they remain underutilized in the broader research community. To enhance the precision and reliability of findings in mega-analyses, it is imperative to disseminate these methodologies more widely and integrate them into commonly used analytical tools. Future research should focus on developing and employing increasingly sophisticated techniques to model and understand heteroscedasticity, thereby improving the robustness of statistical assessments in large-scale neuroimaging studies.

4.1 Limitations of Current Study

First, this study relies on a registration-based method for brain segmentation in the b0 space. Despite rigorous quality assurance, the labels for each brain region may not correspond flawlessly with the true anatomical regions. Consequently, the standard deviation of DTI scalars extracted from each region combines both voxel-wise modeling factors and image analysis factors from neighboring regions. Second, we used backward model selection for the fixed-effect terms of the linear mixed-effects models. Such a method can be unstable according to Breiman[68] Third, we use the variance of DTI scalar values in the ROI as a proxy for measuring noise. This is not an ideal proxy, because signal intensities may not be homogeneously distributed within each ROI, and the ROI-based variance captures not only the image noise but also the spatial variability in voxel intensity due to microstructural variations. This makes it suboptimal to reflect DTI noise. Fourth, the motion value used in this study is based on movement calculated by FMRIB Software Library's eddy,[40] which approximates true head motion. In addition, sensors or motion tracking sequences might be necessary to quantify head motion during scanning more accurately.

5 Conclusion

The notion of harnessing variance to enhance the reliability of analysis is universally applicable. Having a better understanding of variance is pivotal in mega-analyses, in which heteroscedasticity is often an inherent challenge. Our study illuminates the complex and heterogeneous effects of covariates including baseline age, interval, motion, sex, and rescan on DTI variance across ROIs. More comprehensive efforts are required to fully characterize the variance. In the meantime, we encourage researchers to consider models of heteroscedasticity in their analyses and to include their estimates of variance when sharing data. As highlighted in the introduction, the application of the whitening matrix, constructed using the variance of the data, significantly reduces statistical errors. We anticipate that more sophisticated methods can further unlock the potential benefits derived from a nuanced understanding of variance, thereby bolstering the accuracy and reliability of future research.

Disclosures

The authors declare no conflicts of interest.

Code and Data Availability

The dataset used in this study can be accessed upon approval from the Baltimore Longitudinal Study of Aging team at https://www.blsa.nih.gov/ The code for the experiments can be found on GitHub: https://github.com/MASILab/Variance-Aging-Diffusion

Ethics Approval

IRB of Vanderbilt University waived ethical approval for de-identified access to the human subject data (IRB #172072, PI: Bennett A. Landman).

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author Chenyu Gao used ChatGPT to check grammar and improve the readability of the original draft. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Acknowledgments

This work was supported in part by the National Institute of Health (NIH) [Grants Nos. 1R01EB017230-01A1 (Bennett A. Landman), 5-K01-EB032898-02 (Kurt Schilling), 1U24AG074855-01 (Timothy J. Hohman), and ViSE/VICTR VR3029] and by the Intramural Research Program of the National Institute on Aging, NIH. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, Tennessee, United States. We appreciate the NIH S10 Shared Instrumentation grants [Nos. 1S10OD020154-01 (Smith) and 1S10OD023680-01 (Vanderbilt's High-Performance Computer Cluster for Biomedical Research)]. The Vanderbilt Institute for Clinical and Translational Research (VICTR) is funded by the National Center for Advancing Translational Sciences (NCATS) Clinical Translational Science Award (CTSA) Program (Award No. 5UL1TR002243-03). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

The BLSA is supported by the Intramural Research Program, National Institute on Aging, NIH. The BIOCARD study is supported by the National Institute on Aging (Grant No. U19-AG03365). The BIOCARD study consists of seven cores and two projects with the following members: (1) the administrative core (Marilyn Albert, Corinne Pettigrew, Barbara Rodzon); (2) the clinical core (Marilyn Albert, Anja Soldan, Rebecca Gottesman, Corinne Pettigrew, Leonie Farrington, Maura Grega, Gay Rudow, Rostislav Brichko, Scott Rudow, Jules Giles, Ned Sacktor); (3) the imaging core (Michael Miller, Susumu Mori, Anthony Kolasny, Hanzhang Lu, Kenichi Oishi, Tilak Ratnanather, Peter vanZijl, Laurent Younes); (4) the biospecimen core (Abhay Moghekar, Jacqueline Darrow, Alexandria Lewis, Richard O'Brien); (5) the informatics core (Roberta Scherer, Ann Ervin, David Shade, Jennifer Jones, Hamadou Coulibaly, Kathy Moser, Courtney Potter); (6) the biostatistics core (Mei-Cheng Wang, Yuxin Zhu, Jiangxia Wang); (7) the neuropathology core (Juan Troncoso, David Nauen, Olga Pletnikova, Karen Fisher); (8) project 1 (Paul Worley, Jeremy Walston, Mei-Fang Xiao); and (9) project 2 (Mei-Cheng Wang, Yifei Sun, Yanxun Xu). Data collection and sharing for ADNI were supported by the National Institutes of Health (Grant No. U01-AG024904) and the Department of Defense (Award No. W81XWH-12-2-0012). ADNI is also funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through 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. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References

1. J. Fan, F. Han, and H. Liu, "Challenges of big data analysis," *Natl. Sci. Rev.* **1**(2), 293–314 (2014).
2. M. Monti, "Statistical analysis of fMRI time-series: a critical review of the GLM approach," *Front. Hum. Neurosci.* **5**, 28 (2011).
3. A. Kessy, A. Lewin, and K. Strimmer, "Optimal whitening and decorrelation," *Amer. Stat.* **72**(4), 309–314 (2018).
4. C. Pierpaoli et al., "Diffusion tensor MR imaging of the human brain," *Radiology* **201**(3), 637–648 (1996).
5. C. Beaulieu, "The basis of anisotropic water diffusion in the nervous system–a technical review," *NMR Biomed.* **15**(7–8), 435–455 (2002).
6. P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.* **66**(1), 259–267 (1994).

7. K. D. Merboldt, W. Hanicke, and J. Frahm, "Self-diffusion NMR imaging using stimulated echoes," *J. Magn. Reson.* **64**(3), 479–486 (1985).
8. D. G. Taylor and M. C. Bushell, "The spatial mapping of translational diffusion coefficients by the NMR imaging technique," *Phys. Med. Biol.* **30**(4), 345 (1985).
9. D. Le Bihan et al., "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," *Radiology* **161**(2), 401–407 (1986).
10. J. M. Soares et al., "A Hitchhiker's guide to diffusion tensor imaging," *Front. Neurosci.* **7**, 31 (2013).
11. C. J. Cascio, G. Gerig, and J. Piven, "Diffusion tensor imaging: application to the study of the developing brain," *J. Amer. Acad. Child Adolesc. Psychiatry* **46**(2), 213–223 (2007).
12. C. Metzler-Baddeley et al., "Temporal association tracts and the breakdown of episodic memory in mild cognitive impairment," *Neurology* **79**(23), 2233 (2012).
13. K. G. Schilling et al., "Aging and white matter microstructure and macrostructure: a longitudinal multi-site diffusion MRI study of 1218 participants," *Brain Struct. Funct.* **227**(6), 2111–2125 (2022).
14. B. A. Landman, P. L. Bazin, and J. L. Prince, "Estimation and application of spatially variable noise fields in diffusion tensor imaging," *Magn. Reson. Imaging* **27**(6), 741–751 (2009).
15. K. Tabelow, H. U. Voss, and J. Polzehl, "Local estimation of the noise level in MRI using structural adaptation," *Med. Image Anal.* **20**(1), 76–86 (2015).
16. B. Landman, P. L. Bazin, and J. Prince, "Diffusion tensor estimation by maximizing Rician likelihood," in *IEEE 11th Int. Conf. on Comput. Vis.*, pp. 1–8 (2007).
17. L. C. Chang, D. K. Jones, and C. Pierpaoli, "RESTORE: robust estimation of tensors by outlier rejection," *Magn. Reson. Med.* **53**(5), 1088–1095 (2005).
18. D. K. Jones and P. J. Basser, "'Squashing peanuts and smashing pumpkins': how noise distorts diffusion-weighted MR data," *Magn. Reson. Med.* **52**(5), 979–993 (2004).
19. M. E. Bastin, P. A. Armitage, and I. Marshall, "A theoretical study of the effect of experimental noise on the measurement of anisotropy in diffusion imaging," *Magn. Reson. Imaging* **16**(7), 773–785 (1998).
20. J. A. D. Farrell et al., "Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging–derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T," *J. Magn. Reson. Imaging* **26**(3), 756–767 (2007).
21. S. Chung, Y. Lu, and R. G. Henry, "Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters," *NeuroImage* **33**(2), 531–541 (2006).
22. M. M. Covell et al., "Automated analysis of multiple performance characteristics in magnetic resonance imaging systems," *Med. Phys.* **13**(6), 815–823 (1986).
23. B. A. Landman et al., "Robust estimation of spatially variable noise fields," *Magn. Reson. Med.* **62**(2), 500–509 (2009).
24. L. Kaufman et al., "Measuring signal-to-noise ratios in MR imaging," *Radiology* **173**(1), 265–267 (1989).
25. B. Whitcher et al., "Using the wild bootstrap to quantify uncertainty in diffusion tensor imaging," *Hum. Brain Mapp.* **29**(3), 346–362 (2008).
26. H. R. Pardoe et al., "Motion and morphometry in clinical and nonclinical populations," *NeuroImage* **135**, 177–185 (2016).
27. H. K. Hausman et al., "The association between head motion during functional magnetic resonance imaging and executive functioning in older adults," *NeuroImage: Rep.* **2**(2), 100085 (2022).
28. N. M. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics* **38**(4), 963–974 (1982).
29. L. Y. Cai et al., "PreQual: an automated pipeline for integrated preprocessing and quality assurance of diffusion weighted MRI images," *Magn. Reson. Med.* **86**(1), 456–470 (2021).
30. N. W. Shock et al., *Normal Human Aging: The Baltimore Longitudinal Study of Aging*, US Department of Health and Human Services, National Institute on Aging, National Institutes of Health (1984).
31. L. Ferrucci, "The Baltimore Longitudinal Study of Aging (BLSA): a 50-year-long journey and plans for the future," *J. Gerontol. A Biol. Sci. Med. Sci.* **63**(12), 1416–1419 (2008).
32. V. K. Venkatraman et al., "Region of interest correction factors improve reliability of diffusion imaging measures within and across scanners and field strengths," *NeuroImage* **119**, 406–416 (2015).
33. S. Mori et al., "Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template," *NeuroImage* **40**(2), 570–582 (2008).
34. K. Oishi et al., "Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and Alzheimer's disease participants," *NeuroImage* **46**(2), 486–499 (2009).
35. Y. Huo et al., "3D whole brain segmentation using spatially localized atlas network tiles," *NeuroImage* **194**, 105–119 (2019).
36. C. B. Hansen et al., "Pandora: 4-D white matter bundle population-based atlases derived from diffusion MRI fiber tractography," *Neuroinformatics* **19**(3), 447–460 (2021).
37. R Core Team, "R: a language and environment for statistical computing," (2022). https://www.R-project.org

38. D. Bates et al., "Fitting linear mixed-effects models using LME4," *J. Stat. Softw.* **67**(1), 1–48 (2015).
39. A. Kuznetsova, P. B. Brockhoff, and R. H. Christensen, "lmerTest package: tests in linear mixed effects models," *J. Stat. Softw.* **82**(13), 1–26 (2017).
40. J. L. R. Andersson and S. N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging," *NeuroImage* **125**, 1063–1078 (2016).
41. Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *J. R. Stat. Soc.: Ser. B-Methodol.* **57**(1), 289–300 (1995).
42. C. R. Jack, Jr. et al., "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *J. Magn. Reson. Imaging* **27**(4), 685–691 (2008).
43. N. Sacktor et al., "The BIOCARD index: a summary measure to predict onset of mild cognitive impairment (P1.095)," *Neurology* **88**(16 Supplement), P1.095 (2017).
44. T. Zhu et al., "An optimized wild bootstrap method for evaluation of measurement uncertainties of DTI-derived parameters in human brain," *NeuroImage* **40**(3), 1144–1156 (2008).
45. D. H. Salat et al., "Age-related alterations in white matter microstructure measured by diffusion tensor imaging," *Neurobiol. Aging* **26**(8), 1215–1227 (2005).
46. S. W. Davis et al., "Assessing the effects of age on long white matter tracts using diffusion tensor tractography," *NeuroImage* **46**(2), 530–541 (2009).
47. I. J. Bennett et al., "Age-related differences in multiple measures of white matter integrity: a diffusion tensor imaging study of healthy aging," *Hum. Brain Mapp.* **31**(3), 378–390 (2010).
48. C. E. Sexton et al., "A meta-analysis of diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease," *Neurobiol. Aging* **32**(12), 2322.e5–2322.e18 (2011).
49. C. Lebel et al., "Diffusion tensor imaging of white matter tract evolution over the lifespan," *NeuroImage* **60**(1), 340–352 (2012).
50. M. de Groot et al., "White matter degeneration with aging: longitudinal diffusion MR imaging analysis," *Radiology* **279**(2), 532–541 (2015).
51. S. M. Grieve et al., "Cognitive aging, executive function, and fractional anisotropy: a diffusion tensor MR imaging study," *Amer. J. Neuroradiol.* **28**(2), 226–235 (2007).
52. W. Meier-Ruge et al., "Age-related white matter atrophy in the human brain," *Ann. N. Y. Acad. Sci.* **673**(1), 260–269 (1992).
53. L. Bronge, N. Bogdanovic, and L. O. Wahlund, "Postmortem MRI and histopathology of white matter changes in Alzheimer brains: a quantitative, comparative study," *Dement. Geriatr. Cogn. Disord.* **13**(4), 205–212 (2002).
54. N. H. Stricker et al., "Decreased white matter integrity in late-myelinating fiber pathways in Alzheimer's disease supports retrogenesis," *NeuroImage* **45**(1), 10–16 (2009).
55. J. S. Oh et al., "Tractography-guided statistics (TGIS) in diffusion tensor imaging for the detection of gender difference of fiber integrity in the midsagittal and parasagittal corpora callosa," *NeuroImage* **36**(3), 606–616 (2007).
56. F. Liu et al., "Sex differences in the human corpus callosum microstructure: a combined T2 myelin-water and diffusion tensor magnetic resonance imaging study," *Brain Res.* **1343**, 37–45 (2010).
57. K. Menzler et al., "Men and women are different: diffusion tensor imaging reveals sexual dimorphism in the microstructure of the thalamus, corpus callosum and cingulum," *NeuroImage* **54**(4), 2557–2562 (2011).
58. A. Yendiki et al., "Spurious group differences due to head motion in a diffusion MRI study," *NeuroImage* **88**, 79–90 (2014).
59. G. L. Baum et al., "The impact of in-scanner head motion on structural connectivity derived from diffusion MRI," *NeuroImage* **173**, 275–286 (2018).
60. L. L. Zeng et al., "Neurobiological basis of head motion in brain imaging," *Proc. Natl. Acad. Sci.* **111**(16), 6058–6062 (2014).
61. R. D. Cook and S. Weisberg, "Diagnostics for heteroscedasticity in regression," *Biometrika* **70**(1), 1–10 (1983).
62. H. G. Müller and U. Stadtmüller, "Estimation of heteroscedasticity in regression analysis," *Ann. Stat.* **15**(2), 610–625 (1987).
63. H. Dette and A. Munk, "Testing heteroscedasticity in nonparametric regression," *J. R. Stat. Soc. Ser. B Stat. Methodol.* **60**(4), 693–708 (1998).
64. J. S. Long and L. H. Ervin, "Using heteroscedasticity consistent standard errors in the linear regression model," *Amer. Stat.* **54**(3), 217–224 (2000).
65. C. Seiler and S. Holmes, "Multivariate heteroscedasticity models for functional brain connectivity," *Front. Neurosci.* **11**, 696 (2017).
66. A. Zugman et al., "Mega-analysis methods in ENIGMA: the experience of the generalized anxiety disorder working group," *Hum. Brain Mapp.* **43**(1), 255–277 (2022).
67. A. Lutti et al., "Restoring statistical validity in group analyses of motion-corrupted MRI data," *Hum. Brain Mapp.* **43**(6), 1973–1983 (2022).

68. L. Breiman, "Heuristics of instability and stabilization in model selection," *Ann. Stat.* **24**(6), 2350–2383 (1996).

Chenyu Gao is a PhD student in electrical and computer engineering at Vanderbilt University, collaborating with Bennett Landman on the harmonization of diffusion MRI. His research interest is focused on image processing and computer vision with application to medical image analysis. He received his BS degree in biomedical engineering from Sun Yat-sen University and his MS degree in biomedical engineering from Johns Hopkins University, collaborating with Jerry Prince.

Biographies of the other authors are not available.