Characterizing patterns of DTI 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]

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

Approach: We analyze DTI data from 1035 subjects in the Baltimore Longitudinal Study of Aging (BLSA), with ages ranging from 22.4 to 103 years old. For each subject, up to 12 longitudinal sessions were conducted. In each session, a scan and a rescan of DTI were performed. We assess 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 scan-rescan pairs.

Results: Covariate effects are heterogeneous and bilaterally symmetric across ROIs. The interval is positively related to FA variance in the cuneus and occipital gyrus, but negatively in the caudate nucleus. Males show higher FA variance in the right putamen, thalamus, body of the corpus callosum, and cingulate gyrus. In certain ROIs, an increase in motion is associated with a decrease in FA variance. Head motion increases during the rescan of DTI.

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.

Keywords: brain, aging, DTI, variance, motion.

*Address all correspondence to Chenyu Gao, E-mail: chenyu.gao@vanderbilt.edu

1 Introduction

The Autism Brain Imaging Data Exchange (ABIDE),[1] Alzheimer's Disease Neuroimaging Initiative (ADNI),[2] Human Connectome Project (HCP),[3] Open Access Series of Imaging Studies

(OASIS),[4] and UK (United Kingdom) Biobank[5] are making large volumes of their data available to the wider scientific community. These "big data" studies create the possibility of exploring questions that would be impractical with smaller- or moderate-sized datasets[6] while giving rise to the development and application of powerful machine learning technologies which are capable of assimilating and evaluating large volumes of complex healthcare data.[7] One prevalent challenge is that large datasets typically 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.[8]

Diffusion tensor imaging (DTI)[9–12] is a modeling approach used in diffusion-weighted imaging (DWI) [13–15], a variant of conventional magnetic resonance imaging (MRI) based on the tissue water diffusion rate.[16] 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.[17–20] DTI is inherently subject to low signal-to-noise ratios (SNR) and multiple sources of artifacts,[21,22] which can have negative impact on the estimation of the diffusion anisotropy[23] and the structural connectivity[24], among others[25–27]. Moreover, the noise structure in DTI exhibits spatial variability and correlation, primarily attributed to fast imaging and noise suppression techniques.[28,29] Understanding the statistical nature of DTI variance has been proven to be beneficial for diffusion tensor estimation[29,30], outlier detection[31], and reproducibility assessment[23,25,26]. Considerable efforts have been dedicated to estimating DTI variance[28,29,32] and enhancing the reliability of analyses by leveraging the variance[33] or considering models of heteroscedasticity[34]. To gain a better understanding of DTI variance, it is important to characterize the role of physiology, subject compliance, and the interaction between the subject and the scanner.

In the context of many mega-analyses, we are interested in relationship modeling. The general linear model, a structured and widely used framework for relationship modeling, allows us to illustrate the importance of understanding variance and variability. The general linear model is assessed through linear regression. A linear regression can be expressed by $\mathrm{Y} = \mathrm{X}\beta + \varepsilon$, where the response variable Y, the covariate matrix X, and the regression coefficients β are conventionally represented in matrix forms given by:

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

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 estimation of Eq. (1) by transforming ε into a diagonal form with a whitening matrix[35], W:

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

Note $\mathrm{W}\varepsilon \sim \mathcal{N}(0, \mathrm{I}_M)$, 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). The simulations use two populations, each generated from a different linear model, with β set to 0 and 1, respectively. Noise is introduced to Y, following normal distribution with mean zero and standard deviation proportional to X. In each iteration, we sample 50 data points from the population data and conduct ordinary least squares (OLS) regressions using both the standard linear regression equation, $\mathrm{Y} = \mathrm{X}\beta + \varepsilon$, and Eq. (3), where W is constructed

using the sample standard deviations of the noise. Following 10,000 experiments, each comprising 100 iterations, the number of both false positives and false negatives is reduced when using Eq. (3) compared to the standard linear regression equation.

Our approach is driven by two fundamental questions (Fig. 2): Which factor is associated with DTI variance? Where and how does this association manifest? We assess 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 scan-rescan pairs (whether it is the first or the second scan within the session) using linear mixed effects models[36].

2 Methods

We use the PreQual[37] pipeline for preprocessing and quality assurance of the DTI data in the Baltimore Longitudinal Study of Aging (BLSA)[38,39] dataset. We consider all subjects with at least one session comprising both T1-weighted (T1w) 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 resists 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 y/o at baseline) with 2751 sessions (1497 F/ 1254 M). Out of the 2751 sessions, both scan and rescan DTI data (acquired in the same session) are available for 2678 sessions. 4 female subjects and 10 male subjects were diagnosed with Alzheimer's disease during subsequent 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")[40,41] and automated whole-brain segmentation by SLANT[42]. For the Eve atlas, there are three types of parcellations available, each with different regional focus.[40] For SLANT segmentation, labels for 132 regions covering the whole brain are provided.[42] We use the method by Hansen et al.[43] 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[36] to analyze the association between DTI standard deviation and covariates. (R program, version 4.2.2 [44]; Ubuntu 20.04.5 LTS; R package lme4, version 1.1.31 [45]; R package lmerTest, version 3.1.3 [46].)

We study linear mixed-effects models of the form:

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

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 to as "baseline") is the age of subject i at baseline session, $\mathrm{Age}_{\mathrm{interval},ij}$ (hereafter referred to as "interval") is the time between the current session, j, and the baseline session. 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, in millimeters)[47], 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 subject and scanner as two random intercepts, respectively denoted by $r_{1,i}$ and $r_{2,k}$. Prior to fitting the models, we standardize the dependent variable σ. To allow an artificially "amplified" effect of aging in the result, facilitating comparison with other covariates, we convert the units of baseline and interval from years to decades.

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),[40,41] and SLANT (132 ROIs)[42]. Each model starts with a full model, with all fixed effects and random effects, followed by an implementation of backward model selection.[46] The p-values for the fixed-effect terms are calculated based on the associated F tests.[46] 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.[48]

To evaluate the effects of each independent variable on DTI variance, we visualize p-values and coefficients using seaborn[49] package's clustermaps. We use grey cells to indicate missing values, which can arise either from term removal during model selection or from a corresponding p-value above 0.05. We optimize the hierarchical clustering and the ordering of rows based on Euclidean distance. There are 16 such figures, each corresponding to a DTI scalar-segmentation pair. Only 2 are included and discussed in the main text, with the remaining in the Supplementary Materials. ROI names are abbreviated, with a reference table in the Supplementary Materials. We visualize the coefficients of motion and interval in 3D for enhanced interpretation. The transparency of each region is determined by the absolute value of its coefficient so that regions with smaller coefficients appear more transparent than those with larger coefficients.

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 with $p \ll 0.001$ in ROIs such as the cuneus (left: $\beta = 0.951$; right: $\beta = 1.061$), middle occipital gyrus (left: $\beta = 0.957$; right: $\beta = 0.726$), superior occipital gyrus (left: $\beta = 0.734$; right: $\beta = 0.644$), medulla (left: $\beta = 0.537$; right: $\beta = 0.519$), precuneus white matter (left: $\beta = 0.549$; right: $\beta = 0.499$). In contrast, interval is negatively related to FA variance with $p \ll 0.001$ in ROIs such as the caudate nucleus (left: $\beta = -0.383$; right: $\beta = -0.317$), posterior thalamic radiation (left: $\beta = -0.091$; right: $\beta = -0.414$), and superior fronto-occipital fasciculus (left: $\beta = -0.358$; right: $\beta = -0.220$). Males have higher FA variance in the right putamen ($\beta = 0.602, p \ll 0.001$), thalamus (left: $\beta = 0.388$, $p \ll 0.001$; right: $\beta = 0.450$, $p \ll 0.001$), body of corpus callosum (left: $\beta = 0.287$, $p \ll 0.001$; right: $\beta = 0.353$, $p \ll 0.001$), and cingulum (cingulate gyrus) (left: $\beta = 0.509$, $p \ll 0.001$; right: $\beta = 0.471$, $p \ll 0.001$), but lower FA variance in the left middle frontal gyrus ($\beta = -0.370, p \ll 0.001$). In the right inferior temporal gyrus, an increase of 1 millimeter 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 counterintuitively, in several ROIs, including the medulla (left: $\beta = -0.744$; right: $\beta = -0.750$), middle occipital white matter (left: $\beta = -0.501$; right: $\beta = -0.564$), cingulum (Cingulate Gyrus) (left: $\beta = -0.630$; right: $\beta = -0.476$), an increase in motion is linked with a decrease in FA variance ($p \ll 0.001$).

In the lateral fronto-orbital gyrus (left: $\beta = 1.481$ for motion, $\beta = 0.427$ for interval; right: $\beta = 2.011$ for motion, $\beta = 0.457$ for interval), left insular ($\beta = 0.935$ for motion, $\beta = 1.055$ for interval), gyrus rectus (left: $\beta = 0.887$ for motion, $\beta = 0.636$ for interval; right: $\beta = 0.818$ for

motion, $\beta = 0.787$ for interval), and inferior occipital gyrus (left: $\beta = 0.422$ for motion, $\beta = 0.746$ for interval; right: $\beta = 0.597$ for motion, $\beta = 0.505$ for interval), both motion and interval exhibit a positive association with FA variance with $p \ll 0.001$. In the left caudate nucleus ($\beta = -0.337, p = 0.002$ for motion; $\beta = -0.383, p \ll 0.001$ for interval), and right posterior thalamic radiation ($\beta = -0.188, p = 0.018$ for motion; $\beta = -0.414, p \ll 0.001$ for interval), they both show a negative association with FA variance. 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 1). Results from the left ROI closely align with those from the corresponding right ROI. For instance, interval coefficient in the left cuneus ($\beta = 0.951, p \ll 0.001$) is close to the coefficient in the right cuneus ($\beta = 1.061, p \ll 0.001$). There are some ROIs where interval is significantly ($p \ll 0.001$) associated with FA variance, while motion either gets removed during the model selection or shows weak associations ($p \geq 0.05$).

On data extracted from ROIs defined by SLANT segmentation, which has different regional focus and delineation than Eve type-1 segmentation, the aforementioned patterns of effects can also be observed (Fig. 6). For instance, in the right cerebellum exterior, both motion and interval are positively associated with FA variance, with motion's coefficient ($\beta = 1.166, p \ll 0.001$) higher than that of interval ($\beta = 0.265, p \ll 0.001$). This parallels the 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 ($\beta = -0.358, p < 0.001$ for motion; $\beta = 0.768, p \ll 0.001$ for interval), left precuneus ($\beta = -0.230$, $p = 0.018$ for motion; $\beta = 0.826$, $p \ll 0.001$ for interval), right superior occipital gyrus ($\beta = -0.207$, $p = 0.036$ for motion; $\beta = 0.977$, $p \ll 0.001$ for interval), and right middle occipital gyrus ($\beta = -0.227$, $p = 0.028$ for motion; $\beta =$

0.469, $p \ll 0.001$ for interval), motion shows a negative association with FA variance, while interval shows a positive association. This pattern also persists in the counterpart ROIs defined by the Eve type-1 segmentation (Fig. 5, Table 1).

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 clustermaps of coefficients (Fig. 5, Fig. 6), where the cell in either the motion or the rescan column is colored grey. 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 $\Delta\mu = 0.045$ millimeters, relative mean shift $\Delta\mu/\mu = 17.0\%$, coefficient of determination $R^2 = 0.065$).

4 Discussion

While many studies have estimated and shown the spatial variability of DTI variance (or noise),[28,29,32] we characterize how DTI variance is associated with physiological and behavioral factors across brain regions. We answer the questions: Which factor is 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. Within the cuneus, for instance, FA variance is positively associated with interval, meaning that the FA values become "noisier" as interval increases (which corresponds to individuals getting older). In contrast, within the caudate nucleus, FA values become more "monotonous" as interval increases. Long-standing research has demonstrated that there is a decline in white matter microstructure with aging,[50–55] 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 approximately 3% per decade starting at ~35 years of age.[56] Although our study focuses on the standard deviation of FA, we believe 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 change in these areas, potential culprits include the change of uniformity of fiber orientations and fiber density.[57–59]

Previous studies[60–62] have shown differences in FA between genders across brain regions. Oh et al. 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.[60] Menzler et al. found that males show higher FA values in the thalamus, corpus callosum and cingulum.[62] 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.

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.[24,27] 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 doesn't reflect the motion's impact in the image after preprocessing. Instead, it primarily reflects the subject's motion during image acquisition, so a higher motion value does not necessarily correspond to a noisier image post-preprocessing. Furthermore, Zeng et al. 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.[63] These points, taken together, provide explanations from image processing and biological perspectives, respectively, for why FA variance can decrease as motion increases.

4.1 Limitations of Current Study

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 method can be unstable according to Breiman et. al.[64] Additionally, we did not eliminate the random-effect terms during model selection. Lange and Laird showed that the inclusion of unnecessary random effects could lead to a near singular random effect covariance matrix.[65] These two potential pitfalls could compromise the robustness and reliability of our analysis. Third, our study only includes the BLSA dataset. Incorporating data from additional sources could make the findings more convincing. Fourth, the motion value used in this study is based on movement calculated by FSL's eddy[47], which approximates true head motion.

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, where heteroscedasticity is often an inherent challenge. Our study illuminates the complex and heterogeneous effects of

covariates including baseline age, interval, motion, sex, and scan-rescan pairs 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.

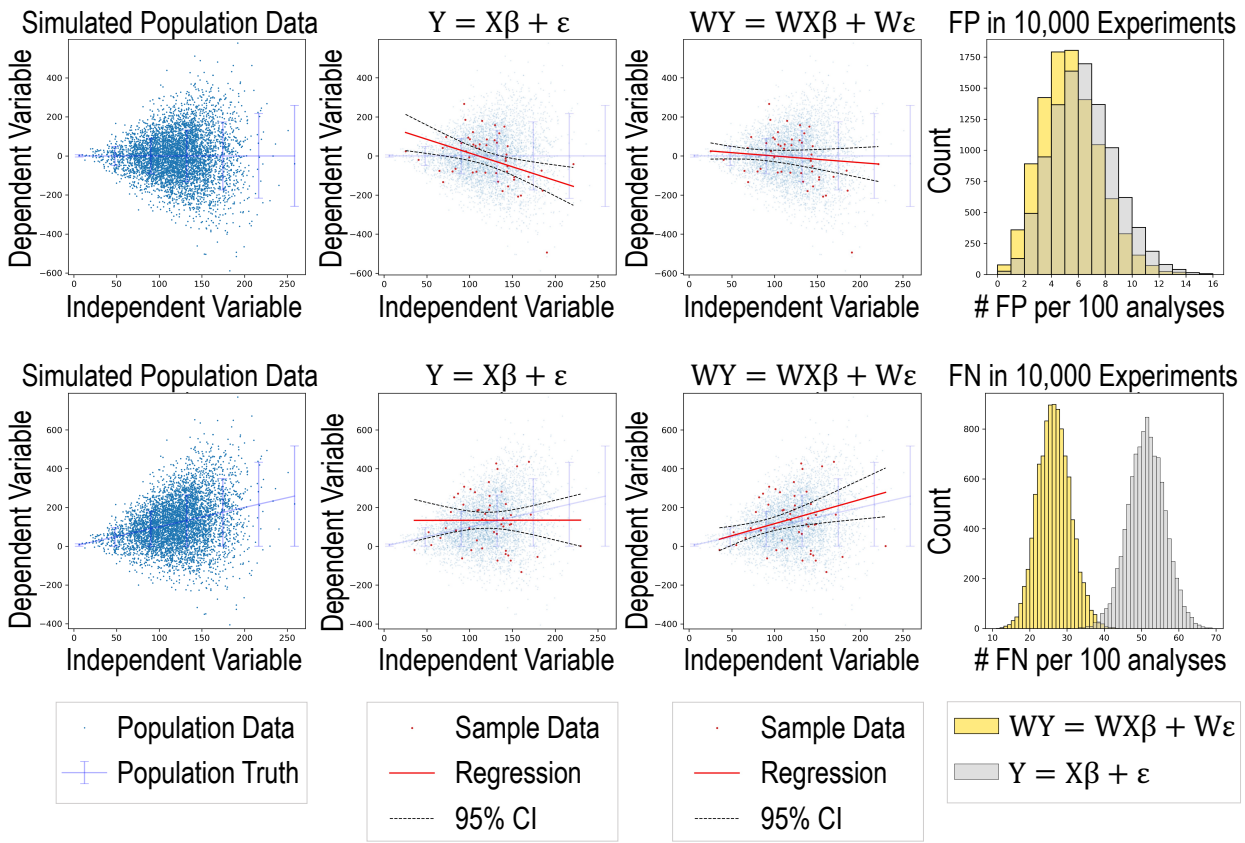


Fig. 1 Simulation shows that applying the whitening matrix to the standard linear regression equation reduces the number of false positives (FP) and false negatives (FN) under heteroskedasticity. In the top row, the population truth has zero slope. In data sampled from the synthetic population data, 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.

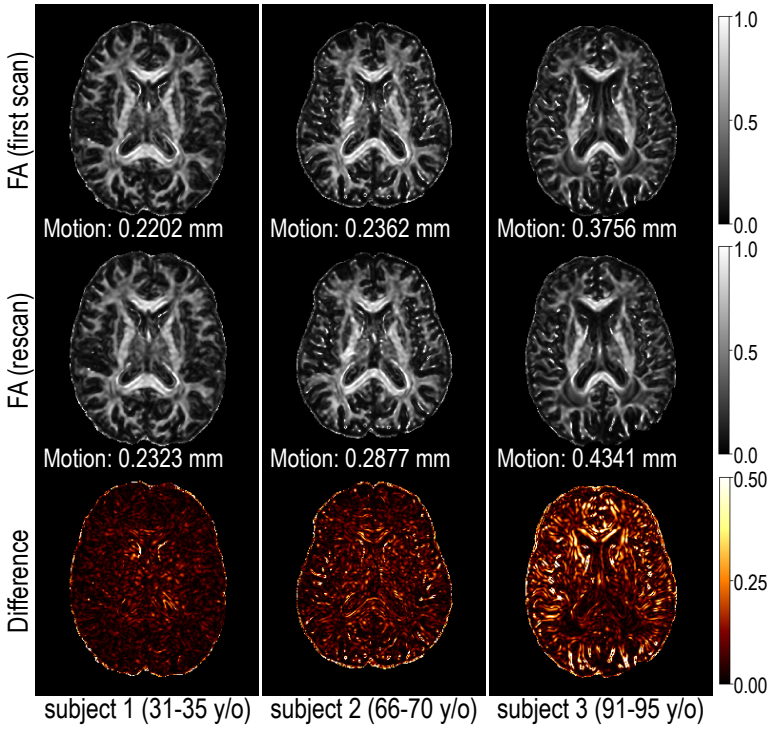


Fig. 2 Noise in DTI scalar images such as fractional anisotropy (FA) images generally increase with age. But motion is also considered to increase with age.[66,67] Here we show three typical subjects with both their motion levels and noise levels in DTI scans increasing.

	Female	Male
# Subjects	562	473
Age at baseline		
Range	24.2 - 94.4	22.4 - 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

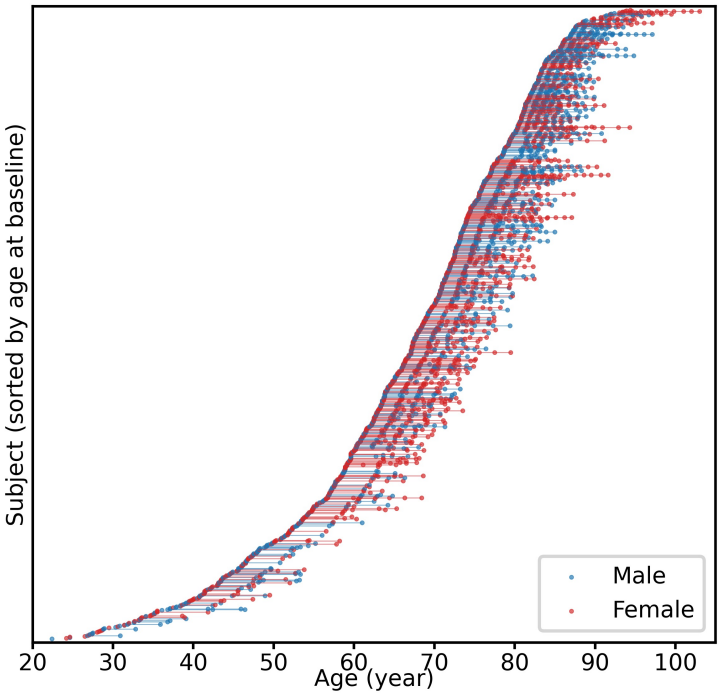
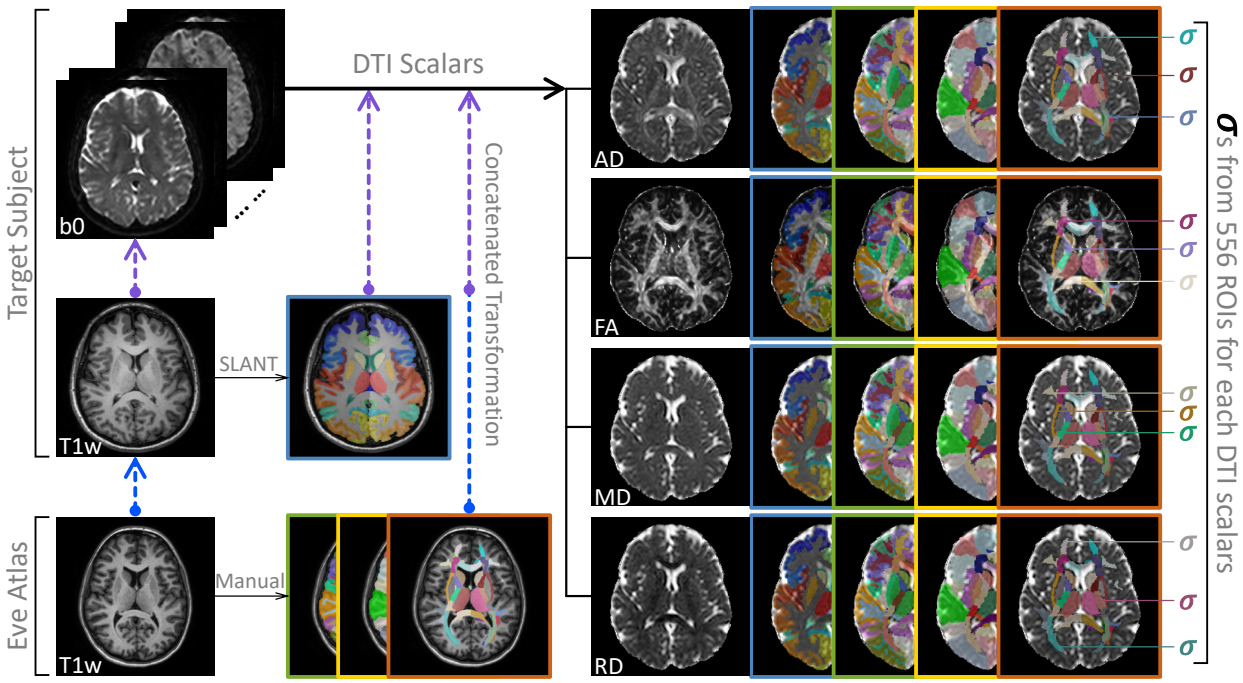


Fig. 3 The BLSA dataset we use has a slight imbalance between the number of females and males, but it is well-matched and appropriate for our research objectives in other aspects: i) the age ranges for both females and males align closely; ii) rescan DTI data were acquired in most sessions; iii) there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


Target Subject
b0
T1w
SLANT
Eve Atlas
T1w
Manual
DTI Scalars
Concatenated Transformation
AD
FA
MD
RD
σ
σ
σ

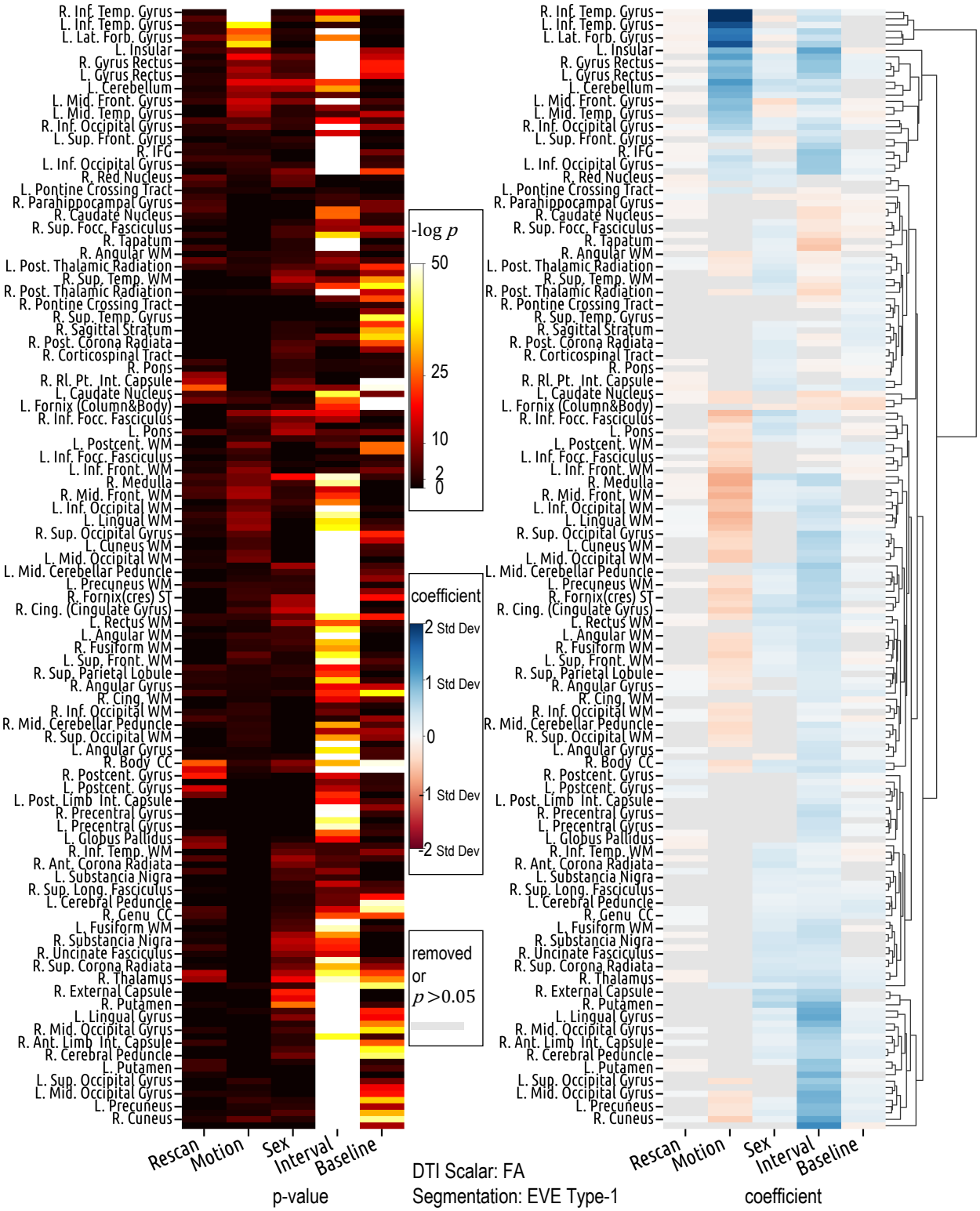


Fig. 5 Covariate effects on FA variance, represented as standardized standard deviation, are region-specific and bidirectional across ROIs. Gender differences exist in multiple ROIs. Motion and interval exhibit opposite effect directions in many regions. Counterintuitively, motion is negatively related to FA variance in many ROIs.

ROI	$\text{Age}_{\text{interval}}$		Motion	
	β	$p-value$	β	$p-value$
Left Body of Corpus Callosum	0.439	7.6e-60	-0.242	2.4e-02
Right Body of Corpus Callosum	0.336	2.6e-32	-0.357	1.1e-03
Left Inferior Cerebellar Peduncle	0.560	2.9e-55	-0.469	2.6e-04
Right Inferior Cerebellar Peduncle	0.605	1.7e-61	-0.346	1.1e-02
Left Cingulum (Cingulate Gyrus)	0.247	1.9e-17	-0.630	2.0e-10
Right Cingulum (Cingulate Gyrus)	0.498	3.0e-57	-0.476	1.2e-05
Left Cuneus	0.951	1.1e-288	-0.318	3.2e-04
Right Cuneus	1.061	≤ 2.3e-308	-0.471	5.4e-07
Left Cuneus WM	0.555	2.0e-96	-0.383	4.0e-05
Right Cuneus WM	0.585	1.4e-95	-0.393	7.9e-05
Left Fornix(cres) Stria Terminalis	0.340	5.4e-41	-0.313	1.4e-03
Right Fornix(cres) Stria Terminalis	0.509	9.0e-60	-0.433	9.9e-05
Left Middle Frontal WM	0.203	1.2e-15	-0.623	8.0e-12
Right Middle Frontal WM	0.275	5.4e-22	-0.716	2.5e-12
Left Superior Frontal WM	0.341	2.4e-48	-0.316	1.6e-04
Right Superior Frontal WM	0.338	3.5e-39	-0.440	7.3e-07
Left Lateral Fronto-orbital WM	0.372	7.3e-28	-0.566	1.7e-06
Right Lateral Fronto-orbital WM	0.410	1.0e-33	-0.376	2.5e-03
Left Lingual WM	0.366	2.3e-36	-0.640	6.4e-10
Right Lingual WM	0.426	6.1e-45	-0.695	1.5e-10
Left Medulla	0.537	1.9e-48	-0.744	1.6e-08
Right Medulla	0.519	3.9e-45	-0.750	1.6e-08
Left Superior Occipital Gyrus	0.734	4.7e-163	-0.311	1.1e-03
Right Superior Occipital Gyrus	0.644	1.3e-134	-0.445	1.7e-06
Left Middle Occipital WM	0.544	3.5e-102	-0.501	2.2e-08
Right Middle Occipital WM	0.308	2.3e-38	-0.564	3.5e-11
Left Precuneus WM	0.549	3.0e-96	-0.341	6.2e-04
Right Precuneus WM	0.499	1.7e-76	-0.508	1.4e-07

Table 1 Selected ROIs (defined by Eve type-1segmentation) where interval coefficient and motion coefficient show opposite signs. Within these ROIs, FA variance increases with interval but decreases with motion.

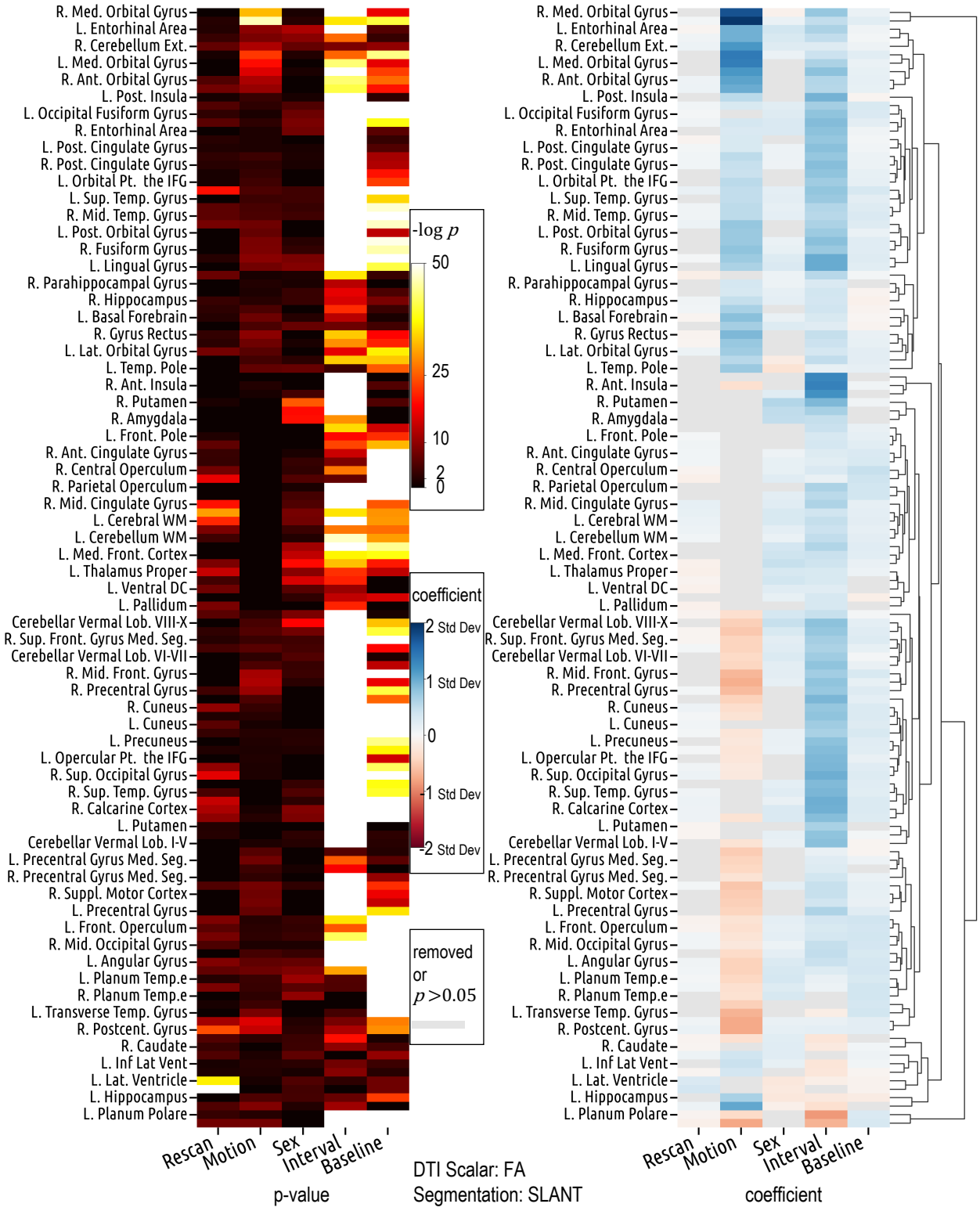


Fig. 6 The region-specific and bidirectional patterns of covariate effects are similarly observed in the results derived from SLANT segmentation, despite its differing definitions and delineations of ROIs compared to Eve type-1 segmentation (Fig. 5).

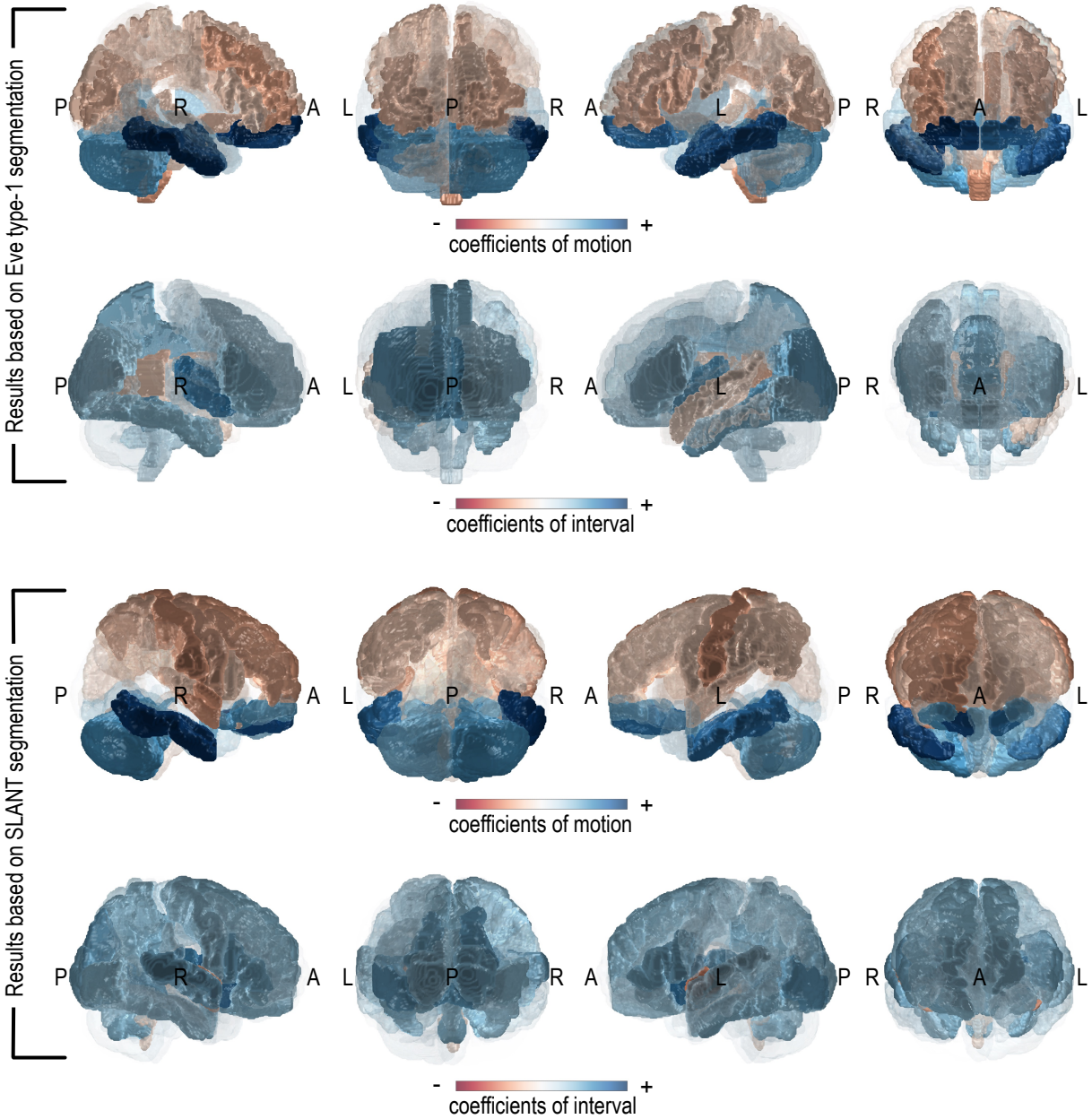


Fig. 7 Despite the different definitions and delineations of ROIs between Eve type-1 and SLANT segmentations, results based on the two segmentation methods are largely similar (comparable regions are colored similarly) and both show that the effects of motion and interval on FA variance vary across ROIs.

Disclosures

No conflicts of interest.

Code, Data, and Materials 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 at https://github.com/MASILab/Variance-Aging-Diffusion.

Acknowledgments

This work was supported in part by the National Institute of Health through grants 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, TN. We appreciate the National Institute of Health S10 Shared Instrumentation grant 1S10OD020154-01 (Smith), and grant 1S10OD023680-01 (Vanderbilt's High-Performance Computer Cluster for Biomedical Research).

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

References

1. A. Di Martino et al., "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism," Mol Psychiatry **19**(6), 659–667 (2014) [doi:10.1038/mp.2013.78].

2. C. R. Jack Jr. et al., “The Alzheimer’s disease neuroimaging initiative (ADNI): MRI methods,” Journal of Magnetic Resonance Imaging **27**(4), 685–691, John Wiley & Sons, Ltd (2008) [doi:https://doi.org/10.1002/jmri.21049].
3. D. C. Van Essen et al., “The Human Connectome Project: A data acquisition perspective,” Neuroimage **62**(4), 2222–2231 (2012) [doi:https://doi.org/10.1016/j.neuroimage.2012.02.018].
4. P. J. LaMontagne et al., “OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease,” medRxiv, 2019.12.13.19014902 (2019) [doi:10.1101/2019.12.13.19014902].
5. C. Sudlow et al., “UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age,” PLoS Med **12**(3), e1001779-, Public Library of Science (2015).
6. R. A. Poldrack and K. J. Gorgolewski, “Making big data open: data sharing in neuroimaging,” Nat Neurosci **17**(11), 1510–1517 (2014) [doi:10.1038/nn.3818].
7. K. Y. Ngiam and I. W. Khor, “Big data and machine learning algorithms for health-care delivery,” Lancet Oncol **20**(5), e262–e273 (2019) [doi:https://doi.org/10.1016/S1470-2045(19)30149-4].
8. J. Fan, F. Han, and H. Liu, “Challenges of Big Data analysis,” Natl Sci Rev **1**(2), 293–314 (2014) [doi:10.1093/nsr/nwt032].
9. P. J. Basser, J. Mattiello, and D. Lebihan, “Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo,” J Magn Reson B **103**(3), 247–254 (1994) [doi:https://doi.org/10.1006/jmrb.1994.1037].
10. P. J. Basser, J. Mattiello, and D. LeBihan, “MR diffusion tensor spectroscopy and imaging,” Biophys J **66**(1), 259–267 (1994) [doi:https://doi.org/10.1016/S0006-3495(94)80775-1].
11. C. Pierpaoli et al., “Diffusion tensor MR imaging of the human brain.,” Radiology **201**(3), 637–648, Radiological Society of North America (1996) [doi:10.1148/radiology.201.3.8939209].
12. C. Beaulieu, “The basis of anisotropic water diffusion in the nervous system – a technical review,” NMR Biomed **15**(7–8), 435–455 (2002) [doi:https://doi.org/10.1002/nbm.782].
13. K.-D. Merboldt, W. Hanicke, and J. Frahm, “Self-diffusion NMR imaging using stimulated echoes,” Journal of Magnetic Resonance (1969) **64**(3), 479–486 (1985) [doi:https://doi.org/10.1016/0022-2364(85)90111-8].
14. 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) [doi:10.1088/0031-9155/30/4/009].
15. D. Le Bihan et al., “MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders.,” Radiology **161**(2), 401–407, Radiological Society of North America (1986) [doi:10.1148/radiology.161.2.3763909].
16. J. M. Soares et al., “A hitchhiker’s guide to diffusion tensor imaging,” Front Neurosci(7 MAR) (2013) [doi:10.3389/fnins.2013.00031].
17. C. J. CASCIO, G. GERIG, and J. PIVEN, “Diffusion Tensor Imaging: Application to the Study of the Developing Brain,” J Am Acad Child Adolesc Psychiatry **46**(2), 213–223 (2007) [doi:https://doi.org/10.1097/01.chi.0000246064.93200.e8].
18. R. A. Charlton et al., “White matter pathways associated with working memory in normal aging,” Cortex **46**(4), 474–489 (2010) [doi:https://doi.org/10.1016/j.cortex.2009.07.005].

19. C. Metzler-Baddeley et al., “Temporal association tracts and the breakdown of episodic memory in mild cognitive impairment,” Neurology **79**(23), 2233 (2012) [doi:10.1212/WNL.0b013e31827689e8].
20. 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) [doi:10.1007/s00429-022-02503-z].
21. D. Le Bihan et al., “Artifacts and pitfalls in diffusion MRI,” Journal of Magnetic Resonance Imaging **24**(3), 478–488, John Wiley & Sons, Ltd (2006) [doi:https://doi.org/10.1002/jmri.20683].
22. D. Perrone et al., “The effect of Gibbs ringing artifacts on measures derived from diffusion MRI,” Neuroimage **120**, 441–455 (2015) [doi:https://doi.org/10.1016/j.neuroimage.2015.06.068].
23. 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, John Wiley & Sons, Ltd (2004) [doi:https://doi.org/10.1002/mrm.20283].
24. G. L. Baum et al., “The impact of in-scanner head motion on structural connectivity derived from diffusion MRI,” Neuroimage **173**, 275–286, Academic Press Inc. (2018) [doi:10.1016/j.neuroimage.2018.02.041].
25. 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) [doi:https://doi.org/10.1016/S0730-725X(98)00098-8].
26. 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,” Journal of Magnetic Resonance Imaging **26**(3), 756–767, John Wiley & Sons, Ltd (2007) [doi:https://doi.org/10.1002/jmri.21053].
27. A. Yendiki et al., “Spurious group differences due to head motion in a diffusion MRI study,” Neuroimage **88**, 79–90, Academic Press (2014) [doi:10.1016/J.NEUROIMAGE.2013.11.027].
28. 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) [doi:https://doi.org/10.1016/j.mri.2009.01.001].
29. 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) [doi:https://doi.org/10.1016/j.media.2014.10.008].
30. B. Landman, P.-L. Bazin, and J. Prince, “Diffusion Tensor Estimation by Maximizing Rician Likelihood,” in 2007 IEEE 11th International Conference on Computer Vision, pp. 1–8 (2007) [doi:10.1109/ICCV.2007.4409140].
31. L.-C. Chang, D. K. Jones, and C. Pierpaoli, “RESTORE: Robust estimation of tensors by outlier rejection,” Magn Reson Med **53**(5), 1088–1095, John Wiley & Sons, Ltd (2005) [doi:https://doi.org/10.1002/mrm.20426].
32. 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) [doi:https://doi.org/10.1016/j.neuroimage.2008.01.016].
33. 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) [doi:https://doi.org/10.1016/j.neuroimage.2015.06.078].

34. B. Wegmann, A. Eklund, and M. Villani, "Bayesian Heteroscedastic Regression for Diffusion Tensor Imaging," in Modeling, Analysis, and Visualization of Anisotropy, T. Schultz, E. Özarslan, and I. Hotz, Eds., pp. 257–282, Springer International Publishing, Cham (2017).
35. A. Kessy, A. Lewin, and K. Strimmer, "Optimal Whitening and Decorrelation," Am Stat **72**(4), 309–314, Taylor & Francis (2018) [doi:10.1080/00031305.2016.1277159].
36. N. M. Laird and J. H. Ware, "Random-Effects Models for Longitudinal Data," Biometrics **38**(4), 963–974, [Wiley, International Biometric Society] (1982) [doi:10.2307/2529876].
37. 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, John Wiley & Sons, Ltd (2021) [doi:https://doi.org/10.1002/mrm.28678].
38. N. W. Shock, *Normal human aging: The Baltimore longitudinal study of aging*, US Department of Health and Human Services, Public Health Service, National … (1984).
39. L. Ferrucci, "The Baltimore Longitudinal Study of Aging (BLSA): a 50-year-long journey and plans for the future," in The Journals of Gerontology Series A: Biological Sciences and Medical Sciences **63**(12), pp. 1416–1419, Oxford University Press (2008).
40. S. Mori et al., "Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template," Neuroimage **40**(2), 570–582, Academic Press (2008) [doi:10.1016/J.NEUROIMAGE.2007.12.035].
41. 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, Academic Press (2009) [doi:10.1016/J.NEUROIMAGE.2009.01.002].
42. Y. Huo et al., "3D whole brain segmentation using spatially localized atlas network tiles," Neuroimage **194**, 105–119 (2019) [doi:https://doi.org/10.1016/j.neuroimage.2019.03.041].
43. 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) [doi:10.1007/s12021-020-09497-1].
44. R Core Team, "R: A Language and Environment for Statistical Computing," Vienna, Austria (2022).
45. D. Bates et al., "Fitting Linear Mixed-Effects Models Using lme4," J Stat Softw **67**(1), 1–48 (2015) [doi:10.18637/jss.v067.i01].
46. A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," J Stat Softw **82**(13), 1–26 (2017) [doi:10.18637/jss.v082.i13].
47. 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) [doi:https://doi.org/10.1016/j.neuroimage.2015.10.019].
48. Y. Benjamini and Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," Journal of the Royal Statistical Society: Series B (Methodological) **57**(1), 289–300, John Wiley & Sons, Ltd (1995) [doi:https://doi.org/10.1111/j.2517-6161.1995.tb02031.x].
49. M. L. Waskom, "seaborn: statistical data visualization," J Open Source Softw **6**(60), 3021, The Open Journal (2021) [doi:10.21105/joss.03021].

50. D. H. Salat et al., “Age-related alterations in white matter microstructure measured by diffusion tensor imaging,” Neurobiol Aging **26**(8), 1215–1227 (2005) [doi:https://doi.org/10.1016/j.neurobiolaging.2004.09.017].
51. 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) [doi:https://doi.org/10.1016/j.neuroimage.2009.01.068].
52. 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, John Wiley & Sons, Ltd (2010) [doi:https://doi.org/10.1002/hbm.20872].
53. 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) [doi:https://doi.org/10.1016/j.neurobiolaging.2010.05.019].
54. C. Lebel et al., “Diffusion tensor imaging of white matter tract evolution over the lifespan,” Neuroimage **60**(1), 340–352 (2012) [doi:https://doi.org/10.1016/j.neuroimage.2011.11.094].
55. M. de Groot et al., “White Matter Degeneration with Aging: Longitudinal Diffusion MR Imaging Analysis,” Radiology **279**(2), 532–541, Radiological Society of North America (2015) [doi:10.1148/radiol.2015150103].
56. S. M. Grieve et al., “Cognitive Aging, Executive Function, and Fractional Anisotropy: A Diffusion Tensor MR Imaging Study,” American Journal of Neuroradiology **28**(2), 226 (2007).
57. W. MEIER-RUGE et al., “Age-Related White Matter Atrophy in the Human Brain,” Ann N Y Acad Sci **673**(1), 260–269, John Wiley & Sons, Ltd (1992) [doi:https://doi.org/10.1111/j.1749-6632.1992.tb27462.x].
58. 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) [doi:10.1159/000057698].
59. 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) [doi:https://doi.org/10.1016/j.neuroimage.2008.11.027].
60. 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) [doi:https://doi.org/10.1016/j.neuroimage.2007.03.020].
61. 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) [doi:https://doi.org/10.1016/j.brainres.2010.04.064].
62. 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) [doi:https://doi.org/10.1016/j.neuroimage.2010.11.029].
63. L.-L. Zeng et al., “Neurobiological basis of head motion in brain imaging,” Proceedings of the National Academy of Sciences **111**(16), 6058–6062, Proceedings of the National Academy of Sciences (2014) [doi:10.1073/pnas.1317424111].
64. L. Breiman, “Heuristics of instability and stabilization in model selection,” The Annals of Statistics **24**(6), 2350–2383 (1996) [doi:10.1214/aos/1032181158].

65. N. Lange and N. M. Laird, "The Effect of Covariance Structure on Variance Estimation in Balanced Growth-Curve Models with Random Parameters," J Am Stat Assoc **84**(405), 241–247, Taylor & Francis (1989) [doi:10.1080/01621459.1989.10478761].
66. H. R. Pardoe, R. Kucharsky Hiess, and R. Kuzniecky, "Motion and morphometry in clinical and nonclinical populations," Neuroimage **135**, 177–185 (2016) [doi:https://doi.org/10.1016/j.neuroimage.2016.05.005].
67. H. K. Hausman et al., "The association between head motion during functional magnetic resonance imaging and executive functioning in older adults," Neuroimage: Reports **2**(2), 100085 (2022) [doi:https://doi.org/10.1016/j.ynirp.2022.100085].

Chenyu Gao is a Ph.D. student in electrical and computer engineering at Vanderbilt University, working with Prof. 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, working with Prof. Jerry Prince.

Bennett A. Landman, Ph.D. is professor and department chair of electrical and computer engineering at Vanderbilt University. He received his BS and MEng degree in electrical engineering and computer science from the Massachusetts Institute of Technology, and his Ph.D. degree in biomedical engineering from Johns Hopkins University. His research concentrates on applying image-processing technologies to leverage large-scale imaging studies to improve understanding of individual anatomy and personalize medicine.

Caption List

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.

Fig. 2 Noise in DTI scalar images such as fractional anisotropy (FA) images generally increase with age. But motion is also considered to increase with age.66,67 Here we show three typical subjects with both their motion levels and noise levels in DTI scans increasing.

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 for both females and males align closely; ii) rescan DTI data were acquired in most sessions; iii) there is a balanced distribution of sessions between females and males.

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.

Fig. 5 Covariate effects on FA variance, represented as standardized standard deviation, are region-specific and bidirectional across ROIs. Gender differences exist in multiple ROIs. Motion and interval exhibit opposite effect directions in many regions. Counterintuitively, motion is negatively related to FA variance in many ROIs.

Table 1 Selected ROIs (defined by Eve type-1segmentation) where interval coefficient and motion coefficient show opposite signs. Within these ROIs, FA variance increases with interval but decreases with motion.

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.