

## RESEARCH ARTICLE

# White matter microstructural metrics are sensitively associated with clinical staging in Alzheimer's disease

Yisu Yang<sup>1</sup> | Kurt Schilling<sup>2,3</sup> | Niranjana Shashikumar<sup>1</sup> | Varuna Jasodanand<sup>1</sup> | Elizabeth E. Moore<sup>1</sup> | Kimberly R. Pechman<sup>1</sup> | Murat Bilgel<sup>4</sup> | Lori L. Beason-Held<sup>4</sup> | Yang An<sup>4</sup> | Andrea Shafer<sup>4</sup> | Shannon L. Risacher<sup>5,6</sup> | Bennett A. Landman<sup>1,2,3,7,8</sup> | Angela L. Jefferson<sup>1,9,10</sup> | Andrew J. Saykin<sup>5,6</sup> | Susan M. Resnick<sup>4</sup> | Timothy J. Hohman<sup>1,9</sup> | Derek B. Archer<sup>1,9</sup> | for the Alzheimer's Disease Neuroimaging Initiative

<sup>1</sup>Vanderbilt Memory and Alzheimer's Center, Vanderbilt University School of Medicine, Nashville, Tennessee, USA

<sup>2</sup>Vanderbilt University Institute of Imaging Science, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>3</sup>Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>4</sup>Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health, Baltimore, Maryland, USA

<sup>5</sup>Indiana University School of Medicine, Indianapolis, Indiana, USA

<sup>6</sup>Indiana Alzheimer's Disease Research Center, Indianapolis, Indiana, USA

<sup>7</sup>Department of Biomedical Engineering, Vanderbilt University, Nashville, Tennessee, USA

<sup>8</sup>Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, Tennessee, USA

<sup>9</sup>Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>10</sup>Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA

## Correspondence

Derek B. Archer, PhD, Department of Neurology, Vanderbilt University Medical Center, 1207 17th Ave South, Suite 204, Nashville, TN 37212, USA.  
 Email: [derek.archer@vumc.org](mailto:derek.archer@vumc.org)

Collaborators: Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](https://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: [https://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

## Abstract

**Introduction:** White matter microstructure may be abnormal along the Alzheimer's disease (AD) continuum.

**Methods:** Diffusion magnetic resonance imaging (dMRI) data from the Alzheimer's Disease Neuroimaging Initiative (ADNI,  $n = 627$ ), Baltimore Longitudinal Study of Aging (BLSA,  $n = 684$ ), and Vanderbilt Memory & Aging Project (VMAP,  $n = 296$ ) cohorts were free-water (FW) corrected and conventional, and FW-corrected microstructural metrics were quantified within 48 white matter tracts. Microstructural values were subsequently harmonized using the *Longitudinal ComBat* technique and inputted as independent variables to predict diagnosis (cognitively unimpaired [CU], mild cognitive impairment [MCI], AD). Models were adjusted for age, sex, race/ethnicity, education, apolipoprotein E (APOE)  $\epsilon 4$  carrier status, and APOE  $\epsilon 2$  carrier status.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* published by Wiley Periodicals, LLC on behalf of Alzheimer's Association.

**Funding information**

NIA, Grant/Award Numbers: R01-AG034962, R01-AG056534, R01-AG062826, F30-AG064847, K01-AG073584, U24-AG074855, R01-AG059716, K24-AG046373; Alzheimer's Association, Grant/Award Number: IIRG-08-88733; National Institutes of Health, Grant/Award Number: U01 AG024904; Department of Defense, Grant/Award Number: W81XWH-12-2-0012; NIBIB, Grant/Award Number: K01-EB032898; NIGMS, Grant/Award Number: T32-GM007347; Intramural NIH, Grant/Award Number: 75N95D22P00141; NCATS, Grant/Award Numbers: UL1-TR000445, UL1-TR002243; OD, Grant/Award Number: 510-OD023680

**Results:** Conventional dMRI metrics were associated globally with diagnostic status; following FW correction, the FW metric itself exhibited global associations with diagnostic status, but intracellular metric associations were diminished.

**Discussion:** White matter microstructure is altered along the AD continuum. FW correction may provide further understanding of the white matter neurodegenerative process in AD.

**KEYWORDS**

diagnosis, diffusion MRI, free-water, harmonization, white matter

**Highlights**

- *Longitudinal ComBat* successfully harmonized large-scale diffusion magnetic resonance imaging (dMRI) metrics.
- Conventional dMRI metrics were globally sensitive to diagnostic status.
- Free-water (FW) correction mitigated intracellular associations with diagnostic status.
- The FW metric itself was globally sensitive to diagnostic status.

Multivariate conventional and FW-corrected models may provide complementary information.

**1 | INTRODUCTION**

Although significant emphasis has been appropriately placed on hippocampal volume in Alzheimer's disease (AD),<sup>1,2</sup> emerging evidence has also demonstrated that measures of white matter microstructure can distinguish diagnosis along the AD continuum (i.e., cognitively unimpaired [CU], mild cognitive impairment [MCI], AD). Diffusion magnetic resonance imaging (dMRI) is one method that allows us to quantify white matter microstructure in vivo, and many studies using this method have focused on fractional anisotropy (FA) and diffusivity (mean diffusivity [MD], axial diffusivity [AxD], radial diffusivity [RD]). Thus far, literature has demonstrated that there are global changes in white matter in AD, with prominent abnormalities in tracts projecting from the medial temporal lobe.<sup>3-11</sup> Due to recent advances in dMRI post-processing,<sup>12</sup> the availability of spatially precise white matter tract templates,<sup>13-16</sup> and data harmonization,<sup>17</sup> there is an unprecedented opportunity to better understand changes in white matter along the AD continuum, which will further our knowledge into the neurodegenerative processes in AD.

Although several prior studies that used dMRI in AD have been pivotal to our understanding the role of white matter in AD, many of these studies used conventional measures of dMRI ( $FA_{CONV}$ ,  $MD_{CONV}$ ,  $AxD_{CONV}$ ,  $RD_{CONV}$ ), which are subject to a well-established partial volume limitation,<sup>8,12,18-20</sup> whereby there is mixture of both tissue and fluid compartments within each voxel. Extracting the fluid compartment from the tissue compartment within each voxel would provide more accurate assessment of tissue-specific changes, and the fluid compartment may also be related to critical biological drivers of dis-

ease in AD (e.g., inflammation, atrophy). Accordingly, post-processing techniques exist that allow us to separate these compartments. One technique, called free-water (FW) elimination, allows us to segment a conventional dMRI scan into: (1) an extracellular component in which we can quantify the FW measure and (2) an intracellular component in which we can quantify FW-corrected values of conventional dMRI measures ( $FA_{FWCORR}$ ,  $MD_{FWCORR}$ ,  $AxD_{FWCORR}$ ,  $RD_{FWCORR}$ ). Prior studies evaluating intracellular metrics have found that the observed differences in white matter along the AD continuum are altered before and after FW correction.<sup>6,7</sup> One recent study focusing on white matter FW in preclinical AD found that FW is sensitively associated with several cerebrospinal fluid-derived biomarkers, including measures of amyloid, phosphorylated tau, total tau, microglial activation, and inflammation.<sup>21</sup> Even though this technique provides an enhanced insight into brain microstructure, studies thus far have primarily been single-site/cohort studies with limited sample sizes. Advances in data harmonization, such as *Longitudinal ComBat*, now allow for the correction for between-site and between-cohort heterogeneity (e.g., protocol and demographic variability) in longitudinal data sets to perform large-scale pooled analysis.

The present study aims to perform a large-scale, harmonized analysis using conventional and FW-corrected dMRI data derived from three well-established aging cohorts, including the Alzheimer's Disease Neuroimaging Initiative (ADNI), Baltimore Longitudinal Study of Aging (BLSA), and Vanderbilt Memory & Aging Project (VMAP). In our analysis, we will evaluate white matter microstructure within 48 different white matter tract templates, including 3 association, 7 limbic, 9 projection, 6 occipital transcallosal (TC), 5 parietal TC, 6 motor

TC, and 12 prefrontal TC tracts. The goal of our study is two-fold: (1) determine between-group differences (CU vs MCI vs AD) in conventional ( $FA_{CONV}$ ,  $MD_{CONV}$ ,  $AxD_{CONV}$ ,  $RD_{CONV}$ ) and FW-corrected ( $FA_{FWCORR}$ ,  $MD_{FWCORR}$ ,  $AxD_{FWCORR}$ ,  $RD_{FWCORR}$ ) metrics within each white matter tract, and (2) use recursive feature elimination (RFE), in separate conventional and FW-corrected models, to determine the best set of microstructural metrics that are most significantly associated with diagnostic category. We hypothesized that white matter tract microstructure from the limbic tracts would be most significantly associated with diagnosis category, and that FW-corrected metrics would display more pathologically relevant between-group differences.

## 2 | METHODS

### 2.1 | Study sample

The present study used data from three well-established cohorts of aging. The largest cohort was the Neuroimaging substudy of the BLSA<sup>22</sup>—behavioral assessment in this cohort began in 1994 and included dementia-free participants 55 to 85 years of age who had up to 10 years of prospective data collection at baseline. In 2009, the original cohort was expanded to include participants 20 to 85 years of age and 3T MRI data collection began. Data from the BLSA cohort are available upon request by a proposal submission through the BLSA website ([www.blsa.nih.gov](http://www.blsa.nih.gov)). Another cohort leveraged in this study was the well-known ADNI ([adni.loni.usc.edu](http://adni.loni.usc.edu)) cohort<sup>23</sup>—this cohort was launched in 2003 as a public-private partnership led by principal investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. The final cohort used in this study was VMAP<sup>24</sup>—data collection for VMAP began in 2012 and includes participants 60 years of age and older who are considered CU or who have MCI. Data from the VMAP cohort can be accessed freely following data use approval ([www.vmacdata.org](http://www.vmacdata.org)). Within each cohort, informed consent was provided by all participants. For each cohort, several demographic and clinical covariates were required for inclusion, including age, sex, educational attainment, race/ethnicity, apolipoprotein E (APOE) haplotype status ( $\epsilon 2$ ,  $\epsilon 4$ ), and cognitive diagnosis (CU, MCI, AD). In total, this study included 1718 participants across 4614 imaging sessions; however, participants were excluded if they had a conversion in diagnostic status within their longitudinal data collection (e.g., CU to MCI between two timepoints). Furthermore, individuals were only considered if they were 50 years of age or older. The cohort used in the present study was also restricted to only baseline timepoints as we are interested in the differences in cross-sectional diagnostic differences. Given all inclusion/exclusion criteria, the present study included 1607 participants. Sample sizes, demographic information, and health characteristics for each cohort can be found in Table 1. Acquisition parameters for each cohort can be found in Table S1.

### RESEARCH IN CONTEXT

- 1. Systematic review:** The authors used PubMed and Google Scholar to review literature that used conventional and free-water (FW)-corrected diffusion magnetic resonance imaging (dMRI) data to evaluate diagnostic differences along the Alzheimer's disease (AD) continuum. Although several studies have leveraged conventional dMRI to find that white matter is sensitively associated with diagnostic status, there has yet to be a large-scale, FW-corrected analysis.
- 2. Interpretation:** Conventional FW metrics are globally associated with diagnostic status. FW correction, however, mitigated these effects and the FW measure itself was highly associated with diagnostic status. Our multivariate approach demonstrated that although conventional and FW-corrected models perform similarly when associated with diagnostic status, the top neuroimaging features involved in these models differ, suggesting that the incorporation of both models may be useful.
- 3. Future directions:** Genetic studies may add knowledge into which biological pathways drive changes in conventional and FW-corrected dMRI measures and may help in understanding the role white matter has in the neurodegenerative cascade of AD. Future large-scale research is needed to fully understand how white matter microstructural metrics are associated with cognitive decline, AD pathology, and genetic predictors of AD.

### 2.2 | dMRI preprocessing

Preprocessing for all dMRI analyses followed standard procedures. First, we used an automated pipeline (*PreQual*) to correct for eddy currents, motions, and distortions. The data outputted from this step was then inputted into *DTIFIT* to calculate conventional dMRI metrics ( $FA_{CONV}$ ,  $MD_{CONV}$ ,  $AxD_{CONV}$ ,  $RD_{CONV}$ ). The preprocessed data (i.e., dMRI brain image, brain mask, bvec file, bval file) was also inputted into well-established MATLAB code<sup>12</sup> to calculate the FW component and FW-corrected dMRI metrics ( $FA_{FWCORR}$ ,  $MD_{FWCORR}$ ,  $AxD_{FWCORR}$ ,  $RD_{FWCORR}$ ).<sup>14,15,24,25</sup> After all microstructural metrics were quantified, a standard space representation of these maps was created by nonlinearly transforming the  $FA_{CONV}$  map to the FMRIB58\_FA atlas using the Advanced Normalization Tools (ANTs) package.<sup>26</sup> This warp was then applied to all remaining microstructural maps.

### 2.3 | White matter tract microstructure harmonization

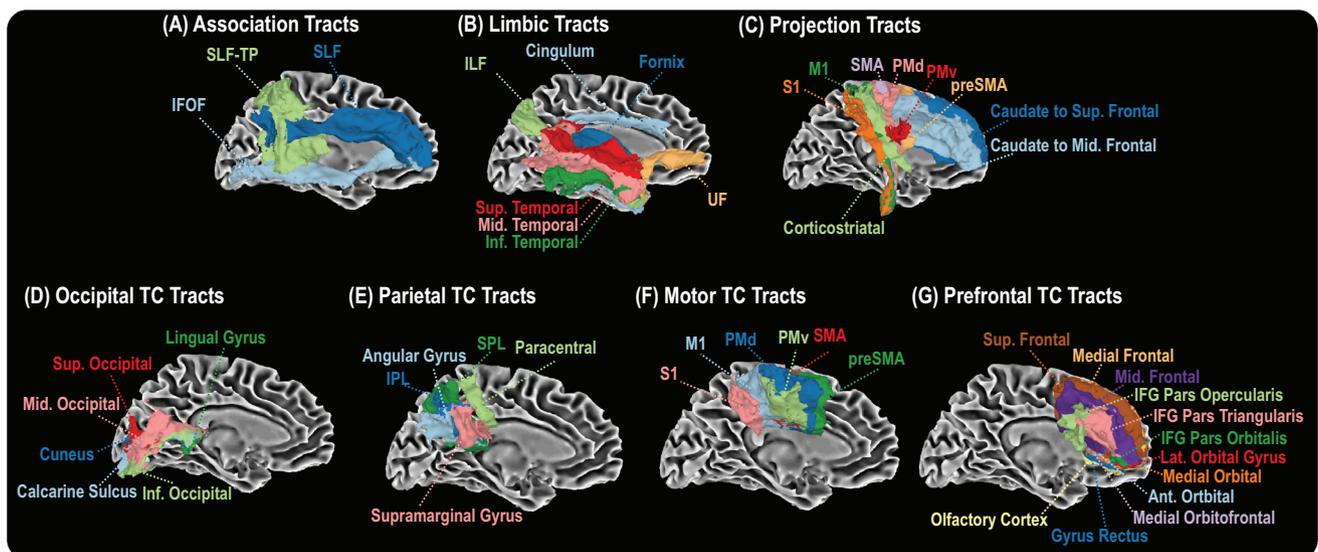
The present study used well-established tractography templates in Montreal Neurological Institute (MNI) space to evaluate white matter

**TABLE 1** Demographic and health characteristics.

Measure	Cohort			p-value
	ADNI (n = 627)	BLSA (n = 684)	VMAP (n = 296)	
Age (y)	73.86 (7.71)	70.82 (10.15)	73.32 (7.26)	<0.001
Sex (% male)	50.56	46.13	57.77	<0.001
Education (y)	16.30 (2.62)	17.03 (2.36)	15.81 (2.68)	<0.001
Race (% non-Hispanic White)	91.23	65.64	86.82	<0.001
APOE $\epsilon$ 2 (% positive)	8.61	17.54	15.20	<0.001
APOE $\epsilon$ 4 (% positive)	40.83	26.17	35.81	<0.001
Cognitive status (% CU)	49.12	98.83	56.76	<0.001

Note: Values denoted as mean (SD) or frequency.

Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; BLSA, Baltimore Longitudinal Study of Aging; CU, cognitively unimpaired; VMAP, Vanderbilt Memory & Aging Project; y, years.



**FIGURE 1** Tractography templates used in the present study. This study leveraged 48 well-established tractography templates of the association (A), limbic (B), projection (C), occipital transcallosal (TC) (D), parietal TC (E), motor TC, and prefrontal TC tracts.

microstructure (Figure 1),<sup>13-16,25,27</sup> all of which are publicly available in a GitHub repository ([https://github.com/VUMC-VMAC/Tractography\\_Templates](https://github.com/VUMC-VMAC/Tractography_Templates)). Mean values for all 48 tracts were calculated for the 4 conventional and 5 FW-corrected dMRI metrics, which were subsequently harmonized using the *Longitudinal ComBat* technique in R (version 4.1.0).<sup>28</sup> Harmonization controlled for several covariates, including age at baseline, education, sex, race/ethnicity, APOE  $\epsilon$ 4 carrier status, APOE  $\epsilon$ 2 carrier status, and the interaction of age at baseline and interval from baseline. Moreover, we included a *site x scanner x protocol* batch variable that would control for all combinations of dMRI acquisition, and random effects for intercept and interval from baseline per participant. It is important to note that the harmonization was conducted using all possible longitudinal data (number of sessions = 4,605), and each cohort differed in number of timepoints (ADNI: 4.04 [2.49]; BLSA: 3.61 [1.95]; VMAP: 3.58 [0.84]) and longitudinal follow-up time

(in years) (ADNI: 1.23 [1.75]; BLSA: 2.42 [2.54]; VMAP: 2.01 [1.86]). Harmonized data were then subset to satisfy the inclusion/exclusion criteria of the present study (number of participants = 1607).

## 2.4 | Statistical analyses

All statistical analyses were performed in R (version 4.1.0) and Python (version 3.9.7). First, we were particularly interested in determining how effectively *Longitudinal ComBat* harmonized the dMRI measures. To evaluate *Longitudinal ComBat*'s ability to harmonize the dMRI measures used in this study, we conducted an analysis using a subset of our CU participants. Specifically, we pulled a single protocol from the ADNI, BLSA, and VMAP cohorts and filtered to only include CU participants. A propensity matching analysis was then conducted to maximize overlap

of age, sex, education, race/ethnicity, APOE  $\epsilon$ 4 carrier status, and APOE  $\epsilon$ 2 carrier status.<sup>29</sup> For the initial ADNI cohort in this analysis ( $n = 131$ ), dMRI data (number of directions: 48;  $b$ -values: 0, 1000; reconstructed resolution: 2 mm  $\times$  2 mm  $\times$  2 mm) was collected on a 3T Siemens scanner. For the initial BLSA cohort in this analysis ( $n = 644$ ), dMRI data (number of directions: 32;  $b$ -values: 0, 700; reconstructed resolution: 0.81 mm  $\times$  0.81 mm  $\times$  2.2 mm) was collected on a 3T Phillips scanner. For the initial VMAP cohort in this analysis ( $n = 167$ ), dMRI data (number of directions: 32;  $b$ -values: 0, 1000; reconstructed resolution: 2 mm  $\times$  2 mm  $\times$  2 mm) was collected on a 3T Philips scanner. Between-protocol effects were found by conducting linear regression analyses and quantifying the main effect of protocol on dMRI measures, covarying for age, sex, education, race/ethnicity, APOE  $\epsilon$ 4 carrier status, and APOE  $\epsilon$ 2 carrier status. Furthermore, we conducted linear regression analyses to determine the association between age and dMRI measures and quantified *age* and *protocol*  $\times$  *age* interactions using the same covariates. We expected harmonized data to yield similar age-effect estimates within each cohort.

We then conducted multinomial logistic regression analyses to determine how each of the four conventional and five FW-corrected dMRI metrics were associated with all diagnostic categories (CU vs MCI vs AD). In this analysis, each white matter metric was evaluated separately as an independent variable, and covariates included age, sex, race/ethnicity, education, APOE  $\epsilon$ 4 carrier status, and APOE  $\epsilon$ 2 carrier status. Significance was set a priori as  $\alpha = 0.05$  and correction for multiple comparisons was made using the false discovery rate (FDR) method. Multiple comparisons were made for the conventional and FW-corrected metrics separately, totaling in the correction 192 and 240 models, respectively. In addition to the multinomial logistic regression analysis, we conducted several binomial logistic regression analyses (CU vs MCI, CU vs AD, MCI vs AD).

RFE was conducted for each diagnostic comparison. Specifically, we input all neuroimaging measures and covariates into each comparison and iteratively removed variables to maximize the weighted precision (i.e., weighted ratio of true positives to sum of true positives and false positives). We conducted this analysis for the conventional and FW-corrected dMRI measures separately for each comparison of interest, totaling in eight RFE analyses. Feature importance (i.e., absolute  $z$ -value) was then quantified for the top 10 neuroimaging measures included in the final models. To reduce collinearity, MD<sub>CONV</sub> and MD<sub>FWcorr</sub> microstructural measures were excluded from all models. Significance was set a priori as  $\alpha = 0.05$  and correction for multiple comparisons was made using the FDR method.

## 3 | RESULTS

### 3.1 | Free-water correction and *Longitudinal Combat* harmonization on dMRI data

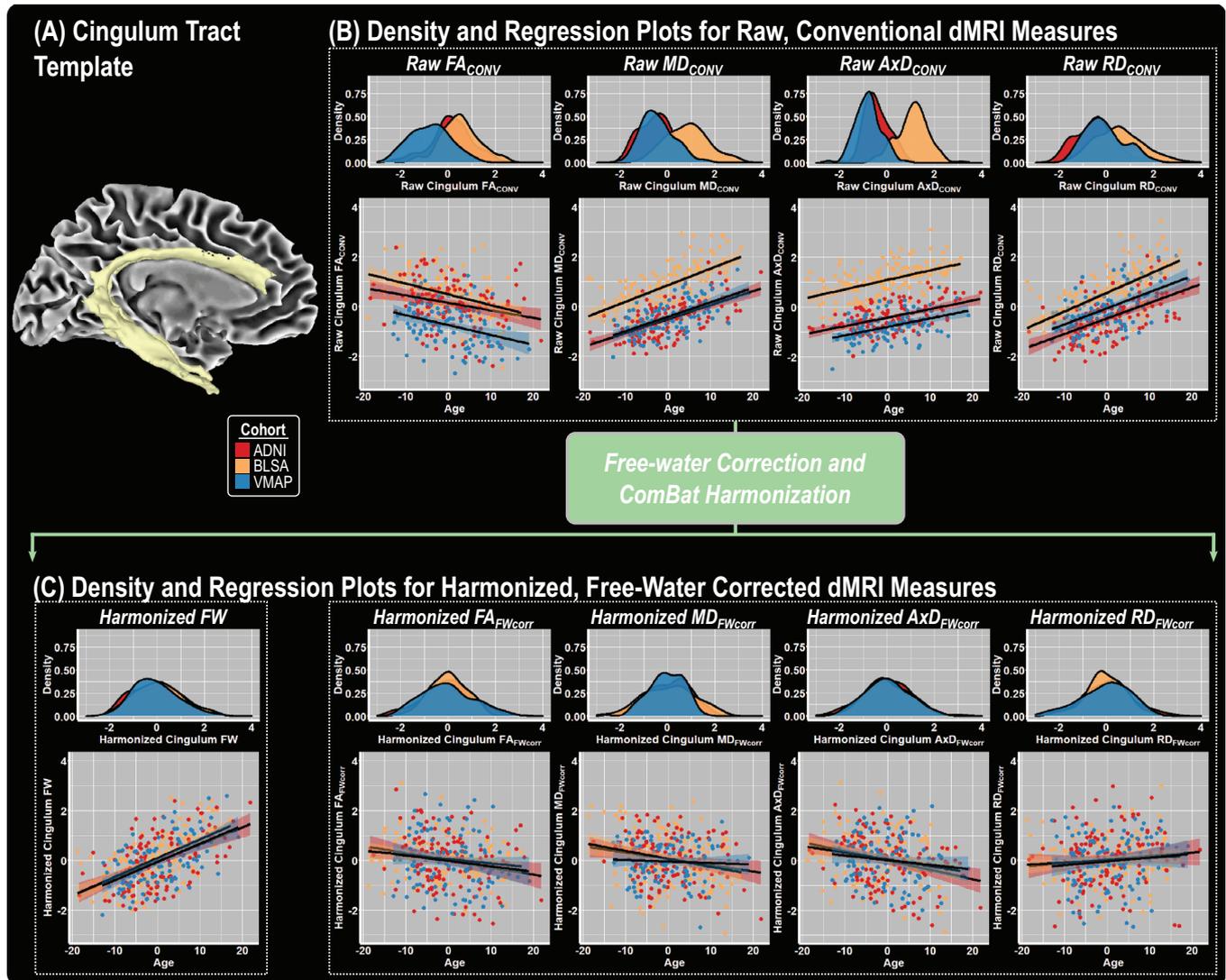
Propensity score-matched cohorts were created for ADNI ( $n = 131$ ), BLSA ( $n = 131$ ), and VMAP ( $n = 108$ ). Table S2 shows the demographic

and health characteristics of these cohorts, in which there were no significant differences in age, sex, education, race, APOE  $\epsilon$ 4 carrier status, or APOE  $\epsilon$ 2 carrier status. Figure 2 illustrates the differences seen before and after FW correction and *Longitudinal Combat* harmonization for the cingulum tract (Figure 2A). Between-group analysis of cingulum distributions of raw, conventional dMRI metrics revealed highly significant differences in FA<sub>CONV</sub>, MD<sub>CONV</sub>, Ax<sub>CONV</sub>, and RD<sub>CONV</sub> (all  $p$ 's  $< 0.001$ ) (Figure 2B). Next, we conducted between-group analysis of cingulum bundle distributions of harmonized, FW-corrected dMRI metrics (Figure 2C) and found no significant effects (all  $p$ 's  $> 0.05$ ). We then evaluated if there were *protocol*  $\times$  *age* interactions in the association with dMRI metrics, and there were no significant interactions for any conventional or FW-corrected variables before or after harmonization (all  $p$ 's  $> 0.05$ ).

### 3.2 | White matter tract microstructure relationship with diagnosis category

The conventional dMRI microstructure effects with all diagnostic categories (CU vs MCI vs AD) are shown in Figure 3 and all relevant statistics are found in Table S3. As shown in Figure 3A (top row), there were nearly global associations with conventional dMRI microstructure on diagnostic category, in which there were increases in MD<sub>CONV</sub>, Ax<sub>CONV</sub>, and RD<sub>CONV</sub> accompanied with decreases in FA<sub>CONV</sub>. As expected, effects were especially pronounced in the limbic tracts, particularly for the cingulum and fornix. For the cingulum, there were significant effects for all metrics, including higher MD<sub>CONV</sub> ( $\beta = 13.37 \pm 1.10$ ;  $p_{\text{FDR}} < 2.2 \times 10^{-16}$ ), Ax<sub>CONV</sub> ( $\beta = 11.58 \pm 1.12$ ;  $p_{\text{FDR}} < 2.20 \times 10^{-16}$ ), and RD<sub>CONV</sub> ( $\beta = 12.78 \pm 1.07$ ;  $p_{\text{FDR}} < 2.20 \times 10^{-16}$ ) in addition to lower FA<sub>CONV</sub> ( $\beta = -5.56 \pm 0.83$ ;  $p_{\text{FDR}} = 2.82 \times 10^{-11}$ ), for increased severity in diagnostic category. For the fornix, there was higher MD<sub>CONV</sub> ( $\beta = 13.82 \pm 1.11$ ;  $p_{\text{FDR}} < 2.20 \times 10^{-16}$ ), Ax<sub>CONV</sub> ( $\beta = 13.00 \pm 1.08$ ;  $p_{\text{FDR}} < 2.20 \times 10^{-16}$ ), and RD<sub>CONV</sub> ( $\beta = 14.28 \pm 1.15$ ;  $p_{\text{FDR}} < 2.20 \times 10^{-16}$ ) in addition to lower FA<sub>CONV</sub> ( $\beta = -2.55 \pm 1.03$ ;  $p_{\text{FDR}} = 0.014$ ), for increased severity in diagnostic category. Illustrations of the effects of conventional dMRI microstructure on diagnostic category are shown in Figure 3B, which shows that the most pronounced differences are within MD<sub>CONV</sub>, Ax<sub>CONV</sub>, and RD<sub>CONV</sub>. Tables S4–S6 show relevant statistics for all other comparisons (CU vs MCI, CU vs AD, MCI vs AD).

The FW-corrected dMRI microstructure effects with all diagnostic categories (CU vs MCI vs AD) are shown in Figure 3 and relevant statistics are shown in Table S7. Although there were near global effects for the conventional dMRI metrics, Figure 3A (bottom row) shows that there is a near global effect for the FW metric itself, but also shows that FW correction mitigated the significance shown in the intracellular dMRI metrics. Nevertheless, significance was still pronounced within the limbic tracts, particularly the cingulum and fornix. For the cingulum, there were significant effects for three metrics, including higher FW ( $\beta = 14.13 \pm 1.13$ ;  $p_{\text{FDR}} < 2.20 \times 10^{-16}$ ), lower MD<sub>FWcorr</sub> ( $\beta = -5.13 \pm 0.82$ ;  $p_{\text{FDR}} = 4.81 \times 10^{-10}$ ), and lower Ax<sub>FWcorr</sub> ( $\beta = -3.45 \pm 0.76$ ;  $p_{\text{FDR}} = 5.43 \times 10^{-6}$ ), with increased severity along

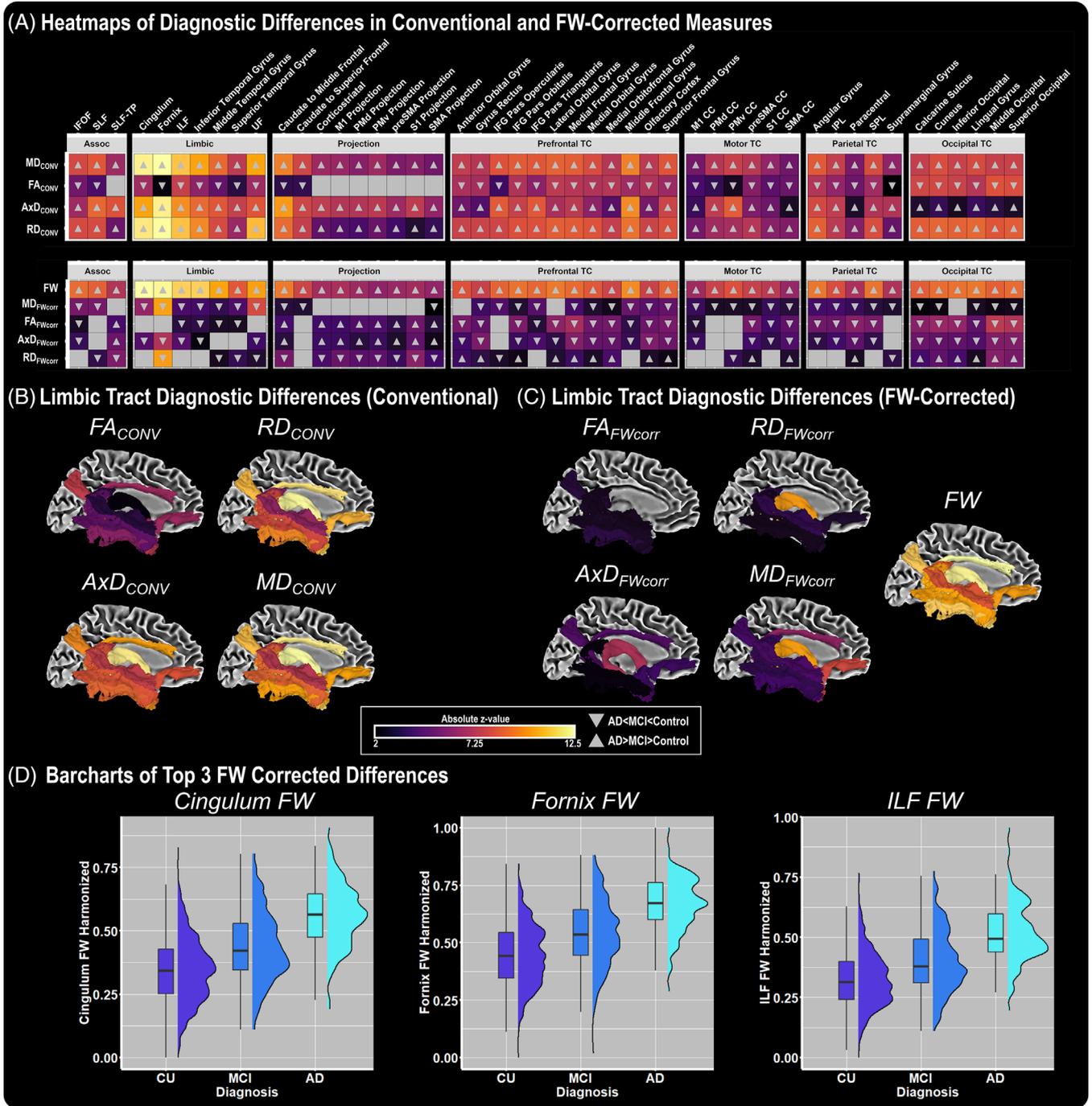


**FIGURE 2** Demonstration of free-water (FW) correction and *Longitudinal ComBat* harmonization in covariate-matched participants. Propensity score matching was conducted to find covariate-matched participants in the Alzheimer's Disease Neuroimaging Initiative (ADNI), Baltimore Longitudinal Study of Aging (BLSA), and Vanderbilt Memory and Aging Project (VMAP) cohorts. It is notable that the propensity score matching was conducted on a single protocol for each cohort and participants were matched to balance age, sex, education race/ethnicity, diagnosis, apolipoprotein E (APOE)  $\epsilon 4$  carrier status, and APOE  $\epsilon 2$  carrier status. Conventional dMRI measures are shown for the cingulum tract template (A), and distributions of the raw, conventional measures in addition to the associations with age are shown (B). The dMRI scans were subsequently FW corrected and *Longitudinal ComBat* harmonization was conducted on the microstructural values. Distributions of the harmonized, FW-corrected measures in addition to the associations with age are shown (C).

the AD continuum. For the fornix, there were significant effects for four metrics, including higher FW ( $\beta = 15.17 \pm 1.23$ ;  $p_{FDR} < 2.20 \times 10^{-16}$ ) in addition to lower  $MD_{FWcorr}$  ( $\beta = -9.18 \pm 0.89$ ;  $p_{FDR} < 2.20 \times 10^{-16}$ ),  $AxD_{FWcorr}$  ( $\beta = -6.38 \pm 0.92$ ;  $p_{FDR} = 3.90 \times 10^{-12}$ ), and  $RD_{FWcorr}$  ( $\beta = -9.07 \pm 0.89$ ;  $p_{FDR} < 2.20 \times 10^{-16}$ ). Illustrations of the effects of FW-corrected microstructure on diagnostic category are shown in Figure 3C, which shows that the most pronounced differences are within the FW metric, and bar charts of the most pronounced FW-corrected differences are shown in Figure 3D. Tables S8–S10 show relevant statistics for all other comparisons (CU vs MCI, CU vs AD, MCI vs AD).

### 3.3 | Multivariate regression to maximize association with diagnosis category

The multivariate logistic regression analysis results for the set of conventional dMRI microstructural metrics that best predict diagnostic category (CU vs MCI vs AD) are shown in Figure 4 and relevant statistics are shown in Table 2. Although Table 2 (left panel) shows that the top conventional dMRI measures selected using RFE analysis significantly predicted diagnostic category ( $R_{adj}^2 = 0.433$ ;  $p_{FDR} = 4.76 \times 10^{-143}$ ), Figure 4A (left panel) shows that significance was most pronounced within the uncinate fasciculus (UF), with other

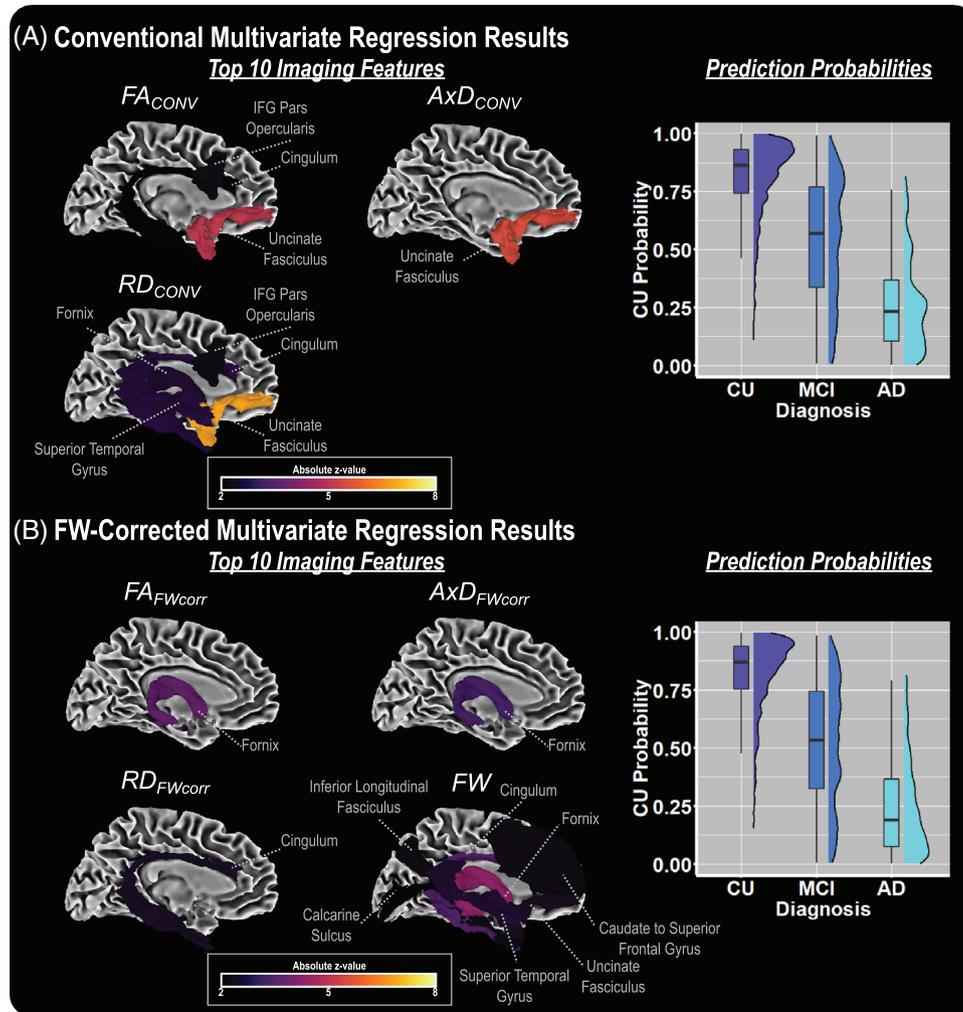


**FIGURE 3** Diagnostic differences in white matter microstructure. The strengths (i.e., absolute z-values) are shown for the relationship between diagnostic category and conventional (A, top row) and free-water (FW) corrected (A, second row) white matter microstructure. Limbic tract diagnostic differences for the conventional (B) and FW-corrected (C) measures are illustrated, which show prominent differences in the fornix and cingulum. FW for the cingulum, fornix, and inferior longitudinal fasciculus (ILF) are shown (D).

contributions from the cingulum, fornix, inferior frontal gyrus (IFG) pars opercularis, and TC superior temporal gyrus. Illustration of prediction probabilities for the CU diagnosis using conventional dMRI metrics are shown in Figure 4A (right panel), which shows that individuals with CU and AD diagnoses displayed the highest and lowest probabilities of obtaining a CU diagnosis, respectively, whereas those with MCI diagnosis varied largely in their probability of obtaining

a CU diagnosis. Table 2 also shows relevant statistics for all other comparisons (CU vs MCI, CU vs AD, MCI vs AD) using conventional measures.

The multivariate logistic regression analysis results for the set of FW-corrected dMRI microstructural metrics that best predict diagnostic category (CU vs MCI vs AD) are shown in Figure 4 and relevant statistics are shown in Table 2. Although Table 2 (right panel)



**FIGURE 4** Multivariate logistic regression on diagnostic category. The top 10 neuroimaging features and prediction probabilities from a recursive feature elimination analysis are shown for the conventional (A) and free-water (FW)-corrected (B) multivariate regression analyses.

**TABLE 2** Multivariate regression statistics.

	Conventional measures			FW-corrected measures		
	$R_{adj}^2$	F-statistic	$p_{FDR}$	$R_{adj}^2$	F-statistic	$p_{FDR}$
CU vs MCI vs AD	0.433	13.01	$4.76 \times 10^{-143}$	0.418	28.38	$1.13 \times 10^{-160}$
CU vs MCI	0.267	5.40	$5.05 \times 10^{-58}$	0.243	19.50	$1.48 \times 10^{-76}$
CU vs AD	0.414	20.74	$8.12 \times 10^{-119}$	0.445	21.55	$1.87 \times 10^{-130}$
MCI vs AD	0.242	5.03	$6.13 \times 10^{-17}$	0.216	9.34	$4.21 \times 10^{-19}$

Note: Several recursive feature elimination analyses were conducted using the conventional and FW-corrected dMRI measures. Abbreviations: AD, Alzheimer's disease; CU, cognitively unimpaired; MCI, mild cognitive impairment.

shows that the top FW-corrected dMRI measures selected using RFE analysis significantly predicted diagnostic category ( $R_{adj}^2 = 0.418$ ;  $p = 1.13 \times 10^{-160}$ ) and were similar in overall model performance to conventional measures. Figure 4B (left panel) shows that, as expected, significance was most pronounced in the FW metric of the fornix, which is part of the limbic tract group. However, the prediction probabilities plot for the CU diagnosis using FW-corrected dMRI metrics

shown in Figure 4B (right panel) is nearly identical to that using conventional dMRI metrics, where individuals with CU and AD diagnoses displayed the highest and lowest probabilities of obtaining a CU diagnosis, respectively, whereas those with MCI diagnosis varied largely in their probability of obtaining a CU diagnosis. Table 2 also shows relevant statistics for all other comparisons (CU vs MCI, CU vs AD, MCI vs AD).

## 4 | DISCUSSION

This study leveraged dMRI data within 48 white matter tracts in conjunction with the FW-correction post-processing technique and *Longitudinal ComBat* harmonization to investigate the relationship between white matter microstructure alterations and progression along the AD continuum. Specifically, we identified white matter metrics that were most robustly correlated with an individual's diagnosis as CU, MCI, or AD using conventional and FW-corrected dMRI measures and evaluated these two methods regarding their predictive strengths. First, we found that consistent with the previous literature and current understanding of AD neuropathology, microstructural metrics of limbic tracts white matter (e.g., cingulum, fornix) were identified as the most robust predictors of diagnosis in both conventional and FW-corrected analyses. Second, we found that the top neuroimaging features involved in conventional and FW-corrected multivariate analyses differed drastically, suggesting that the incorporation of both conventional and FW-corrected measures in studies of AD may provide complementary information.

We found global associations with conventional dMRI metrics, in which increased clinical severity along the AD continuum was associated with higher  $MD_{CONV}$ ,  $AxD_{CONV}$ , and  $RD_{CONV}$  in addition to lower  $FA_{CONV}$ . In addition, we found that the most significant effects were localized to the limbic tracts, including the cingulum, fornix, ILF, UF as well as TC projections of the inferior, middle, and superior temporal gyri. These results confirm long-standing research that white matter microstructure within the limbic tracts can differentiate diagnostic categories.<sup>3-11</sup> For example, one study found that widespread regions within the cingulum bundle exhibited lower  $FA_{CONV}$  and higher  $MD_{CONV}$  in amnesic MCI ( $n = 23$ ) and AD ( $n = 31$ ) participants compared to CU ( $n = 14$ ).<sup>10</sup> Another study using ADNI data found that although there were global changes in conventional dMRI metrics between CU ( $n = 44$ ), MCI ( $n = 88$ ), and AD ( $n = 23$ ) participants, tracts projecting from the temporal and posterior brain regions were particularly sensitive to diagnostic status and sensitively associated with cognitive impairment.<sup>11</sup> Finally, a recent study leveraged a support vector, machine-learning algorithm to distinguish between CU ( $n = 15$ ), MCI ( $n = 15$ ), and AD ( $n = 15$ ); investigators found that ventral portions of the cingulum were capable of distinguishing categories with high accuracy (AD vs controls: 90%, AD vs MCI: 87%). Although our analysis using conventional metrics provided results that were similar to those of prior studies, *Longitudinal ComBat* harmonization allowed us to perform this analysis on a substantially larger cohort of CU ( $n = 1,152$ ), MCI ( $n = 350$ ), and AD ( $n = 105$ ) individuals.

In our FW-corrected analysis, we found a global association with the FW metric itself, such that increased clinical severity along the AD continuum was associated with elevated FW. This confirms smaller FW studies studying the AD continuum.<sup>6,7</sup> For example, one study used a whole-brain white matter mask in 81 CU participants, 103 MCI patients, and 42 AD patients to investigate if the FW measure itself can determine diagnostic category.<sup>7</sup> They found that FW was elevated in the MCI and AD participants even after removing white matter

hyperintensities from the analyses. In addition to global FW effects, the present study found that there were significant effects within the intracellular metrics, although the effect sizes were mitigated compared to their conventional dMRI counterparts. These results are in line with a prior study that evaluated white matter microstructural patterns in CU ( $n = 20$ ) and MCI ( $n = 25$ ) patients before and after FW correction.<sup>6</sup> Specifically, they found that using conventional dMRI metrics made analyses susceptible to both false-positive and false-negative between-group differences. In this prior study, they suggest that cerebrospinal fluid (i.e., partial volume) contamination is particularly pronounced within the fornix, and therefore conventional dMRI metrics are vulnerable to this contamination. Our results confirm this hypothesis by demonstrating that  $MD_{CONV}$ ,  $AxD_{CONV}$ , and  $RD_{CONV}$  are elevated along the AD continuum, but after FW correction, the elevation is confined within the FW metric itself and  $MD_{FWCORR}$ ,  $AxD_{FWCORR}$ , and  $RD_{FWCORR}$  are lower along the AD continuum.

We also conducted separate multivariate logistic regression analyses to determine the best set of conventional and FW-corrected variables that predict status along the AD continuum. Of interest, we found comparable performance for the conventional ( $R_{adj}^2 = 0.433$ ) and FW-corrected ( $R_{adj}^2 = 0.418$ ) models, in which prediction probability plots for CU diagnosis look almost identical between the conventional and FW-corrected analyses (see Figure 4). However, differences emerge between models when evaluating feature importance. Given these findings, we suggest that the incorporation of both conventional and FW-corrected metrics may provide useful information when studying the neurodegenerative process in AD.<sup>30,31</sup>

The present study has several strengths. First, we leveraged dMRI data of 48 white matter tracts from a combined sample of 1607 individuals spanning three well-established cohorts. This large sample size allows us to examine the predictive strengths of relevant dMRI measures more systematically and makes results more generalizable across studies. Second, we employed the FW-correction technique on these neuroimaging results, which allowed for the separation and individual analysis of extracellular and intracellular components of a diffusion image. Third, we effectively harmonized dMRI data across dozens of *site x scanner x protocol* combinations. Despite these strengths, our cohort is mostly non-Hispanic White and is highly educated. Although we included only non-converting (i.e., no change in clinical diagnosis) participants in our study, it is possible that further examination of these participants would lead to diagnostic changes. In addition, we did not consider the high heterogeneity across AD, and consequently, it is possible that our findings may not generalize to other studies. Future studies evaluating differential white matter patterns between AD subtypes (e.g., typical, limbic-predominant, hippocampal-sparing, minimal atrophy)<sup>33</sup> are necessary. Although prior studies have evaluated amyloidosis and its role on white matter abnormalities in AD,<sup>34-36</sup> we did not consider this as a covariate—future large-scale studies evaluating amyloidosis and white matter microstructure in preclinical AD would be useful. We found that FW was sensitive along the AD continuum; however, it is currently unclear what exact biological mechanisms are associated with this metric—additional attention on this variable is necessary to determine which mechanisms in AD are associated with

FW abnormalities. Replication of our results in a more diverse, diagnostically balanced cohort that has longer longitudinal follow-up is needed to further evaluate conventional and FW-corrected dMRI measures in their abilities to distinguish between different stages of AD and to identify highest contributing white matter microstructural metrics.

## ACKNOWLEDGMENTS

This study was supported by several funding sources, including K01-EB032898 (K.G.S.), T32-GM007347 (E.E.M.), F30-AG064847 (E.E.M.), K01-AG073584 (D.B.A.), U24-AG074855 (T.J.H.), 75N95D22P00141 (T.J.H.), R01-AG059716 (T.J.H.), UL1-TR000445 and UL1-TR002243 (Vanderbilt Clinical Translational Science Award), and S10-OD023680 (Vanderbilt's High-Performance Computer Cluster for Biomedical Research). The research was supported in part by the Intramural Research Program of the National Institutes of Health, the National Institute on Aging. Study data were obtained from the Vanderbilt Memory and Aging Project (VMAP). VMAP data were collected by Vanderbilt Memory and Alzheimer's Center Investigators at Vanderbilt University Medical Center. This work was supported by NIA grants R01-AG034962 (PI: Jefferson), R01-AG056534 (PI: Jefferson), R01-AG062826 (PI: Gifford), and Alzheimer's Association IIRG-08-88733 (PI: Jefferson). This research was supported in part by the Intramural Research Program of the National Institutes of Health, the National Institute on Aging. Data collection and sharing for this project was funded (in part) by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is 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](http://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.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the [Supporting Information](#).

## REFERENCES

1. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:280-292.
2. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14:535-562.
3. Sexton CE, Kalu UG, Filippini N, Mackay CE, Ebmeier KP. A meta-analysis of diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease. *Neurobiol Aging*. 2011;32:2322 e5-18.
4. Stricker NH, Schweinsburg BC, Delano-Wood L, et al. Decreased white matter integrity in late-myelinating fiber pathways in Alzheimer's disease supports retrogenesis. *Neuroimage*. 2009;45:10-16.
5. Walsh R, Bergamino M, Stokes A. Free-water diffusion tensor imaging (DTI) improves the accuracy and sensitivity of white matter analysis in Alzheimer's disease. *Scientific Reports*. 2021;11:6990.
6. Berlot R, Metzler-Baddeley C, Jones DK, O'Sullivan MJ. CSF contamination contributes to apparent microstructural alterations in mild cognitive impairment. *Neuroimage*. 2014;92:27-35.
7. Dumont M, Roy M, Jodoin PM, et al. Free water in white matter differentiates MCI and AD from control subjects. *Front Aging Neurosci*. 2019;11:270.
8. Schouten TM, Koini M, Vos F, et al. Individual classification of Alzheimer's disease with diffusion magnetic resonance imaging. *Neuroimage*. 2017;152:476-481.
9. Dalboni da Rocha JL, Bramati I, Coutinho G, Tovar Moll F, Sitaram R. Fractional anisotropy changes in parahippocampal cingulum due to Alzheimer's Disease. *Sci Rep*. 2020;10:2660.
10. Bozzali M, Giulietti G, Basile B, et al. Damage to the cingulum contributes to Alzheimer's disease pathophysiology by deafferentation mechanism. *Hum Brain Mapp*. 2012;33:1295-1308.
11. Nir TM, Jahanshad N, Villalon-Reina JE, et al. Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging. *Neuroimage Clin*. 2013;3:180-195.
12. Pasternak O, Sochen N, Gur Y, Intrator N, Assaf Y. Free water elimination and mapping from diffusion MRI. *Magn Reson Med*. 2009;62:717-730.
13. Archer DB, Vaillancourt DE, Coombes SA. A template and probabilistic atlas of the human sensorimotor tracts using diffusion MRI. *Cereb Cortex*. 2018;28:1685-1699.
14. Archer DB, Moore EE, Shashikumar N, et al. Free-water metrics in medial temporal lobe white matter tract projections relate to longitudinal cognitive decline. *Neurobiol Aging*. 2020;94:15-23.
15. Archer DB, Coombes SA, McFarland NR, DeKosky ST, Vaillancourt DE. Development of a transcallosal tractography template and its application to dementia. *Neuroimage*. 2019;200:302-312.
16. Brown CA, Johnson NF, Anderson-Mooney AJ, et al. Development, validation and application of a new fornix template for studies of aging and preclinical Alzheimer's disease. *Neuroimage Clin*. 2017;13:106-115.
17. Radua J, Vieta E, Shinohara R, et al. Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *Neuroimage*. 2020;218:116956.
18. Metzler-Baddeley C, O'Sullivan MJ, Bells S, Pasternak O, Jones DK. How and how not to correct for CSF-contamination in diffusion MRI. *Neuroimage*. 2012;59:1394-1403.
19. Alexander AL, Hasan KM, Lazar M, Tsuruda JS, Parker DL. Analysis of partial volume effects in diffusion-tensor MRI. *Magn Reson Med*. 2001;45:770-780.
20. Vos SB, Jones DK, Viergever MA, Leemans A. Partial volume effect as a hidden covariate in DTI analyses. *Neuroimage*. 2011;55:1566-1576.
21. Hoy AR, Ly M, Carlsson CM, et al. Microstructural white matter alterations in preclinical Alzheimer's disease detected using

- free water elimination diffusion tensor imaging. *PLoS One*. 2017;12:e0173982.
22. Ferrucci L. The Baltimore Longitudinal Study of Aging (BLSA): a 50-year-long journey and plans for the future. *J Gerontol: Series A*. 2008;63:1416-1419.
  23. Jack CR, Bernstein MA, Fox NC, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging*. 2008;27:685-691.
  24. Jefferson AL, Gifford KA, Acosta LM, et al. The Vanderbilt Memory & Aging Project: study design and baseline cohort overview. *J Alzheimer's Dis*. 2016;52:539-559.
  25. Archer DB, Bricker JT, Chu WT, et al. Development and validation of the automated imaging differentiation in parkinsonism (AID-P): a multicentre machine learning study. *Lancet Digit Health*. 2019;1:e222-e231.
  26. Avants B, Epstein C, Grossman M, Gee J. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal*. 2008;12:26-41.
  27. Mori S, Oishi K, Jiang H, et al. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage*. 2008;40:570-582.
  28. Beer JC, Tustison NJ, Cook PA, et al. Longitudinal ComBat: a method for harmonizing longitudinal multi-scanner imaging data. *Neuroimage*. 2020;220:117129.
  29. Ho DE, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. *J Stat Softw*. 2011;42:1-28.
  30. Jack CR, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol*. 2013;12:207-216.
  31. Whitwell JL, Josephs KA, Murray ME, et al. MRI correlates of neurofibrillary tangle pathology at autopsy: a voxel-based morphometry study. *Neurology*. 2008;71:743-749.
  32. Ferreira D, Nordberg A, Westman E. Biological subtypes of Alzheimer disease: a systematic review and meta-analysis. *Neurology*. 2020;94:436-448.
  33. Molinuevo JL, Ripolles P, Simó M, et al. White matter changes in pre-clinical Alzheimer's disease: a magnetic resonance imaging-diffusion tensor imaging study on cognitively normal older people with positive amyloid  $\beta$  protein 42 levels. *Neurobiol Aging*. 2014;35:2671-2680.
  34. Racine AM, Adluru N, Alexander AL, et al. Associations between white matter microstructure and amyloid burden in preclinical Alzheimer's disease: a multimodal imaging investigation. *Neuroimage Clin*. 2014;4:604-614.
  35. Collij LE, Ingala S, Top H, et al. White matter microstructure disruption in early stage amyloid pathology. *Alzheimers Dement (Amst)*. 2021;13:e12124.
  36. Vipin A, Ng KK, Ji F, et al. Amyloid burden accelerates white matter degradation in cognitively normal elderly individuals. *Hum Brain Mapp*. 2019;40:2065-2075.

### SUPPORTING INFORMATION

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

**How to cite this article:** Yang Y, Schilling K, Shashikumar N, et al. White matter microstructural metrics are sensitively associated with clinical staging in Alzheimer's disease. *Alzheimer's Dement*. 2023;15:e12425. <https://doi.org/10.1002/dad2.12425>