

## Cross-sectional and longitudinal Biomarker extraction and analysis for multicentre FLAIR brain MRI<sup>☆</sup>

J. DiGregorio<sup>a</sup>, A. Gibicar<sup>a</sup>, H. Khosravani<sup>b</sup>, P. Jabejdar Maralani<sup>c</sup>, J.-C. Tardif<sup>d,e</sup>, P. N. Tyrrell<sup>f,c</sup>, A.R. Moody<sup>c</sup>, A. Khademi<sup>a,g,h,\*</sup>

<sup>a</sup> Electrical, Computer and Biomedical Engineering Dept., Ryerson University, Toronto, ON, Canada

<sup>b</sup> Division of Neurology, Department of Medicine, University of Toronto, Toronto, ON, Canada

<sup>c</sup> Department of Medical Imaging, University of Toronto, Toronto, ON, Canada

<sup>d</sup> Montreal Heart Institute, Montreal, QU, Canada

<sup>e</sup> Department of Medicine, Université de Montréal, QU, Canada

<sup>f</sup> Department of Statistical Science, University of Toronto, Toronto, ON, Canada

<sup>g</sup> Keenan Research Center for Biomedical Science, St. Michael's Hospital, Unity Health Network, Toronto, ON, Canada

<sup>h</sup> Institute for Biomedical Engineering, Science and Technology (iBEST), A Partnership Between St. Michael's Hospital and Ryerson University, Canada

### ARTICLE INFO

#### Keywords:

FLAIR MRI  
Cerebral biomarkers  
Alzheimer's disease  
Cerebrovascular disease

### ABSTRACT

Fluid-attenuated inversion recovery (FLAIR) MRI has emerged as an important sequence for the analysis of cerebrovascular (CVD) and Alzheimer's disease (AD). Large-scale, automated cross-sectional and longitudinal cerebral biomarker extraction from FLAIR datasets could progress disease characterization, improve disease monitoring, and help to determine optimal intervention times. Despite this, most automated biomarker extraction algorithms are designed for T1-weighted or multi-modal inputs. In this work, automated tools were used to extract biomarkers from large, FLAIR-only datasets to evaluate the feasibility of this sequence to characterize healthy, AD, and CVD subjects in a similar manner to traditional approaches. Total brain volume (TBV), cerebrospinal fluid (CSF) volume, and white matter lesion (WML) volume was measured for the cross-sectional biomarkers and the corresponding annual rates of change over multiple scans represented the longitudinal biomarkers. Biomarkers were extracted from two dementia cohorts (4356 vol, 162 233 images) and one vascular disease cohort (869 vol, 42 850 images) using deep learning-based segmentation algorithms designed specifically for FLAIR. Biomarkers from all cohorts were summarized using descriptive statistics, correlation analysis, and ANCOVA to assess differences in diagnostic labels while accounting for demographic and acquisition factors. Biomarkers from FLAIR MRI had similar trends with those extracted from traditional modalities in the literature for characterizing healthy, AD, and CVD subjects. This demonstrates that FLAIR MRI can be used for end-to-end analysis of large AD and CVD datasets, which can lower acquisition costs, simplify clinical translation, and reduce measurement error associated with multi-modal approaches.

### 1. Introduction

Neurodegenerative disease research is a growing priority due to the significant burden placed on healthcare systems. It is estimated that by 2031, one million Canadians will be living with dementia with an annual cost of care exceeding \$16 billion (Chambers et al., 2016). Alzheimer's disease (AD) is the leading cause of dementia (Mayeux and

Stern, 2012) with cerebrovascular disease (CVD) identified as the second most common contributor to advanced cognitive impairment (Smith et al., 2017). Currently, there are limited options for prevention and treatment (Iadecola, 2013). Cognitive impairment in early stages of AD is commonly referred to as mild cognitive impairment (MCI) and represents individuals who have significant memory deficits but do not meet AD criteria (Evans et al., 2010; Jack et al., 2008a; Silbert et al.,

<sup>☆</sup> \* Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://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 analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

\* Corresponding author. Electrical, Computer and Biomedical Engineering Dept., Ryerson University, Toronto, ON, Canada.

E-mail address: [akhademi@ryerson.ca](mailto:akhademi@ryerson.ca) (A. Khademi).

<https://doi.org/10.1016/j.ynirp.2022.100091>

Received 23 November 2021; Received in revised form 18 February 2022; Accepted 28 February 2022

Available online 8 March 2022

2666-9560/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Table 1**  
Summary of data in the CAIN, ADNI, and CCNA databases.

Disease Volumes per Disease	CAIN (total = 869, subjects = 400)						ADNI (total = 4102, subjects = 900)					CCNA (subjects = 254)			
	Ischemic Disease 869						CN 1270	MCI 2296	AD 536	SCI 50	MCI 98	V-MCI 63	AD 43		
Age Group Volumes per Age Group	40s	50s	60s	70s	80s	90s	50s	60s	70s	80s	90s	60s	70s	80s	
	8	56	288	398	111	10	131	1180	1990	757	44	86	120	48	
Total Image Slices	42 850						150 000					12 233			
Male/Female (%)	62%/38%						53%/47%					54%/46%			
Centers	8						58					18			
Scanner Vendors	GE, Philips, Siemens						GE, Philips, Siemens					GE, Philips, Siemens			
Magnetic Field (T)	3						3					3			
TR (ms)	800–11000						6000–11900					9000–9840			
TE (ms)	117–150						90–192					117–148			
TI (ms)	2200–2800						2250–2800					2250–2500			
Pixel Spacing (mm)	0.43–1.09						0.78–1.02					0.94			
Slice Thickness (mm)	3–5						2–6					3			

2009). The MCI phase represents a treatment intervention window before advanced cognitive impairment and irreversible neurodegeneration occurs (Evans et al., 2010). However, determining whom to treat and when requires reliable biomarkers that stratify patients, monitor disease progression, and establish optimal intervention times. As MRI provides near real-time clues related to neurodegeneration and aging, MRI based biomarkers have been strong candidates for these applications.

There are many cerebral biomarkers studied for neurodegenerative disease characterization. For example, white matter lesions (WML) have been associated with cognitive impairment and increased risk of stroke and dementia (DeBette and Markus, 2010). Total brain volume (TBV), primarily composed of gray matter (GM) and white matter (WM), has also been studied, and while TBV naturally declines with age at a rate of about 5% per decade after the age of 40 (N. C. Fox and Schott, 2004; Scahill et al., 2003; Svennerholm et al., 1997), AD and CVD increase this rate (N. Fox et al., 1999; Scahill et al., 2002). As GM and WM regions atrophy, cerebrospinal fluid (CSF) spaces in the ventricles and cortical sulci expand. As such, ventricular enlargement and increased CSF volume have also been correlated with brain atrophy and cognitive decline (Freeborough and Fox, 1997; Nestor et al., 2008). Manual analysis of these MRI biomarkers is laborious and prone to inter-observer variability (Bilello et al., 2010; Wilke et al., 2011). In contrast, automated tools can extract neuroimaging biomarkers from large cohorts efficiently and objectively.

There are many automated frameworks for neurological biomarker extraction designed for T1- and T2-weighted images (Evans et al., 2010; Hansen et al., 2015; Heinen et al., 2019; Narayana et al., 2020; Nordenskjöld et al., 2013). However, the fluid-attenuated inversion recovery (FLAIR) MRI sequence has emerged as an important modality for AD and CVD imaging (García-Lorenzo et al., 2013; Heinen et al., 2019; Khademi et al., 2011; Wardlaw et al., 2015). WML visualization is ideal in FLAIR MRI since the CSF signal is suppressed and WML are hyperintense (Soltanian-Zadeh and Peck, 2001). WML have typically been analyzed by registering T2-weighted or FLAIR images to their corresponding T1-weighted space (Wardlaw et al., 2015). As these methods are dependent on a secondary sequence they cannot be directly translated to FLAIR alone (Reiche et al., 2019). Additionally, registration can distort lesion location and size in volumes with small lesion loads due to partial volume averaging, especially when there is a mismatch in voxel sizes (De Boer et al., 2007). In (Narayana et al., 2020), the authors showed that using FLAIR as a sole input to a deep learning model for WML segmentation provided similar performance to models trained with multimodal sequences and FLAIR on its own gave way to lower false positive rates for low lesion loads. In (Khademi et al., 2021), we showed that deep learning-based WML segmentation on multicentre FLAIR MRI provided the highest similarity to ground truth WML delineations compared to traditional thresholding and machine learning

methods. These studies strongly support the use of FLAIR-only systems for automated WML segmentation.

Brain atrophy has traditionally been quantified using T1-weighted images (Messina and Patti, 2014; Morgen et al., 2006), however, the nulled CSF signal makes FLAIR a suitable candidate for atrophy quantification. In (DiGregorio et al., 2021), it was shown that deep learning-based systems for intracranial volume (ICV) segmentation in multicentre FLAIR MRI achieved high segmentation performance as compared to manual annotations. In (Reiche et al., 2019) and (Khademi et al., 2020), it was also shown that ICV segmented, and intensity standardized FLAIR can be robustly thresholded to isolate the TBV and CSF tissues.

As FLAIR is routinely acquired, there is value in biomarker extraction algorithms that operate solely on this sequence for AD and CVD cohorts. Additionally, new studies are emerging that demonstrate other measurements from FLAIR could be a valuable addition to biomarker pipelines (Maillard et al., 2013) and measurements from retrospectively collected FLAIR could substantially improve the ability to detect treatment related differences (Maillard et al., 2013). Therefore, future works will continue to develop novel biomarkers for the FLAIR sequence and unified biomarker systems would be of value. Not only are integration hurdles reduced by extracting metrics from a single sequence, but also, measurement variability across biomarkers would be reduced as well. Each set of images and tools are subject to measurement variabilities and biases. If different sequences are used to compute biomarkers, the sources of variability are inherently different, making it difficult to elucidate small changes. Extracting features from the same sequence (and employing the same preprocessing) has the benefit of lower and more consistent measurement variability across biomarkers. Therefore, comparison across biomarkers and time points can be more reliable.

This work presents analysis of TBV, CSF, and WML volumetric biomarkers derived solely from neurological FLAIR MRI to establish FLAIR only cross-sectional and longitudinal benchmarks in large image cohorts of subjects with AD and CVD. Biomarkers are extracted using tools previously validated for multicentre FLAIR MRI. It is hypothesized that cross-sectional and longitudinal TBV, CSF, and WML volumes from CVD and AD FLAIR MRI databases can replicate previously demonstrated trends measured with standard MRI techniques.

## 2. Methods

Cerebral biomarkers were extracted using only FLAIR MRI volumes. Cross-sectional biomarkers included TBV, CSF volume, and WML volume and were calculated using the first time point of each patient's scans. All volume measurements were calculated in millilitres (mL or cm<sup>3</sup>). Longitudinal biomarkers included rates of change to extracted tissue volumes.

**Table 2**

Summary of longitudinal biomarker estimates from the literature. Estimates are shown as means  $\pm$  standard deviation (when reported).

Biomarker	Disease	Estimate	MRI Sequence	Reference
TBV Change	CN	-0.27%/year $\pm$ 0.15	T1	De Stefano et al. (2016)
	CN	-0.50%/year $\pm$ 0.22	T1	Scahill et al. (2003)
	CN	-0.45%/year	T1	Fotinos et al. (2005)
	CN	-7.35 mL/year	T1	Driscoll et al. (2009)
	CVD	-0.95%/year	T1	Seghier et al. (2014)
	MCI	-1.05%/year $\pm$ 0.92	T1	Evans et al. (2010)
	MCI	-0.98%/year	T1	Fotinos et al. (2005)
	MCI	-8.71 mL/year	T1	Driscoll et al. (2009)
	AD	-1.5%/year $\pm$ 0.92	T1	Evans et al. (2010)
	Ventricular CSF Change	CN	3.1%/year	T1
CN		1.31 mL/year	T1	Driscoll et al. (2009)
MCI		3.09 mL/year $\pm$ 2.72	T1	Evans et al. (2010)
		8.2%/year		
MCI		1.86 mL/year	T1	Driscoll et al. (2009)
AD		4.40 mL/year $\pm$ 2.99	T1	Evans et al. (2010)
WML Change	CN	0.8 mL/year $\pm$ 1.10	T1, FLAIR	Silbert et al. (2009)
	CN	0.23-1.33 mL/year	T2	Schmidt et al. (2016)
	CVD	1.59 mL/year	T1, FLAIR	Seghier et al. (2014)
		6.80%/year		
	CVD	3.96 mL/year	T2, DTI	Schmidt et al. (2016)
	MCI	2.4 mL/year $\pm$ 2.20	T1, FLAIR	Silbert et al. (2009)
	MCI	6.5%/year	T1, FLAIR	Silbert et al. (2012)
	AD	10%/year	T1, T2, FLAIR	Richard et al. (2010)
	AD	7-11.6%/year	T1, FLAIR	Stephen et al. (2019)

### 2.1. Experimental data

The three multicentre FLAIR MRI datasets used in this work are summarized in Table 1. The first data repository was from the Canadian Atherosclerosis Imaging Network (CAIN) (Tardif et al., 2013), a pan-Canadian vascular disease study. CAIN contains 400 unique subjects with one to two follow-up scans totalling 869 imaging volumes (42 850 image slices). The mean amount of elapsed time between scans for CAIN patients was 1.00 years  $\pm$  0.08. The CAIN dataset was mainly used to analyze the effect of demographic factors on biomarkers and the mean age is 71.0 years  $\pm$  7.9. The second dataset was from the Alzheimer's Disease Neuroimaging Initiative (ADNI), one of the largest open repositories for the study of AD and dementia (Jack et al., 2008a). ADNI contains 900 unique subjects with a mean age of 73.5 years  $\pm$  7.4 with one to five follow-up scans for 4102 imaging volumes (150 000 image slices). The mean amount of elapsed time between scans for ADNI patients was 0.69 years  $\pm$  0.41. In ADNI, there were cognitive diagnosis labels for 1270 cognitively normal (CN) scans, 2296 early or late MCI scans, and 536 AD scans. The third dataset was from the Canadian Consortium on Neurodegeneration in Aging (CCNA), a pan-Canadian study investigating different types of dementia (Chertkow et al., 2019; Mohaddes et al., 2018). The CCNA subset used in this work contains cross-sectional scans for 50 subjective cognitive impairment (SCI) subjects, 98 MCI subjects, 63 vascular MCI (V-MCI subjects), and 43 AD subjects. The mean age in CCNA is 73.2 years  $\pm$  6.9. Diagnosis for both CCNA and ADNI were obtained through a battery of clinical instruments (Jack et al., 2008a; Mohaddes et al., 2018). The distribution of patients across acquisition centers for each site is shown graphically in the Appendix (Fig. F).

### 2.2. Biomarker extraction

FLAIR MRI volumes were first preprocessed using bias field correction and intensity standardization (Reiche et al., 2019). Intensity standardization normalizes and aligns intensity ranges for each of the prominent tissue classes, which eases segmentation and improves

generalization across datasets (Reiche et al., 2019). The intracranial volume (ICV) was segmented using a MultiResUNet convolutional neural network (CNN) architecture developed for FLAIR MRI in (DiGregorio et al., 2021). This MultiResUNet was trained, validated, and tested using 125 CAIN, 25 ADNI, and 25 CCNA volumes with gold standard ground truth labels (DiGregorio et al., 2021). Dice loss, Adam optimizer, and a batch size of 16 over 50 epochs was used in the MultiResUNet. WML were segmented from the standardized and ICV segmented images using a skip connection U-Net (SC U-Net) CNN designed in (Khademi et al., 2021) for FLAIR MRI. The SC U-Net for lesion segmentation was also developed using subsets of CAIN, ADNI, and CCNA volumes with gold-standard lesion labels. A generalized dice loss, Adam optimizer, and a batch size of 64 over 75 epochs was utilized for the SC U-Net. The volumes used for model training are distinct from those analyzed in this work and both automated tools have been extensively tested in terms of their ability to generalize to unseen data (DiGregorio et al., 2021; Khademi et al., 2021).

To compute CSF and TBV, the ICV was segmented by thresholding the intensity standardized images (Reiche et al., 2019). Since the histograms of imaging volumes were aligned via standardization, the same threshold was applied over all images. TBV, CSF, and WML volumes were calculated by summing the number of pixels in each class and multiplying by voxel dimensions. To compensate for pre-morbid brain size, gender differences, and inter-subject head size variation (Hansen et al., 2015; Nordenskjöld et al., 2013), ICV was used to normalize volumetric biomarkers (Hansen et al., 2015; Nordenskjöld et al., 2013).

For longitudinal biomarkers, the normalized, annual volume changes were computed in %/year:

$$\Delta vol \% = \left( \frac{vol_2 - vol_1}{vol_1} \right) \times 100$$

where  $vol_1$  is the volume of a structure at time point #1,  $vol_2$  is the volume of the same structure at time point #2, and  $TBS$  is the time between scans in years. The average rate of change between pairs of consecutive scans was computed for subjects with longitudinal scans.

### 2.3. Statistical analysis

Biomarkers extracted from all cohorts were first summarized with descriptive statistics (means (SD), medians (IQR), and frequencies as counts (%)). Pearson's correlation coefficient was used to investigate the relationship between cross-sectional biomarkers and subject age for all subjects, men, and women to observe the impact of sex. To quantitatively examine differences in biomarker means across groups (i.e., disease, age), ANCOVA was used for cross-sectional and longitudinal biomarkers. Normality of residuals were visualized for each ANCOVA model and homogeneity in variance was assessed using Levene's test (Table B,C,D). Covariates (i.e., sex, center, scanner) were included where appropriate. For instances where ANCOVA tests were significant, Tukey's Honest Significant Difference (HSD) post-hoc tests, which corrects for multiple comparisons, were used to examine the sources of difference. Statistical significance was considered for  $\alpha = 0.05$ .

### 2.4. Biomarker comparisons

FLAIR biomarkers were analyzed to investigate whether CVD and AD-dementia disease trends were comparable to similar studies using standard MRI techniques. The focus is on comparing to literature estimates for longitudinal biomarkers as inter-study differences between cross-sectional biomarkers are likely cohort related. Studies with similar characteristics to the clinical datasets used in this work were mainly considered. Namely, research that included measuring changes to TBV, CSF, or WML volumes in healthy, ischemic disease, MCI, or AD cohorts were used to establish biomarker benchmarks. All studies had large sample sizes ( $n > 40$  patients), spanned multiple age decades, and had

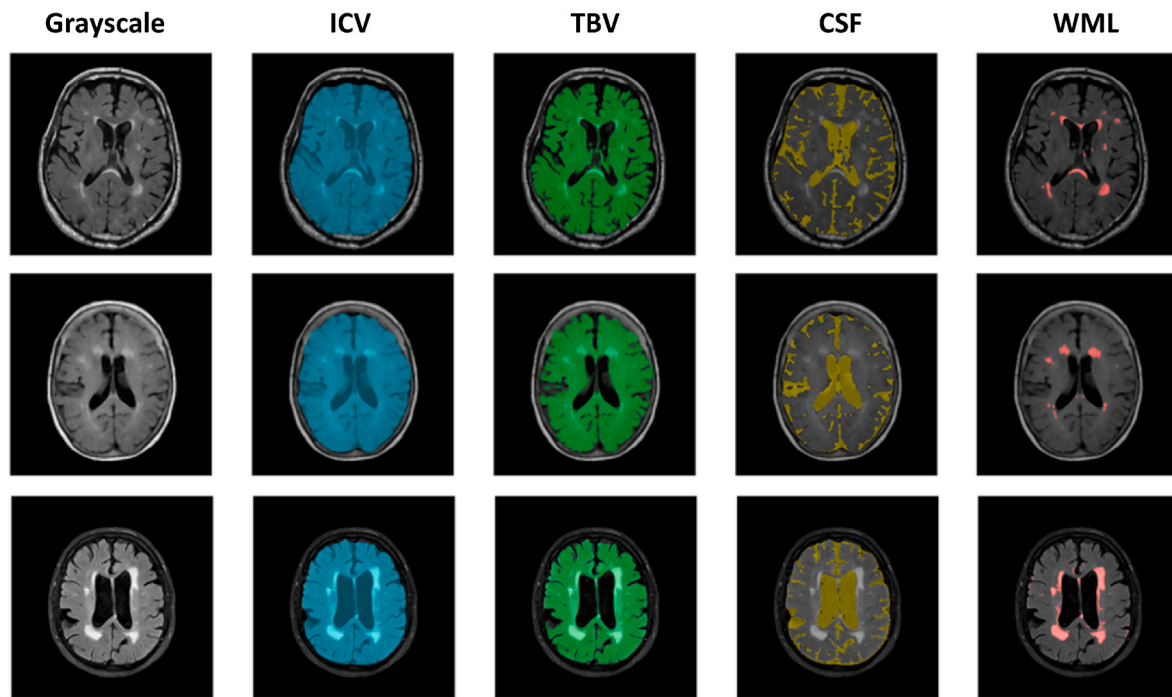


Fig. 1. ICV, TBV, CSF and WML segmentations for volume slices from CAIN (top), ADNI (middle), and CCNA (bottom).

Table 3

Mean ( $\pm$ SD) cross-sectional biomarkers for all disease classifications in each dataset. Raw biomarkers (mL) and ICV normalized biomarkers (%) are shown for each tissue.

Dataset	Disease	ICV (mL)	TBV (mL)	CSF (mL)	WML (mL)	TBV (%)	CSF (%)	WML (%)
CAIN	Ischemic Disease	1362 $\pm$ 144.6	1153 $\pm$ 136.1	211.6 $\pm$ 77.18	11.05 $\pm$ 11.44	85.06 $\pm$ 5.77	14.93 $\pm$ 5.77	0.77 $\pm$ 0.79
ADNI	CN	1345 $\pm$ 149.4	1094 $\pm$ 122.4	259.7 $\pm$ 63.57	7.99 $\pm$ 11.41	81.22 $\pm$ 4.14	18.77 $\pm$ 4.14	0.57 $\pm$ 0.80
	MCI	1371 $\pm$ 145.7	1101 $\pm$ 132.3	269.6 $\pm$ 72.27	8.72 $\pm$ 9.64	80.70 $\pm$ 4.64	19.29 $\pm$ 4.64	0.61 $\pm$ 0.69
	AD	1344 $\pm$ 158.2	1028 $\pm$ 138.6	325.0 $\pm$ 68.95	10.84 $\pm$ 9.82	76.56 $\pm$ 4.53	23.43 $\pm$ 4.53	0.82 $\pm$ 0.76
CCNA	SCI	1346 $\pm$ 118.7	1153 $\pm$ 96.05	193.8 $\pm$ 48.53	6.21 $\pm$ 5.00	85.69 $\pm$ 2.92	14.31 $\pm$ 2.92	0.46 $\pm$ 0.41
	MCI	1378 $\pm$ 133.1	1155 $\pm$ 116.5	222.3 $\pm$ 52.98	6.12 $\pm$ 6.94	83.90 $\pm$ 3.33	16.09 $\pm$ 3.33	0.44 $\pm$ 0.49
	V-MCI	1428 $\pm$ 147.0	1161 $\pm$ 127.8	266.4 $\pm$ 70.47	25.92 $\pm$ 20.52	81.41 $\pm$ 4.33	18.58 $\pm$ 4.33	1.78 $\pm$ 1.35
	AD	1373 $\pm$ 146.7	1088 $\pm$ 114.6	285.1 $\pm$ 73.5	7.07 $\pm$ 4.33	79.36 $\pm$ 4.31	20.63 $\pm$ 4.31	0.51 $\pm$ 0.30

Table 4

Mean ( $\pm$ SD) longitudinal biomarkers for all disease classifications in each dataset.

Dataset	Disease	TBV (mL/year)	CSF (mL/year)	WML (mL/year)	TBV Atrophy (%/year)	CSF Expansion (%/year)	WML Expansion (%/year)
CAIN	Ischemic Disease	-8.07 $\pm$ 11.50	6.03 $\pm$ 8.48	0.68 $\pm$ 1.67	-0.69 $\pm$ 1.01	3.66 $\pm$ 5.46	7.07 $\pm$ 13.46
ADNI	CN	-8.13 $\pm$ 12.59	5.27 $\pm$ 11.59	0.42 $\pm$ 1.27	-0.72 $\pm$ 1.15	2.28 $\pm$ 4.67	5.65 $\pm$ 15.41
	MCI	-11.71 $\pm$ 15.54	8.06 $\pm$ 12.74	0.65 $\pm$ 1.77	-1.01 $\pm$ 1.67	3.23 $\pm$ 5.31	10.97 $\pm$ 24.45
	AD	-19.11 $\pm$ 17.09	11.57 $\pm$ 16.13	1.07 $\pm$ 2.41	-1.84 $\pm$ 1.69	3.95 $\pm$ 5.46	11.54 $\pm$ 22.85

male and female representation. All studies considered used semi- or fully automated techniques for biomarker extraction to enable fair comparison. Table 2 summarizes the literature estimates used to compare to the longitudinal FLAIR cohorts and biomarkers in this work. Majority of the literature for TBV, and CSF volume changes utilized T1-weighted sequences, and CSF analysis was mainly for ventricular CSF enlargement (compared to whole CSF compartment used in this work). Studies measuring changes to WML volumes relied on FLAIR data, though analysis utilized multi-modal inputs.

Estimated rates of TBV volume changes in healthy populations ranged from -0.27%/year (De Stefano et al., 2016) to -0.50%/year (Seahill et al., 2003). One study found a mean of -0.45%/year TBV change, which they compared to various literatures between -0.37%/year to -0.88%/year (Fotinos et al., 2005), while another reports -7.35 mL/year (Driscoll et al., 2009). Ischemic cohorts slightly

exceeded this range (Jouvent et al., 2010) with estimates including -0.95%/year (Seghier et al., 2014). For MCI and AD patients, estimates were as high as -1.05%/year and -1.5%/year respectively (Evans et al., 2010). There was less consensus for estimated measurements of CSF system enlargement, partially because of the variation between studies regarding how much of the ventricular system should be included (Evans et al., 2010). Estimates of ventricular expansion in healthy controls included 3.1%/year (Evans et al., 2010) and 1.31 mL/year (Driscoll et al., 2009). MCI cohort estimates included 3.09 mL/year, 8.2%/year (Evans et al., 2010) and 1.86 mL/year (Driscoll et al., 2009). For AD cohorts, estimates were generally the highest at 4.40 mL/year and 10.4%/year (Evans et al., 2010). WML volume changes in healthy cohorts have been estimated at various ranges with estimates of 0.8 mL/year (Silbert et al., 2009) and 0.23-1.33 mL/year (Schmidt et al., 2016). Estimates for ischemic cohorts were higher at 1.59 mL/year or

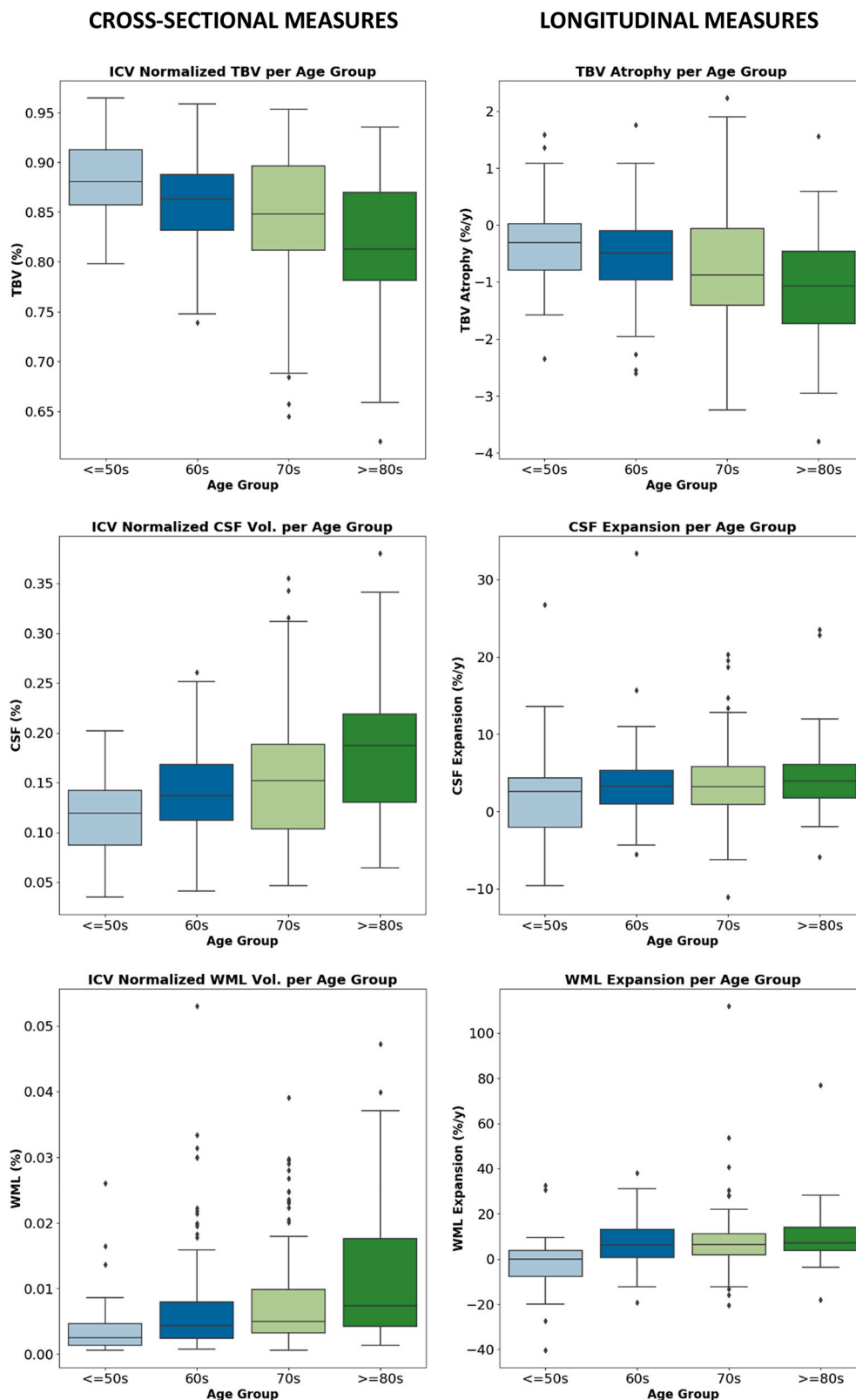


Fig. 2. Cross-sectional and longitudinal biomarker distributions for CAIN versus age group.

6.80%/year (Seghier et al., 2014) and 3.96 mL/year (Schmidt et al., 2016). Estimates for MCI cohorts included 6.5%/year (Silbert et al., 2012) and 2.4 mL/year (Silbert et al., 2009), while estimates for AD have exceeded 10%/year (Richard et al., 2010; Stephen et al., 2019). Differences in estimates are likely attributed to differences in disease

severity between studies, age differences across groups as well as biomarker measurement methodology (Chan et al., 2003; Jack et al., 2008b; Mungas et al., 2005; Scahill et al., 2002).

**Table 5**

ANCOVA results for cross-sectional and longitudinal biomarkers in CAIN between age groups with sex, scanner, and center as covariates. Reported as p-values ( $\alpha = 0.05$ ) where bold indicates significance. Post-hoc testing is reported as effect size (difference in means), p-value.

ANCOVA						
	Age Group	Sex	Scanner	Center		
TBV (%)	$p < 0.001$	$p = 0.002$	$p < 0.001$	$p < 0.001$		
CSF (%)	$p < 0.001$	$p = 0.002$	$p < 0.001$	$p < 0.001$		
WML (%)	$p < 0.001$	$p = 0.588$	$p = 0.342$	$p = 0.092$		
TBV Atrophy (%/y)	$p = 0.019$	$p = 0.713$	$p = 0.593$	$p = 0.081$		
CSF Expansion (%/y)	$p = 0.659$	$p = 0.373$	$p = 0.993$	$p = 0.055$		
WML Expansion (%/y)	$p = 0.024$	$p = 0.666$	$p = 0.136$	$p = 0.024$		
Post-hoc Testing						
	$\leq 50s$ vs $60s$	$\leq 50s$ vs $70s$	$\leq 50s$ vs $\geq 80s$	$60s$ vs $70s$	$60s$ vs $\geq 80s$	$70s$ vs $\geq 80s$
TBV (%)	2.62, $p = 0.079$	<b>3.77</b> , $p = 0.003$	<b>6.66</b> , $p < 0.001$	1.15, $p = 0.281$	<b>4.04</b> , $p < 0.001$	<b>2.89</b> , $p = 0.008$
CSF (%)	2.62, $p = 0.079$	<b>3.77</b> , $p = 0.003$	<b>6.66</b> , $p < 0.001$	1.15, $p = 0.281$	<b>4.04</b> , $p < 0.001$	<b>2.89</b> , $p = 0.008$
WML (%)	0.25, $p = 0.368$	0.35, $p = 0.093$	<b>0.80</b> , $p < 0.001$	0.10, $p = 0.652$	<b>0.55</b> , $p < 0.001$	<b>0.45</b> , $p = 0.021$
TBV Atrophy (%/y)	0.26, $p = 0.756$	0.37, $p = 0.513$	<b>0.81</b> , $p = 0.043$	0.11, $p = 0.883$	<b>0.55</b> , $p = 0.049$	0.44, $p = 0.109$
WML Expansion (%/y)	8.60, $p = 0.092$	8.98, $p = 0.055$	<b>11.66</b> , $p = 0.023$	0.38, $p = 0.900$	3.06, $p = 0.671$	2.68, $p = 0.705$

### 3. Results

#### 3.1. Biomarker extraction

Sample ICV, TBV, CSF and WML segmentations are shown in Fig. 1 for images from CAIN, ADNI, and CCNA, which shows accurate delineations across multicentre, multi-disease datasets. Mean  $\pm$  SD cross-sectional and longitudinal biomarkers over all datasets are shown in Tables 3 and 4, respectively. Mean biomarkers are also visualized for each cohort in Figs. A and B of the Appendix. Moderate standard deviations are likely attributed to the diversity of the cohorts and multitude of acquisition sites since the tools have been previously validated. The magnitude of standard deviation is also comparable to other works measuring cerebral biomarkers (Evans et al., 2010; Scahill et al., 2003; Stephen et al., 2019).

Cross-sectional biomarker trends from FLAIR MRI showed progression of neurodegenerative disease is associated with less TBV (WM and GM atrophy) and higher CSF volume (larger ventricles, more sub-arachnoid spaces). Specifically, TBV was lower in CCNA and ADNI over disease groups indicating more cerebral tissue loss for worse cognitive condition. CSF volumes followed a similar but opposite trend. For WML volumes, there was a slightly higher load between CN, MCI, and AD in ADNI, with similar trends in CCNA. CAIN subjects had moderate lesion loads that were slightly higher or comparable to subjects with AD. The largest WML load was for the V-MCI group ( $1.78\% \pm 1.35$ ). Within the ADNI cohort, AD subjects had higher normalized WML volumes compared to MCI ( $0.82\% \pm 0.76$  versus  $0.61\% \pm 0.69$ ). This was also found to a lesser extent in the CCNA cohort ( $0.51\% \pm 0.30$  versus  $0.44\% \pm 0.49$ ). Healthy subjects had the lowest mean WML volumes for ADNI and were comparable to MCI subjects in CCNA.

Longitudinal biomarker trends showed that ADNI subjects with AD experienced the highest rates of TBV change ( $-1.84\%/year \pm 1.69$ ) and CSF change ( $3.95\%/year \pm 5.46$ ) followed by MCI ( $-1.01\%/year \pm 1.67$ ,  $3.23\%/year \pm 5.31$ ) and CN ( $-0.72\%/year \pm 1.15$ ,  $2.28\%/year \pm 4.67$ ) subjects. In the CAIN dataset, longitudinal TBV biomarkers were comparable to non-impaired subjects from ADNI while longitudinal CSF biomarkers were comparable to MCI subjects from ADNI. With respect to WML volume change, AD ( $11.54\%/year \pm 22.85$ ) and MCI subjects ( $10.97\%/year \pm 24.45$ ) had the highest volume changes. Vascular disease subjects in CAIN had less WML expansion ( $7.07\%/year \pm 13.46$ ) but were still beyond the rate of change in non-impaired cohorts ( $5.65\%/year \pm 15.41$ ).

#### 3.2. CAIN analysis

Analysis on the CAIN dataset focused on biomarker differences

between age groups. Age groups were defined by step sizes of 10 years and differences in biomarker means between age groups were analyzed using ANCOVA. To balance sample sizes, subjects in their 50s or younger, along with subjects in their 80s or older were grouped into single categories.

Cross-sectional and longitudinal biomarkers per age group are summarized in Table A (Appendix) and biomarker distributions are shown in Fig. 2. TBV decreased for increasing age while TBV atrophy increased with each progressive decade. The highest rate of TBV atrophy was for subjects in their 80s or older ( $-1.13\%/year$ ). CSF volumes also increased as a function of age, with the highest rate of CSF expansion in oldest age group ( $4.79\%/year$ ). WML volumes also increased with age, with the largest increase in mean lesion load between the 70s and 80+ age groups. This trend was also observed for longitudinal WML biomarkers where rates of expansion were  $7.39\%/year$  and  $10.07\%/year$  for subjects in their 70s and 80s + respectively.

Results of ANCOVA testing across age groups (adjusting for sex, scanner, and center) showed significant differences for all three cross-sectional biomarkers ( $p < 0.05$ ), with all covariates having a significant effect for TBV and CSF volume. Post-hoc testing in Table 5 shows that TBV and CSF volumes had significant differences ( $p < 0.01$ ) between all age groups outside of the  $\leq 50s$  versus  $60s$  ( $p = 0.079$ ) and  $60s$  versus  $70s$  ( $p = 0.281$ ) comparisons. WML volume had significant differences between all groups except for the  $\leq 50s$  versus  $60s$  ( $p = 0.368$ ),  $\leq 50s$  versus  $70s$  ( $p = 0.093$ ), and  $60s$  versus  $70s$  ( $p = 0.652$ ) comparisons. For longitudinal biomarkers, significant differences ( $p < 0.05$ ) were found between age groups for rates of TBV and WML volume change with center also having an effect for WML expansion. Post-hoc testing for TBV atrophy revealed differences between  $\leq 50s$  versus  $\geq 80s$  years of age ( $p = 0.043$ ) and  $60s$  versus  $\geq 80s$  years of age ( $p = 0.049$ ). For WML expansion, significant differences only existed for the  $\leq 50s$  versus  $\geq 80s$  age group comparison ( $p = 0.023$ ).

Correlation plots between subject age and cross-sectional biomarkers are shown in Fig. C (Appendix). Significant ( $p < 0.05$ ) negative correlations were found between TBV and age for all subjects ( $r = -0.29$ ), men ( $r = -0.32$ ), and women ( $r = -0.26$ ). The same trends were observed between CSF volumes and age but with positive correlations. Significant ( $p < 0.05$ ) positive correlations were found between WML volume and age for all subjects ( $r = 0.22$ ), men ( $r = 0.18$ ), and women ( $r = 0.29$ ).

#### 3.3. ADNI analysis

In the ADNI cohort, biomarkers were analyzed as a function of cognitive diagnosis along with correlation analysis between subject age and cross-sectional biomarkers. The distribution of cross-sectional and

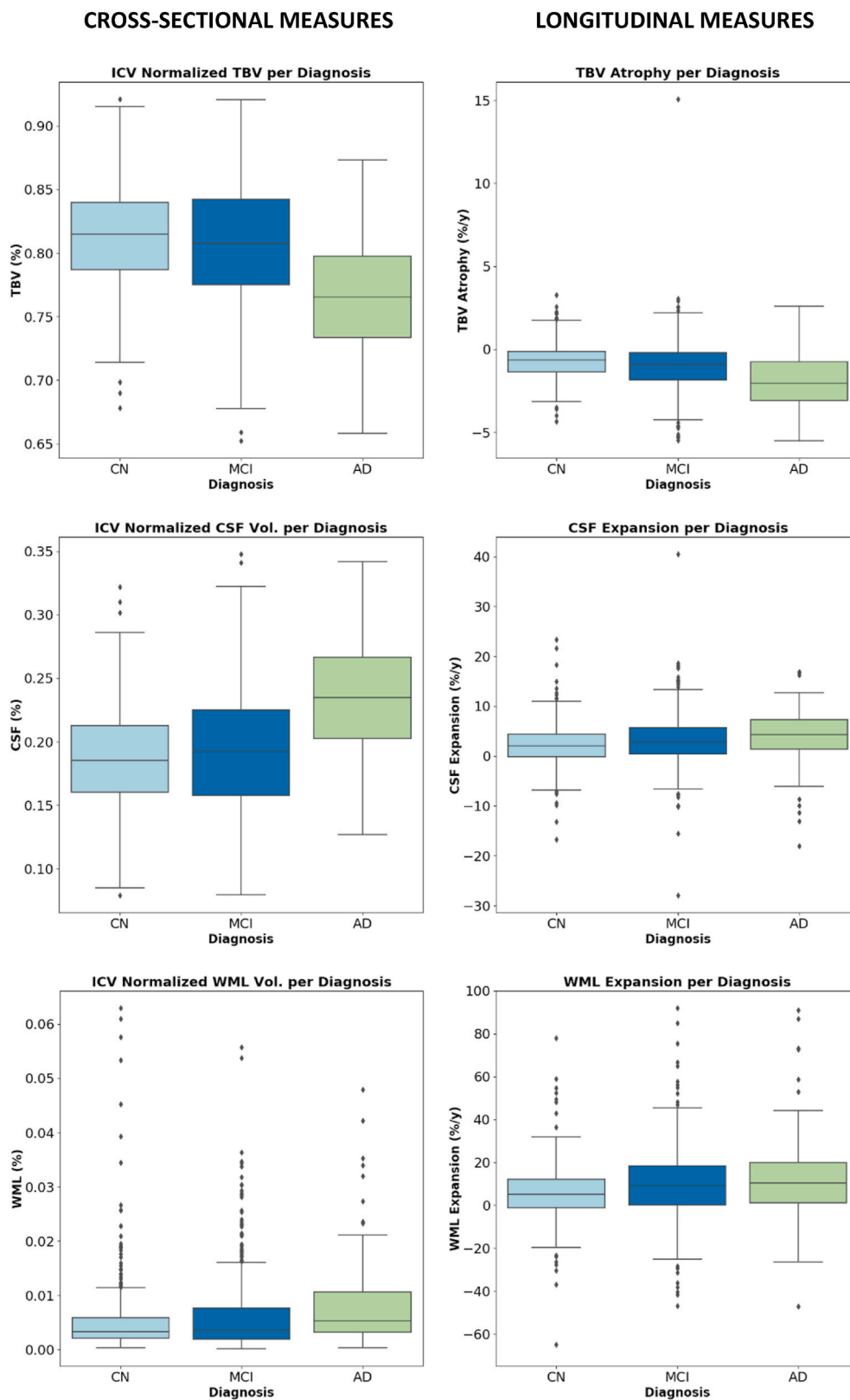


Fig. 3. Distributions of cross-sectional and longitudinal biomarkers in ADNI versus disease classification.

longitudinal biomarker measurements as a function of cognitive label can be seen in Fig. 3. With increasing cognitive impairment, TBV decreased while WML and CSF volumes increased. Similarly, for the longitudinal biomarkers, the rates of atrophy, CSF expansion and WML expansion increased with worse cognitive outcome.

ANCOVA analysis (with age, sex, scanner, and center as covariates) in Table 6 showed significant differences ( $p < 0.05$ ) between disease classifications for all cross-sectional biomarkers, with age, sex, and scanner typically having an effect. Post-hoc testing of TBV, CSF, and WML volume showed significant differences ( $p < 0.01$ ) when AD was

**Table 6**

ANCOVA results for cross-sectional and longitudinal biomarkers in ADNI between disease categories with age, sex, scanner, and center as covariates. Reported as p-values ( $\alpha = 0.05$ ) where bold indicates significance. Post-hoc testing is reported as effect size (difference in means), p-value.

ANCOVA					
	Disease	Age	Sex	Scanner	Center
TBV (%)	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.464$
CSF (%)	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.464$
WML (%)	$p = 0.006$	$p < 0.001$	$p = 0.107$	$p = 0.002$	$p = 0.975$
TBV Atrophy (%/y)	$p < 0.001$	$p = 0.022$	$p = 0.417$	$p = 0.002$	$p = 0.989$
CSF Expansion (%/y)	$p = 0.003$	$p = 0.929$	$p = 0.112$	$p = 0.071$	$p = 0.550$
WML Expansion (%/y)	$p = 0.003$	$p = 0.378$	$p = 0.297$	$p = 0.029$	$p = 0.838$
Post-hoc Testing					
	CN vs MCI	CN vs AD	MCI vs AD		
TBV (%)	0.52, $p = 0.211$	<b>4.6, <math>p &lt; 0.001</math></b>	<b>4.1, <math>p &lt; 0.001</math></b>		
CSF (%)	0.52, $p = 0.211$	<b>4.6, <math>p &lt; 0.001</math></b>	<b>4.1, <math>p &lt; 0.001</math></b>		
WML (%)	0.03, $p = 0.772$	<b>0.24, <math>p = 0.001</math></b>	<b>0.21, <math>p = 0.004</math></b>		
TBV Atrophy (%/y)	<b>0.285, <math>p = 0.024</math></b>	<b>1.12, <math>p &lt; 0.001</math></b>	<b>0.83, <math>p &lt; 0.001</math></b>		
CSF Expansion (%/y)	0.95, $p = 0.077$	<b>1.67, <math>p = 0.003</math></b>	0.72, $p = 0.185$		
WML Expansion (%/y)	<b>5.33, <math>p = 0.004</math></b>	<b>5.89, <math>p = 0.003</math></b>	0.56, $p = 0.900$		

one of the cognitive categories. Disease was found to have an effect ( $p < 0.05$ ) for all longitudinal biomarkers, with age and scanner often having effects. Post-hoc testing showed significant differences ( $p < 0.01$ ) between all classifications with respect to atrophy. Regarding CSF expansion, post-hoc testing only showed significant differences for CN versus AD ( $p = 0.003$ ). Post-hoc testing of WML expansion found differences for all comparisons except for MCI versus AD ( $p = 0.900$ ). For all biomarkers, the acquisition center did not have an effect.

Correlation plots between subject age and cross-sectional biomarkers can be found in Fig. D (Appendix). Significant ( $p < 0.05$ ) negative correlations were found between TBV and age for all subjects ( $r = -0.62$ ), men ( $r = -0.66$ ), and women ( $r = -0.57$ ). The same trends were observed between CSF volumes and age but with positive correlations. Significant ( $p < 0.05$ ) positive correlations were found between WML volume and age for all subjects ( $r = 0.38$ ), men ( $r = 0.36$ ), and women ( $r = 0.42$ ).

### 3.4. CCNA analysis

Analysis of CCNA included ANCOVA to examine differences in biomarker means across disease categories and correlation analysis with subject age. The distribution of cross-sectional biomarker measurements as a function of disease classification can be seen in Fig. 4. There was decreasing TBV volume with worse cognitive diagnosis with a corresponding increase in WML and CSF volumes. The TBV and CSF volumes for V-MCI on average had intermediate values between MCI and AD. However, the WML loads were highest in the V-MCI group.

ANCOVA results between disease groups (with age, sex, scanner, and center as covariates) are summarized in Table 7. Significant differences ( $p < 0.05$ ) were found between disease groups for all cross-sectional biomarkers, with age, sex, and scanner also having significant effects (except for WML volume). Post-hoc testing of TBV and CSF volume showed significant differences ( $p < 0.01$ ) between all cognitive diagnoses. For WML volume, post-hoc testing only revealed differences ( $p < 0.01$ ) when V-MCI was one of the groups.

Correlation plots between subject age and cross-sectional biomarkers can be found in Fig. E of the Appendix. Significant ( $p < 0.05$ ) negative correlations were found between TBV and age for all subjects ( $r = -$

0.56), men ( $r = -0.55$ ), and women ( $r = -0.53$ ). The same trends were observed for correlations between CSF volumes and age but with positive correlations. Significant ( $p < 0.05$ ) positive correlations were found between WML volume and age for all subjects ( $r = 0.40$ ), men ( $r = 0.39$ ), and women ( $r = 0.40$ ).

### 3.5. Biomarker comparison

Table 8 contains the literature estimates for TBV and WML change, as well as those found by the FLAIR-only biomarkers in this work. Note that literature estimates were only found for ventricle (CSF) change which does not apply to the whole CSF compartment measurement in this work, so only TBV and WML change were compared. Trends in the longitudinal biomarkers (see Table 4, Fig. B) were comparable to those found in the literature (see Section 2.4). For instance, in (Seghier et al., 2014), a T1-based study, TBV atrophy and WML expansion in CVD populations were reported at  $-0.95\%/year$  and  $6.8\%/year$ , respectively. In the CVD dataset (CAIN) analyzed in this work, the mean annual rates of TBV atrophy and WML expansion were measured to be  $-0.69\%/year$  and  $7.07\%/year$ , respectively. Another study utilizing T1 and FLAIR sequences to measure WML expansion reported rates exceeding  $10\%/year$  in AD patients (Stephen et al., 2019). In this work, findings were similar with WML expansion rates of  $10.9\%/year$  for MCI and  $11.54\%/year$  for AD in the ADNI cohort. For atrophy comparisons, TBV atrophy in the CN cohort was slightly higher than the range reported in the literature for T1 ( $-0.72\%/year$  versus  $-0.50\%/year$ ) and were similar for the MCI ( $-1.01\%/year$  versus  $-1.05\%/year$ ) and AD ( $-1.84\%/year$  versus  $-1.5\%/year$ ) groups. Note that for the CAIN dataset estimates were on the lower side of the estimates for TBV and this may be due to the fact that the cohort was largely cognitively intact.

## 4. Discussion

This is one of the first works that extracts and analyzes cross-sectional and longitudinal TBV, CSF, and WML biomarkers using multicentre, FLAIR-only data for AD and CVD cohorts. The tools to segment objects and measure volumetric biomarkers have been previously developed and validated on multicentre FLAIR MRI datasets (DiGregorio et al., 2021; Khademi et al., 2020; Reiche et al., 2019). It was hypothesized that biomarkers extracted from FLAIR MRI could characterize longitudinal changes in TBV, CSF, and WML volume in AD and CVD populations similarly to common MR techniques. Comparison to literature estimates for TBV and WML change show that there are similarities between multimodal or T1 approaches, and those obtained with the FLAIR sequence in this work. Variation in estimates come from cohort differences, differences in mean age/age ranges of the cohorts and differences in biomarker extraction tools and methodologies. The cohorts are large, highly diverse, and come from several acquisition sites with unique imaging parameters. The age ranges varied quite a bit for the cohorts, including CAIN that had subjects in their 40s all the way until their 90s. Similarly, ADNI was 50–90+ years of age and CCNA had subjects between 60 and 80+ years of age. Variability in patient age can cause variability in the measurements and can create challenges for directly comparing to results from other cohorts with different ages, since neurodegeneration manifests differently depending on the age.

Also note some of the literature estimates have slightly higher WML yearly estimates (in mL/year) and lower % changes for some of the groups (CVD, MCI). Although much of this is likely related to cohort differences, some may be related to the fact that using other modalities for WML quantification is more susceptible to false positives which can inflate volume estimates (whereas FLAIR alone is shown to reduce FP rates (Narayana et al., 2020)). Despite variations in literature estimates, multiple studies have shown the same trend of increased TBV atrophy, CSF expansion, and WML expansion in AD subjects compared to healthy controls, with MCI groups having intermediate measurements (Dickie et al., 2016a,b; Evans et al., 2010; Jouvent et al., 2010). Ischemic disease



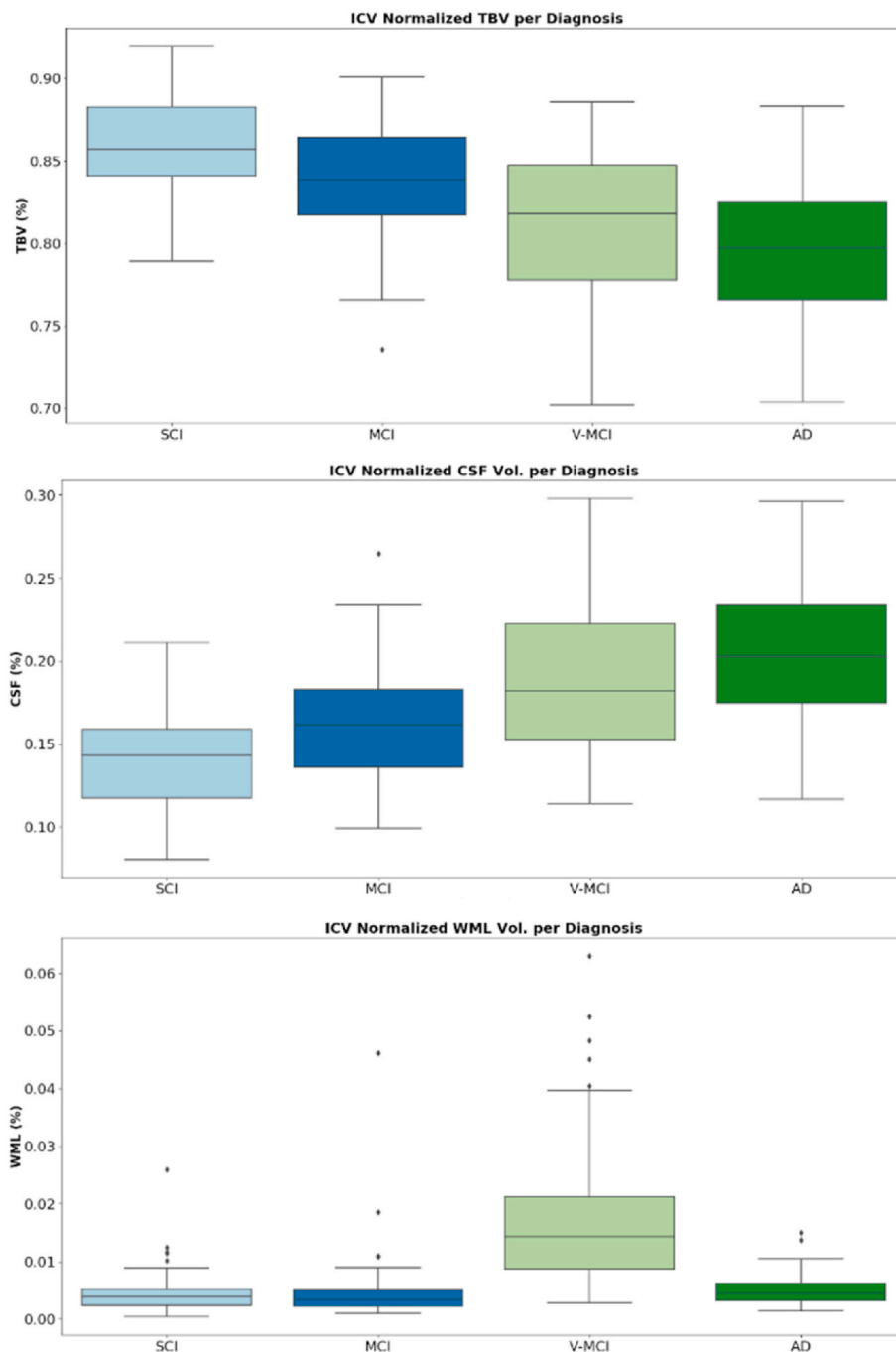


Fig. 4. Distributions of cross-sectional biomarkers in CCNA versus disease classification.

cohorts have been characterized by greater lesion presence and accelerated brain tissue loss following a vascular event compared to healthy controls (Seghier et al., 2014). This concurred with our ANCOVA testing which revealed all biomarkers to be statistically different between cohorts while controlling for demographic (age, sex) and acquisition factors (scanner, center). Rates of TBV atrophy, CSF expansion, and WML expansion all increased with worsening cognitive classifications and post-hoc testing consistently differentiated between healthy and AD patients. Our results suggest that FLAIR can independently be used to obtain longitudinal biomarkers from CVD and AD cohorts.

FLAIR-derived biomarkers also provided insight into AD and CVD disease characterization. For correlations between TBV and CSF with age, stronger correlation coefficients were reported for men. This may indicate a more prominent rate atrophy in men compared to women

which concurs with the literature (Chan et al., 2003). Correlations between age and TBV were also stronger in the two dementia cohorts (ADNI and CCNA) than in the ischemic cohort (CAIN). This may be attributed to the increased rate of cerebral atrophy associated with AD-type dementias. Cross-sectional WML volumes were the highest in the CAIN and CCNA V-MCI cohorts (see Fig. A), surpassing AD subjects in ADNI and CCNA. This aligns with other studies (Evans et al., 2010) implying that WML volume is more closely related to vascular disease and that atrophy biomarkers may be stronger predictors of AD-type dementia in cross-sectional studies. However, rates of WML expansion were higher in MCI and AD cohorts than in patients with vascular disease (CAIN, V-MCI). It has also been found that there is a significant increase in WML progression up to a decade before the onset of MCI (Silbert et al., 2009). This implies that analysis of WML may be a better

**Table 7**

ANCOVA results for cross-sectional biomarkers in CCNA between disease categories with age, sex, scanner, and center as covariates. Reported as p-values ( $\alpha = 0.05$ ) where bold indicates significance. Post-hoc testing is reported as effect size (difference in means), p-value.

ANCOVA						
	Disease	Age	Sex	Scanner	Center	
TBV (%)	$p < 0.001$	$p < 0.001$	$p = 0.004$	$p = 0.003$	$p = 0.336$	
CSF (%)	$p < 0.001$	$p < 0.001$	$p = 0.004$	$p = 0.003$	$p = 0.336$	
WML (%)	$p = 0.009$	$p < 0.001$	$p = 0.595$	$p = 0.250$	$p = 0.919$	
Post-hoc Testing						
	SCI vs MCI	SCI vs AD	SCI vs AD	MCI vs V-MCI	MCI vs AD	V-MCI vs AD
TBV (%)	<b>1.79, p = 0.029</b>	<b>4.55, p &lt; 0.001</b>	<b>6.33, p &lt; 0.001</b>	<b>2.49, p &lt; 0.001</b>	<b>4.54, p &lt; 0.001</b>	<b>2.05, p = 0.028</b>
CSF (%)	<b>1.78, p = 0.029</b>	<b>4.27, p &lt; 0.001</b>	<b>6.32, p &lt; 0.001</b>	<b>2.49, p = 0.009</b>	<b>4.54, p &lt; 0.001</b>	<b>2.05, p = 0.028</b>
WML (%)	0.02, p = 0.900	<b>1.32, p &lt; 0.001</b>	0.05, p = 0.900	<b>1.34, p &lt; 0.001</b>	0.07, p = 0.900	<b>1.27, p &lt; 0.001</b>

**Table 8**

Comparison between literature estimates and mean FLAIR measurements for TBV and WML volume change.

Biomarker	Disease	Literature Estimates	Our Estimate	Dataset
TBV Change	CVD	-0.95%/year	-0.69%/year	CAIN
	CN	-0.27%/year to -0.50%/year	-0.72%/year	ADNI
	MCI	-7.35 mL/year	-8.13 mL/year	ADNI
	AD	-0.98%/year to -1.05%/year	-1.01%/year	ADNI
WML Change	CVD	-8.71 mL/year	-11.71 mL/year	ADNI
	CN	-1.50%/year	-1.84%/year	ADNI
	MCI	6.80%/year	7.07%/year	CAIN
	AD	1.59-3.96 mL/year	0.68 mL/year	ADNI
	CN	0.23-1.33 mL/year	0.42 mL/year	ADNI
	MCI	6.5%/year	10.97%/year	ADNI
	AD	2.4 mL/year	0.65 mL/year	ADNI
	AD	7-11.6%/year	11.54%/year	ADNI

indicator of disease progression and can be used to determine the intervention window compared to cognitive testing. Also notable, was the fact that CSF expansion rates were higher in CAIN vascular disease subjects than in MCI (see Fig. B), which may indicate that vascular risk factors are also linked to atrophy. Recent evidence is pointing towards a “two-hit” vascular hypothesis for AD and vascular disease, where hit one includes CVD risk factors that lead to blood brain barrier dysfunction and reduced cerebral blood flow that precedes dementia, followed by hit two caused by an increase in beta-amyloid amplifying neuronal dysfunction, neurodegeneration and disease (Lamar et al., 2020). In (Meng et al., 2017), the authors suggest that vascular and neurodegenerative pathological processes have a supra-additive effect on cognitive performance. Perhaps these biomarkers are demonstrating this effect. In the future, further investigation into the differences (and similarities) across biomarker profiles for vascular disease and dementia will be conducted.

Deviations in biomarker measurements can likely be attributed to 1) differences in contrast mechanisms between FLAIR and T1 MRI sequences, 2) measurement error and 3) cohort differences. Imaging techniques and tissue characteristics differ between FLAIR and T1 which creates differences in volumetric biomarkers from the two sequences. For example, CSF spaces can appear smaller in FLAIR than in T1 since the CSF signal is nulled (Narayana et al., 2020; Soltanian-Zadeh and Peck, 2001). Since most studies utilize T1 data, there could be a bias towards T1-based measurements. Also, the slice thickness of FLAIR MRI is often higher than in T1, creating another source of measurement difference. Despite these differences, there is a strong similarity between FLAIR measurements found in this work, and those found in the

literature. As with any automated biomarker study, measurement errors are a source of limitation. First, if there is insufficient time between follow-up scans it may be difficult to elucidate subtle changes to tissue volumes (Narayanan et al., 2020). Adequate time between scans (Narayanan et al., 2020) reports at least 1 year) is necessary for the anatomical changes to be larger than measurement error. If this is not the case, there can be positive TBV atrophy, or negative CSF expansion which has also been found in other studies using popular brain tissue measurement tools (Evans et al., 2010; Scahill et al., 2003; Storelli et al., 2018). These effects are less noticeable in mean trends especially for large datasets, such as the one used in this work. There is also a strong dependency on relative brain position in the scanner (Narayanan et al., 2020). Small changes to subject’s head location during follow-up scans can induce measurement error and dilute the ability to detect subtle changes. These challenges can be seen in Table A of the Appendix where a negative annual rate of change in WML volumes for subjects in their 50s and younger was found. Compared to the other age groups, mean WML volumes for this group are substantially lower, indicating smaller and possibly harder to detect lesions. Small lesion segmentation is a known challenge of WML segmentation algorithms and future works could aim at improving this aspect of performance. Small lesions in the deep WM or subcortical regions may have different etiological origins (DeBette and Markus, 2010; Wardlaw et al., 2015) and should not be missed. Additionally, the sample size in this cohort ( $\leq 50s$ ) was small making measurement error more noticeable. In the future, we plan to examine WML lesion loads in a normative population to analyze trends and these effects further. While sources of measurement error can dilute the signal, the biomarkers strongly resemble known trends across neurodegenerative and vascular disease populations supporting the use of FLAIR for WML, TBV and CSF volume measurement. There are also differences between the cohorts analyzed in this work, and those found in the literature. The mean age of the FLAIR datasets differed in comparison to many of the studies referenced, which could impact disease progression measurements, since older subjects are more susceptible to rapid aging.

There are significant advantages of using a single modality for multiple biomarkers. While measurement error is closely related to strength of biomarker extraction methodology, a single modality can help keep measurement variability in a specific range and better elucidate more subtle patterns between diseases and cognitive conditions. The intensity standardization (preprocessing) is the same for both ICV and WML segmentation, and U-Net (Ronneberger et al., 2015) variants were used to segment both tissues which have similar generalization capabilities and performance. Therefore, the measurement variability is similar across biomarkers and could lead to better comparisons. Technically, the barriers to integrating AI tools in clinical workflows would be less for a single sequence as well. FLAIR MRI is routinely used, and such tools can automatically measure biomarkers and populate reports while providing visualization of biomarker values as compared to longitudinal and cross-sectional trends from the population. Future work may include deep-learning based TBV and CSF biomarker detection, enhanced lesion segmentation for fine lesions and analysis that examines the relationship of the proposed cross-sectional and longitudinal FLAIR-MRI biomarkers with respect to vascular risk factors and other clinical data. To further boost the validity of FLAIR-derived AD and CVD biomarkers, future works could include the implementation of a predictive classifier that uses biomarkers as features for estimating disease or cognitive group. Perhaps such analysis could uncover neuroanatomical profiles that differentiate between vascular and AD-type neurodegeneration in the brain, such that patients can be stratified into homogeneous groups for clinical trials, or for mechanistic understanding of disease etiology.

## 5. Conclusion

In this work, algorithms designed for neurological FLAIR MRI were

used to measure cross-sectional and longitudinal biomarkers in large multicenter, AD and CVD datasets. Biomarkers were related to volumes of the GM and WM (TBV), CSF spaces, and WML and results aligned with known trends typically gleaned from T1-based analysis. There were significant differences across cognitive diagnosis groups for the dementia cohorts (ADNI and CCNA) and differences across age groups in CAIN for cross-sectional and longitudinal biomarkers while controlling for demographic and acquisition variables. Assuming the availability of validated biomarker extraction methodology, these results demonstrate that FLAIR alone can be used for end-to-end analysis of AD and CVD datasets, which can reduce acquisition costs, ease clinical translation, and reduce measurement error associated with multi-modal analysis approaches.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

- We would like to thank and acknowledge, Dr. Sandra E. Black,

Neurologist and Professor of Medicine at the University of Toronto, for her contributions to this research as Principal Investigator (PI) of CAIN, and, for providing feedback and guidance on the work and data.

- We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) through the NSERC Discovery Grant program.

- The Canadian Atherosclerosis Imaging Network (CAIN) was established through funding from a Canadian Institutes of Health Research Team Grant for Clinical Research Initiatives (CIHR-CRI 88057). Funding for the infrastructure was received from the Canada Foundation for Innovation (CFI-CAIN 20099), with matching funds provided by the governments of [H]ta, Ontario, and Quebec.

- Data collection and sharing for this project was partially funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defence award number W81XWH-12-2-0012).

- We acknowledge the support from the Canadian Consortium on Neurodegeneration and Aging (CCNA) for data from the COMPASS-ND study, data storage, and management and standardized MRI protocols.

**Appendix**

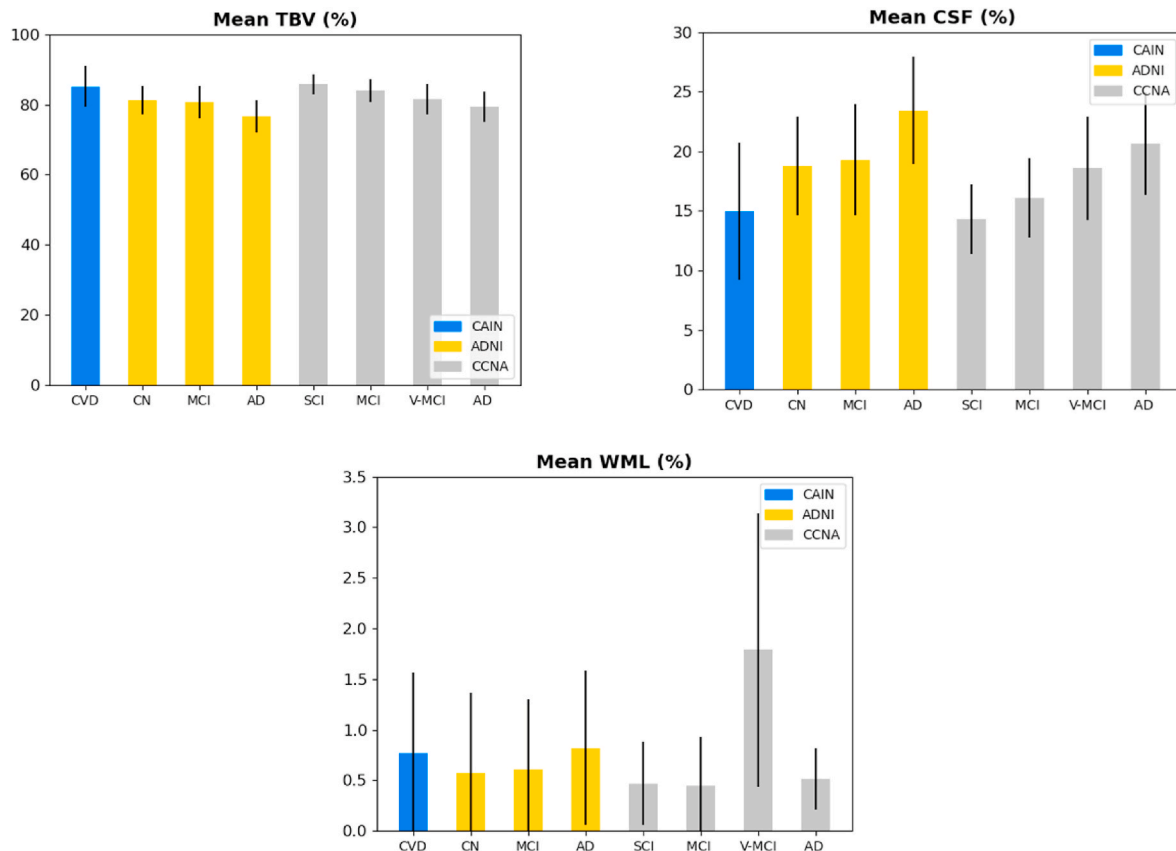


Fig. A. Bar plots of mean cross-sectional biomarkers per database and cohort with standard deviation shown as error bars.

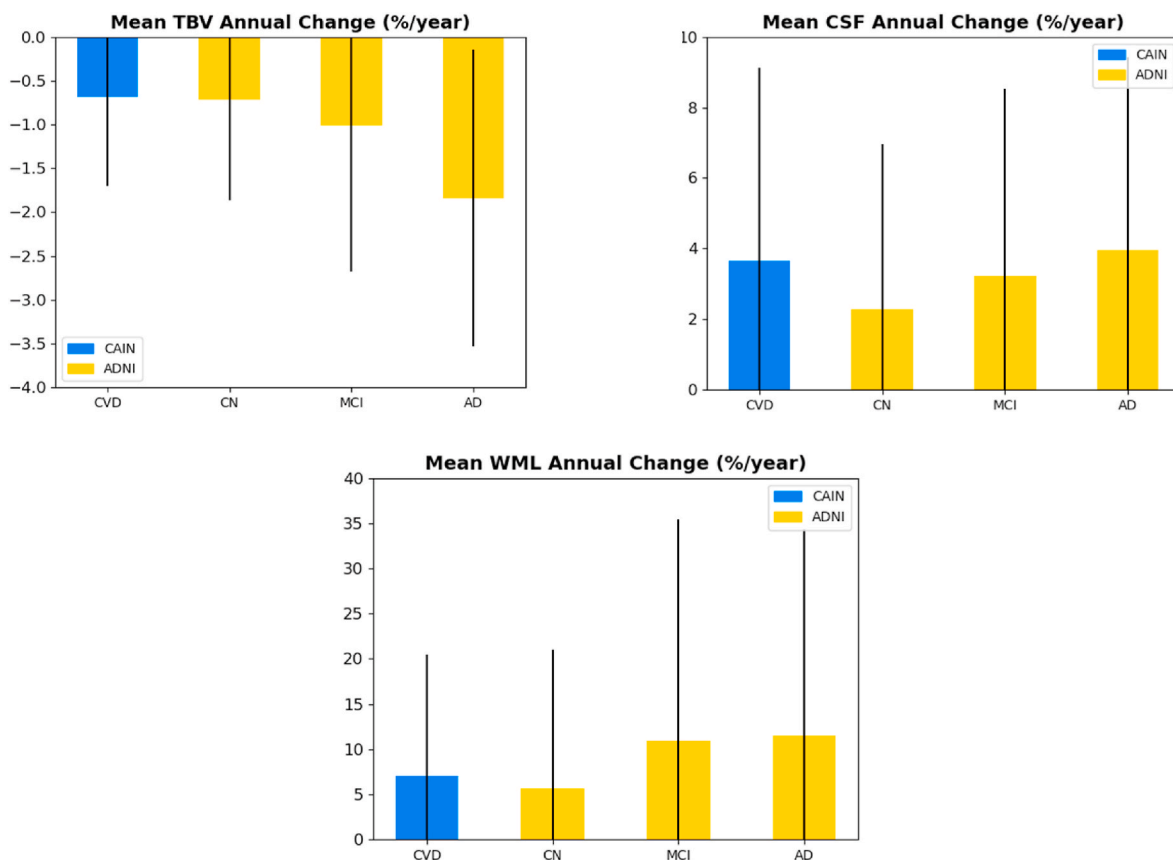


Fig. B. Bar plots of mean longitudinal biomarkers per database and cohort with standard deviation shown as error bars.

Table A

Summary of ICV normalized CAIN biomarkers organized by age group for all subjects, men, and women. Shown as mean ± standard deviation.

	≤50s			60s			70s			≥80s		
	All	M	F	All	M	F	All	M	F	All	M	F
# Subjects	32	17	15	135	86	49	180	108	72	53	32	21
TBV (%)	88.84 ± 4.05	88.26 ± 4.32	88.74 ± 3.85	85.87 ± 4.59	85.08 ± 4.81	87.25 ± 3.87	84.71 ± 6.02	84.51 ± 5.23	85.07 ± 7.23	81.82 ± 6.99	80.54 ± 7.58	83.92 ± 5.45
CSF (%)	11.51 ± 4.05	11.73 ± 4.32	11.25 ± 3.85	14.12 ± 4.59	14.91 ± 4.81	12.74 ± 3.87	15.28 ± 6.02	15.48 ± 5.23	14.92 ± 7.23	18.17 ± 6.99	19.45 ± 7.58	16.07 ± 5.45
WML (%)	0.43 ± 0.53	0.56 ± 0.67	0.28 ± 0.14	0.68 ± 0.76	0.66 ± 0.76	0.71 ± 0.75	0.78 ± 0.70	0.82 ± 0.76	0.71 ± 0.58	1.23 ± 1.11	1.12 ± 1.01	1.41 ± 1.26
TBV Atrophy (%/year)	-0.32 ± 1.02	-0.61 ± 1.01	-0.03 ± 1.03	-0.58 ± 1.02	-0.65 ± 1.01	-0.48 ± 1.03	-0.69 ± 1.02	-0.69 ± 1.01	-0.69 ± 1.03	-1.13 ± 1.01	-0.80 ± 1.01	-1.48 ± 1.03
CSF Expansion (%/year)	2.86 ± 5.46	4.16 ± 5.13	1.57 ± 5.94	3.54 ± 5.46	3.58 ± 5.14	3.47 ± 5.94	3.55 ± 5.46	3.19 ± 5.14	4.27 ± 5.94	4.79 ± 5.46	3.15 ± 5.13	6.54 ± 5.94
WML Expansion (%/year)	-1.59 ± 13.46	-7.91 ± 13.75	4.73 ± 13.05	7.01 ± 13.46	6.41 ± 13.75	7.86 ± 13.06	7.39 ± 13.47	8.23 ± 13.75	5.66 ± 13.05	10.07 ± 13.46	7.57 ± 13.75	12.73 ± 13.05

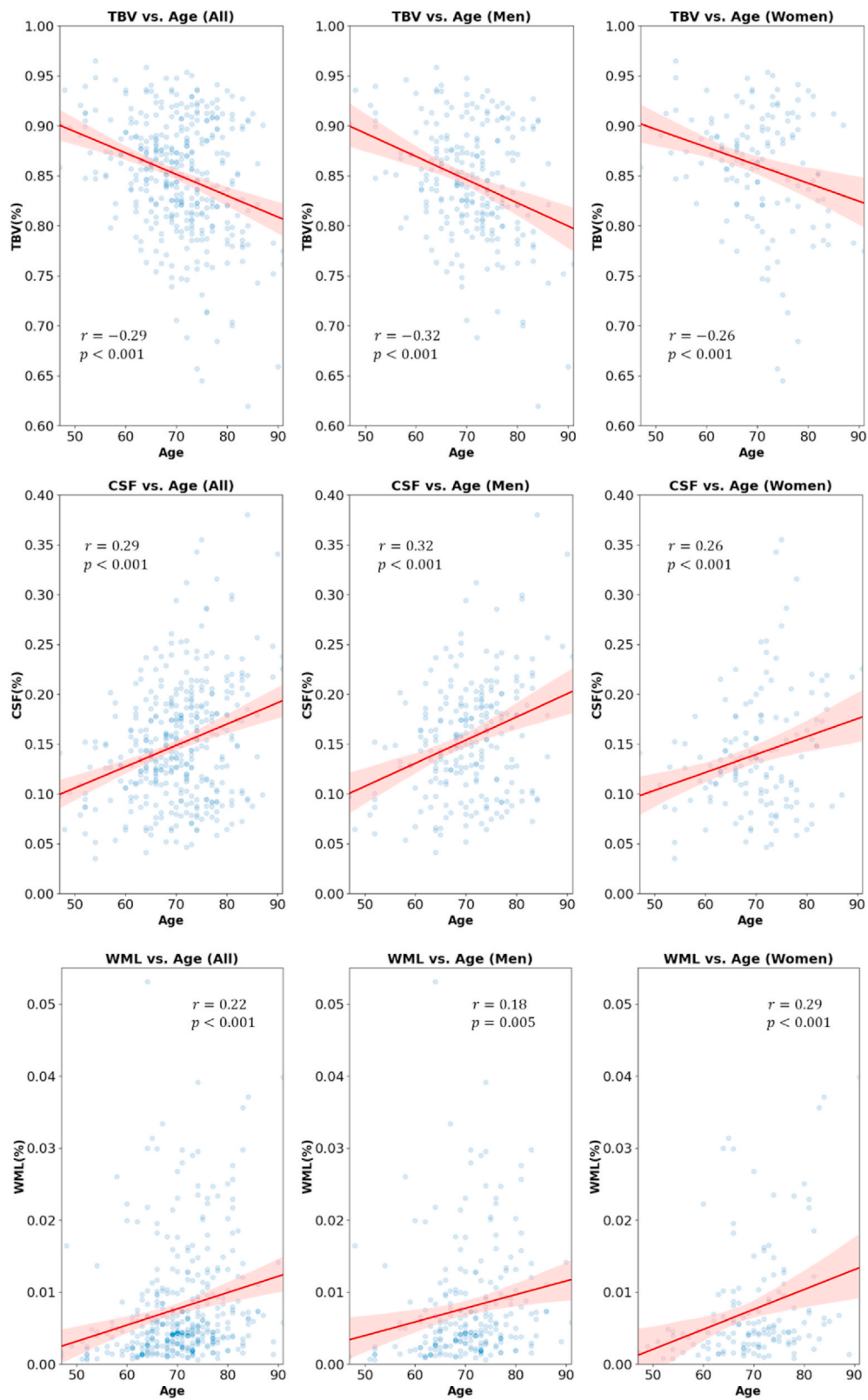


Fig. C. Correlations between cross-sectional biomarkers and age for all subjects, men, and women within the CAIN dataset.

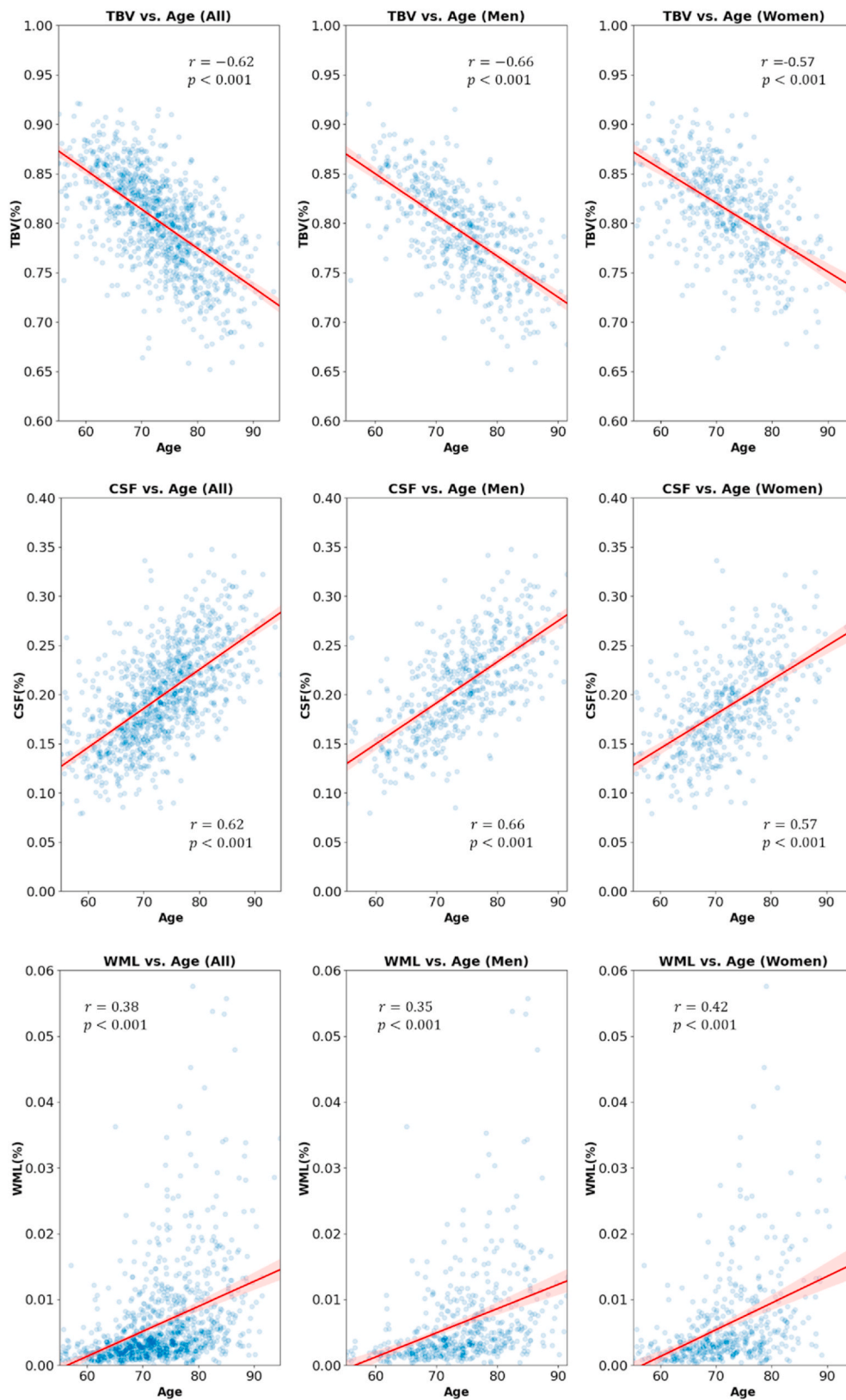


Fig. D. Correlations between cross-sectional biomarkers and age for all subjects, men, and women within the ADNI dataset.

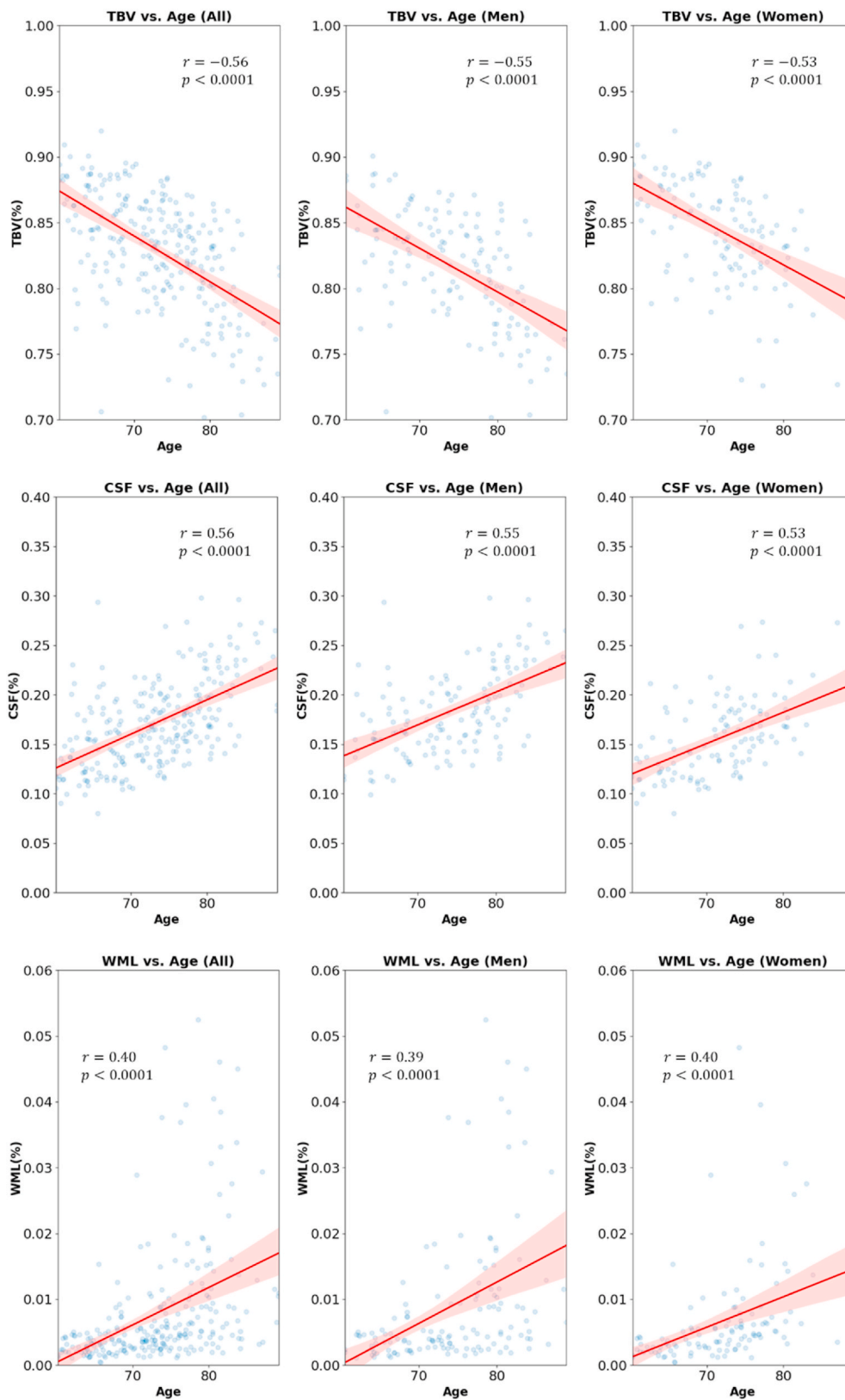


Fig. E. Correlations between cross-sectional biomarkers and age for all subjects, men, and women within the CCNA dataset.

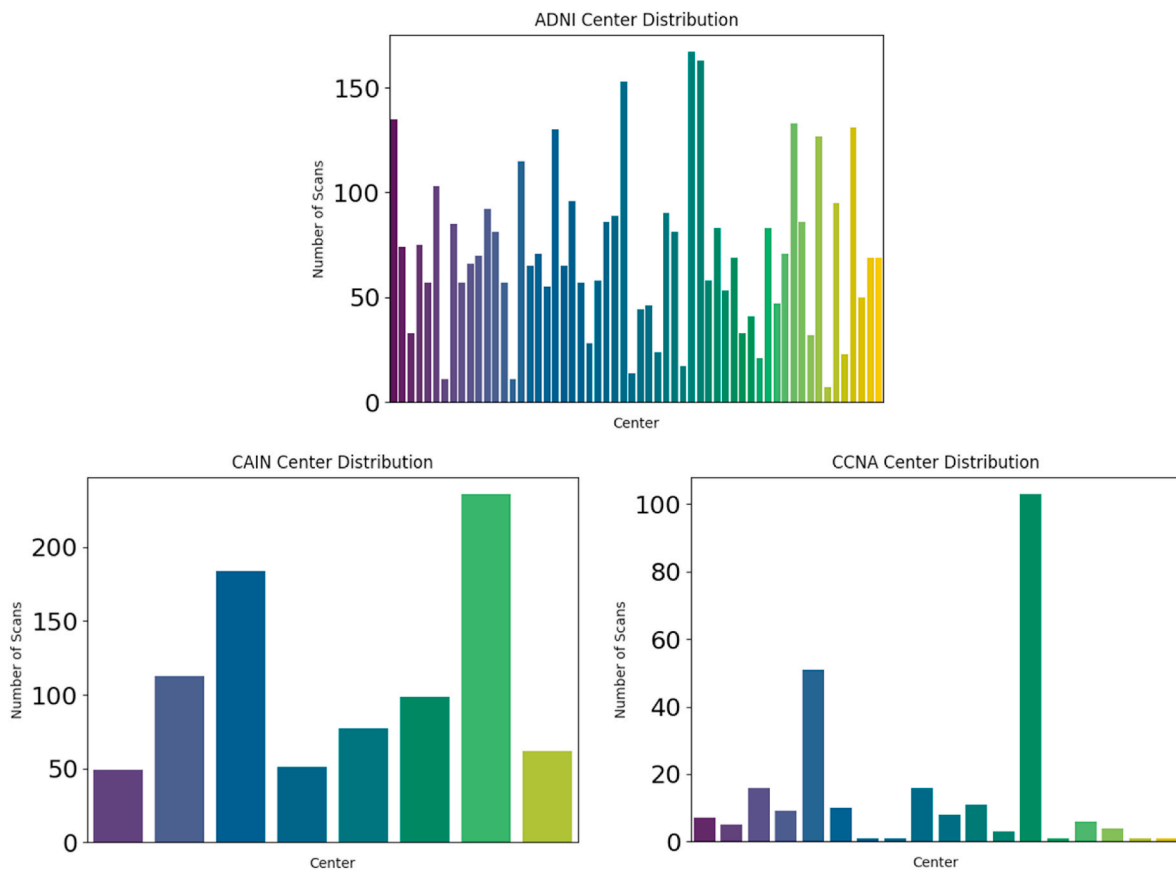
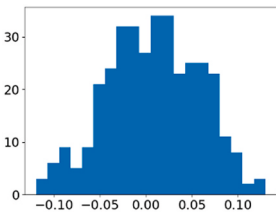
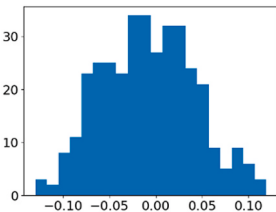


Fig. F. Distributions of scans across acquisition centers for each clinical dataset.

**Table B**

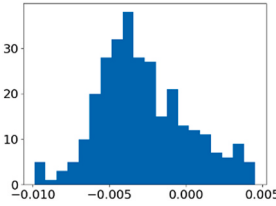
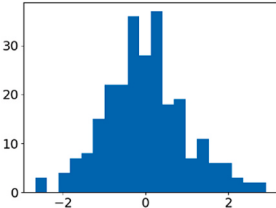
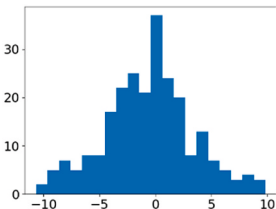
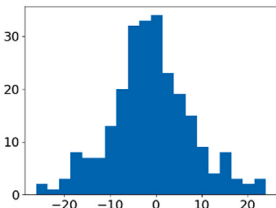
Distribution of residuals of CAIN biomarkers from each ANCOVA model. Levene's test for equality of variances (null hypothesis that variances are equal) was conducted between age groups using a significance level of  $\alpha = 0.05$ .

Biomarker	Distribution of Residuals	Levene test (statistic, p-value)
TBV (%)		2.52, p = 0.08
CSF (%)		2.35, p = 0.07
WML (%)		2.41, p = 0.06

(continued on next page)

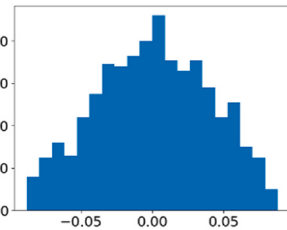


**Table B (continued)**

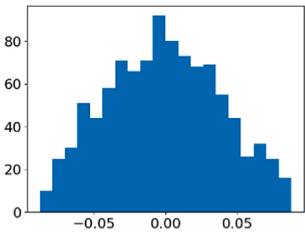
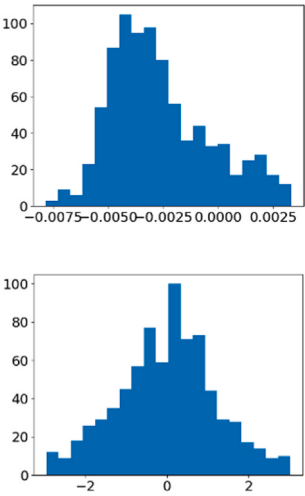
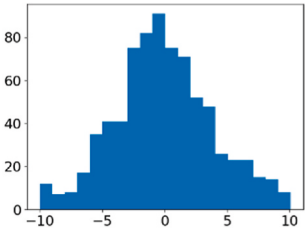
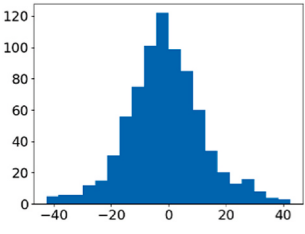
Biomarker	Distribution of Residuals	Levene test (statistic, p-value)
		
TBV Atrophy (%/y)		2.13, p = 0.09
CSF Expansion (%/y)		1.17, p = 0.32
WML Expansion (%/y)		0.93, p = 0.43

**Table C**

Distribution of residuals of ADNI biomarkers from each ANCOVA model. Distribution of residuals was derived from each biomarker's ANCOVA model. Levene's test for equality of variances (null hypothesis that variances are equal) was conducted between disease classifications using a significance level of  $\alpha = 0.05$ .

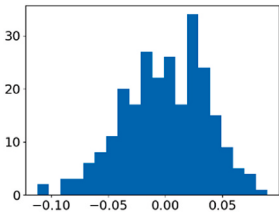
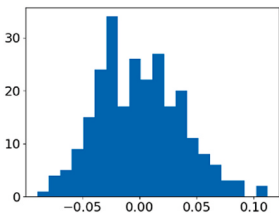
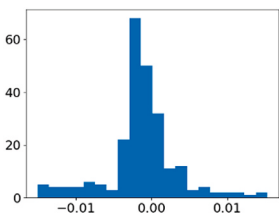
Biomarker	Distribution of Residuals	Levene test (statistic, p-value)
TBV (%)		1.06, p = 0.35
CSF (%)		1.06, p = 0.35 <i>(continued on next page)</i>

**Table C (continued)**

Biomarker	Distribution of Residuals	Levene test (statistic, p-value)
WML (%)		2.36, p = 0.09
TBV Atrophy (%/y)		2.95, p = 0.05
CSF Expansion (%/y)		1.63, p = 0.20
WML Expansion (%/y)		2.95, p = 0.05

**Table D**

Distribution of residuals of CCNA biomarkers from each ANCOVA model. Distribution of residuals was derived from each biomarker's ANCOVA model. Levene's test for equality of variances (null hypothesis that variances are equal) was conducted between disease classifications using a significance level of  $\alpha = 0.05$ .

Biomarker	Distribution of Residuals	Levene test (statistic, p-value)
TBV (%)		0.65, $p = 0.58$
CSF (%)		0.65, $p = 0.58$
WML (%)		0.04, $p = 0.98$

**References**

Bilello, M., Suri, N., Krejza, J., Woo, J.H., Bagley, L.J., Mamourian, A.C., Vossough, A., Chen, J.Y., Millian, B.R., Mulderink, T., 2010. An approach to comparing accuracies of two FLAIR MR sequences in the detection of multiple sclerosis lesions in the brain in the absence of gold standard. *Acad. Radiol.* 17 (6), 686–695.

Chambers, L., Bancej, C., McDowell, I., 2016. Prevalence and Monetary Costs of Dementia in Canada: Population Health Expert Panel. The Alzheimer Society of Canada. Public Health Agency of Canada, Toronto, ON, Canada.

Chan, D., Janssen, J.C., Whitwell, J.L., Watt, H.C., Jenkins, R., Frost, C., Rossor, M.N., Fox, N.C., 2003. Change in rates of cerebral atrophy over time in early-onset Alzheimer's disease: longitudinal MRI study. *Lancet* 362 (9390), 1121–1122.

Chertkow, H., Borrie, M., Whitehead, V., Black, S.E., Feldman, H.H., Gauthier, S., Hogan, D.B., Masellis, M., McGilton, K., Rockwood, K., 2019. The comprehensive assessment of neurodegeneration and dementia: Canadian cohort study. *Can. J. Neurol. Sci.* 46 (5), 499–511.

De Boer, R., Van Der Lijn, F., Vrooman, H.A., Vernooij, M.W., Ikram, M.A., Breteler, M. M., Niessen, W.J., 2007. Automatic Segmentation of Brain Tissue and White Matter Lesions in MRI, pp. 652–655.

De Stefano, N., Stromillo, M.L., Giorgio, A., Bartolozzi, M.L., Battaglini, M., Baldini, M., Portaccio, E., Amato, M.P., Sormani, M.P., 2016. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatr.* 87 (1), 93–99.

Debette, S., Markus, H., 2010. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. *BMJ* 341.

Dickie, D.A., Karama, S., Ritchie, S.J., Cox, S.R., Sakka, E., Royle, N.A., Aribisala, B.S., Hernández, M.V., Maniega, S.M., Pattie, A., 2016a. Progression of white matter disease and cortical thinning are not related in older community-dwelling subjects. *Stroke* 47 (2), 410–416.

Dickie, D.A., Ritchie, S.J., Cox, S.R., Sakka, E., Royle, N.A., Aribisala, B.S., Hernández, M., del, C.V., Maniega, S.M., Pattie, A., Corley, J., 2016b. Vascular risk factors and progression of white matter hyperintensities in the Lothian Birth Cohort 1936. *Neurobiol. Aging* 42, 116–123.

DiGregorio, J., Arezza, G., Gibicar, A., Moody, A.R., Tyrrell, P.N., Khademi, A., 2021. Intracranial volume segmentation for neurodegenerative populations using multicentre flair mri. *Neuroimage: Report 1* (1), 100006.

Driscoll, I., Davatzikos, C., An, Y., Wu, X., Shen, D., Kraut, M., Resnick, S., 2009. Longitudinal pattern of regional brain volume change differentiates normal aging from MCI. *Neurology* 72 (22), 1906–1913.

Evans, M.C., Barnes, J., Nielsen, C., Kim, L.G., Clegg, S.L., Blair, M., Leung, K.K., Douiri, A., Boyes, R.G., Ourselin, S., 2010. Volume changes in Alzheimer's disease and mild cognitive impairment: cognitive associations. *Eur. Radiol.* 20 (3), 674–682.

Fotinos, A.F., Snyder, A.Z., Girton, L.E., Morris, J.C., Buckner, R.L., 2005. Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology* 64 (6), 1032–1039.

Fox, N.C., Schott, J.M., 2004. Imaging cerebral atrophy: normal ageing to Alzheimer's disease. *Lancet* 363 (9406), 392–394.

Fox, N., Scallin, R., Crum, W., Rossor, M., 1999. Correlation between rates of brain atrophy and cognitive decline in AD. *Neurology* 52 (8), 1687–1687.

Freeborough, P.A., Fox, N.C., 1997. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. *IEEE Trans. Med. Imag.* 16 (5), 623–629.

García-Lorenzo, D., Francis, S., Narayanan, S., Arnold, D.L., Collins, D.L., 2013. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Med. Image Anal.* 17 (1), 1–18.

Hansen, T., Brezova, V., Eikenes, L., Håberg, A., Vangberg, T., 2015. How does the accuracy of intracranial volume measurements affect normalized brain volumes? Sample size estimates based on 966 subjects from the HUNT MRI cohort. *Am. J. Neuroradiol.* 36 (8), 1450–1456.

Heinen, R., Steenwijk, M.D., Barkhof, F., Biesbroek, J.M., van der Flier, W.M., Kuijff, H.J., Prins, N.D., Vrenken, H., Biessels, G.J., de Bresser, J., 2019. Performance of five automated white matter hyperintensity segmentation methods in a multicenter dataset. *Sci. Rep.* 9 (1), 1–12.

Iadecola, C., 2013. The pathobiology of vascular dementia. *Neuron* 80 (4), 844–866.

Jack Jr., C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J.L., Ward, C., 2008b. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imag.: J. Int. Soc. Magnetic Resonance. Med.* 27 (4), 685–691.

Jack, C., Weigand, S.D., Shiung, M.M., Przybelski, S.A., O'Brien, P.C., Gunter, J.L., Knopman, D.S., Boeve, B.F., Smith, G.E., Petersen, R.C., 2008a. Atrophy rates accelerate in amnesic mild cognitive impairment. *Neurology* 70 (19 Part 2), 1740–1752.

Jouvent, E., Viswanathan, A., Chabriat, H., 2010. Cerebral atrophy in cerebrovascular disorders. *J. Neuroimaging* 20 (3), 213–218.

- Khademi, A., Gibicar, A., Arezza, G., DiGregorio, J., Tyrrell, P.N., Moody, A., 2021. Segmentation of white matter lesions in multicentre FLAIR MRI. *Neuroimage: Report* 1 (4), December 2021.
- Khademi, A., Reiche, B., DiGregorio, J., Arezza, G., Moody, A.R., 2020. Whole volume brain extraction for multi-centre, multi-disease FLAIR MRI datasets. *Magn. Reson. Imag.* 66, 116–130.
- Khademi, A., Venetsanopoulos, A., Moody, A.R., 2011. Robust white matter lesion segmentation in FLAIR MRI. *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 59 (3), 860–871.
- Lamar, M., Boots, E.A., Arfanakis, K., Barnes, L.L., Schneider, J.A., 2020. Common brain structural alterations associated with cardiovascular disease risk factors and Alzheimer's dementia: future directions and implications. *Neuropsychol. Rev.* 1–12.
- Maillard, P., Carmichael, O., Harvey, D., Fletcher, E., Reed, B., Mungas, D., DeCarli, C., 2013. FLAIR and diffusion MRI signals are independent predictors of white matter hyperintensities. *Am. J. Neuroradiol.* 34 (1), 54–61.
- Mayeux, R., Stern, Y., 2012. Epidemiology of alzheimer disease. *Cold Spring Harbor Perspectives in Medicine* 2 (8), a006239.
- Meng, D., Hosseini, A.A., Simpson, R.J., Shaikh, Q., Tench, C.R., Dineen, R.A., Auer, D.P., 2017. Lesion topography and microscopic white matter tract damage contribute to cognitive impairment in symptomatic carotid artery disease. *Radiology* 282 (2), 502–515.
- Messina, S., Patti, F., 2014. Gray Matters in Multiple Sclerosis: Cognitive Impairment and Structural MRI. *Multiple Sclerosis International*, 2014.
- Mohaddes, Z., Das, S., Abou-Haidar, R., Safi-Harab, M., Blader, D., Callegaro, J., Henri-Bellemare, C., Tunteng, J.-F., Evans, L., Campbell, T., 2018. National neuroinformatics framework for canadian consortium on neurodegeneration in aging (CCNA). *Front. Neuroinf.* 12, 85.
- Morgen, K., Sammer, G., Courtney, S.M., Wolters, T., Melchior, H., Blecker, C.R., Oschmann, P., Kaps, M., Vaitl, D., 2006. Evidence for a direct association between cortical atrophy and cognitive impairment in relapsing–remitting MS. *Neuroimage* 30 (3), 891–898.
- Mungas, D., Harvey, D., Reed, B.R., Jagust, W., DeCarli, C., Beckett, L., Mack, W., Kramer, J., Weiner, M., Schuff, N., 2005. Longitudinal volumetric MRI change and rate of cognitive decline. *Neurology* 65 (4), 565–571.
- Narayana, P.A., Coronado, I., Sujit, S.J., Sun, X., Wolinsky, J.S., Gabr, R.E., 2020. Are multi-contrast magnetic resonance images necessary for segmenting multiple sclerosis brains? A large cohort study based on deep learning. *Magn. Reson. Imag.* 65, 8–14.
- Narayanan, S., Nakamura, K., Fonov, V.S., Maranzano, J., Caramanos, Z., Giacomini, P. S., Collins, D.L., Arnold, D.L., 2020. Brain volume loss in individuals over time: source of variance and limits of detectability. *Neuroimage* 214, 116737.
- Nestor, S.M., Rupsingh, R., Borrie, M., Smith, M., Accomazzi, V., Wells, J.L., Fogarty, J., Bartha, R., Alzheimer's Disease Neuroimaging Initiative., 2008. Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. *Brain* 131 (9), 2443–2454.
- Nordenskjöld, R., Malmberg, F., Larsson, E.-M., Simmons, A., Brooks, S.J., Lind, L., Ahlström, H., Johansson, L., Kullberg, J., 2013. Intracranial volume estimated with commonly used methods could introduce bias in studies including brain volume measurements. *Neuroimage* 83, 355–360.
- Reiche, B., Moody, A., Khademi, A., 2019. Pathology-preserving intensity standardization framework for multi-institutional FLAIR MRI datasets. *Magn. Reson. Imag.* 62, 59–69.
- Richard, E., Gouw, A.A., Scheltens, P., van Gool, W.A., 2010. Vascular care in patients with Alzheimer disease with cerebrovascular lesions slows progression of white matter lesions on MRI: the evaluation of vascular care in Alzheimer's disease (EVA) study. *Stroke* 41 (3), 554–556.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional Networks for Biomedical Image Segmentation, pp. 234–241.
- Scahill, R.I., Frost, C., Jenkins, R., Whitwell, J.L., Rossor, M.N., Fox, N.C., 2003. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch. Neurol.* 60 (7), 989–994.
- Scahill, R.I., Schott, J.M., Stevens, J.M., Rossor, M.N., Fox, N.C., 2002. Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. *Proc. Natl. Acad. Sci. Unit. States Am.* 99 (7), 4703–4707.
- Seghier, M.L., Ramsden, S., Lim, L., Leff, A.P., Price, C.J., 2014. Gradual lesion expansion and brain shrinkage years after stroke. *Stroke* 45 (3), 877–879.
- Silbert, L.C., Dodge, H.H., Perkins, L.G., Sherbakov, L., Lahna, D., Erten-Lyons, D., Woltjer, R., Shinto, L., Kaye, J.A., 2012. Trajectory of white matter hyperintensity burden preceding mild cognitive impairment. *Neurology* 79 (8), 741–747.
- Silbert, L.C., Howieson, D.B., Dodge, H., Kaye, J.A., 2009. Cognitive impairment risk: white matter hyperintensity progression matters. *Neurology* 73 (2), 120–125.
- Schmidt, R., Seiler, S., Loitfelder, M., 2016. Longitudinal change of small-vessel disease-related brain abnormalities. *J. Cerebr. Blood Flow Metabol.* 36 (1), 26–39.
- Smith, E.E., Cieslak, A., Barber, P., Chen, J., Chen, Y., Donnini, I., Edwards, J.D., Frayne, R., Field, T.S., Hegedus, J., 2017. Therapeutic strategies and drug development for vascular cognitive impairment. *J. Am. Heart Assoc.* 6 (5), e005568.
- Soltanian-Zadeh, H., Peck, D.J., 2001. Feature space analysis: effects of MRI protocols. *Med. Phys.* 28 (11), 2344–2351.
- Stephen, R., Liu, Y., Ngandu, T., Antikainen, R., Hulkkonen, J., Koikkalainen, J., et al., 2019. Brain volumes and cortical thickness on MRI in the Finnish geriatric intervention study to prevent cognitive impairment and disability (FINGER). *Alzheimer's Res. Ther.* 11 (1), 1–10.
- Storelli, L., Rocca, M.A., Pagani, E., Van Hecke, W., Horsfield, M.A., De Stefano, N., Rovira, A., Sastre-Garriga, J., Palace, J., Sima, D., 2018. Measurement of whole-brain and gray matter atrophy in multiple sclerosis: assessment with MR imaging. *Radiology* 288 (2), 554–564.
- Svennerholm, L., Boström, K., Jungbjer, B., 1997. Changes in weight and compositions of major membrane components of human brain during the span of adult human life of Swedes. *Acta Neuropathol.* 94 (4), 345–352.
- Tardif, J.-C., Spence, J.D., Heinonen, T.M., Moody, A., Pressacco, J., Frayne, R., L'allier, P., Chow, B.J., Friedrich, M., Black, S.E., 2013. Atherosclerosis imaging and the Canadian atherosclerosis imaging network. *Can. J. Cardiol.* 29 (3), 297–303.
- Wardlaw, J.M., Valdés Hernández, M.C., Muñoz-Maniega, S., 2015. What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. *J. Am. Heart Assoc.* 4 (6), e001140.
- Wilke, M., de Haan, B., Juenger, H., Karnath, H.-O., 2011. Manual, semi-automated, and automated delineation of chronic brain lesions: a comparison of methods. *Neuroimage* 56 (4), 2038–2046.