RESEARCH ARTICLE

Validation of T1w-based segmentations of white matter hyperintensity volumes in large-scale datasets of aging

Mahsa Dadar[1] | Josefina Maranzano[2] | Simon Ducharme[2] |

Owen T. Carmichael[3] | Charles Decarli[4] |

D. Louis Collins,[1]* Alzheimer's Disease Neuroimaging Initiative

[1]NeuroImaging and Surgical Tools Laboratory, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

[2]Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

[3]Pennington Biomedical Research Center, Baton Rouge, Louisiana

[4]University of California, Davis, California

Correspondence
D. Louis Collins, Magnetic Resonance Imaging (MRI), Montreal Neurological Institute, 3801 University Street, Room WB315, Montréal, QC H3A 2B4, Canada.
Email: louis.collins@mcgill.ca

Abstract

Introduction: Fluid-attenuated Inversion Recovery (FLAIR) and dual T2w and proton density (PD) magnetic resonance images (MRIs) are considered to be the optimum sequences for detecting white matter hyperintensities (WMHs) in aging and Alzheimer's disease populations. However, many existing large multisite studies forgo their acquisition in favor of other MRI sequences due to economic and time constraints.

Methods: In this article, we have investigated whether FLAIR and T2w/PD sequences are necessary to detect WMHs in Alzheimer's and aging studies, compared to using only T1w images. Using a previously validated automated tool based on a Random Forests classifier, WMHs were segmented for the baseline visits of subjects from ADC, ADNI1, and ADNI2/GO studies with and without T2w/PD and FLAIR information. The obtained WMH loads (WMHLs) in different lobes were then correlated with manually segmented WMHLs, each other, age, cognitive, and clinical measures to assess the strength of the correlations with and without using T2w/PD and FLAIR information.

Results: The WMHLs obtained from T1w-Only segmentations correlated with the manual WMHLs (ADNI1: $r = .743$, $p < .001$, ADNI2/GO: $r = .904$, $p < .001$), segmentations obtained from T1w + T2w + PD for ADNI1 ($r = .888$, $p < .001$) and T1w + FLAIR for ADNI2/GO ($r = .969$, $p < .001$), age (ADNI1: $r = .391$, $p < .001$, ADNI2/GO: $r = .466$, $p < .001$), and ADAS13 (ADNI1: $r = .227$, $p < .001$, ADNI2/GO: $r = .190$, $p < 0.001$), and NPI (ADNI1: $r = .290$, $p < .001$, ADNI2/GO: $r = 0.144$, $p < .001$), controlling for age.

Conclusion: Our results suggest that while T2w/PD and FLAIR provide more accurate estimates of the true WMHLs, T1w-Only segmentations can still provide estimates that hold strong correlations with the actual WMHLs, age, and performance on various cognitive/clinical scales, giving added value to datasets where T2w/PD or FLAIR are not available.

1 | INTRODUCTION

White matter hyperintensities (WMHs), defined as regions of higher signal than the surrounding normal appearing white matter (NAWM) on T2w or FLAIR MR images, are one of the most common findings in structural MR imaging in older adults, reflecting demyelination and axonal loss (Prins & Scheltens, 2015). While sensitive as an expression of abnormality in the white matter (WM) tissue, the etiology of WMHs is quite varied, with ischemia due to cerebral small vessel disease playing an important role in the majority of older subjects (Gouw et al., 2010; Yoshita, Fletcher, & DeCarli, 2005). This age-related ischemic small vessel disease is also referred to as arteriolosclerosis, hypertension-related, or vascular-risk-factor-related small vessel disease (Pantoni, 2010). However, the term small vessel disease is also related to other pathologies that affect small arteries, arterioles, venules, and capillaries, such as cerebral amyloid angiopathy, genetic small vessel disease distinct from amyloid angiopathy (e.g., cerebral autosomal dominant/recessive arteriopathy with subcortical infarcts and leukoencephalopathy or CADASIL/CARASIL), inflammatory-mediated small vessel disease (e.g., primary angiitis of central nervous system CNS, Wegener's granulomatosis), and venous collagenosis (Pantoni, 2010).

Some of the consequences of small vessel disease include lacunar infarcts, WMHs, micro and macro bleeding. The first two phenomena are easily detected on MR images. In contrast, small vessels cannot be seen using MRI, so the term small vessel disease on MRI has been used for (and become equivalent to) WMHs and lacunar infarcts (Pantoni, 2010). Unfortunately, there is great heterogeneity across neuropathological centers regarding the definition of MRI small vessel disease, with overall agreement lower than 50% (Pantoni et al., 2006). As there is no conclusive data showing the levels of specificity and sensitivity of WMHs on MRI as a reflection of a specific etiology of small vessel disease, we consider, in our cases, that the two major groups (arteriolosclerosis and amyloid angiopathy) are probably the main substrates of the WMHs. These two etiologies on their own have a crucial role in three major clinical areas: stroke, neurocognitive disorders (dementia), and aging-related cognitive decline (Pantoni, 2010).

The location and load of WMHs have been shown to correlate with age, a history of hypertension, hyperinsulinemia (Hawkins et al., 2017), and cognitive deficits (Biesbroek, Weaver, & Biessels, 2017; DeCarli et al., 1995; Dubois et al., 2014). Therefore, WMHs constitute a clinically meaningful biomarker of cognitive decline related to general aging and pathological vascular processes, which are known contributors to multifactorial neurodegenerative diseases (Iturria-Medina, Sotero, Toussaint, Mateos-Pérez, & Evans, 2016). They are a particularly important clinical measure in the elderly populations (Carmichael et al., 2010; De Groot et al., 2002; DeCarli et al., 2001; Dubois et al., 2014; Longstreth et al., 1996; van Straaten et al., 2008).

WMHs are generally assessed using FLAIR or T2w/PD scans which have optimum contrast for detecting such lesions (Caligiuri et al., 2015). T2w/PD and FLAIR WMHs have been shown to correspond to myelin stain lesions in postmortem histology studies (Fernando et al., 2004; Takao et al., 1999). However, WMHs can also be detected on T1w scans to some extent. The characteristic bright WMH signal of FLAIR, T2w and PD, manifests in the T1w sequence as a hypointense area, heterogeneous in the value of the lower signal, ranging from iso-intense to hypo-intense in relation to the surrounding NAWM. In other words, a FLAIR, T2w/PD homogeneous hyperintense area would correspond, in the T1w modality, to a similar area of heterogeneous hypointense signal, ranging from values close to fluid, to isointense in relation to the surrounding NAWM. This phenomenon is probably determined by the different types and degrees of change occurring in the WM at the same time (e.g., more/less intense demyelination and axonal loss). This range of T1w hypointensities in a region of WM tissue is more homogeneously represented by the bright signal of T2w and FLAIR sequences. As the quantification of lesion volumes in a given MRI modality depends on the contrast between the lesional area and the surrounding NAWM, these volumes will always be larger if the detection considers the bright signal of FLAIR or T2w scans, as opposed to using only T1w images. The hypointensity seen on T1w images presumably reflects the most severe spectrum of WM injury.

Although FLAIR or T2w/PD scans are the optimal sequences to detect WMHs, many especially large-scale studies forgo acquisition of either one or all of the optimal modalities because of time and financial constraints. There can also be differences in WMH volume levels when comparing T2w/PD with FLAIR scans, with FLAIR scans tending to give higher overall levels. Consequently, it would be extremely useful if one can get an estimate of the load and location of WMHs without requiring these optimal modalities. While there have been other studies that define and use T1w white matter signal abnormality (WMSA) detected by Freesurfer (Fischl, 2012) as a measure of WMH in aging and AD populations (Jacobs et al., 2013; Leritz et al., 2014; Salat et al., 2010), to our knowledge, no studies have investigated and validated the relationship between these T1w hypointensities and FLAIR or T2w/PD-based WMH segmentations and whether there is a significant difference in their relationships with clinical measures.

In this article, we aimed to compare the ability of T1w, T2w/PD and FLAIR scans in differentiating between healthy tissue and WMHs, both in terms of (a) detection in comparison with manually segmented labels and (b) correlation with a variety of clinical measures. Our goal is to determine if WMHs can be partially but accurately segmented based only on T1w images, and how reliable T1w-based assessments are in comparison with the more accurate estimates obtained based on FLAIR or T2w/PD sequences.

2 | MATERIALS AND METHODS

2.1 | Subjects

The WMHs were segmented both manually and automatically in three different datasets to ensure generalizability of the results. Table 1 summarizes the information for each dataset.

1. The first dataset (ADC) consists of 70 elderly individuals who received a full clinical workup and structural MR scans including T1w, double-echo PD/T2w, and FLAIR scans at their enrollment into the University of California, Davis Alzheimer's Disease Center (ADC) (Hinton et al., 2010). Subjects were 70–90 years old with normal cognition, mild cognitive impairment, or AD. All subjects were manually segmented by an expert rater.

2. The second dataset included subjects selected from ADNI study. This data was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI 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 magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). ADNI was carried out with the goal of recruiting 800 adults aged from 55 to 90, and consists of ∼200 cognitively normal, 400 MCI, and 200 AD subjects. ADNIGO is a later study that followed ADNI participants that were in cognitively normal or early MCI stages (http://www.adcs.org/studies/imagineadni.aspx). ADNI2 study followed patients in the same categories as well as recruiting 550 new subjects (http://www.adcs.org/studies/ImagineADNI2.aspx). Baseline visit data from ADNI1 and ADNI2/GO subjects were used in this study (Table 1). Forty-six subjects with T1w and FLAIR scans and different loads of WMHs were selected from ADNI2/GO study for manual segmentation. To ensure that the datasets used for training and validation of the method have a wide range of WMHs, segmentation techniques generally make sure to include subjects with small, medium, and large WMH loads (Dadar et al., 2017a; Griffanti et al., 2016; Schmidt et al., 2012; Simões et al., 2013). Here, subjects were selected from different sites and scanners and a preliminary assessment was performed to evaluate their WMH load with the goal of acquiring subjects with different scanner information as well as different loads of WMHs. For each scanner model, we selected datasets that had low (<5 CCs), medium (5–20 CCs), and high lesion loads (>20 CCs). Equal numbers of male and female subjects were selected. The age of the subjects was also considered for the selection, with the aim of achieving a normal distribution. Using a similar strategy, 53 subjects with T1w, T2w, and PD scans and different WMH loads were selected from the ADNI1 study for manual segmentation.

2.2 | Clinical evaluations

We did not have the clinical evaluations available for the ADC study. The clinical assessment and cognitive testing of ADNI study followed a

TABLE 1 Descriptive statistics for the ADNI subjects enrolled in this study

Test	ADNI1	ADNI2/GO
Number (male)	669 (393)	481 (298)
Age	75.26 ± 6.84	72.62 ± 7.54
MMSE	26.71 ± 2.70	27.58 ± 2.76
ADAS11	11.63 ± 6.24	9.89 ± 6.80
ADAS13	18.45 ± 9.08	15.29 ± 9.66
FAQ	4.88 ± 6.47	3.44 ± 5.70
RAVLTI	32.09 ± 11.17	37.72 ± 13.08
RAVLTF	4.31 ± 2.38	4.29 ± 2.62
RAVLTL	3.61 ± 2.62	4.65 ± 2.80
Executive	0.71 ± 0.52	0.91 ± 0.60
NPI	10.30 ± 0.51	7.39 ± 7.77

Note. Data are number or mean ± standard deviation.
ADNI = Alzheimer's Disease Neuroimaging Initiative; MMSE = Mini-Mental State Examination; ADAS = Alzheimer's Disease Assessment Scale; FAQ = Functional Assessment Questionnaire; RAVLT = Rey Auditory Verbal Learning Task (I = Immediate, F = Forgetting, L = Learning); Executive = Executive Function; NPI = Total Neuropsychiatric Inventory Score.

standardized protocol that has been described previously (Petersen et al., 2010). At each visit, the participants underwent a standardized clinical evaluation and cognitive tests including Mini-Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog), Functional Assessment Questionnaire (FAQ), Neuropsychiatric Inventory (NPI), a composite score for executive function (Gibbons et al., 2012), and Immediate, Forgetting and Learning subscores from Rey Auditory Verbal Learning Task (RAVLT). Each individual had a self-reported history of hypertension, and cardiovascular risk factors available. Table 1 summarizes this information for the subjects that were used in this study. For details on the administration and scoring, see http://www.adni-info.org/Scientists/ADNIData.html.

2.3 | MR imaging

This section describes the scanner information and image acquisition parameters for the abovementioned datasets. Table 2 shows the summary of this information for each sequence in each dataset.

2.3.1 | ADC dataset

MRI data were acquired on two 1.5 T MRI scanners: a GE MEDICAL SYSTEMS Signa scanner located at UCD Medical Center (Sacramento, CA) and a Philips Eclipse scanner located at the Veterans Administration Northern California Health Care System (Martinez, CA). T1w scans were acquired with an FSPGR pulse sequence with 1.5 mm slice thickness, 128 slices covering the entire brain, a 250 × 250 mm field of view, a 256 × 256 scan matrix, voxel size of 0.9765 × 0.9765 × 1.5 mm, repetition time (TR) of 9 ms, and echo time (TE) of 2.9 ms. FLAIR scans were acquired with a fast spin echo (FSE) sequence with

TABLE 2 Scanner information and MRI acquisition parameters for ADC, ADNI1, and ADNI2/GO datasets

Modality	Dataset	ADC	ADNI1	ADNI2/GO
	Scanner manufacturer	GE, Philips	GE, Philips, SIEMENS	Philips, SIEMENS
T1w	Slice thickness	1.5 mm	1.2 mm	1.2 mm
	No. of slices	128	160	196
	Field of view	250×250 cm^2	192×192 cm^2	256×256 cm^2
	Scan matrix	256×256 cm^2	192×192 cm^2	256×256 cm^2
	Repetition time (TR)	9 ms	3000 ms	7.2 ms
	Echo time (TE)	2.9 ms	3.55 ms	3.0 ms
	Pulse sequence	FSPGR	MPRAGE	MPRAGE
Other	Contrast	T2w/PD/FLAIR	T2w/PD	FLAIR
	Slice thickness	3/3/3 mm	3 mm	5 mm
	No. of slices	42/42/48	56	42
	Field of view	$240 \times 240/240 \times 240/220 \times 220$ cm^2	256×256 cm^2	256×256 cm^2
	Scan matrix	$256 \times 256/256 \times 256/256 \times 192$ cm^2	256×256 cm^2	256×256 cm^2
	Repetition time (TR)	2420/2420/11000 ms	3000/3000 ms	11000 ms
	Echo time (TE)	90/20/144 ms	95.2/10.5 ms	150 ms
	Pulse sequence	DSE/DSE/FSE	FSE	SE/IR

3 mm slice thickness, a 220×220 mm field of view, and a 256×192 scan matrix, voxel size of $0.9765 \times 0.9765 \times 3$ mm, TR = 11000 ms, and TE = 144 ms. Analogous sequences were installed on both the GE and Philips scanners.

2.3.2 | ADNI1 dataset

The MRI data used were acquired on scanners from three different manufacturers, GE, Philips, and SIEMENS. All patients had similar MRI protocols for T1w and T2w/PD scans. T1w scans were acquired in 3D
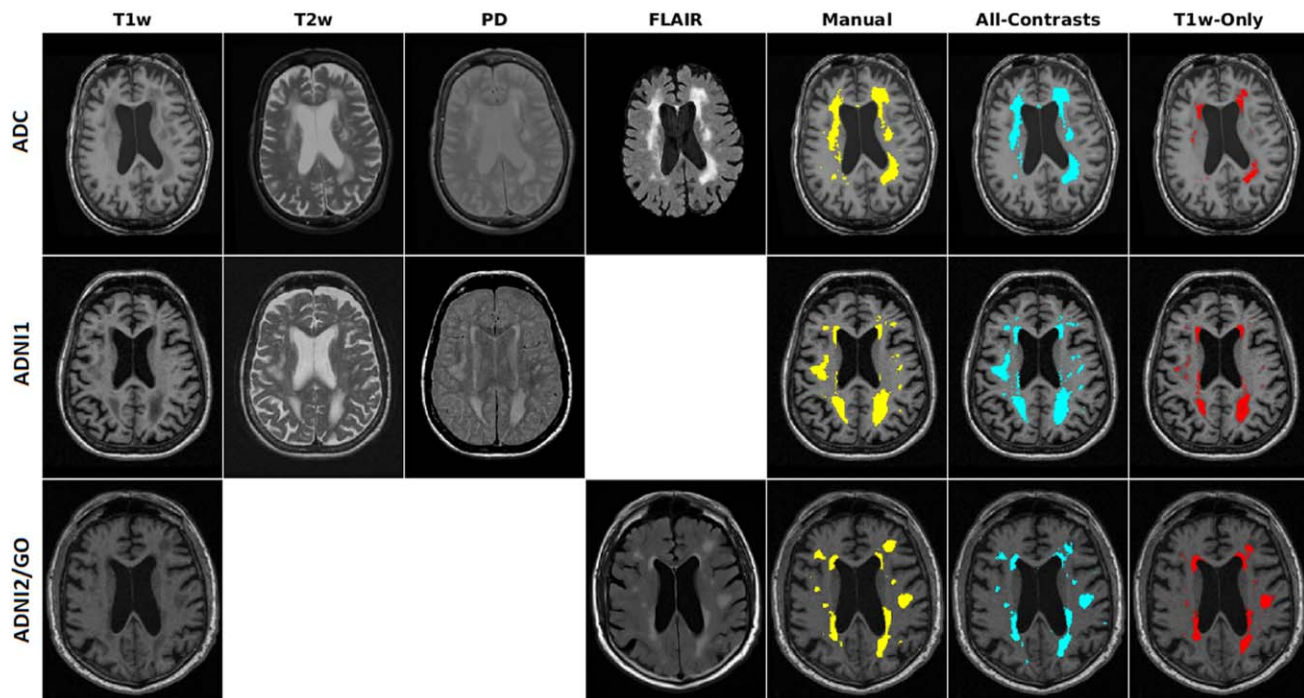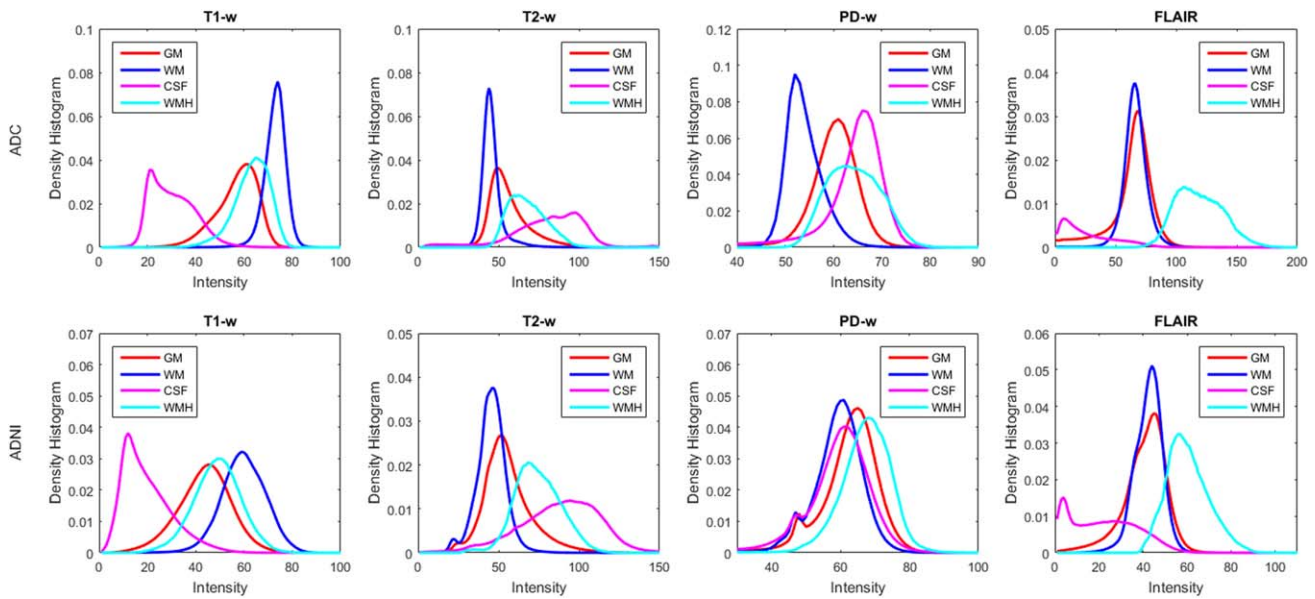

FIGURE 1 Axial slices showing T1w, T2w, PD, and FLAIR images and manual (yellow), all-contrasts (cyan), and T1w-Only (red) segmentations of subjects from ADC (top), ADNI1 (middle), and ADNI2/GO (bottom) datasets [Color figure can be viewed at wileyonlinelibrary.com]

FIGURE 2 Intensity histograms of white matter (WM), grey matter (GM), cerebrospinal fluid (CSF), and white matter hyperintensities (WMHs) for (a) ADC and (b) ADNI datasets [Color figure can be viewed at wileyonlinelibrary.com]

with a gradient recalled sequence with 1.2 mm slice thickness, 160 sagittal slices, covering the entire brain, a 192 × 192 mm field of view, and a 192 × 192 scan matrix, voxel size of 1.2 × 0.9375 ×0.9375 mm, TR = 3000 ms, and TE = 3.55 ms. T2w/PD scans were acquired in 2D with an FSE sequence with 3.0 mm slice thickness, 56 axial slices covering the entire brain, a 256 × 256 mm field of view, and a 256 × 256 scan matrix, a voxel size of 0.8594 × 0.8594 × 5 mm, with TR = 3000 ms, TE = 95.2 ms for T2w, and TE = 10.5 ms for PD images.

2.3.3 | ADNI2/GO dataset

The MRI data used were acquired on scanners from three different manufacturers, Philips, GE and SIEMENS. All patients had similar MRI protocols for T1w and FLAIR scans. T1w scans were acquired in 3D with a gradient recalled sequence with 1.2 mm slice thickness, 196 sagittal slices, covering the entire brain, a 256 × 256 mm field of view,

and a 256 × 256 scan matrix, voxel size of 1 × 1 × 1.2 mm, TR = 7.2 ms, and TE = 3.0 ms. FLAIR scans were acquired in 2D with a spin echo inversion recovery sequence with 5.0 mm slice thickness, 42 axial slices covering the entire brain, a 256 × 256 mm field of view, and a 256 × 256 scan matrix, voxel size of 0.8594 × 0.8594 × 5 mm, with TR = 11000 ms and TE = 150 ms.

2.4 | Preprocessing

All MRI scans were preprocessed using our standardized pipeline. Images were denoised (Manjón, Coupé, Martí-Bonmatí, Collins, & Robles, 2010), corrected for image intensity inhomogeneity (Sled, Zijdenbos, & Evans, 1998) and intensity scaled (Fonov et al., 2011). The T2w, PD, and FLAIR scans were then co-registered to the structural T1w scan of the same subject using a six-parameter rigid body registration (Collins, Neelin, Peters, & Evans, 1994). The T1w scans were nonlinearly registered to the ADNI template based on intensity correlation coefficient (Collins & Evans, 1997). The quality of the nonlinear registrations was visually assessed and the results that did not pass this quality control were discarded. Using the T1w-to-template transformations (i.e., linear + nonlinear), the other modalities (e.g., FLAIR, T2w, PD) were registered to the ADNI template as well. The manually segmented lesion maps were also registered to the ADNI template using the transformations of their corresponding FLAIR images.

2.5 | Manual segmentation

In ADC and ADNI2/GO datasets, the WMHs were manually segmented by an expert with more than 12 years of experience in reading MRI and developing standardized MRI guidelines to detect WM lesions using different image modalities (Maranzano, Rudko, Arnold, & Narayanan, 2016) with FLAIR used as the primary contrast and with T1w used to aid in the decision process to include or exclude a voxel from

TABLE 3 Percentage of overlap between density histograms of WMHs and GM, WM, and CSF

Dataset	Dataset	WM	GM	CSF
ADC	T1w	22.33	37.71	5.33
	T2w	9.02	29.80	30.75
	PD	15.77	36.47	56.91
	FLAIR	2.73	3.76	1.02
ADNI1	T1w	30.06	42.49	9.72
	T2w	9.32	23.95	32.95
	PD	28.96	39.90	32.37
ADNI2/GO	T1w	28.72	41.87	8.97
	FLAIR	12.41	14.53	5.00

Note. WMH = white matter hyperintensity; GM = gray matter; WM = white matter.
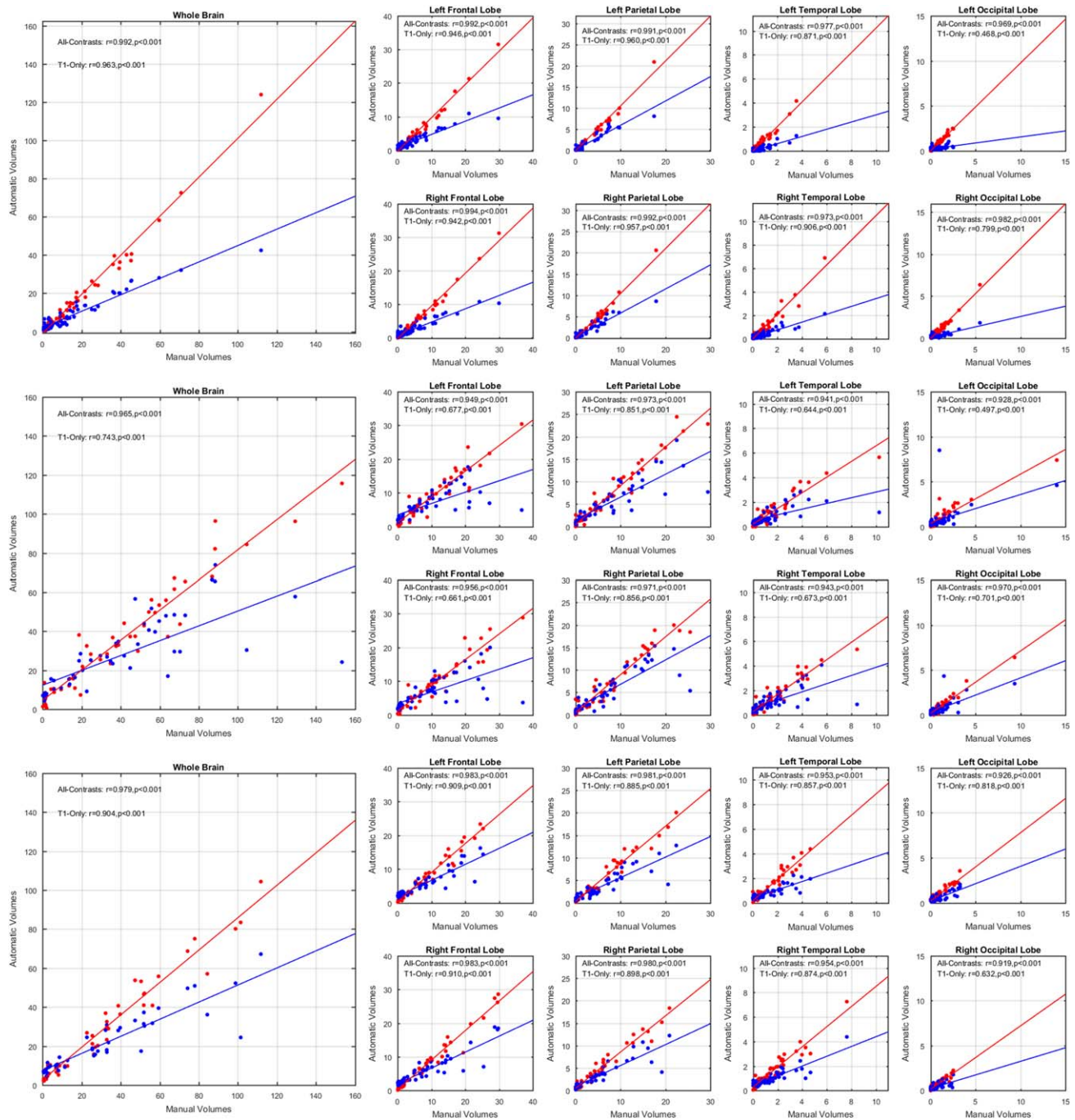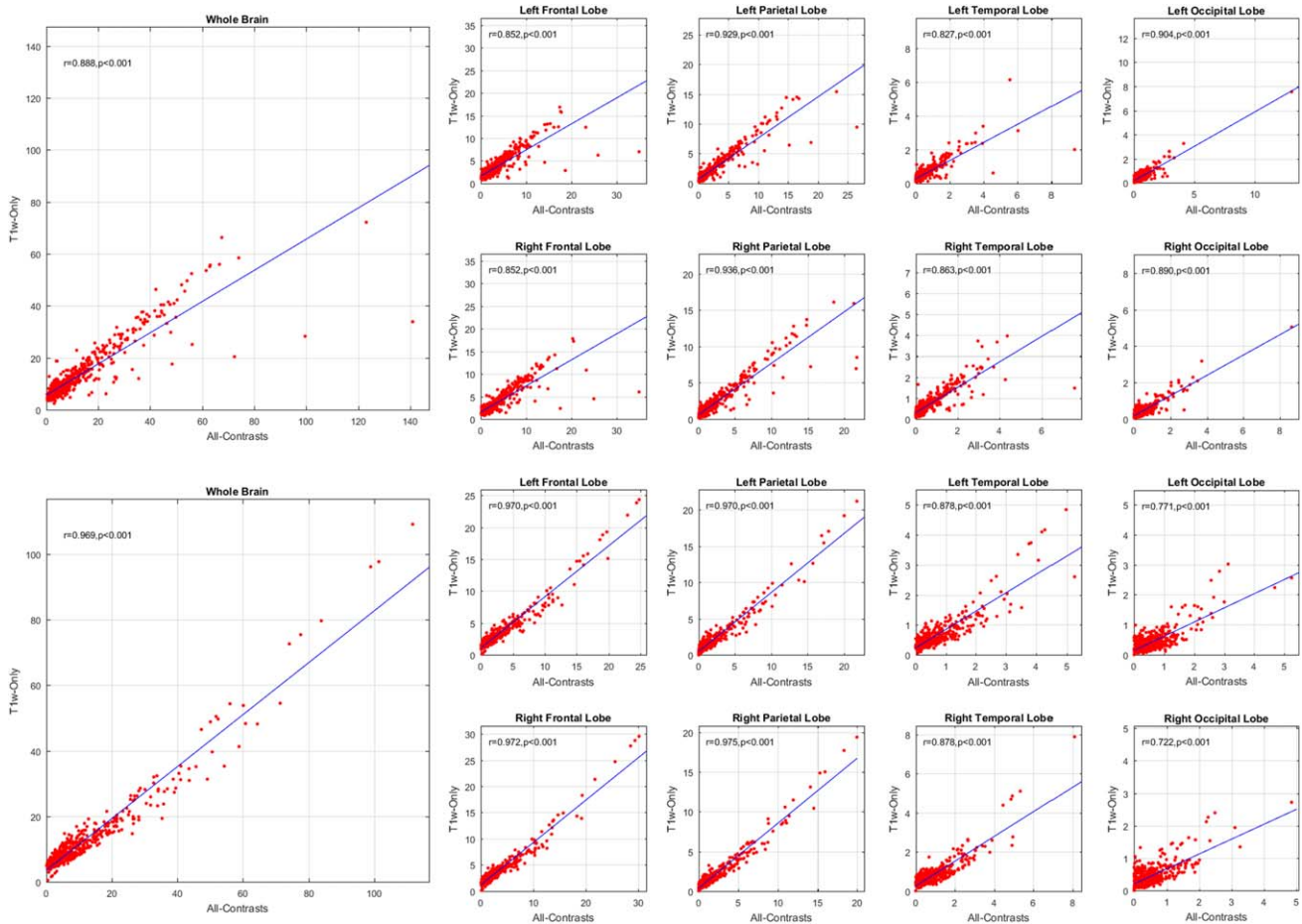
FIGURE 3 Total brain and per lobe correlations for automatic versus manually segmented WMH volumes (CCs), using all contrasts available (red) and using only T1w contrast (blue) for (a) ADC, (b) ADNI1, and (c) ADNI2/GO datasets [Color figure can be viewed at wileyonlinelibrary.com]

the lesion mask. In ADNI1, T2w images were used as the primary contrast with T1w and PD used to aid in the decision process. The rater defined the voxel of interest according to anatomical location and intensity information in all given MRI modalities. To be considered a WMH, a given voxel had to be hyperintense in relation to the surrounding NAWM on T2w, PD, or FLAIR. The same voxel had to be iso to hypointense in relation to the NAWM on T1w images. Previous work (Dadar et al., 2017a) showed that intrarater Dice Kappa was 0.72, 0.80, and 0.86 for ADC, ADNI1, and ADNI2/GO, respectively.

2.6 | Automatic segmentation

A previously validated fully automatic WMH segmentation technique was used to automatically segment the WMHs in all three datasets using a set of intensity and spatial features and a Random Forest classifier (Dadar et al., 2017a, 2017b). The intensity features include voxel intensity for all available modalities, the probability of a specific intensity value being a WMH (p_{WMH}) or non-WMH ($p_{non-WMH}$) for each available modality, and the ratio of these two probabilities for each available modality.

FIGURE 4 Total brain and per lobe correlations for automatically segmented WMHs volumes (CCs) based on all available contrast images versus only T1w contrast image for ADNI1 (left) and ADNI2/GO (right) datasets [Color figure can be viewed at wileyonlinelibrary.com]

The spatial features include a spatial WMH probability map indicating the probability of a voxel at a specific location being a WMH and the average intensity of a non-WMH voxel at that specific voxel location for each available modality. After preprocessing and co-registration of all available modalities, these spatial and intensity features are calculated for each modality. The Random Forest classifier is then trained using these features to segment the WMHs (Dadar et al., 2017a).

For each dataset, the automatic technique was first trained and validated based on the manual segmentations. Two sets of automatic segmentations were completed, first with all available modalities (referred to as All-Contrasts segmentations) and second without using the T2w, PD, and FLAIR information (referred to as T1w-Only segmentations). The trained classifiers were then used to segment the entire ADC, ADNI1, and ADNI2/GO datasets. The quality of the segmentations was then assessed and verified by a human expert. The volumes of the WMHs for the left and right frontal, parietal, temporal, and occipital lobes and the entire brain were calculated by nonlinearly warping the Hammers atlas (Hammers et al., 2003) to the T1w scans of individual subjects and normalized for head size to make population comparisons possible.

The WMH volumes obtained from the T1w-Only segmentations were then correlated with All-Contrasts volumes and clinical scores. False discovery rate (FDR) correction was performed to correct for all multiple comparisons separately for each dataset (significance threshold = 0.05). All the correlations with clinical scores were performed with log transformed WMHs (to achieve normal distribution), controlling for age.

The key concern when using T1w-Only segmentations is whether the WMH portions that are not captured by T1w-Only segmentations are clinically significant. To assess whether the difference between the WMH volumes provided by T1w-Only and All-Contrast segmentations is statistically significant, a general linear model was used to regress the contrast (WMHL$_{\text{All-Contrast}}$-WMHL$_{\text{T1w-Only}}$) and each measure.

Figure 1 shows the axials slices of T1w, T2w, PD, and FLAIR images for a subject from ADC, T1w, T2w, and PD images for a subject from ADNI1 and T1w and FLAIR images for a subject from ADNI2/GO, along with the manual segmentations, as well as T1w-Only and All-Contrasts automated segmentations. While All-Contrasts segmentations conform very well with the manual labels, the T1w-Only segmentations seem to mostly capture the brightest of the WMHs.

3 | RESULTS

3.1 | Comparison of tissue histograms

Here we compare the histograms of WMHs with white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF) density
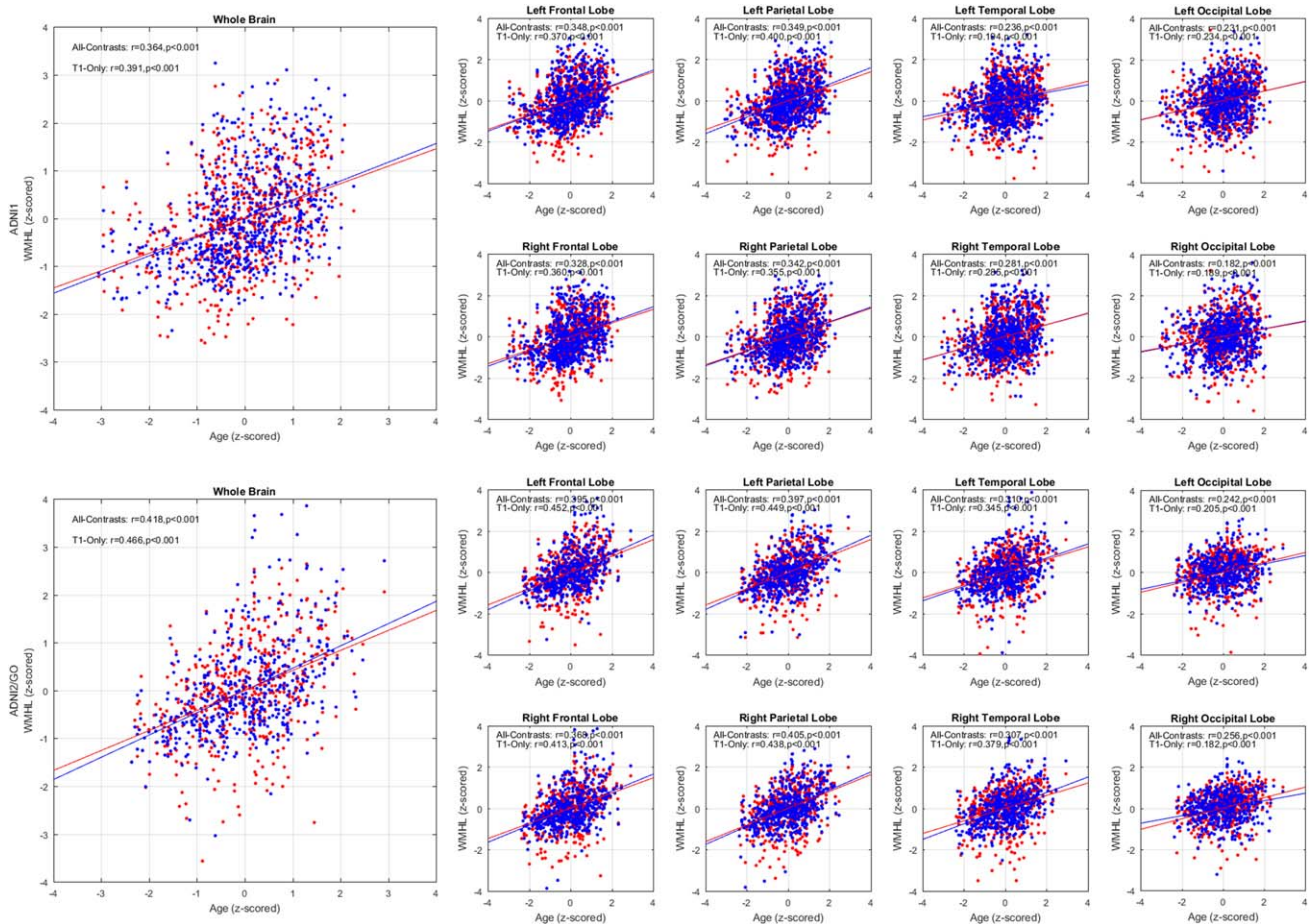
FIGURE 5 Total brain and per lobe correlations for *z*-scored log-transformed automatically segmented WMH volumes correlated with age (*z*-scored) based on all available contrasts (red) versus only T1w images (blue) for ADNI1 (left) and ADNI2/GO (right) datasets. WMHL = white matter hyperintensity load [Color figure can be viewed at wileyonlinelibrary.com]

histograms using the manually segmented labels for each dataset (Fig. 2). Table 3 shows the percentage of overlap between the density histograms of WMHs and WM, GM, and CSF. The tissue histograms show the greatest separation of WMH with GM, WM, and CSF in FLAIR contrast images, followed by T2w, PD, and T1w.
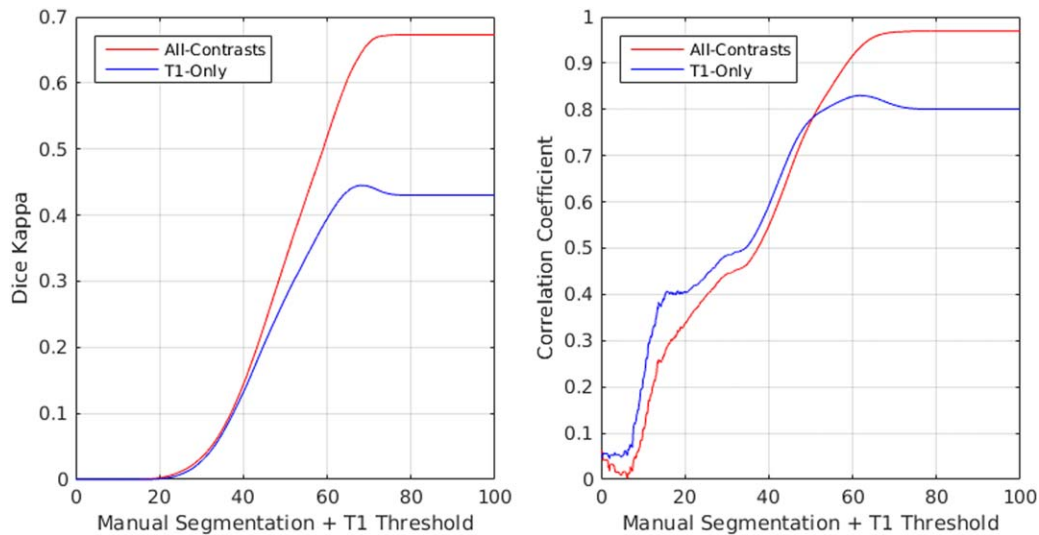
FIGURE 6 Dice kappa and volumetric correlation between T1w-Only and All-contrasts segmentations and thresholded manual labels based on T1w image intensity [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 4 Coefficients of correlation between WMH loads in different lobes and cognitive measures for ADNI1 subjects, controlling for age

	Test	LFL	RFL	LPL	RPL	LTL	RTL	LOL	ROL	WB
T1+T2+PD	MMSE	−0.165*	−0.171*	−0.139*	−0.145*	−0.081*	−0.090*	−0.125*	−0.134*	−0.174*
	ADAS11	0.203*	0.199*	0.183*	0.192*	0.108*	0.131*	0.129*	0.172*	0.217*
	ADAS13	0.219*	0.217*	0.201*	0.199*	0.127*	0.151*	0.141*	0.190*	0.236*
	FAQ	0.209*	0.215*	0.171*	0.192*	0.109*	0.128*	0.139*	0.208*	0.216*
	RAVLTI	−0.174*	−0.179*	−0.161*	−0.163*	−0.119*	−0.168*	−0.130*	−0.157*	−0.197*
	RAVLTF	0.074	0.051	0.062	−0.028	−0.005	−0.005	−0.023	−0.020	−0.061
	RAVLTL	−0.084*	−0.100*	−0.079	−0.079	−0.070	−0.085*	−0.084*	−0.095*	−0.103*
	Executive	−0.111*	−0.110*	−0.121*	−0.107*	−0.112*	−0.113*	−0.182*	−0.139*	−0.130*
	NPI	0.210*	0.199*	0.279*	0.273*	0.270*	0.239*	0.242*	0.221*	0.277*
Only T1	MMSE	−0.203*	−0.211*	−0.127*	−0.147*	−0.031	−0.057	−0.144*	−0.162*	−0.187*
	ADAS11	0.231*	0.230*	0.149*	0.180*	0.054	0.083*	0.132*	0.146*	0.217*
	ADAS13	0.251*	0.251*	0.168*	0.186*	0.069	0.092*	0.148*	0.156*	0.236*
	FAQ	0.226*	0.238*	0.150*	0.182*	0.065	0.089*	0.160*	0.224*	0.218*
	RAVLTI	−0.220*	−0.217*	−0.166*	−0.179*	−0.100*	−0.130*	−0.133*	−0.141*	−0.217*
	RAVLTF	0.074	0.067	0.054	0.043	0.053	0.033	0.036	0.044	0.073
	RAVLTL	−0.109*	−0.113*	−0.059	−0.070	−0.020	−0.037	−0.070	−0.084*	−0.097*
	Executive	−0.128*	−0.135*	−0.117*	−0.113*	−0.018	−0.077	−0.124*	−0.092*	−0.130*
	NPI	0.228*	0.237*	0.308*	0.300*	0.211*	0.200*	0.200*	0.228*	0.290*

Note. WMH = white matter hyperintensity; LFL/RFL = left/right frontal lobe; LPL/RPL = left/right parietal lobe; LTL/RTL = left/right temporal lobe; LOL/ROL = left/right occipital lobe; WB = whole brain; MMSE = Mini-Mental State Examination; ADAS = Alzheimer's Disease Assessment Scale; FAQ = Functional Assessment Questionnaire; RAVLT = Rey Auditory Verbal Learning Task (I = Immediate, F = Forgetting, L = Learning); Executive = executive function (Gibbons et al., 2012); NPI = Total Neuropsychiatric Inventory Score.

3.2 | Comparisons with manual segmentations

To assess the importance of using the T2w, PD, and FLAIR information, the WMH loads obtained from segmentations with and without the information of the optimal modality (i.e., T2w/PD and FLAIR sequences for ADNI1 and ADNI2/GO datasets, respectively) were correlated with the equivalent volumes obtained from the manual segmentations. Correlations were computed for total brain WMH, and for lobar WMH loads. Figure 3 shows this information for the three datasets. While the T1w-Only volumes systematically underestimate the gold standard volumes (obtained from the manual segmentations), they are still able to retain high correlations in all regions and datasets ($r = .963$, $p < .001$ for ADC, $r = .743$, $p < .001$ for ADNI1, and $r = .904$, $p < .001$ for ADNI2/GO for whole brain T1w-Only correlations). The WMH volume range (i.e., the range of the WMH volumes obtained from T1w-Only and All-Contrasts segmentations) is also substantially reduced (e.g., 0–42.5 CCs vs 0–124 CCs for ADC, 0–74 CCs vs 0–116 CCs for ADNI1, and 0–67 CCs vs 0–104 CCs for ADC, for T1w-Only and All-Contrasts segmentations, respectively).

3.3 | Large-scale correlations

A correlation analysis of the WMH loads for the whole brain as well as different lobes was performed between the automatic segmentations

obtained with and without T2w/PD and FLAIR information (All-Contrasts and T1w-Only, respectively) for the subjects from ADNI1 and ADNI2/GO datasets. Figure 4 shows the results of these comparisons for each dataset. The correlations were significant for every lobe in both datasets ($r = .888$, $p < .001$ for ADNI1, and $r = .969$, $p < .001$ for ADNI2/GO for whole brain correlations). The results show a trend of under-segmentation that remains consistent with the change in the WMH load. The amount of this underestimation is also highly correlated with the total WMH load for both datasets ($r = .790$, $p < .00001$ for ADNI1 and $r = .717$, $p < .00001$ for ADNI2/GO).

The log-transformed WMH loads were also correlated with age. Figure 5 shows the results of the correlations with age for ADNI1 and ADNI2/GO datasets.

If one considers the T1w intensity profile of the tissue that is segmented as WMH based on FLAIR or T2w images, the intensities range from hypointense to isointense in relation to the neighboring tissue. To investigate whether the T1w-Only segmentations have different sensitivity levels for detecting WMHs (i.e., whether the T1w-Only segmentations label the hypo-intense portion of the WMHs and not the isointense areas), the manual WMH masks created by the expert were thresholded based on T1w intensity of NAWM at different values, reflecting different levels of sensitivity. The resulting WMH masks were then compared with the T1w-Only and All-Contrasts WMH

TABLE 5 Coefficients of correlation between WMH loads in different lobes and cognitive measures for ADNI2/GO subjects, controlling for age

	Test	LFL	RFL	LPL	RPL	LTL	RTL	LOL	ROL	WB
T1+FLAIR	MMSE	−0.153*	−0.124*	−0.145*	−0.137*	−0.095	−0.086	−0.147*	−0.083	−0.152*
	ADAS11	0.210*	0.191*	0.199*	0.189*	0.166*	0.130*	0.231*	0.135*	0.219*
	ADAS13	0.197*	0.184*	0.184*	0.180*	0.161*	0.119*	0.227*	0.138*	0.208*
	FAQ	0.160*	0.148*	0.155*	0.148*	0.132*	0.104*	0.124*	0.099*	0.167*
	RAVLTI	−0.108*	−0.115*	−0.096	−0.104*	−0.064	−0.024	−0.180*	−0.070	−0.121*
	RAVLTF	−0.069	−0.048	−0.077	−0.046	−0.036	−0.087	−0.022	−0.061	−0.072
	RAVLTL	−0.124*	−0.143*	−0.126*	−0.138*	−0.096	−0.102*	−0.088	−0.011	−0.137*
	Executive	−0.057	−0.078	−0.056	−0.074	−0.036	−0.013	−0.104	−0.100	−0.080
	NPI	0.179*	0.204*	0.175*	0.213*	0.200*	0.247*	0.175*	0.195*	0.215*
Only T1	MMSE	−0.126*	−0.094	−0.109*	−0.106*	−0.055	−0.067	−0.097	−0.041	−0.118*
	ADAS11	0.209*	0.183*	0.186*	0.158*	0.136*	0.102*	0.140*	0.097	0.199*
	ADAS13	0.192*	0.165*	0.162*	0.145*	0.120*	0.084	0.126*	0.083	0.179*
	FAQ	0.151*	0.118*	0.113*	0.096	0.071	0.037	0.055	0.051	0.127*
	RAVLTI	−0.115*	−0.105*	−0.101*	−0.112*	−0.057	−0.033	−0.072	−0.012	−0.111*
	RAVLTF	−0.057	−0.056	−0.062	−0.024	−0.037	−0.042	−0.050	−0.073	−0.061
	RAVLTL	−0.158*	−0.156*	−0.094	−0.105*	−0.134*	−0.108*	−0.067	−0.021	−0.142*
	Executive	−0.075	−0.077	−0.069	−0.095	−0.016	−0.055	−0.046	−0.031	−0.083
	NPI	0.125*	0.130*	0.113*	0.152*	0.145*	0.153*	0.034	0.108*	0.144*

Note. WMH = white matter hyperintensity; LFL/RFL = left/right frontal lobe; LPL/RPL = left/right parietal lobe; LTL/RTL = left/right temporal lobe; LOL/ROL = left/right occipital lobe; WB = whole brain; MMSE = Mini-Mental State Examination; ADAS = Alzheimer's Disease Assessment Scale; FAQ = Functional Assessment Questionnaire; RAVLT = Rey Auditory Verbal Learning Task (I = Immediate, F = Forgetting, L = Learning); Executive = executive function; NPI = Total Neuropsychiatric Inventory Score.

segmentations for different threshold values using Dice Kappa similarity measure and volumetric correlation (Fig. 6). The results show higher Dice Kappa values as well as stronger correlations with T1w-Only segmentations at lower threshold values, confirming that the T1w-Only segmentations correspond to the more hypointense regions of the manual WMH masks, with the optimal intensity threshold at ~63 (note that the intensity range of all the images has been normalized to the range of 0–100).

3.3.1 | WMHs and clinical measures

Tables 4 and 5 summarize the results of correlating different cognitive measures and the log-transformed WMH loads in different lobes obtained from All-Contrasts and T1w-Only segmentations, controlling for age for ADNI1 and ADNI2/Go datasets, respectively. The "*" indicates significant correlations after FDR correction, controlling for age for ADNI1 and ADNI2/GO datasets. Figure 7 shows an example of these correlations for ADAS13 score.

To assess whether the difference between the T1w-Only and All-Contrast WMH volumes is statistically significant, a general linear model was used to regress the contrast and each measure. None of these contrasts was significant after correcting for multiple comparisons using FDR.

3.3.2 | WMHs and other measures

The log-transformed total WMH loads were significantly different between subjects with and without cardiovascular risk factors and history of hypertension, for both T1w-Only and All-Contrasts segmentations ($p < .0001$). Figure 8 shows the boxplots of the log-transformed WMH loads in normal control (NC), mild cognitive impairment (MCI), and dementia groups, separately for T1w-Only and All-Contrasts loads in ADNI1 and ADNI2/GO datasets. In all cases, the dementia cohort has significantly larger WMH loads than the other two groups (note that the log-transformed values are plotted here, and that a 0.4 difference in the log-transformed values is equivalent to approximately 3 CCs of WMHs).

Tables 6 summarizes the results of correlating different risk factors with the log-transformed total WMH loads from T1w-Only images and All-Contrasts segmentations, for ADNI1 and ADNI2/Go datasets, respectively. The "*" indicates significant correlations, after multiple comparisons correction using false discovery rate (FDR). None of the contrasts (WMHL_{All-Contrast}-WMHL_{T1w-Only}) was significant after correcting for multiple comparisons.

4 | DISCUSSION

White matter hyperintensities are an important clinical marker of small vessel disease in aging, and patients suffering from stroke and
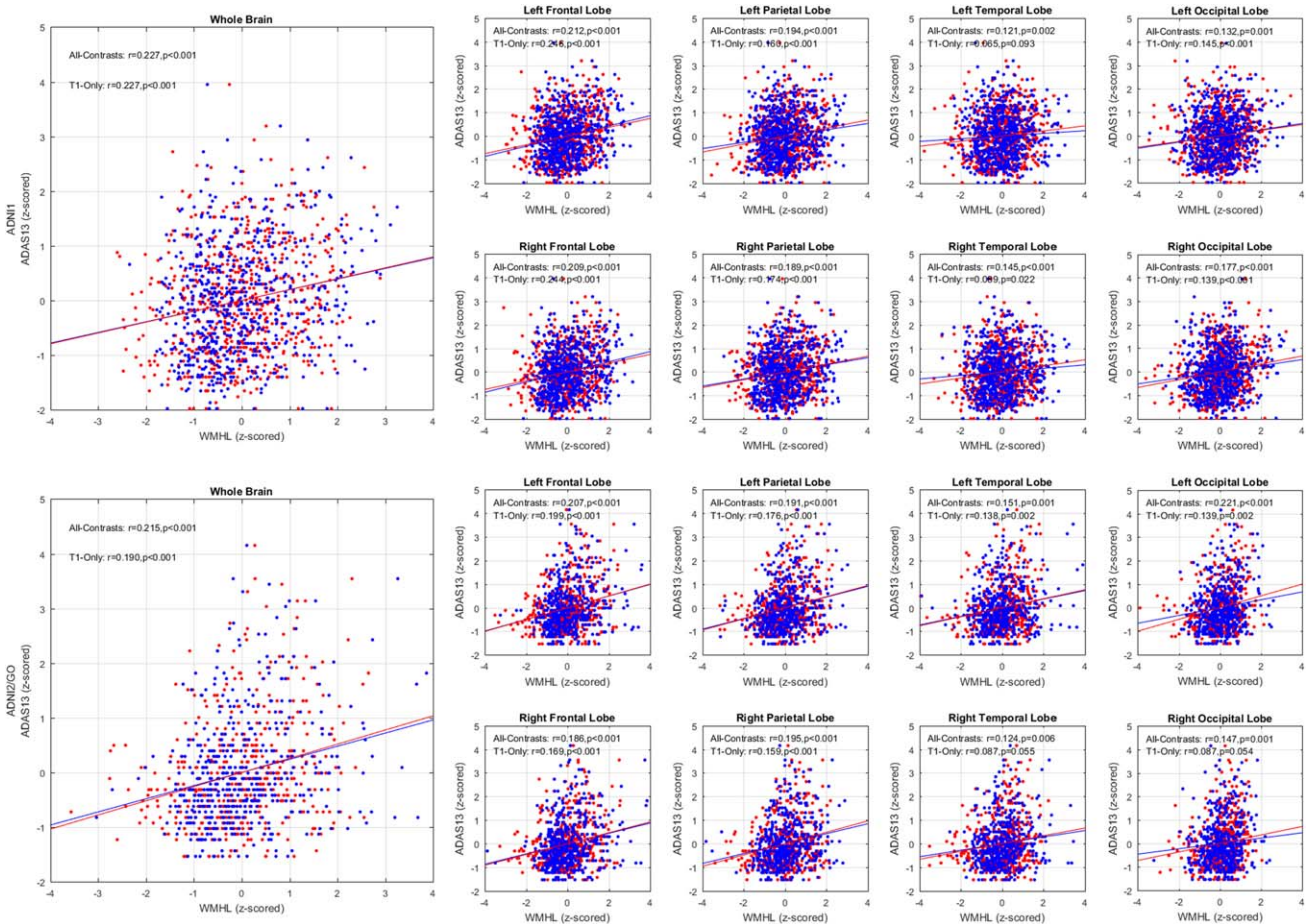
FIGURE 7 Total brain and per lobe correlations for log transformed automatically segmented WMH volumes (CCs) versus ADAS13 while controlling for age, based on all available contrasts (red) and based only on T1w images (blue) for ADNI1 (left) and ADNI2/GO (right) datasets [Color figure can be viewed at wileyonlinelibrary.com]

dementia (Carmichael et al., 2010; DeCarli et al., 1995; Pantoni et al., 2006; van Straaten et al., 2008). In recent years, there has been an increasing interest in using WMHs as an outcome in clinical trials investigating cerebral small vessel disease in the context of stroke and dementia (Debette & Markus, 2010). They reflect the burden of the disease in relation to a small-vessel component (Pantoni, 2010), and are associated with decline in different cognitive domains. Specifically in AD, WMHs are emerging as a potential biomarker of the preclinical vulnerability risk for the disease (Brickman et al., 2012; Deoni, Matthews, & Kolind, 2013; Provenzano et al., 2013).

The optimal MRI sequences for detecting WMHs are FLAIR and T2w/PD scans. However, many previous large-scale datasets have forgone the acquisition of these sequences in favor of other imaging modalities or due to time and financial concerns. WMHs are also visible as hypointense regions on T1w images, but their range of hypointensity is more spread out when compared to the bright signal obtained in FLAIR and T2w/PD, with large isointense areas, or almost isointense in relation with the surrounding NAWM (Fig. 9). This lower contrast between the NAWM tissue and WMHs on T1w sequence makes accurate detection of the full extent of WMHs more challenging in this MRI modality. However, by assessing, even to some extent, the load and location of WMHs using only T1w data, it is possible to study them in

other retrospective datasets, where T2w/PD, or FLAIR information are not available. The key concern when using T1w-Only segmentations would be whether the WMH portions that are not captured by T1w-Only segmentations are clinically significant.

In our study, a previously validated automated technique was used to segment the WMHs with and without the optimal FLAIR and T2w/PD information. In a previous study, we have shown that this automated tool was able to detect WMHs using different combinations of MRI sequences (Dadar et al., 2017a). Here, a random forest classifier was chosen to report our results based on our previous experiments and validations as it had the best performance in detecting WMHs among the studied linear and nonlinear classifiers. However, similar experiments were also performed using the other available classifiers such as AdaBoost which showed similar results regardless of the choice of classifier.

Our results suggest that the measures obtained from using only T1w images underestimate the actual WMH burden, since they only capture the portion of the lesions that co-localizes with the lower intensity signal of the overall T1w hypointense area (Fig. 9). We speculate that these deeper hypointense portions are probably due to more severely affected tissue (i.e., more extensive demyelination, and/or more axonal loss) which are likely to be more clinically relevant (hence
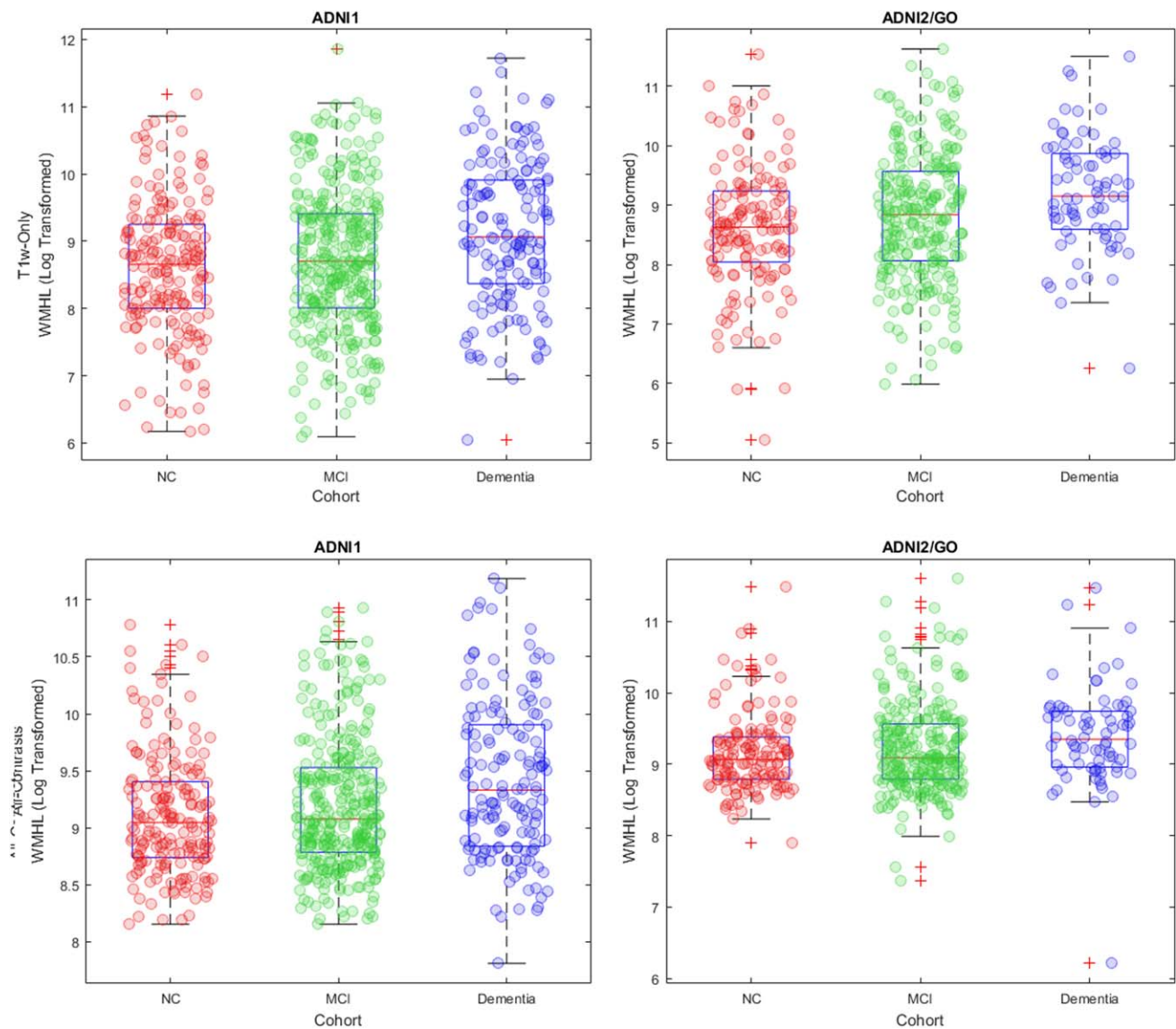
FIGURE 8 Boxplots of log transformed WMHLs for NC, MCI, and dementia cohorts. WMHL = white matter hyperintensity load; NC = normal control; MCI = mild cognitive impairment [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 6 Coefficients of correlation between total WMH loads and measures for ADNI1 and ADNI2/GO subjects

Dataset	ADNI1			ADNI2/GO		
Measure	N	T1 + T2 + PD	T1	N	T1 + FLAIR	T1
Systolic blood pressure	148	0.002	−0.014	409	0.162*	0.199*
Diastolic blood pressure	148	−0.052	−0.027	409	0.103	0.089
Hyperhomocysteinemia	667	0.124*	0.146*	47	0.016	0.048
FDG	342	0.183*	0.135*	448	0.284*	0.261*
AV45	0	-	-	443	0.172*	0.162
Serum glucose	626	0.119*	0.112*	417	0.048	0.043
Cholesterol	557	0.065	0.080	382	0.153*	0.105
CSF protein	333	0.209*	0.125*	73	0.232*	0.247*

Note. WMH = white matter hyperintensity; FDG = fluorodeoxyglucose PET; AV45 = Florbetapir; PET = positron emission tomography; CSF = cerebrospinal fluid.
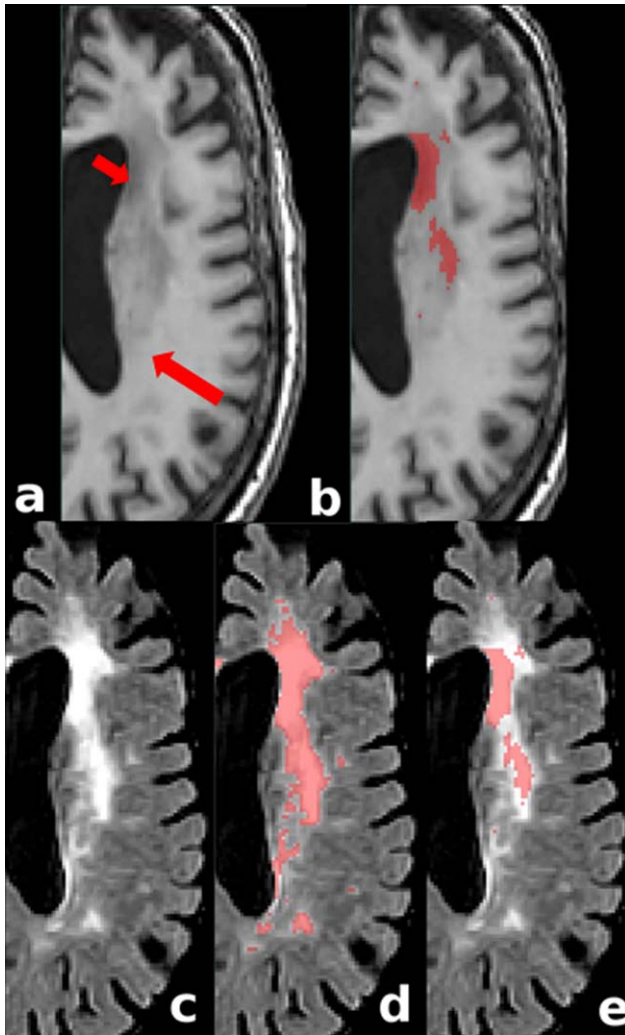
FIGURE 9 Comparing FLAIR hyperintensities and T1w hypointensities. (a) T1w image showing different hypointense values: short and long arrows indicate lower and higher signals, respectively. (b) T1w manual mask corresponding to a given level of hypointensity in T1w image. Note that only the lower signal areas are labeled as WMHs. (c) FLAIR image (d) FLAIR and manual mask corresponding to WMHs detected based on FLAIR image (d) FLAIR and T1w-based mask. Note how the hypointense region of the T1w sequence colocalizes only with a portion of the WMHs on FLAIR even though the signal on the FLAIR sequence is mostly homogeneous [Color figure can be viewed at wileyonlinelibrary.com]

the maintained correlations). Future studies correlating histological specimens and MR T1w information would be necessary to clarify the specific histological substrate of the full range of hypointense signal captured by T1w sequence and the more homogeneous bright signal on FLAIR/T2w/PD. Nevertheless, despite the absence of a histopathology gold standard, T1w-Only WMHs volumes are still able to hold strong correlations with both manual and automatic segmentations obtained using the optimal modalities. Additionally, the assessment of colocalization shows higher Dice Kappa values for the T1w-Only classifications when they are compared with the manual mask thresholded to a percentage value of the T1w NAWM, confirming that the partially

detected WMH area colocalizes with the lower T1w signal voxels (Fig. 6). Similarly, the volumetric correlation of the T1w-Only volumes with the thresholded manual expert volumes reinforces the concept of specific partial detection of hypointense voxels on T1w images.

Regarding the assessment of clinical outcomes, the T1w-Only WMH volumes mostly correlate with age, cognitive, and clinical measures as strongly as the WMH volumes determined using the optimal modalities of FLAIR or T2w/PD. This is true for WMHs in whole brain and in different single brain lobes. Similarly, vascular risk factors show a significant correlation with T1w-Only WMHs volumes, similar to those obtained with all the optimal modalities. This suggest that, although the WMH burden might be underestimated in T1w-Only segmentations, the identified lesions can be used for clinical correlations in datasets where the optimal sequences are not available. This will enable us to generate and use the WMH loads obtained from the T1w images as an estimate of the actual WMH loads in datasets where the FLAIR or T2w/PD information is unavailable, and also in T1w scans that do not have full FLAIR or T2w/PD brain coverage.

Finally, many studies acquire T2w/PD or FLAIR information in their baseline visit, or at longer intervals compared with the T1w scans, which are acquired at every MRI visit. Even though the T1w-Only segmentations systematically underestimate the volume of WMHs, having an estimate of the accurate WMH load from the baseline T2w/PD or FLAIR images, one may correct for this bias to obtain more accurate estimates of WMH loads, for the visits that only have T1w acquisitions. In addition, T1w images are generally acquired at higher resolutions (i.e., 1-mm-thick slices) than the T2w/PD or FLAIR scans which are generally acquired at 3–5 mm slice thicknesses. The higher spatial resolution of T1w images can also be used to obtain more spatially accurate segmentations, where such information may not available due to the higher slice thickness of the T2w/PD and FLAIR images.

Our study indicates that datasets that lack T2w/PD and FLAIR modalities may still benefit from the estimation of WMHs using our T1w-Only segmentation approach, to correlate this MRI data with other clinical variables available for the subjects.

ORCID

Mahsa Dadar http://orcid.org/0000-0003-4008-2672

Owen T. Carmichael http://orcid.org/0000-0002-0576-0047

REFERENCES

Biesbroek, J. M., Weaver, N. A., & Biessels, G. J. (2017). Lesion location and cognitive impact of cerebral small vessel disease. *Clinical Science (London, England: 1979)*, 131, 715–728.

Brickman, A. M., Provenzano, F. A., Muraskin, J., Manly, J. J., Blum, S., Apa, Z., ... Mayeux, R. (2012). Regional white matter hyperintensity volume, not hippocampal atrophy, predicts incident Alzheimer disease in the community. *Archives of Neurology*, 69, 1621–1627.

Caligiuri, M. E., Perrotta, P., Augimeri, A., Rocca, F., Quattrone, A., & Cherubini, A. (2015). Automatic detection of white matter hyperintensities in healthy aging and pathology using magnetic resonance imaging: A review. *Neuroinformatics*, 13, 261–276.

Carmichael, O., Schwarz, C., Drucker, D., Fletcher, E., Harvey, D., Beckett, L., ... DeCarli, C. (2010). Longitudinal changes in white matter disease and cognition in the first year of the Alzheimer disease neuroimaging initiative. *Archives of Neurology*, 67, 1370–1378.

Collins, D. L., & Evans, A. C. (1997). Animal: Validation and applications of nonlinear registration-based segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, 11, 1271–1294.

Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, 18, 192–205.

Dadar, M., Maranzano, J., Misquitta, K., Anor, C. J., Fonov, V. S., Tartaglia, M. C., ... Collins, D. L. and Alzheimer's Disease Neuroimaging Initiative (2017a). Performance comparison of 10 different classification techniques in segmenting white matter hyperintensities in aging. *NeuroImage*, 157, 233–249.

Dadar, M., Pascoal, T., Manitsirikul, S., Misquitta, K., Tartaglia, C., Brietner, J., ... Collins, D. L. (2017b). Validation of a regression technique for segmentation of white matter hyperintensities in Alzheimer's disease. *IEEE Transactions on Medical Imaging*.

De Groot, J. C., De Leeuw, F.-E., Oudkerk, M., Van Gijn, J., Hofman, A., Jolles, J., & Breteler, M. (2002). Periventricular cerebral white matter lesions predict rate of cognitive decline. *Annals of Neurology*, 52, 335–341.

Debette, S., & Markus, H. S. (2010). The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: Systematic review and meta-analysis. *British Medical Journal*, 341, c3666.

DeCarli, C., Murphy, D. G. M., Tranh, M., Grady, C. L., Haxby, J. V., Gillette, J. A., ... Rapoport, S. I. (1995). The effect of white matter hyperintensity volume on brain structure, cognitive performance, and cerebral metabolism of glucose in 51 healthy adults. *Neurology*, 45, 2077–2084.

DeCarli, C., Miller, B. L., Swan, G. E., Reed, T., Wolf, P. A., & Carmelli, D. (2001). Cerebrovascular and brain morphologic correlates of mild cognitive impairment in the National Heart, Lung, and Blood Institute Twin Study. *Archives of Neurology*, 58, 643–647.

Deoni, S. C., Matthews, L., & Kolind, S. H. (2013). One component? Two components? Three? The effect of including a nonexchanging "free" water component in multicomponent driven equilibrium single pulse observation of T1 and T2. *Magnetic Resonance in Medicine*, 70, 147–154.

Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., ... Cummings, J. L. (2014). Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *Lancet Neurology*, 13, 614–629.

Fernando, M. S., O'brien, J. T., Perry, R. H., English, P., Forster, G., McMeekin, W., ... Ince, P. G. (2004). Comparison of the pathology of cerebral white matter with post-mortem magnetic resonance imaging (MRI) in the elderly brain. *Neuropathology and Applied Neurobiology*, 30, 385–395.

Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62, 774–781.

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., & Collins, D. L. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54, 313–327.

Gibbons, L. E., Carle, A. C., Mackin, R. S., Harvey, D., Mukherjee, S., Insel, P., ... Crane, P. K. (2012). A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. *Brain Imaging and Behaviour*, 6, 517–527.

Gouw, A. A., Seewann, A., Van Der Flier, W. M., Barkhof, F., Rozemuller, A. M., Scheltens, P., & Geurts, J. J. (2010). Heterogeneity of small vessel disease: A systematic review of MRI and histopathology correlations. *Journal of Neurology, Neurosurgery, and Psychiatry*, 2009.

Griffanti, L., Zamboni, G., Khan, A., Li, L., Bonifacio, G., Sundaresan, V., ... Jenkinson, M. (2016). BIANCA (Brain Intensity AbNormality Classification Algorithm): A new tool for automated segmentation of white matter hyperintensities. *NeuroImage*, 141, 191–205.

Hammers, A., Allom, R., Koepp, M. J., Free, S. L., Myers, R., Lemieux, L., ... Duncan, J. S. (2003). Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping*, 19, 224–247.

Hawkins, K. A., Emadi, N., Pearlson, G. D., Winkler, A. M., Taylor, B., Dulipsingh, L., ... Blank, K. (2017). Hyperinsulinemia and elevated systolic blood pressure independently predict white matter hyperintensities with associated cognitive decrement in the middle-aged offspring of dementia patients. *Metabolic Brain Disease*, 1–9.

Hinton, L., Carter, K., Reed, B. R., Beckett, L., Lara, E., DeCarli, C., & Mungas, D. (2010). Recruitment of a community-based cohort for research on diversity and risk of dementia. *Alzheimer Disease and Associated Disorders*, 24, 234.

Iturria-Medina, Y., Sotero, R. C., Toussaint, P. J., Mateos-Pérez, J. M., & Evans, A. C. (2016). Early role of vascular dysregulation on late-onset

Alzheimer's disease based on multifactorial data-driven analysis. *Nature Communications*, 7, 11934.

Jacobs, H. I. L., Leritz, E. C., Williams, V. J., Van Boxtel, M. P. J., Elst, W. V D., Jolles, J., . . . Salat, D. H. (2013). Association between white matter microstructure, executive functions, and processing speed in older adults: The impact of vascular health. *Human Brain Mapping*, 34, 77–95.

Leritz, E. C., Shepel, J., Williams, V. J., Lipsitz, L. A., McGlinchey, R. E., Milberg, W. P., & Salat, D. H. (2014). Associations between T1 white matter lesion volume and regional white matter microstructure in aging. *Human Brain Mapping*, 35, 1085–1100.

Longstreth, W. T., Manolio, T. A., Arnold, A., Burke, G. L., Bryan, N., Jungreis, C. A., . . . Fried, L. (1996). Clinical correlates of white matter findings on cranial magnetic resonance imaging of 3301 elderly people. *Stroke*, 27, 1274–1282.

Manjón, J. V., Coupé, P., Martí-Bonmatí, L., Collins, D. L., & Robles, M. (2010). Adaptive non-local means denoising of MR images with spatially varying noise levels. *Journal of Magnetic Resonance Imaging*, 31, 192–203.

Maranzano, J., Rudko, D. A., Arnold, D. L., & Narayanan, S. (2016). Manual segmentation of MS cortical lesions using MRI: A comparison of 3 MRI reading protocols. *American Journal of Neuroradiology*, 37, 1623–1628.

Pantoni, L. (2010). Cerebral small vessel disease: From pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurology*, 9, 689–701.

Pantoni, L., Sarti, C., Alafuzoff, I., Jellinger, K., Munoz, D. G., Ogata, J., & Palumbo, V. (2006). Postmortem examination of vascular lesions in cognitive impairment. *Stroke*, 37, 1005–1009.

Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., . . . Weiner, M. W. (2010). Alzheimer's disease neuroimaging initiative (ADNI) clinical characterization. *Neurology*, 74, 201–209. Alzheimer's

Prins, N. D., & Scheltens, P. (2015). White matter hyperintensities, cognitive impairment and dementia: An update. *Nature Reviews. Neurology*, 11, 157–165.

Provenzano, F. A., Muraskin, J., Tosto, G., Narkhede, A., Wasserman, B. T., Griffith, E. Y., . . . Brickman, A. M. (2013). White matter

hyperintensities and cerebral amyloidosis: Necessary and sufficient for clinical expression of Alzheimer disease? *JAMA Neurology*, 70, 455–461.

Salat, D. H., Tuch, D. S., van der Kouwe, A. J. W., Greve, D. N., Pappu, V., Lee, S. Y., . . . Rosas, H. D. (2010). White matter pathology isolates the hippocampal formation in Alzheimer's disease. *Neurobiology of Aging*, 31, 244–256.

Schmidt, P., Gaser, C., Arsic, M., Buck, D., Förschler, A., Berthele, A., . . . Mühlau, M. (2012). An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. *NeuroImage*, 59, 3774–3783.

Simões, R., Mönninghoff, C., Dlugaj, M., Weimar, C., Wanke, I., van Walsum, A.-M., . . . Lump, C. (2013). Automatic segmentation of cerebral white matter hyperintensities using only 3D FLAIR images. *Magnetic Resonance Imaging*, 31, 1182–1189.

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, 17, 87–97.

van Straaten, E. C., Harvey, D., Scheltens, P., Barkhof, F., Petersen, R. C., Thal, L. J., . . . DeCarli, C. (2008). Periventricular white matter hyperintensities increase the likelihood of progression from amnestic mild cognitive impairment to dementia. *Journal of Neurology*, 255, 1302.

Takao, M., Koto, A., Tanahashi, N., Fukuuchi, Y., Takagi, M., & Morinaga, S. (1999). Pathologic findings of silent hyperintense white matter lesions on MRI. *Journal of the Neurological Sciences*, 167, 127–131.

Yoshita, M., Fletcher, E., & DeCarli, C. (2005). Current concepts of analysis of cerebral white matter hyperintensities on magnetic resonance imaging. *Topics in Magnetic Resonance Imaging*, 16, 399.