White matter signal abnormality quality differentiates mild cognitive impairment that converts to Alzheimer's disease from nonconverters

Emily R. Lindemer [a,b,c,*], David H. Salat [b,c,d], Eric E. Smith [e], Khoa Nguyen [b,c], Bruce Fischl [a,b,c,f], Douglas N. Greve [b,c], for the Alzheimer's Disease Neuroimaging Initiative[1]

[a] Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA
[b] Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA
[c] Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[d] VA Boston Healthcare System, Boston, MA, USA
[e] Department of Clinical Neurosciences and Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada
[f] Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology (MIT), Cambridge, MA, USA

ARTICLE INFO

ABSTRACT

The objective of this study was to assess how longitudinal change in the quantity and quality of white matter signal abnormalities (WMSAs) contributes to the progression from mild cognitive impairment (MCI) to Alzheimer's disease (AD). The Mahalanobis distance of WMSA from normal-appearing white matter using T1-, T2-, and proton density–weighted MRI was defined as a quality measure for WMSA. Cross-sectional analysis of WMSA volume in 104 cognitively healthy older adults, 116 individuals with MCI who converted to AD within 3 years (mild cognitive impairment converter [MCI-C]), 115 individuals with MCI that did not convert in that time (mild cognitive impairment nonconverter [MCI-NC]), and 124 individuals with AD from the Alzheimer's Disease Neuroimaging Initiative revealed that WMSA volume was substantially greater in AD relative to the other groups but did not differ between MCI-NC and MCI-C. Longitudinally, MCI-C exhibited faster WMSA quality progression but not volume compared with matched MCI-NC beginning 18 months before MCI-C conversion to AD. The strongest difference in rate of change was seen in the time period starting 6 months before MCI-C conversion to AD and ending 6 months after conversion ($p < 0.001$). The relatively strong effect in this time period relative to AD conversion in the MCI-C was similar to the relative rate of change in hippocampal volume, a traditional imaging marker of AD pathology. These data demonstrate changes in white matter tissue properties that occur within WMSA in individuals with MCI that will subsequently obtain a clinical diagnosis of AD within 18 months. Individuals with AD have substantially greater WMSA volume than all MCI suggesting that there is a progressive accumulation of WMSA with progressive disease severity, and that quality change predates changes in this total volume. Given the timing of the changes in WMSA tissue quality relative to the clinical diagnosis of AD, these findings suggest that WMSAs are a critical component for this conversion and are a critical component of this clinical syndrome.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

White matter (WM) damage is a common occurrence in older adults and is often incidentally detected through magnetic resonance imaging (MRI). These patches of tissue damage are increasingly recognized as a substantial correlate of age-associated cognitive decline (de Leeuw et al., 2001; Frisoni et al., 2007; Grueter and Schulz, 2012). Although most often recognized as a hyperintense signal on T2 and fluid-attenuated inversion recovery MRI, this damage can also appear as hypointense on T1-weighted images and therefore, are

* Corresponding author at: Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA. Tel.: (617) 726 3197; fax: (617) 726 7422.
   E-mail address: lindemer@mit.edu (E.R. Lindemer).

more generically referred to as white matter signal abnormalities (WMSAs). Vascular dysfunction is a primary mechanism of WM damage in older adults, and this damage is likely to contribute to the development of cognitive impairment and dementia (Barker et al., 2014; Bocti et al., 2005; Brickman et al., 2009a, 2009b; Debette and Markus, 2010; DeCarli et al., 2005; Delano-Wood et al., 2009; Frisoni et al., 2007; Gurol et al., 2006; Iadecola, 2013; Levy-Cooperman et al., 2008; Tullberg et al., 2004; Yoshita et al., 2006). Little is known, however, about the development and time course of WM tissue damage in patients transitioning from normal cognition to dementia. Additionally, little is known regarding the degree to which WMSAs contribute to different states of cognitive decline—for example, whether WMSAs are associated with conversion from a cognitively healthy status to mild cognitive impairment (MCI), or from MCI to dementia such as Alzheimer's disease (AD) (DeCarli et al., 1996; Medina et al., 2006; Smith et al., 2008; Wolf et al., 2000).

There is currently a strong interest in predicting individuals that will eventually develop AD as it is now assumed that trials for novel therapeutics will require intervention before substantial neurodegeneration (Jack et al., 2013). Studies using structural imaging have demonstrated a greater total WM damage burden in individuals with AD compared with those with MCI (Huang et al., 2012; Pievani et al., 2010; Prins et al., 2004). Several groups have highlighted promising evidence of WMSAs potentially being a structural predictor of AD development (Carmichael et al., 2010); some suggesting that this measure is on par with hippocampal volume (a traditional imaging marker of AD) (Brickman et al., 2012; Canu et al., 2012).

Neuropathology studies demonstrate histopathologic and imaging heterogeneity within and across WM lesions (Gouw et al., 2011; Iadecola, 2013; Maillard et al., 2014; Pettersen et al., 2008; Viswanathan, 2014; Wardlaw et al., 2013; Young et al., 2008), suggesting that quantifying WMSA on MRI by measuring total volume, as is usually done, may not accurately reflect the total severity of the damage. For example, on a T2-weighted scan some damaged tissue may evince signal intensities as bright as fluid, while other locations exhibit just a slight brightening relative to normal-appearing white matter (NAWM) intensities. These irregular signal properties contribute to poorly-defined boundaries and create difficulties in automatically segmenting damaged tissue from healthy tissue. Previous attempts at automatic segmentation have obtained increased sensitivity and specificity using a combination of T1-weighted, T2-weighted, proton density (PD), and fluid-attenuated inversion recovery imaging, demonstrating advantages of a multi-spectral approach (Maillard et al., 2008; Schwarz et al., 2009). Additionally, the current "gold standard" for WMSA quantification and validation is the manual delineation of WMSA labels. This is labor intensive and suffers from poor levels of both inter-rater and intrarater reliability (Grimaud et al., 1996; Zijdenbos et al., 1994). Thus, improvements in the automated segmentation and quantification of WMSAs would contribute to increased reliability, and potentially enhance the clinical utility of this marker of tissue damage (García-Lorenzo et al., 2013; Mortazavi et al., 2012).

In the present study, we examine how the quality of WMSAs change over time with respect to NAWM via a novel image quantification technique that implements the Mahalanobis distance (MD) of WMSA to NAWM. The implementation of WM lesion quality to study cognitive decline is relatively novel but is well-supported by studies describing heterogeneity in healthy WM as well as within lesions in the context of vascular integrity, normal aging (Chen et al., 2013; Spilt et al., 2006), and cognitive impairment (Delano-Wood et al., 2009; Viswanathan, 2014). We demonstrate here for the first time that WM damage and within-lesion quality in individuals with MCI changes in a manner that is closely timed to their conversion to a clinical diagnosis of AD. Furthermore, associations between WMSA quality and MCI conversion are similar to those exhibited by hippocampal volume, a known structural imaging indicator of AD (Convit et al., 1997; Gosche et al., 2002; Jack et al., 1999). These novel findings provide critical information for understanding the pathophysiology of the clinically manifested AD dementia syndrome, as WM damage is not considered in traditional models (e.g., Jack et al., 2013) of AD pathology and may provide an important and tenable mechanism for therapeutic targeting.

2. Materials and methods

2.1. Data

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and nonprofit organizations, as a $60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The principal investigator of the ADNI initiative is Michael W. Weiner, MD, VA Medical Center and University of California—San Francisco. ADNI is the result of efforts of many coinvestigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the United States and Canada. The initial goal of ADNI was to recruit 800 subjects, but ADNI has been followed by ADNI Grand Opportunity (ADNI-GO) and ADNI-2. To date, these 3 protocols have recruited >1500 adults, ages 55—90 years, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow-up duration of each group is specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

All data in the present study are taken from ADNI-1. A cross-sectional analysis was conducted using images across 4 diagnostic groups, and a longitudinal analysis was then conducted on 2 of these groups. The first set of images comprised a single scanning time point for 459 individuals. These data encompass individuals who fall into 1 cognitive status categories: (1) older controls without clinical diagnosis during the study (other controls [OC], n = 104), (2) MCI without conversion to AD during the course of the study (mild cognitive impairment nonconverters [MCI-NC], n = 116), (3) MCI with conversion to AD during the study (mild cognitive impairment converters [MCI-C], n = 115), and (4) those diagnosed with AD throughout the study (AD, n = 124) as described by ADNI (www.adni-info.org). Briefly, all MCI participants have reported a subjective memory concern either autonomously or via an informant or clinician but do not have significant levels of impairment in other cognitive domains and have essentially preserved activities of daily living with no signs of dementia (i.e., all MCI individuals are amnestic MCI only). AD participants were evaluated and met the National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association criteria for probable AD. Through this evaluation process, ADNI aims to reduce the risk of including subjects with vascular and other types of dementia.

The data for the longitudinal analysis comprised all 115 MCI-C individuals and 115 of the MCI-NC individuals from the cross-sectional analysis and was used to study the progression (we use the term "progression" to refer to the change of a voxel from NAWM to WMSA. In contrast, we use the term "conversion" to refer to the change of diagnosis of a subject from MCI to AD) of WMSAs. These individuals all had between 3 and 6 sets of longitudinal imaging data that were acquired at months 0, 6, 12, 18, 24, and 36 during the ADNI study. Each MCI-C individual converted to AD at some point along this time course. For the longitudinal analysis, all MCI-C individuals were aligned to each other based on time of AD conversion. Next, each MCI-C individual was age-matched to an MCI-NC individual and these were also aligned to the AD conversion time point of the corresponding MCI-C individuals, for a total-study timeline of 66 months (36 months before MCI-C conversion to AD until 30 months after AD conversion; Fig. 1). This matching procedure was done to analyze structural changes that occur in the time surrounding AD conversion in individuals with MCI while controlling for the strong effects that age is known to have on brain structure. For each time point, data were only included if both the MCI-C individual and the matched MCI-NC individual had imaging data, leading to a differing number of data points for each time point. Because of this, the 60-month and 66-month time points only had 7 and 4 subjects per group, respectively and were not included in the subsequent analyses. Subjects in the present study were included from ADNI-1 based on the availability of at least 3 longitudinal T1, T2, and PD data sets that were of high enough quality for preprocessing through the robust FreeSurfer longitudinal registration stream (Reuter et al., 2012).

2.2. MRI acquisition

All data were acquired on a 1.5-T scanner at rigorously validated sites, which all followed a previously described standardized protocol (Jack et al., 2008). The protocol included a high-resolution, T1-weighted sagittal volumetric magnetization prepared rapid gradient echo sequence and axial PD and/or T2-weighted fast spin echo sequence. The ADNI MRI core optimized the acquisition parameters of these sequences for each make and model of scanner included in the study. All scanner sites were required to pass a strict scanner validation test before being allowed to scan ADNI participants. Additionally, each scan of ADNI participants included a scan of the phantom, which was required to pass strict validation tests.

2.3. MRI preprocessing

The WMSA segmentation procedure was developed within the FreeSurfer image analysis suite, which allowed this procedure to be performed in the context of robust whole-brain labeling. Cortical reconstruction and volumetric segmentation was performed using FreeSurfer (surfer.nmr.mgh.harvard.edu/, version 5.1). The technical details of these procedures are described in prior publications (Dale and Sereno, 1993; Dale et al., 1999; Fischl and Dale, 2000; Fischl et al., 1999, 2001, 2002, 2004a, 2004b; Han et al., 2006; Jovich et al., 2006).

Data were then processed by the longitudinal FreeSurfer stream (Reuter et al., 2012). Specifically, an unbiased within-subject template space and image was created using robust, inverse consistent registration (Reuter et al., 2010). Several processing steps, such as skull stripping, Talairach transforms, atlas registration, and spherical surface maps and parcellations were then initialized with common information from the within-subject template which significantly increases reliability and statistical power (Reuter et al., 2012).

2.4. Atlas creation

The WMSA atlas was created as an extension of FreeSurfer's unimodal T1-weighted atlas used for standard automatic segmentation of healthy gray and WM structures (Fischl et al., 2002). WMSAs were manually labeled on 7 subjects by a trained expert (Eric E. Smith) using T1, T2, and PD images as a guide. WMSAs were defined based on consensus guidelines for measurement of WM hyperintensities of presumed vascular origin (Wardlaw et al., 2013). Using these new WMSA labels in conjunction with all standard FreeSurfer labels, a multimodal Gaussian classifier array (MMGCA)
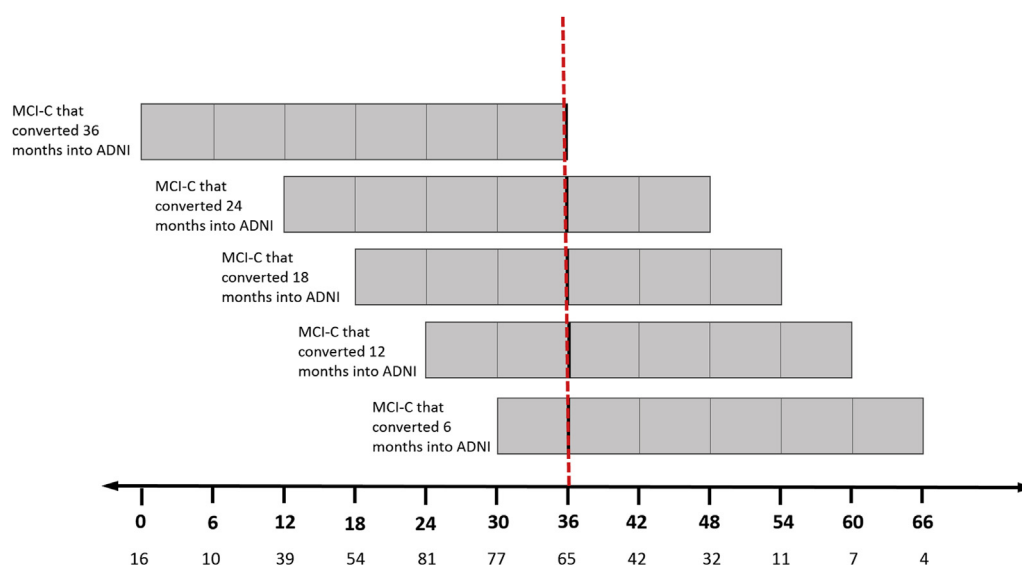

Fig. 1. MCI-C individual alignment to point of AD conversion on a total-study timeline. Bold numbers on the x-axis are months. Light numbers represent the number of data sets from MCI-C individuals at each time point. Each gray bar represents the set of all MCI-C individuals who converted at the same time point on the ADNI timeline, and each dividing line represents a data collection point during the ADNI study. The red dashed line indicates the time of AD conversion. Each MCI-C was then matched for age and sex with an MCI-NC, and this MCI-NC was aligned with its MCI-C counterpart on the total-study timeline. Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; AD, Alzheimer's disease; MCI-C, mild cognitive impairment converters. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

was created that contained a 3-dimensional covariance matrix (T1, T2, and PD) for each structure at each voxel in addition to spatial and neighborhood prior information. The following results present volumes in terms of number of voxels, where 1 voxel corresponds to 1 mm^3 of tissue.

2.5. Automatic WMSA segmentation, classification, and quantification

Visual guides to supplement the following automatic segmentation methodology are provided in Fig. 2. A voxel is first classified based on the maximum a posteriori probability that it is in a given class given the voxel's T1/T2/PD intensity profile and the MMGCA probabilities at that location. We note that, this first step extends prior work with T1/T2/PD modalities in WMSA labeling (Maillard et al., 2008; Schwarz et al., 2009). Although this simple atlas approach successfully labels many WMSAs, it is not sufficient generally because WMSAs can occur anywhere in WM (not just in likely atlas locations) and because the intensities are highly variable (and so not fit well by a Gaussian model; Fig. 2C.3). We, therefore, follow up the MMGCA procedure with several refinements designed to catch unlabeled WMSAs. These refinements rely heavily on the MD (Mahalanobis, 1936) of a WMSA voxel from NAWM. The MD gives a voxel's distance away from a distribution and is defined as:

$$MD = (y_v - \mu_H)^T \Sigma_H^{-1} (y_v - \mu_H) \qquad (1)$$

where, y_v is the vector of multimodal signal intensities at voxel v, μ_H is the vector of mean multimodal intensities in NAWM, and $\sum_H$ is

the covariance matrix of the multimodal intensities in NAWM (Fig. 2B). We interpret the MD as a measure of WM damage.

In the first refinement step, non-WMSA neighbors of voxels initially labeled as WMSA were examined. Using a heuristic that combined information regarding the voxel's number of WMSA neighbors, its intensity values in relation to the MMGCA intensity values for other tissue types, and the MD of the intensity values from other tissue types defined in the MMGCA, they were relabeled as WMSA or left unchanged. In the second refinement step, the MMGCA atlas parameters were abandoned in favor of using statistics from the subject's own NAWM and WMSA as a reference. It follows a similar region-growing procedure as in the first step but uses these individual-based values instead (Fig. 2A). This yields a final binary labeling of WMSAs. Although the final labeling retains some "false" positive (FP) labeling (Fig. 2C.4) compared with the manual label, analyses described in the following Section 2.6 suggest that a portion of these voxels are in fact unique and are representative of changing tissue. Difference in signal tissue intensity can be seen in the subject's T1-weighted image in these FP regions (Fig. 2C.1).

2.6. Validation

We used 2 methods to validate the automatic procedure. First, a set of manual labels was created from 9 of the 459 ADNI subjects by a trained labeler (Khoa Nguyen; example in Fig. 2C.2). This data was used as an independent test set for the WMSA segmentation algorithm. For each of these data sets, the overlap between the
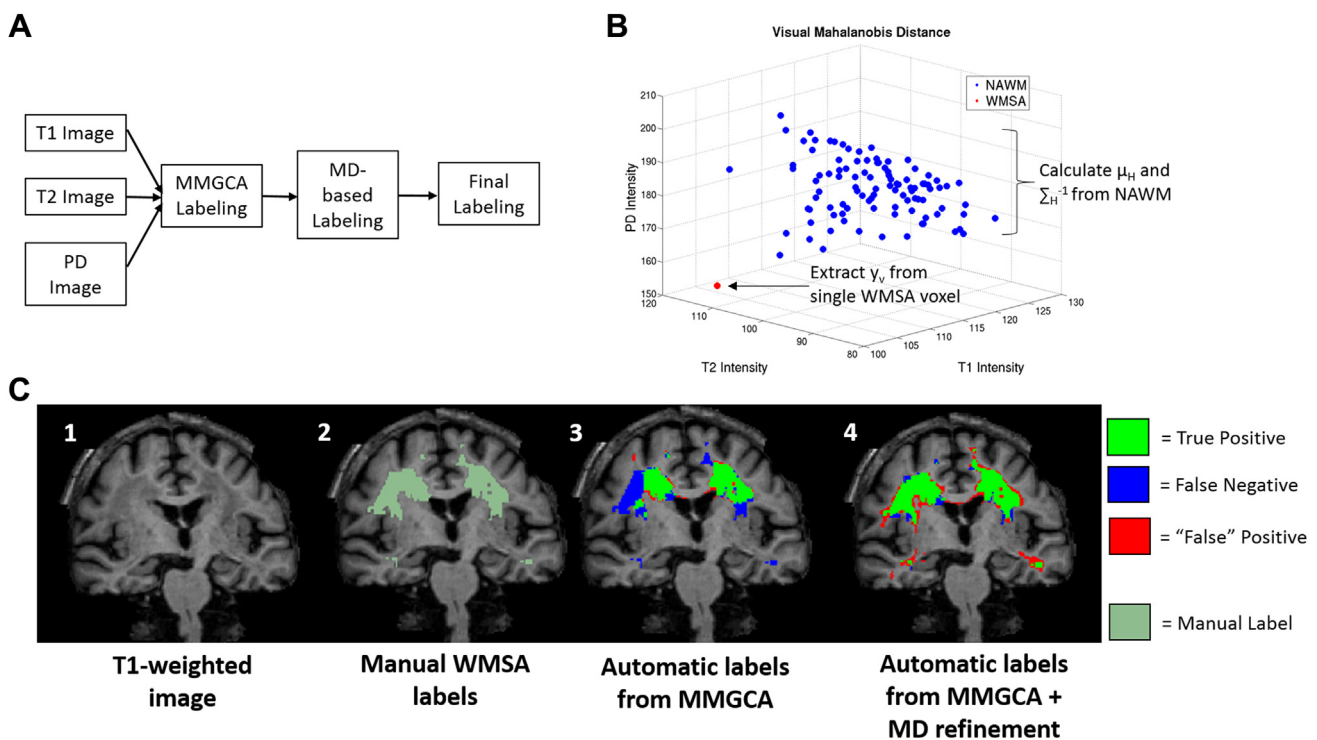

Fig. 2. (A) Flowchart of major steps involved in automatic WMSA labeling pipeline. (B) Visual representation of Mahalanobis distance (MD), using a single WMSA voxel and 100 randomly selected NAWM voxels from an Alzheimer's Disease Neuroimaging Initiative data set. Red point indicates a single WMSA voxel's position in T1/T2/PD intensity space, and blue dots indicate positions of NAWM voxels. The WMSA voxel's MD is measured from NAWM by extracting y_v, μ_H, and Σ_H^{-1} and plugging them into Equation 1. (C) Comparison of the automatic WMSA labeling with only the MMGCA step to the manual labels (C.3) and comparison using the automatic labeling using MMGCA + MD refinement (C.4). This single subject is a representation of the general trend seen in the ADNI data, where the MMGCA labeling frequently failed to label connected sections of WMSA voxels that were detected by a manual labeler, and these sections were corrected with the MD refinement step. We additionally note that the "false" positives seen in (C.4) are addressed in the Section 2.6, and actually reflect damaged tissue that was undetected by the human rater. Subtle differences in tissue signal intensity in these regions can be seen in the T1-weighted image (C.1). Abbreviations: MMGCA, multimodal Gaussian classifier array; NAWM, normal-appearing white matter; PD, proton density; WMSA, white matter signal abnormality. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Demographics for the 4 diagnostic groups

	Older controls (OC)	Mild cognitive impairment nonconverters (MCI-NC)	Mild cognitive impairment converters (MCI-C)	Alzheimer's disease (AD)
Number of subjects	104	116	115	124
Age (y), mean (SD)	76.6 (5.8)	75.6 (6.8)	75.2 (6.9)	77.0 (5.7)
Sex (% male)	50	62.93	55.65	50.81
Years of education, mean (SD)	15.9 (2.7)[a]	15.7 (3.2)[a]	15.7 (3.0)	14.7 (3.1)
MMSE score, mean (SD)	29.1 (0.9)[b]	27.1 (1.8)	26.5 (1.7)	23.2 (2.0)[b]
History of hypertension (% of total)	42.31	47.41	50.43	46.77
History of hypercholesterolemia (% of total)	11.54	13.79	13.04	16.94
History of endocrine-metabolic disorder (% of total)	40.38	31.90	39.13	47.58

Key: MMSE, Mini-Mental State Examination; SD, standard deviation.
[a] Significantly different from AD, $p < 0.05$.
[b] Significantly different from MCI-C, $p < 0.05$.

manual and automatic WMSA labels was quantified using the Dice coefficient:

$$Dice = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{2}$$

where, A corresponds to the set of voxels labeled by the manual rater, and B corresponds to the voxels labeled by the automatic procedure.

Low Dice scores were predominantly because of voxels being labeled as WMSA by the automated procedure that were not labeled by the manual rater (i.e., putative FPs Fig. 2C.4). On closer visual inspection, these FP did not appear to be healthy WM; yet, they were not labeled as WMSA by the manual rater. This motivated a second validation procedure to determine whether the voxels labeled WMSA by the automated procedure, but not by the manual rater, were indeed different from NAWM. The basic idea is that, if these voxels are truly healthy, then they should appear healthy in future time points and should not progress like WMSAs. To test our hypothesis, we identified FP voxels and extracted their T1/T2/PD intensity profiles at baseline (the time of manual labeling) as well as at 6 and 12 months post labeling, for a total of 9 intensity values for each FP voxel. The mean T1/T2/PD intensities of NAWM at baseline were also calculated. For each FP voxel, the mean baseline NAWM intensity was subtracted from each of the 3 time points, for a total of 9 FP–NAWM difference values. A 1-sample t test was then performed with the 9 values for each voxel, with the null hypothesis that these values came from a 0 mean distribution, and were therefore not significantly different from NAWM. All voxels resulting with $p < 0.01$ were used as new "true positives" in the recalculation of an updated Dice score.

We note that, the purpose of this longitudinal validation method is to demonstrate the cross-sectional validity of the automatic

segmentation tool, and longitudinal data is not necessary for accurate WMSA segmentation. Additionally, for all clinical analyses in this study only automatic segmentations were used, not manual labels.

2.7. Statistical analyses

All statistical analyses were conducted in MATLAB version R2013b (2013). Cross-sectional analyses with age, years of education, Mini-Mental State Examination, and baseline WMSA volume were conducted using a 1-way analysis of variance (ANOVA) with follow-up Tukey tests in cases where the ANOVA resulted in a significant group difference. Cross-sectional group differences in sex, history of hypertension, hypercholesterolemia, and history of endocrine-metabolic disorders were determined using a χ^2 test. Tests for group differences in WMSA were subsequently controlled for the effects of all other demographic variables.

For each time point of the MCI-C v. MCI-NC longitudinal analyses of WMSA volume, WMSA quality, and hippocampal volume, a 2-sample t test was conducted to determine if there existed a difference between the 2 groups. Next, for each set of 3 consecutive time points (1 year), a repeated-measures ANOVA was conducted to determine if there was a significant difference between groups in the rate of change in the variable of interest.

3. Results

3.1. Group demographics

The 4 groups (OC, MCI-NC, MCI-C, and AD) did not differ significantly in age, sex distribution, history of hypertension, hypercholesterolemia, or endocrine-metabolic disorder, but there existed a significant difference in years of education between AD and MCI-NC as well as between AD and OC ($p < 0.05$) (Table 1). As expected, significant differences in Mini-Mental State Examination score existed between AD, OC, and MCI groups but not between the MCI-C and MCI-NC ($p < 0.05$). Distribution of sex, history of hypertension, history of hypercholesterolemia, and history of endocrine-metabolic disorder did not differ between groups. Although data presented in this study comprise only a subset of all available ADNI-1 data, this subset did not differ significantly from the overall ADNI-1 cohort in terms of age, sex, years of education, history of hypertension, history of hypercholesterolemia, history of endocrine-metabolic disorder, or cognitive scores (data not shown).

3.2. Baseline WMSA volume group findings

WMSA volume was evaluated both as an absolute number and as a ratio of total WM volume. Total WMSA was significantly different across the 4 groups ($p < 0.0001$) as was total WMSA/total

Table 2
Dice coefficients between all WMSAs labeled by automatic procedure and human rater for 9 data sets with manual labels available (column 2); percent of FPs whose longitudinal NAWM-subtracted intensities resulted in a p-value <0.01 with a 1-sample t test (column 3); updated Dice coefficients using significant FP voxels as new true positives (column 4)

Subject	Dice coefficient	% FPs with $p < 0.01$	Updated Dice coefficient
1	0.7897	82.4	0.8741
2	0.7008	81.95	0.9260
3	0.7368	81.93	0.9186
4	0.6934	74.32	0.8840
5	0.1613	83.32	0.9086
6	0.7180	65.63	0.8913
7	0.6691	64.16	0.7887
8	0.5916	71.38	0.8891
9	0.7033	78.09	0.9053

Key: FPs, false positives; NAWM, normal-appearing white matter; WMSAs, white matter signal abnormalities.

WM ($p < 0.0001$) (Fig. 3). These results remained at their significance levels after controlling for all demographic variables in Table 1. Total WM did not differ between groups (including between MCI-C and MCI-NC). Post hoc Tukey tests demonstrated that OC, MCI-NC, and MCI-C had significantly lower values compared with AD for both WMSA measures.

3.3. Longitudinal WMSA volume findings in MCI-C and MCI-NC

We next examined whether WMSA volume differed between MCI-C and MCI-NC across the time frame 36 months before MCI-C conversion to AD until 18 months after conversion. Volume was measured both in raw mm^3 units, as well as a ratio of WMSA volume to total WM volume. The time courses of these 2 measurements were extremely similar and so only the ratio of WMSA volume to total WM is shown in Fig. 4.

No single time point demonstrated a significant difference in volume or ratio between MCI-C and MCI-NC, but there was a trend-level difference at 18 months and at 30 months ($p = 0.07$). Additionally, no consecutive 3 time points demonstrated a significant difference in rate of growth. We note, however, that of the 32 MCI-C individuals depicted at time point 48 months, 18 individuals actually had an increase in volume ratio from 42 months, and we attribute the dip in Fig. 4 to a loss of data from individuals with the highest ratios at time point 42 months.

3.4. Longitudinal hippocampal volume findings in MCI-C and MCI-NC

As hippocampal volume is a known marker of AD, we analyzed its volumetric trajectory in both MCI-C and MCI-NC for a comparison to WMSA trajectories. Hippocampal volume measurements were corrected for total intracranial volume and were produced automatically by FreeSurfer with no manual intervention. These time courses are demonstrated in Fig. 5. Hippocampal volume was significantly different at 12 and 18 months ($p < 0.05$) and at all consecutive time points until 48 months ($p < 0.001$). A significant difference in the rate of hippocampal volume change was observed between 30 months and 42 months ($p < 0.01$).

3.5. Longitudinal WMSA quality findings in MCI-C and MCI-NC

For subjects in the longitudinal group, WM damage expanded outward from an initial lesion over a 3-year period. This expansion consisted of the progression of NAWM to WMSA, generally in voxels that neighbored an existing lesion. To analyze the qualitative changes of WMSAs over time, we tracked 2 different sets of voxels: (1) voxels that started out as WMSA in each individual's first scanning session (enduring WMSA) and (2) voxels that started out as NAWM in the first scanning session but progressed to WMSA by the end of the individual's enrollment in ADNI (incident WMSA). Voxels that remained NAWM from the first scanning session through the last scanning session were also isolated (enduring NAWM), and their baseline T1/T2/PD values alone were extracted as a reference NAWM distribution with which to calculate MD values for the 2 WMSA groups. Both WMSA sets were followed over time, and at every available time point the mean MD of these voxels to the reference NAWM distribution was calculated. The time courses of these changes in enduring WMSA MD and incident WMSA MD are depicted in Fig. 6.

The MD of enduring WMSA from enduring NAWM was significantly different between groups at the point of AD conversion and at 42 months ($p < 0.01$) as well as at 48 months ($p < 0.05$). A significant difference between groups in the rate of enduring WMSA
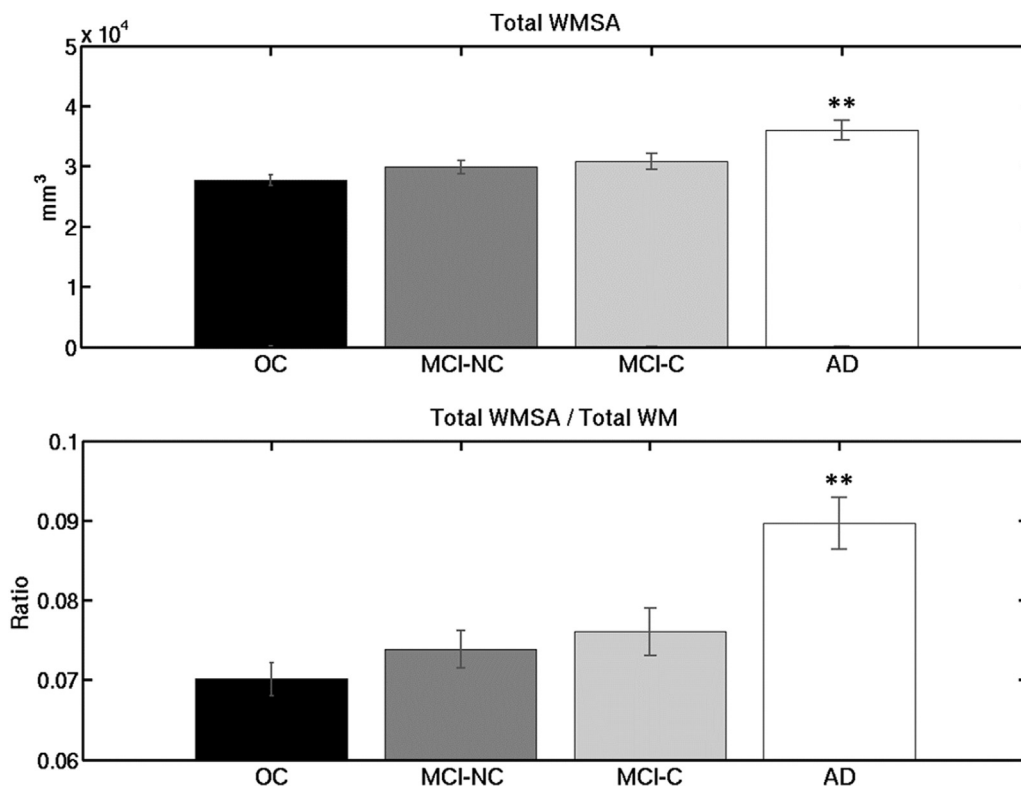

Fig. 3. Cross-sectional group differences in total WMSA volume and WMSA to total WM volume ratio. **Significantly different from OC, MCI-NC, and MCI-C ($p < 0.0001$). Error bars are standard error of the mean. Abbreviations: AD, Alzheimer's disease; MCI-C, mild cognitive impairment converters; MCI-NC, mild cognitive impairment nonconverters; OC, older controls; WM, white matter; WMSA, white matter signal abnormality.
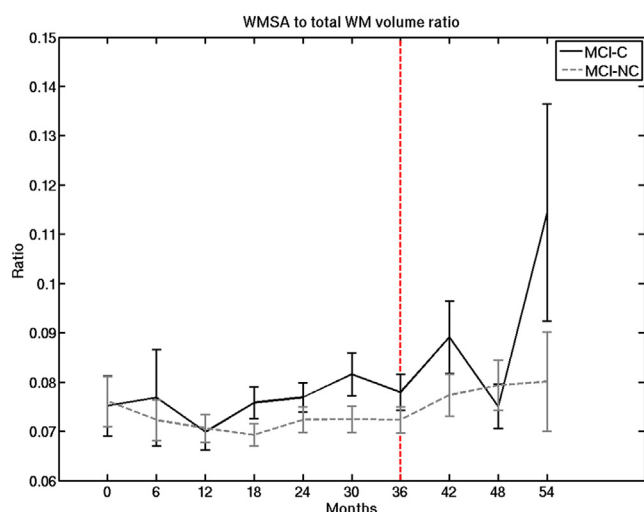
Fig. 4. Time courses of WMSA to total WM volume ratio in MCI-C and MCI-NC individuals. Error bars are standard error of the mean. Red vertical line indicates time of Alzheimer's disease conversion in MCI-C group. Abbreviations: MCI-C, mild cognitive impairment converters; MCI-NC, mild cognitive impairment nonconverters; WM, white matter; WMSA, white matter signal abnormality. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

MD change was observed between 18 months and 30 months ($p < 0.05$), between 24 months and AD conversion ($p < 0.05$), and between 30 months and 42 months ($p < 0.001$). The MD of incident WMSA from enduring NAWM was significantly different between groups at the point of AD conversion, at 42 months and at 54 months ($p < 0.05$). A significant difference in the rate of MD change in incident WMSA was observed between 24 months and AD conversion ($p < 0.05$) as well as between 30 months and 42 months ($p < 0.01$).

We note that, the T1/T2/PD trajectory of change in both NAWM and WMSA is generally consistent within subject; that is, the compression of the intensity vector into a single MD value does not exaggerate differences between NAWM and WMSA.
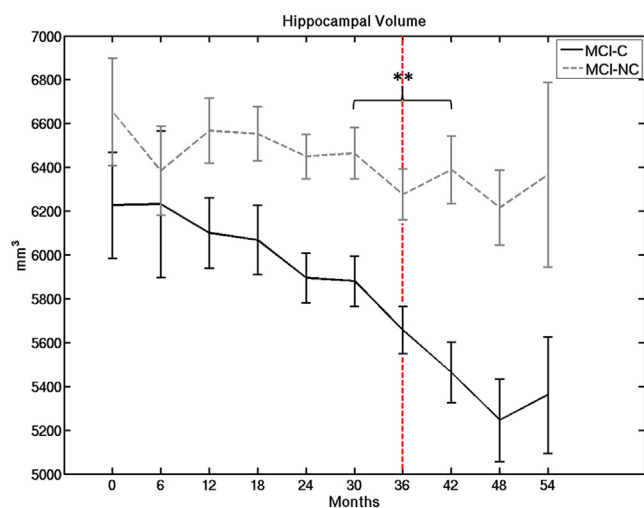

Fig. 5. Time courses of hippocampal volume change in MCI-NC and MCI-C individuals. **Significant interaction in rate of change ($p < 0.01$). Error bars are standard error of the mean. Red vertical line indicates time of Alzheimer's disease conversion in MCI-C group. Abbreviations: MCI-C, mild cognitive impairment converters; MCI-NC, mild cognitive impairment nonconverters. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.6. Manual label agreements

The Dice coefficients between manual and automated labels are reported in Table 2. There was variation in the overlap across subjects. A 1-sample t test was used to compare FP voxels over time to baseline NAWM as described in the Section 2. Under the assumption that these voxels truly were NAWM and were mislabeled by the automatic segmentation, the null hypothesis was that the intensity differences of these voxels from baseline NAWM would resemble values from a zero-mean distribution. Table 2 also shows the percent of all FP voxels whose intensity difference from NAWM intensities remained significant at $p < 0.01$, indicating a true difference from NAWM. These significant voxels were then relabeled as "true positives" and a new Dice coefficient was calculated, resulting in a substantial improvement for many subjects.

4. Discussion

Several prior studies have demonstrated increased WM lesion burden in individuals with AD compared with cognitively healthy matched individuals (Barber et al., 1999; Bowen et al., 1990; Tanabe et al., 1997; Yoshita et al., 2006), as well as between individuals with MCI and healthy individuals (Silbert et al., 2012). To date, however, it is unclear if this burden is a simple comorbidity due to vascular disease or if there is an AD-related pathophysiological process within the WM tracts. We employ MD as a novel image quantification procedure to demonstrate that, although WM lesions do not differ at baseline, there is a difference in the trajectory of lesion progression and accumulation in individuals with MCI that convert to AD compared with MCI that do not convert. We additionally show that the progression of healthy WM to damaged WM demonstrates a distinct temporal dynamic synchronized to the period just before conversion to a clinical diagnosis of AD, with greater deviation from normal T1, T2, and PD intensities based on MD, suggesting that this change may be a direct determinant of this conversion. The associations that we present between these labels and clinical profiles indicate that there are subtle changes in the WM that can be detected on MRI before a diagnosis of AD. These data may even suggest that WMSA progression is a robust component of the conversion from MCI to AD. The significant increase in WMSA volume in individuals with an AD diagnosis provides additional support for this interpretation. Finally, we implement a novel tool that has the capability of automatically identifying WMSAs in the context of a whole-brain segmentation procedure, and this tool will be made available in a future public release of FreeSurfer.

Many biological indicators have been proposed for tracking AD progression (Jack et al., 2013), yet little attention has been given to the use of WMSAs as a predictive tool. Furthermore, the only structural measures that have been robustly indicated as a predictor of AD development thus far have been hippocampal and entorhinal volumes (Convit et al., 1997; Dickerson et al., 2001; Gosche et al., 2002; Jack et al., 1999; Schuff et al., 2009; Shi et al., 2009). We believe that the longitudinal information presented here in the context of MCI may be critical to understanding the biological differences between individuals with MCI that are subsequently diagnosed with AD and those who are not. To describe our findings, we present the new predictive model depicted in Fig. 7. This model contains several key components that differentiate MCI-C individuals from MCI-NC individuals. First, the data in Fig. 6 suggest that the rate at which WM tissue becomes damaged differs between the 2 groups during a time frame that begins roughly around 18 months before MCI conversion to AD. We capture this in our model by showing the MCI-NC damage profiles as a slow progression and the MCI-C individuals exhibiting a steeper slope starting
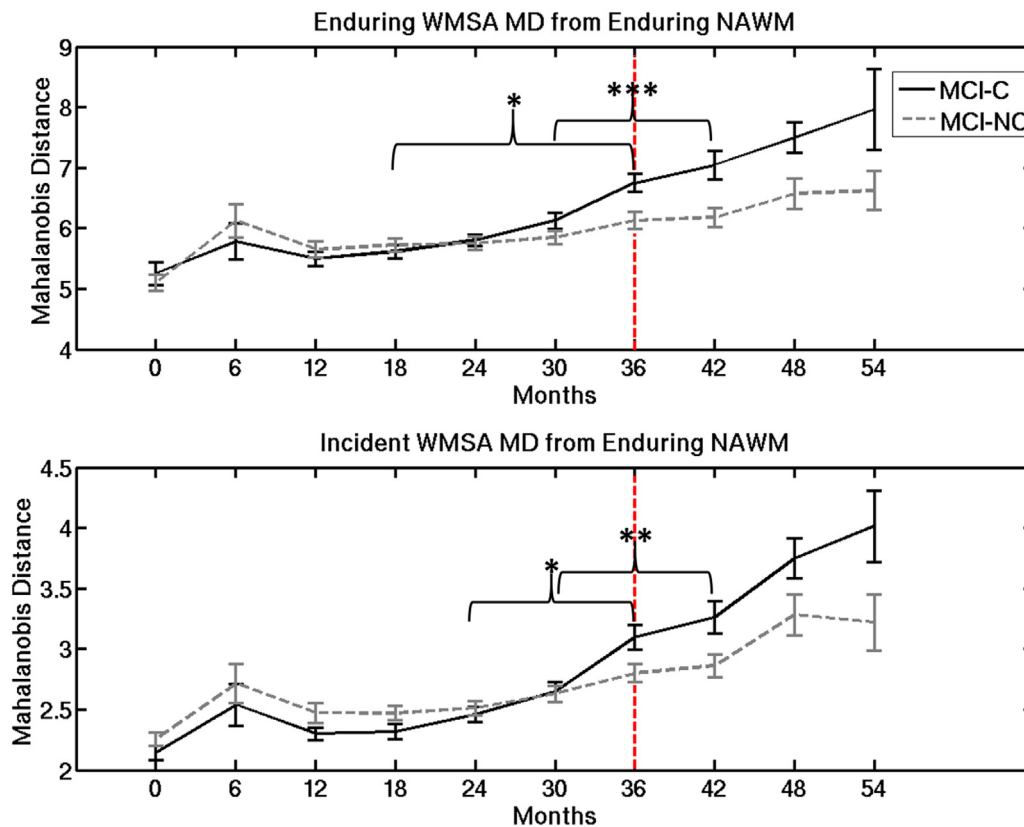
Fig. 6. Time courses of enduring WMSA MD from enduring NAWM (top) and incident WMSA from enduring NAWM (bottom) in MCI-NC and MCI-C individuals. *Significant interaction ($p < 0.05$), **$p < 0.01$, ***$p < 0.0001$. Error bars are standard error of the mean. Red vertical lines indicate time of AD conversion in MCI-C group. Abbreviations: MCI-C, mild cognitive impairment converters; MCI-NC, mild cognitive impairment nonconverters; MD, Mahalanobis distance; NAWM, normal-appearing white matter; WMSA, white matter signal abnormality. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

18 months before AD onset. We demonstrated in Fig. 6 that this difference in trajectory can be seen both in enduring WMSAs as well as in incident WMSAs which are classified as NAWM at an individual's baseline scan. Our model also indicates where the mean MD of WMSA voxels becomes higher in the MCI-C group than the MCI-NC at the point of MCI-C conversion to AD. These results may indicate differing pathophysiological processes in the WM in MCI-C compared with MCI-NC.

The second critical component of this model is the increase in WMSA volume that occurs in the MCI-C group relative to the MCI-NC group after MCI-C conversion to AD. Fig. 4 demonstrates that as a percent of total WM, differences in WMSA volume demonstrate only trend-level significance before MCI-C conversion to AD. In our cross-sectional analyses, shown in Fig. 3, individuals with AD show a significant increase in WMSA volume from individuals with MCI. This is consistent with previous reports in which AD individuals show higher levels of WMSAs than either subjects with MCI or healthy controls (Canu et al., 2012; Pievani et al., 2010; Tanabe et al., 1997; Xie et al., 2006). Although we see a difference in WMSA volume in our cross-sectional analysis between AD and MCI-C individuals, we do not see where this change becomes significant in the MCI-C cohort in the longitudinal analysis. We speculate that the increase in WMSA volume occurs gradually after the onset of AD, as the AD individuals in the ADNI cohort have had a clinical diagnosis of AD for a varying number of years. It is possible that a significant difference in WMSA volume is not present until 3 or more years after AD conversion, and our findings suggest that a change in WMSA quality precedes a change in WMSA quantity. We additionally note that, volume measurements are likely to be noisier

than quality measurements (MD) in the WM as volume can change with basic physiological properties such as hydration level.

We compare our longitudinal WMSA findings to longitudinal hippocampal volume measures, a known marker of AD, in the same MCI-C and MCI-NC individuals. Conflicting evidence has been presented about when a significant difference in hippocampal volume can be seen between these 2 groups before MCI-C conversion to AD (Dickerson et al., 2001; Shi et al., 2009). Others have demonstrated that the combination of hippocampal atrophy and WMSA burden can differentiate between different subtypes of MCI (van de Pol et al., 2009), and furthermore that both of these changes are associated with brain hypoperfusion in MCI (Caroli et al., 2007). Our findings indicate that a difference in hippocampal volume can be seen 24 months before MCI-C conversion to AD. A significant group difference in the rate of hippocampal volume decline is not seen until 6 months before AD conversion, which is when the greatest change in rate of WMSA quality progression is seen. We interpret this finding to mean that some triggering event for hippocampal volume decrease in those with MCI on a trajectory to AD occurs >3 years before AD onset, and a second event potentially occurs 6 months before AD onset in the MCI-C cohort, as seen in Fig. 5. As this second event is timed to the change in WMSA progression rate as well as MCI-C conversion to AD, we speculate that these different biological indicators are both important to AD, and future work will attempt to determine whether these events are independent or related to one another.

Previous studies have demonstrated there is an association between WMSA burden and the rate of cognitive decline in MCI and AD individuals (Smith et al., 2008; Tosto et al., 2014; Wolf et al., 2000).
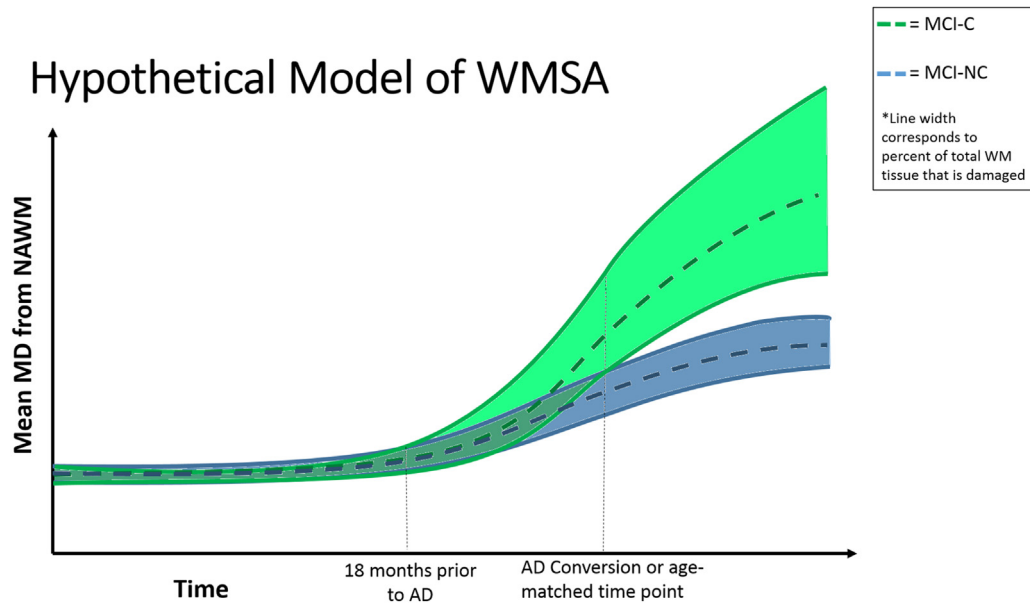
Fig. 7. Hypothetical model of the trajectories of WM damage progression over the course of MCI development in populations that do and do not convert to AD. The width of each bounded region corresponds to the percent of the total WM that is damaged (WMSA to total WM volume). The dashed line in the middle of each region corresponds to the mean MD of all WMSAs from NAWM (degree of damage within lesions). At 18 months before AD conversion, the MCI-C group exhibits a faster increase in WMSA development than the MCI-NC group. After AD conversion in the MCI-C group, volume differences start to be seen between the groups, and the MCI-C group exhibits a larger volume increase. Abbreviations: AD, Alzheimer's disease; MCI-C, mild cognitive impairment converters; MCI-NC, mild cognitive impairment nonconverters; MD, Mahalanobis distance; NAWM, normal-appearing white matter; WM, white matter; WMSA, white matter signal abnormality.

Additionally, the notion of differing levels of WM damage (Maillard et al., 2013), particularly a penumbra encircling a central damage region, have been addressed before (Maillard et al., 2014; Viswanathan, 2014), and there is an extensive body of literature of the heterogeneous histological profiles of WM lesions that appear similar on MRI (Erten-Lyons et al., 2013; Gouw et al., 2011; Young et al., 2008). Taken together, these studies suggest that WMSAs contain more information than just their overall volume, and that the subtle heterogeneities within lesions may provide further information to help differentiate between clinical populations. Our findings regarding the group differences in WMSA progression profiles indicate that different biological processes may be at play. Differences in cerebrovascular integrity may help to explain these differences in WM damage progression, and this issue has gained increasing interest in the AD community (Barker et al., 2014; Iadecola, 2013; Pettersen et al., 2008). WMSAs have long been used as an indirect marker of small-vessel disruption (Huang et al., 2010; Pantoni, 2010; Topakian et al., 2010), and it is possible that the differences in damage progression rate that occur before AD conversion are linked to differences in small-vessel vulnerability to damage. There is also evidence that WM damage burden is related to future amyloid accumulation (Grimmer et al., 2012; Gurol et al., 2006) as well as neurofibrillary tangles (Erten-Lyons et al., 2013), both of which are classic markers of AD. These are particularly novel findings in that they present the possibility that between MCI-C and MCI-NC individuals, different biological processes may occur at the sub-voxel level before clinical symptoms of AD are present, and may be valuable as a prognostic tool or as a marker of therapeutic intervention efficacy.

One of the main sources of difficulty in validating any automatic segmentation tool for WMSAs lies in the lack of a true gold standard for comparison. Manual labeling of WMSAs suffers from low interrater and intrarater reliability. Dice coefficients can vary widely between subjects in the same study, as well as across studies with different manual raters (van Straaten et al., 2006). Additionally, because the Dice coefficient is defined by true positives, false

negatives, and FPs, 2 individuals with the same number of erroneous labels may have very different scores, depending on the amount of true lesion as defined by a manual labeler (i.e., subjects with much larger lesion loads will tend to have significantly higher Dice scores). The Dice coefficients calculated for our data indicate high precision for some subjects, but low precision for others as compared with manual labels. This phenomenon is common in the literature (Cerasa et al., 2012; García-Lorenzo et al., 2008, 2009), and we attempted to investigate the validity of our FPs, which were the major cause of low Dice values, with longitudinal data. Our results suggest that while a manual labeler may not have recognized these voxels as lesions, they contain subtle intensity information that differentiates them from truly NAWM, and can be followed in time to show more profound differences. By following an initial time point's FPs voxels longitudinally, we show compelling evidence that our technique is actually more sensitive to detecting WMSAs than a human rater. We suggest from these findings and the difficulty in manual labeling that future WMSA segmentation procedures be compared with quantitative gold standards derived from signal properties in addition to manual confirmatory procedures.

Future work will better characterize the damage reported here with the use of high-resolution structural sequences. Additionally, the WMSA labeling procedure was developed specifically for these combined image types (T1/T2/PD), which limited the number of ADNI data sets available for WMSA labeling and longitudinal analysis. We are currently extending the procedures described here to be more generalizable to a range of imaging data as well as more quantitative output describing the degree of abnormality from NAWM. Moreover, how MD properties relate to other imaging markers or histological properties is unknown, and this is the focus of ongoing research. Future research will also focus on regionally specific changes in WMSA quality over time, as the current findings are limited to whole-brain analyses and do not specify whether or not these changes are driven by specific types of WMSA, namely periventricular lesions. Additionally, this method of analyzing

WMSA is limited clinically, in that it cannot be used to diagnose a single subject as other methods can (Sheltens et al., 1992). Future work will be directed toward this goal by analyzing WMSA quality with respect to other clinical biomarkers for AD such as Pittsburgh compound B status seen in positron emission tomography. Despite any limitations, this work demonstrates the utility of the developed automated WMSA labeling tool in clinical assessment and demonstrates potentially novel mechanisms governing the conversion of MCI to AD. It is important to note that "conversion" as defined clinically has some ambiguity and does not occur at 1 precise time. The changes that we measure occur along a continuum that predates the actual conversion period and extends beyond, and therefore, we use the conversion point as a guideline for when clinical change is apparent.

Disclosure statement

Acknowledgements

References

Barber, R., Scheltens, P., Gholkar, A., Ballard, C., McKeith, I., Ince, P., Perry, R., O'Brien, J., 1999. White matter lesions on magnetic resonance imaging in dementia with Lewy bodies, Alzheimer's disease, vascular dementia, and normal aging. J. Neurol. Neurosurg. Psychiatry 67, 66–72.

Barker, R., Ashby, E.L., Wellington, D., Barrow, V.M., Palmer, J.C., Kehoe, P.G., Esiri, M.M., Love, S., 2014. Pathophysiology of white matter perfusion in Alzheimer's disease and vascular dementia. Brain 137 (Pt 5), 1524–1532.

Bocti, C., Swartz, R.H., Gao, F.-Q., Sahlas, D.J., Behl, P., Black, S.E., 2005. A new visual rating scale to assess strategic white matter hyperintensities within cholinergic pathways in dementia. Stroke 36, 2126–2131.

Bowen, B.C., Barker, W.W., Loewenstein, D.A., Sheldon, J., Duara, R., 1990. MR signal abnormalities in memory disorder and dementia. AJR Am. J. Roentgenol. 11, 283–290.

Brickman, A.M., Muraskin, J., Zimmerman, M.E., 2009b. Structural neuroimaging in Alzheimer's disease: do white matter hyperintensities matter? Dialogues Clin. Neurosci. 11, 181–190.

Brickman, A.M., Provenzano, F.A., Muraskin, J., Manly, J.J., Blum, S., Apa, Z., Stern, Y., Brown, T.R., Luchsinger, J.A., Mayeux, R., 2012. Regional white matter

hyperintensity volume, not hippocampal atrophy, predicts incident Alzheimer disease in the community. Arch. Neurol. 69, 1621–1627.

Brickman, A.M., Zahra, A., Muraskin, J., Steffener, J., Holland, C.M., Habeck, C., Borogovac, A., Ramos, M.A., Brown, T.R., Asllani, I., Stern, Y., 2009a. Reduction in cerebral blood flow in areas appearing as white matter hyperintensities on magnetic resonance imaging. Psychiatry Res. 172, 117–120.

Canu, E., Frisoni, G.B., Agosta, F., Pievani, M., Bonetti, M., Filippi, M., 2012. Early and late onset Alzheimer's disease patients have distinct patterns of white matter damage. Neurobiol. Aging 33, 1023–1033.

Carmichael, O., Schwarz, C., Drucker, D., Fletcher, E., Harvey, D., Beckett, L., Jack, C.R., Weiner, M., DeCarli, C., 2010. Longitudinal changes in white matter disease and cognition in the first year of the Alzheimer disease neuroimaging initiative. Arch. Neurol. 67, 1370–1378.

Caroli, A., Testa, C., Geroldi, C., Nobili, F., Guerra, U.P., Bonetti, M., Frisoni, G.B., 2007. Brain perfusion correlates of medial temporal lobe atrophy and white matter hyperintensities in mild cognitive impairment. J. Neurol. 254, 1000–1008.

Cerasa, A., Bilotta, E., Augimeri, A., Cherubini, A., Pantano, P., Zito, G., Lanza, P., Valentino, P., Gioia, M.C., Quattrone, A., 2012. A Cellular Neural Network methodology for the automated segmentation of multiple sclerosis lesions. J. Neurosci. Methods 203, 193–199.

Chen, J.J., Rosas, H.D., Salat, D.H., 2013. The relationship between cortical blood flow and sub-cortical white-matter health across the adult age span. PLoS One 8, e56733.

Convit, A., De Leon, M.J., Tarshish, C., De Santi, S., Tsui, W., Rusinek, H., George, A., 1997. Specific hippocampal volume reductions in individuals at risk for Alzheimer's disease. Neurobiol. Aging 18, 131–138.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.

Dale, A.M., Sereno, M.I., 1993. Improved localizadon of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J. Cogn. Neurosci. 5, 162–176.

Debette, S., Markus, H.S., 2010. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. BMJ 341, c3666.

DeCarli, C., Grady, C.L., Clark, C.M., Katz, D.A., Brady, D.R., Murphy, D.G., Haxby, J.V., Salerno, J.A., Gillette, J.A., Gonzalez-Aviles, A., Rapoport, S.I., 1996. Comparison of positron emission tomography, cognition, and brain volume in Alzheimer's disease with and without severe abnormalities of white matter. J. Neurol. Neurosurg. Psychiatry 60, 158–167.

DeCarli, C., Massaro, J., Harvey, D., Hald, J., Tullberg, M., Au, R., Beiser, A., D'Agostino, R., Wolf, P.A., 2005. Measures of brain morphology and infarction in the framingham heart study: establishing what is normal. Neurobiol. Aging 26, 491–510.

Delano-Wood, L., Bondi, M.W., Sacco, J., Abeles, N., Jak, A.J., Libon, D.J., Bozoki, A., 2009. Heterogeneity in mild cognitive impairment: differences in neuropsychological profile and associated white matter lesion pathology. J. Int. Neuropsychol. Soc. 15, 906–914.

De Leeuw, F.E., de Groot, J.C., Achten, E., Oudkerk, M., Ramos, L.M., Heijboer, R., Hofman, A., Jolles, J., van Gijn, J., Breteler, M.M., 2001. Prevalence of cerebral white matter lesions in elderly people: a population based magnetic resonance imaging study. The Rotterdam Scan Study. J. Neurol. Neurosurg. Psychiatry 70, 9–14.

Dickerson, B.C., Goncharova, I., Sullivan, M.P., Forchetti, C., Wilson, R.S., Bennett, D.A., Beckett, L.A., DeToledo-Morrell, L., 2001. MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. Neurobiol. Aging 22, 747–754.

Erten-Lyons, D., Woltjer, R., Kaye, J., Mattek, N., Dodge, H.H., Green, S., Tran, H., Howieson, D.B., Wild, K., Silbert, L.C., 2013. Neuropathologic basis of white matter hyperintensity accumulation with advanced age. Neurology 81, 977–983.

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97, 11050–11055.

Fischl, B., Liu, A., Dale, A.M., 2001. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans. Med. Imaging 20, 70–80.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.

Fischl, B., Salat, D.H., Van Der Kouwe, A., Makris, N., Segonne, F., Quinn, B.T., Dale, A.M., 2004a. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.

Fischl, B., Sereno, M.I., Tootell, R.B.H., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8, 272–284.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004b. Sequence-independent segmentation of magnetic resonance images. Neuroimage 23, S69–S84.

Frisoni, G.B., Galluzzi, S., Pantoni, L., Filippi, M., 2007. The effect of white matter lesions on cognition in the elderly - small but detectable. Nat. Clin. Pract. Neurol. 3, 620–627.

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, 1–18.

García-Lorenzo, D., Lecoeur, J., Arnold, D.L., Collins, D.L., Barillot, C., 2009. Multiple sclerosis lesion segmentation using an automatic multimodal graph cuts. Med. Image Comput. Comput. Assist. Interv. 12 (Pt 2), 584–591.

García-lorenzo, D., Prima, S., Collins, D., Arnold, D., Morrissey, S., Barillot, C., 2008. Combining Robust Expectation Maximization and Mean Shift Algorithms for Multiple Sclerosis Brain Segmentation, pp. 82–91. New York, "Processin.

Gosche, K.M., Mortimer, J.A., Smith, C.D., Markesbery, W.R., Snowdon, D.A., 2002. Hippocampal volume as an index of Alzheimer neuropathology: findings from the Nun Study. Neurology 58, 1476–1482.

Gouw, A.A., Seewann, A., van der Flier, W.M., Barkhof, F., Rozemuller, A.M., Scheltens, P., Geurts, J.J.G., 2011. Heterogeneity of small vessel disease: a systematic review of MRI and histopathology correlations. J. Neurol. Neurosurg. Psychiatry 82, 126–135.

Grimaud, J., Lai, M., Thorpe, J., Adeleine, P., Wang, L., Barker, G.J., Plummer, D.L., Tofts, P.S., McDonald, W.I., Miller, D.H., 1996. Quantification of MRI lesion load in multiple sclerosis: a comparison of three computer-assisted techniques. Magn. Reson. Imaging 14, 495–505.

Grimmer, T., Faust, M., Auer, F., Alexopoulos, P., Förstl, H., Henriksen, G., Perneczky, R., Sorg, C., Yousefi, B.H., Drzezga, A., Kurz, A., 2012. White matter hyperintensities predict amyloid increase in Alzheimer's disease. Neurobiol. Aging 33, 2766–2773.

Grueter, B.E., Schulz, U.G., 2012. Age-related cerebral white matter disease (leukoaraiosis): a review. Postgrad. Med. J. 88, 79–87.

Gurol, M.E., Irizarry, M.C., Smith, E.E., Raju, S., Diaz-Arrastia, R., Bottiglieri, T., Rosand, J., Growdon, J.H., Greenberg, S.M., 2006. Plasma beta-amyloid and white matter lesions in AD, MCI, and cerebral amyloid angiopathy. Neurology 66, 23–29.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., Maguire, P., Rosas, D., Makris, N., Dale, A., Dickerson, B., Fischl, B., 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. Neuroimage 32, 180–194.

Huang, H., Fan, X., Weiner, M., Martin-Cook, K., Xiao, G., Davis, J., Devous, M., Rosenberg, R., Diaz-Arrastia, R., 2012. Distinctive disruption patterns of white matter tracts in Alzheimer's disease with full diffusion tensor characterization. Neurobiol. Aging 33, 2029–2045.

Huang, Y.H., Zhang, W.W., Lin, L., Feng, J., Zhao, X.X., Guo, W.H., Wei, W., 2010. Could changes in arterioles impede the perivascular drainage of interstitial fluid from the cerebral white matter in leukoaraiosis? Neuropathol. Appl. Neurobiol. 36, 237–247.

Iadecola, C., 2013. The pathobiology of vascular dementia. Neuron 80, 844–866.

Jack, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J.L., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L.G., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27, 685–691.

Jack, C.R., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., Shaw, L.M., Vemuri, P., Wiste, H.J., Weigand, S.D., Lesnick, T.G., Pankratz, V.S., Donohue, M.C., Trojanowski, J.Q., 2013. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol. 12, 207–216.

Jack, C.R., Petersen, R.C., Xu, Y.C., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Waring, S.C., Tangalos, E.G., Kokmen, E., 1999. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. Neurology 52, 1397–1403.

Jovicich, J., Czanner, S., Greve, D., Haley, E., Van Der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., MacFall, J., Fischl, B., Dale, A., 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. Neuroimage 30, 436–443.

Levy-Cooperman, N., Ramirez, J., Lobaugh, N.J., Black, S.E., 2008. Misclassified tissue volumes in Alzheimer disease patients with white matter hyperintensities: importance of lesion segmentation procedures for volumetric analysis. Stroke 39, 1134–1141.

Mahalanobis, P.C., 1936. On the generalised distance in statistics. Proc. Natl. Inst. Sci. India 2, 49–55.

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. AJNR Am. J. Neuroradiol. 34, 54–61.

Maillard, P., Delcroix, N., Crivello, F., Dufouil, C., Gicquel, S., Joliot, M., Tzourio-Mazoyer, N., Alpérovitch, A., Tzourio, C., Mazoyer, B., 2008. An automated procedure for the assessment of white matter hyperintensities by multispectral (T1, T2, PD) MRI and an evaluation of its between-centre reproducibility based on two large community databases. Neuroradiology 50, 31–42.

Maillard, P., Fletcher, E., Lockhart, S.N., Roach, A.E., Reed, B., Mungas, D., DeCarli, C., Carmichael, O.T., 2014. White matter hyperintensities and their penumbra lie along a continuum of injury in the aging brain. Stroke 45, 1721–1726.

MATLAB version R2013b, 2013. MATLAB and Statistics Toolbox Release. The MathWorks, Inc., Natick, Massachusetts, United States.

Medina, D., DeToledo-Morrell, L., Urresta, F., Gabrieli, J.D.E., Moseley, M., Fleischman, D., Bennett, D.A., Leurgans, S., Turner, D.A., Stebbins, G.T., 2006. White matter changes in mild cognitive impairment and AD: a diffusion tensor imaging study. Neurobiol. Aging 27, 663–672.

Mortazavi, D., Kouzani, A.Z., Soltanian-Zadeh, H., 2012. Segmentation of multiple sclerosis lesions in MR images: a review. Neuroradiology 54, 299–320.

Pantoni, L., 2010. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. Lancet Neurol. 9, 689–701.

Pettersen, J.A., Sathiyamoorthy, G., Gao, F.-Q., Szilagyi, G., Nadkarni, N.K., St George-Hyslop, P., Rogaeva, E., Black, S.E., 2008. Microbleed topography, leukoaraiosis, and cognition in probable Alzheimer disease from the Sunnybrook dementia study. Arch. Neurol. 65, 790–795.

Pievani, M., Agosta, F., Pagani, E., Canu, E., Sala, S., Absinta, M., Geroldi, C., Ganzola, R., Frisoni, G.B., Filippi, M., 2010. Assessment of white matter tract damage in mild cognitive impairment and Alzheimer's disease. Hum. Brain Mapp. 31, 1862–1875.

Prins, N.D., van Dijk, E.J., den Heijer, T., Vermeer, S.E., Koudstaal, P.J., Oudkerk, M., Hofman, A., Breteler, M.M.B., 2004. Cerebral white matter lesions and the risk of dementia. Arch. Neurol. 61, 1531–1534.

Reuter, M., Rosas, H.D., Fischl, B., 2010. Highly accurate inverse consistent registration: a robust approach. Neuroimage 53, 1181–1196.

Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage 61, 1402–1418.

Schuff, N., Woerner, N., Boreta, L., Kornfield, T., Shaw, L.M., Trojanowski, J.Q., Thompson, P.M., Jack, C.R., Weiner, M.W., 2009. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain 132, 1067–1077.

Schwarz, C., Fletcher, E., DeCarli, C., Carmichael, O., 2009. Fully-automated white matter hyperintensity detection with anatomical prior knowledge and without FLAIR. Inf. Process. Med. Imaging 21, 239–251.

Scheltens, P., Barkhof, F., Valk, J., Algra, P.R., Van der Hoop, R.G., Nauta, J., Wolters, E.C., 1992. White matter lesions on magnetic resonance imaging in clinically diagnosed Alzheimer's disease. Evidence for heterogeneity. Brain 115, 735–748.

Shi, F., Liu, B., Zhou, Y., Yu, C., Jiang, T., 2009. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies. Hippocampus 19, 1055–1064.

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, 741–747.

Smith, E.E., Egorova, S., Blacker, D., Killiany, R.J., Muzikansky, A., Dickerson, B.C., Tanzi, R.E., Albert, M.S., Greenberg, S.M., Guttmann, C.R.G., 2008. Magnetic resonance imaging white matter hyperintensities and brain volume in the prediction of mild cognitive impairment and dementia. Arch. Neurol. 65, 94–100.

Spilt, A., Goekoop, R., Westendorp, R.G.J., Blauw, G.J., de Craen, A.J.M., van Buchem, M.A., 2006. Not all age-related white matter hyperintensities are the same: a magnetization transfer imaging study. AJNR Am. J. Neuroradiol. 27, 1964–1968.

Tanabe, J.L., Amend, D., Schuff, N., DiSclafani, V., Ezekiel, F., Norman, D., Fein, G., Weiner, M.W., 1997. Tissue segmentation of the brain in Alzheimer disease. AJNR Am. J. Neuroradiol. 18, 115–123.

Topakian, R., Barrick, T., Howe, F., Markus, H., 2010. Blood-brain barrier permeability is increased in normal appearing white matter in patients with lacunar stroke and leukoaraiosis. J. Neurol. Neurosurg. Psychiatry 81, 1–26.

Tosto, G., Zimmerman, M.E., Carmichael, O.T., Brickman, A.M., 2014. Predicting aggressive decline in mild cognitive impairment: the importance of white matter hyperintensities. JAMA Neurol. 71, 872–877.

Tullberg, M., Fletcher, E., DeCarli, C., Mungas, D., Reed, B.R., Harvey, D.J., Weiner, M.W., Chui, H.C., Jagust, W.J., 2004. White matter lesions impair frontal lobe function regardless of their location. Neurology 63, 246–253.

Van de Pol, L.A., Verhey, F., Frisoni, G.B., Tsolaki, M., Papapostolou, P., Nobili, F., Wahlund, L.-O., Minthon, L., Frölich, L., Hampel, H., Soininen, H., Knol, D.L., Barkhof, F., Scheltens, P., Visser, P.J., 2009. White matter hyperintensities and medial temporal lobe atrophy in clinical subtypes of mild cognitive impairment: the DESCRIPA study. J. Neurol. Neurosurg. Psychiatry 80, 1069–1074.

Van Straaten, E.C.W., Fazekas, F., Rostrup, E., Scheltens, P., Schmidt, R., Pantoni, L., Inzitari, D., Waldemar, G., Erkinjuntti, T., Mäntylä, R., Wahlund, L.-O., Barkhof, F., 2006. Impact of white matter hyperintensities scoring method on correlations with clinical data: the LADIS study. Stroke 37, 836–840.

Viswanathan, A., 2014. Shades of white: separating degrees of injury in the aging brain. Stroke 45, 1606–1607.

Wardlaw, J.M., Smith, E.E., Biessels, G.J., Cordonnier, C., Fazekas, F., Frayne, R., Lindley, R.I., O'Brien, J.T., Barkhof, F., Benavente, O.R., Black, S.E., Brayne, C., Breteler, M., Chabriat, H., Decarli, C., de Leeuw, F.-E., Doubal, F., Duering, M., Fox, N.C., Greenberg, S., Hachinski, V., Kilimann, I., Mok, V., Oostenbrugge, R.v., Pantoni, L., Speck, O., Stephan, B.C.M., Teipel, S., Viswanathan, A., Werring, D., Chen, C., Smith, C., van Buchem, M., Norrving, B., Gorelick, P.B., Dichgans, M., 2013. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 12, 822–838.

Wolf, H., Ecke, G.M., Bettin, S., Dietrich, J., Gertz, H.J., 2000. Do white matter changes contribute to the subsequent development of dementia in patients with mild cognitive impairment? A longitudinal study. Int. J. Geriatr. Psychiatry 15, 803–812.

Xie, S., Xiao, J.X., Gong, G.L., Zang, Y.F., Wang, Y.H., Wu, H.K., Jiang, X.X., 2006. Voxel-based detection of white matter abnormalities in mild Alzheimer disease. Neurology 66, 1845–1849.

Yoshita, M., Fletcher, E., Harvey, D., Ortega, M., Martinez, O., Mungas, D.M., Reed, B.R., DeCarli, C.S., 2006. Extent and distribution of white matter hyperintensities in normal aging, MCI, and AD. Neurology 67, 2192–2198.

Young, V.G., Halliday, G.M., Kril, J.J., 2008. Neuropathologic correlates of white matter hyperintensities. Neurology 71, 804–811.

Zijdenbos, A.P., Dawant, B.M., Margolin, R.A., Palmer, A.C., 1994. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans. Med. Imaging 13, 716–724.