



Mean diffusivity in cortical gray matter in Alzheimer's disease: The importance of partial volume correction

Judith Henf^{a,b,*}, Michel J. Grothe^a, Katharina Brueggen^a, Stefan Teipel^{a,b}, Martin Dyrba^{a,c}

^a DZNE, German Center for Neurodegenerative Diseases, Rostock, Germany

^b Department for Psychosomatic Medicine, University Medicine Rostock, Rostock, Germany

^c MMIS Group, University of Rostock, Rostock, Germany

ARTICLE INFO

Keywords:

Alzheimer's disease
Mild cognitive impairment
Partial volume effects
Diffusion tensor imaging
Mean diffusivity
Gray matter

ABSTRACT

Mean diffusivity (MD) measured by diffusion tensor imaging can reflect microstructural alterations of the brain's gray matter (GM). Therefore, GM MD may be a sensitive marker of neurodegeneration related to Alzheimer's Disease (AD). However, due to partial volume effects (PVE), differences in MD may be overestimated because of a higher degree of brain atrophy in AD patients and in cases with mild cognitive impairment (MCI). Here, we evaluated GM MD changes in AD and MCI compared with healthy controls, and the effect of partial volume correction (PVC) on diagnostic utility of MD.

We determined region of interest (ROI) and voxel-wise group differences and diagnostic accuracy of MD and volume measures between matched samples of 39 AD, 39 MCI and 39 healthy subjects before and after PVC. Additionally, we assessed whether effects of GM MD values on diagnosis were mediated by volume.

ROI and voxel-wise group differences were reduced after PVC. When using these ROIs for predicting group separation in logistic models, both PVE corrected and uncorrected GM MD values yielded a poorer diagnostic accuracy in single predictor models than regional volume. For the discrimination of AD patients and healthy controls, the effect of GM MD on diagnosis was significantly mediated by volume of hippocampus and posterior cingulate ROIs.

Our results suggest that GM MD measurements are strongly confounded by PVE in the presence of brain atrophy, underlining the necessity of PVC when using these measurements as specific metrics of microstructural tissue degeneration. Independently of PVC, regional MD was not superior to regional volume in separating prodromal and clinical stages of AD from healthy controls.

1. Introduction

Diffusion Tensor Imaging (DTI) allows the examination of microstructural cell damage (Uluğ et al., 1999). During the course of Alzheimer's Disease (AD), these changes are assumed to predate the macrostructural changes that are observable using volumetric magnetic resonance imaging (Weston et al., 2015). Therefore, DTI measures could be used as sensitive markers of AD-related neurodegeneration and help to identify individuals with prodromal AD. Mean diffusivity (MD) reflects the average degree of diffusion of water molecules in all directions. Fractional anisotropy (FA), on the other hand, reflects directionality of diffusion. As opposed to FA that is mostly used for assessing white matter (WM) fiber tract integrity, MD may also be used to assess microstructural alterations of the brain's gray matter (GM). An

increase of MD in GM is assumed to reflect the breakdown of microstructural barriers to diffusion which would predate volumetric changes (Weston et al., 2015). Therefore, GM MD could be a valuable measure of early GM cell damage in AD.

Previous studies on GM MD in AD reported increases of hippocampal MD and whole brain GM MD in subjects with mild cognitive impairment (MCI) as compared to healthy controls (Cherubini et al., 2010; Eustache et al., 2016; Fellgiebel et al., 2004; Müller et al., 2005; Scola et al., 2010), as well as in MCI patients that converted to AD compared to MCI subjects that remained stable over a period of several years (Douaud et al., 2013; Fellgiebel et al., 2006; Scola et al., 2010; van Uden et al., 2016). In some studies, hippocampal diffusivity even demonstrated a higher diagnostic and prognostic accuracy than hippocampal volume (Fellgiebel et al., 2006; Kantarci et al., 2005; Müller

Abbreviations: AD, Alzheimer's Disease; DTI, diffusion tensor imaging; FLAIR, fluid-attenuated inversion recovery; GM, gray matter; MCI, mild cognitive impairment; MD, mean diffusivity; PCC, posterior cingulate cortex; PVC, partial volume correction; PVE, partial volume effects; ROI, region of interest

* Corresponding author at: DZNE German Center for Neurodegenerative Diseases, Rostock, c/o Zentrum für Nervenheilkunde, Gehlsheimer Str. 20, D-18147 Rostock, Germany.

E-mail address: judith.henf@med.uni-rostock.de (J. Henf).

<http://dx.doi.org/10.1016/j.nicl.2017.10.005>

Received 20 February 2017; Received in revised form 6 September 2017; Accepted 3 October 2017

Available online 04 October 2017

2213-1582/ © 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

et al., 2007). Considering MCI a possibly prodromal phase of AD, these findings corroborate the potential utility of GM MD as an earlier biomarker than volumetric changes. However, longitudinal evidence for the assumed succession of microstructural and macrostructural GM alterations is still lacking, even though microstructural alterations have been shown to precede macrostructural changes in the white matter (Ly et al., 2014).

Regarding the findings on GM MD changes in MCI subjects, one needs to acknowledge that extensive hippocampus atrophy is already present in the MCI stage of AD (Shi et al., 2009), probably confounding the measurement of GM MD values due to partial volume effects (PVE). Partial volume effects arise when the signal within a cortical voxel does not purely represent the GM signal at this location, but is confounded by intermixing signals of surrounding cerebrospinal fluid (CSF) or WM tissue. In older people and AD patients, the probability of GM voxels to also contain CSF signal increases due to brain atrophy (Jeon et al., 2012). Thus, given the high MD signal in CSF, PVE may lead to overestimation of differences in GM MD in patients with AD as compared to healthy controls. In this case, MD would represent the joint effect of microstructural changes and macrostructural brain atrophy. However, most of the studies using DTI data of healthy older and AD participants did not consider PVE at all (Cherubini et al., 2010; Douaud et al., 2013; Eustache et al., 2016; Müller et al., 2007). One study stated that PVC was not necessary as the regions of interest (ROIs) were not located at the surface of the brain (Eustache et al., 2016).

Several approaches for correcting PVE in DTI data have been proposed. Kantarci et al. (2001, 2010) used a dedicated DTI sequence based on fluid-attenuated inversion recovery (FLAIR) to suppress the CSF signal during data acquisition. Salminen et al. (2016a, 2016b) recomputed the MD for each region of interest (ROI) using a customized diffusion weighted imaging sequence with multiple non-zero b -values and an extended model to fit the tensor. They could show that the decay of signal intensity was mono-exponential for b -values ≥ 680 , indicating successful suppression of CSF signal after removal of $b \sim 0$ data. Both approaches by Kantarci et al. (2001, 2010) and by Salminen et al. (2016a, 2016b) are well-grounded, but they require additional MR sequences that go beyond current clinical standards, either a FLAIR DTI sequence or a DTI sequence with more than one non-zero b -value. So far, three approaches were proposed for post-acquisition partial volume correction (PVC) of single-shell DTI as used in clinical practice. First, Rose et al. (2008) proposed to exclude all voxels exceeding a certain threshold of MD ($1300 \times 10^{-6} \text{ mm}^2/\text{s}$). This fixed threshold, however, leads to an altered distribution of MD values (Weston et al., 2015). Second, Pasternak et al. (2009) introduced the method of free water elimination, consisting of sub-voxel modeling of two tensors: a tissue and a CSF tensor. However, this method suffers from computational challenges that require defining regularization parameters and constraints on tensor estimates to obtain a unique solution for model fitting. Although Pasternak's approach has been found to reduce CSF contamination and to improve diffusion measures in white matter fiber tracts (Berlot et al., 2014; Metzler-Baddeley et al., 2012; Pasternak et al., 2009), a formal validation of this model is still missing. Additionally, differences of the permeability of the cell bodies in gray matter in comparison to the myelin sheets of the fiber bundles in white matter may make it necessary to adjust some of the model parameters when applying this model to gray matter (Pasternak et al., 2009). Third, Koo et al. (2009) proposed the CSF contamination model. They found a high consistency of their PVC approach with FLAIR DTI, corroborating the validity of this post-hoc data correction procedure. Jeon et al. (2012) successfully applied this CSF contamination model to GM MD data of young and older participants and participants with AD to correct for partial volume effects. They found that PVC stronger reduced GM MD in elder and AD participants than in younger participants, but they did not examine the potential diagnostic value of corrected GM MD for AD.

The aim of the current study was to evaluate the effect of PVC on

cortical and subcortical gray matter MD changes and their potential diagnostic utility in AD and MCI. Therefore, we assessed regional and voxel-wise GM MD changes in AD and MCI compared to cognitively healthy older controls (HC). To correct the MD values for PVE, we chose the CSF contamination model proposed by Koo et al. (2009) because it can be applied on single-shell DTI data post acquisition and it was previously used for PVC of GM (Jeon et al., 2012). We hypothesized that people with AD dementia and people with MCI would have higher MD values than matched healthy controls. Following previous literature (e.g. Fellgiebel et al., 2004), this effect should be most pronounced in the left hippocampus. As we assumed that MD alterations were partly confounded by increased brain atrophy in AD subjects, we expected that PVC would decrease group differences and diagnostic accuracy of MD and reduce the correlation between regional MD and volume measures.

2. Methods

2.1. Sample

We included 117 subjects from the database of the German Center for Neurodegenerative Diseases in Rostock. Among those subjects, there were 39 cognitively healthy controls (HC), 39 individuals with AD and 39 subjects with MCI. The subjects were matched according to age, gender, years of education and imaging protocol, given that two different sequences were used for MRI and DTI acquisition (Table S1 of the Supplementary material). All diagnoses were made in an interdisciplinary team of an experienced neurologist, psychiatrist and neuropsychologist. Diagnosis of AD was made in accordance with the NINCDS-ADRCA criteria (McKhann et al., 2011) and MCI was diagnosed according to the Mayo criteria (Peterson et al., 1999).

All subjects had to fulfill the following criteria: MRI scans that passed the quality control (see Section 2.2 for details), time interval between MRI and neuropsychological assessment of at maximum 3 months, no significant neurological, psychiatric or medical condition (except for MCI or AD), in particular cerebrovascular apoplexy, vascular dementia, subclinical hypothyroidism or substance abuse. Healthy controls were free of cognitive complaints and scored within one standard deviation of the age and education adjusted norm in all subtests of the Consortium to Establish a Registry of Alzheimer's Disease (CERAD) cognitive battery (Morris et al., 1987).

All subjects or their representatives had given informed consent according to the declaration of Helsinki. The study was approved by the ethics committee of the University Medical Center Rostock (HV 2009–0010).

2.2. Image acquisition

All images were acquired using a 3T scanner (SIEMENS MAGNETOM Verio). Two different protocols were used for the T1-weighted MRI and DTI sequences. Detailed imaging parameters can be found in Table S1 of the Supplementary material. Only scans were included that passed quality control by a trained expert who evaluated ghosting effects, blurring, motion and susceptibility artifacts.

2.3. Processing of MRI and DTI data

Deformation-based analysis of the T1-weighted scans was performed using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK, <http://www.fil.ion.ucl.ac.uk/spm/>) implemented in Matlab 2013b (Mathworks, Natwick). First, MRI scans were segmented into GM, WM and CSF partitions using the VBM8 toolbox. Then, the images were normalized to an aging and AD-specific reference template from a previous study (Grothe et al., 2013) using the Diffeomorphic Anatomical Registration Through Exponentiated Lie (DARTEL) algebra algorithm (Ashburner, 2007) with modulation for non-linear

transformation components only. Finally, gray matter maps were smoothed using an 8 mm full width at half maximum (FWHM) kernel. To verify the sensitivity of our results to the choice of segmentation and normalization algorithm, we also applied an alternative segmentation and normalization algorithm on our imaging data using FSL (Version 5.0.9, FMRIB, Oxford, UK, <http://www.fmrib.ox.ac.uk/fsl/>). All further analysis steps were identical in both processing pipelines.

DTI data were preprocessed using the diffusion toolbox of FSL. Automated batch processing of DTI data in FSL was performed using in-house software. Data was corrected for eddy currents and head motion. Then, skull stripping was performed using the Brain Extraction Tool. Diffusion Tensors were fitted to the data with DTIfit, and the resulting FA and MD maps were coregistered to the T1-weighted scans.

To correct for PVE caused by CSF in GM voxels, the CSF contamination model (Koo et al., 2009) was used as adapted by Jeon et al. (2012):

$$\exp(-b\bar{D}(k)) = \lambda_{app-GM}(k) \cdot \exp(-b\bar{D}_{GM}(k)) + \lambda_{app-CSF}(k) \cdot \exp(-b\bar{D}_{CSF}(k)) \quad (1)$$

with $\bar{D}(k)$ being the measured mean diffusivity at a specific voxel k , $\lambda_{app-GM}(k)$ and $\lambda_{app-CSF}(k)$ being the apparent signal fraction weightings of GM and CSF compartments, b being 1000 s/mm² and \bar{D}_{CSF} being defined as a constant, 3.0×10^{-3} mm²/s (Pasternak et al., 2009). The tissue probability maps obtained from the segmentation algorithm of the T1-weighted scans provided the apparent signal fraction weightings $\lambda_{app-GM}(k)$ and $\lambda_{app-CSF}(k)$ (Jeon et al., 2012). Eq. (1) was solved for \bar{D}_{GM} and applied to the coregistered MD maps. Potential WM partial volume effects were excluded by eliminating those voxels with a FA > 0.2.

Coregistered PVE corrected and uncorrected MD maps were normalized to the aforementioned aging and AD-specific template by applying the deformation fields obtained for the T1-weighted scans. For voxel-wise analyses, preprocessed MD and GM maps were spatially smoothed using an 8 mm FWHM kernel.

2.4. Extraction of gray matter mean diffusivity and volume in regions of interest

Left and right hippocampi as well as the posterior cingulate cortex (PCC) were chosen as regions of interest following the literature on peak differences in regional MD and volumes in AD (Cherubini et al., 2010; Eustache et al., 2016; Rose et al., 2008). The Harvard-Oxford structural atlas (Desikan et al., 2006) was warped from MNI space to the aging and AD-specific reference space using non-linear DARTEL transformation. This normalized atlas was used to provide anatomical labels for the examined ROIs. Only those voxels were included that had a gray matter probability of 50% or higher according to the individual normalized tissue probability map of the respective participant. Average regional MD values and volumes were then obtained for each subject using the normalized data.

2.5. Statistical modeling

Whole-brain voxel-wise comparisons were made using analysis of covariance (ANCOVA) models in SPSS, with diagnosis as factor and MRI protocol, age, gender and years of education as covariates. Voxel-wise effects were assessed at a statistical threshold of $p < 0.001$, uncorrected. ROI based analysis of MD and volume was conducted using linear models in SPSS 21, with diagnosis as main outcome and the aforementioned covariates (ANCOVA). To determine the effect size of each factor in the ANCOVA, a partial eta² was computed. Then, post-hoc t -tests were performed.

To compare the accuracy of group separation based on corrected and uncorrected regional MD and regional volume, logistic regression models including MRI protocol, age, gender and years of education as covariates were calculated using the *glm* function in R version 3.3.2 (R

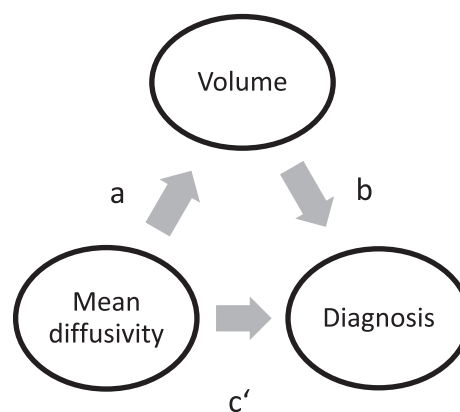


Fig. 1. Mediation model.

Development Core Team, 2008). The binary outcome to predict was the diagnosis (MCI vs. HC or AD vs. HC). Models were arranged to either contain volume, PVE corrected MD or uncorrected MD. To evaluate the diagnostic utility of the predictors in the logistic regression models, we calculated areas under the receiver operator characteristic curve (AUC) for each model and compared the AUC values using bootstrapped CIs ($N = 5000$, *pROC* package in R).

We used partial correlations calculated in SPSS to determine the correlations between regional MD and regional volumes, controlled for the covariates MRI protocol, age, gender and years of education. To compare the resulting correlations, the Williams' test (Williams, 1959) for dependent correlations was used as implemented in the *Psych* package in R. A mediation analysis was performed to determine whether the relationship of GM MD and diagnosis was mediated by regional volume (Fig. 1). Separate mediation models were estimated for the diagnostic subgroups (MCI vs. HC or AD vs. HC) and for PVE corrected and uncorrected MD values. Age, years of education, gender and MRI protocol were included as covariates. Direct (c') and indirect effects ($a * b$) were calculated using the PROCESS tool for SPSS by Hayes (2013), in which logistic regression models were applied for the prediction of diagnostic subgroup. Statistical significance of the direct and indirect effects was determined using bias-corrected bootstrapped confidence intervals (CI) with 10,000 repetitions and a CI of 95%.

To assess the sensitivity of our imaging variables (MD and volume) to the choice of segmentation and normalization algorithm, we correlated the imaging measures resulting from the SPM processing pipeline with the imaging measures resulting from the FSL pipeline. Also, we calculated the Pearson correlation of the AUCs resulting from logistic regression using the SPM imaging values with those resulting from logistic regression using the FSL imaging values. Thus, we could also determine the sensitivity of our results to the segmentation and normalization algorithm.

3. Results

3.1. Sample

The diagnostic subgroups did not differ in age and gender but they did slightly differ in years of education (see Table 1). As expected, the groups differed in MMSE scores with AD patients having the lowest scores and healthy controls having the highest scores.

3.2. Whole brain voxel-wise analysis

Without PVC, MD was elevated across wide parts of the cortex in AD patients compared to healthy controls (Fig. 2). When comparing MCI cases with healthy controls, increased MD was found in both hippocampi. After correcting for partial volume effects, MCI cases and

Table 1
Sample characteristics.

Diagnosis	AD	MCI	Healthy controls (HC)
No. of participants (women)	39 (21)	39 (21)	39 (21)
Age, mean (SD) ¹	74.3 (5.1)	74.9 (5.5)	73.3 (5.3)
Years of education, mean (SD) ²	11.8 (2.7)	12.5 (2.8)	13.2 (2.8)
MMSE, mean (SD) ³	22.56 (4.8)	26.92 (1.7)	28.26 (1.0)

¹ $F(2,114) = 0.92, p = 0.4$, ² $F(2,114) = 2.61, p = 0.08$, ³ $F(2,114) = 39.18, p < 0.001$.

healthy controls did not differ significantly in GM MD. The comparison of AD patients with healthy controls still yielded significant results, with elevated MD in both hippocampi, as well as in the middle and posterior cingulate cortex (Fig. 2).

3.3. Analysis of covariance (ROI based)

Main effects for diagnosis on MD were found in all three ROIs, regardless of PVC (Table 2). Effect sizes were larger in uncorrected MD than in PVE corrected MD data. Post hoc tests revealed that the main effect of diagnosis was driven by the significantly increased MD in the AD subgroup compared to both healthy controls and MCI subjects. Even though mean MD values of the MCI group were numerically greater than in the HC group for all ROIs, none of these differences were statistically significant.

Main effects for regional volume were found in all ROIs and with similar effect sizes across ROIs. Post hoc tests showed that this effect was also mainly driven by decreased volumes in the AD group. However, a significant difference in the MCI group compared to the healthy controls was found for the left hippocampus volume (Table 2).

3.4. Logistic regression models for ROIs

The results of the logistic regression models are shown in Table 3. When discriminating MCI from HC, the models with bilateral hippocampal MD or volume yielded AUCs above chance, whereas the AUCs of the models containing MD or volume of the PCC did not exceed chance level. When comparing the AUCs of the different models per ROI, the differences of AUCs between those models containing uncorrected MD,

corrected MD or volume were non-significant ($p > 0.05$) in the MCI/HC subgroup.

When discriminating AD from HC, all models yielded significant AUCs. Comparing the AUCs in this subgroup, the volume models yielded greater AUCs than the models with PVE corrected MD. Also, the PVE uncorrected MD models reached significantly greater AUCs than PVE corrected MD models.

In all diagnostic group comparisons, the AUCs were numerically greater for the models containing volume as a predictor than those with MD as a predictor, although the AUCs mostly did not differ significantly.

3.5. Correlation analysis (ROI based)

Partial correlations of regional MD values and volumes are shown in Fig. 3. In each diagnostic group and for all ROIs, the correlation between GM MD and GM volume was significantly higher in uncorrected data than in PVE corrected data (all $p < 0.001$, Williams' Test). In some cases, the correlations even became non-significant after PVC was applied. In the PCC, the differences in correlations before and after PVC were numerically smaller than the differences in both hippocampi (range of difference: 0.02 to 0.08 in the PCC and 0.16 to 0.39 in the hippocampi), indicating more severe PVEs in the hippocampus compared to the PCC. In the right hippocampus, the change in correlation was considerably higher in the MCI subgroup than in the AD and HC groups. However, this result could not be reproduced when a different segmentation and normalization algorithm was used (Fig. S2 in the Supplementary material).

3.6. Mediation analysis

The results of the mediation analysis can be found in Table 4. As indicated by the bootstrapped CIs, the indirect effect of regional volume was significant for the discrimination of AD vs. HC, whereas none of the direct effects of MD were significantly different from zero. When separating MCI from HC, a trend (90% CI) for an indirect effect could be found in the left hippocampus, equally for PVE corrected and uncorrected MD. Numerically, the indirect effects were greater for uncorrected MD than for corrected MD in the bilateral hippocampi, although the CIs clearly overlapped.

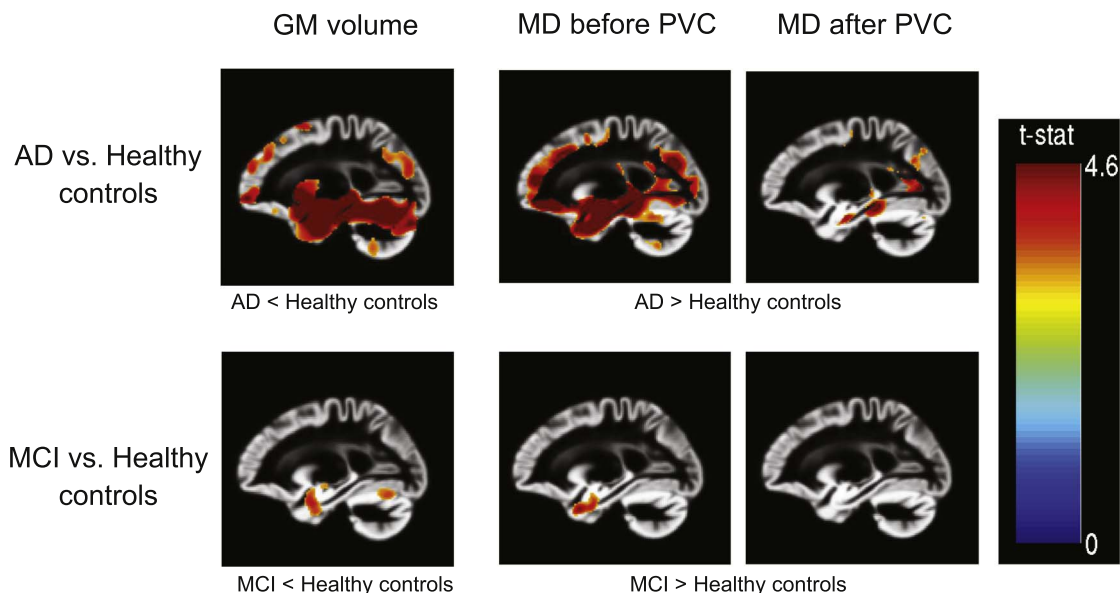


Fig. 2. T-maps of voxel-wise comparisons between the diagnosis subgroups, controlled for age, gender, years of education and MRI protocol. $p < 0.001$ (uncorrected), $x = -22$.

Table 2
ROI Analysis of uncorrected and corrected MD values and volumes, controlled for age, gender, years of education and MRI protocol.

		Mean (SD)			ANCOVA			
		MD in mm ² /s × 10 ⁻³			F	p (unc.)	Partial eta ²	Post hoc t-tests, p < 0.05
		AD	MCI	HC				
Left hippocampus	MD without PVC	1.22 (0.13)	1.15 (0.08)	1.10 (0.08)	14.61	< 0.001	0.21	AD > MCI, AD > HC
	MD with PVC	1.05 (0.10)	1.02 (0.072)	0.99 (0.07)	5.59	< 0.01	0.09	AD > HC AD > MCI†
Right hippocampus	Volume	3.77 (0.67)	4.44 (0.64)	4.82 (0.51)	30.16	< 0.001	0.35	AD < MCI < HC
	MD without PVC	1.26 (0.13)	1.19 (0.07)	1.15 (0.08)	13.98	< 0.001	0.23	AD > MCI, AD > HC
	MD with PVC	1.10 (0.09)	1.07 (0.06)	1.05 (0.06)	4.9	< 0.01	0.08	AD > MCI† AD > HC
	Volume	2.40 (0.48)	2.85 (0.38)	3.02 (0.32)	26.62	< 0.001	0.33	AD < MCI AD < HC
PCC	MD without PVC	1.24 (0.16)	1.14 (0.10)	1.11 (0.09)	19.68	< 0.001	0.26	AD > MCI, AD > HC
	MD with PVC	1.05 (0.15)	0.98 (0.10)	0.95 (0.09)	15.46	< 0.001	0.22	AD > MCI, AD > HC
	Volume	11.93 (1.16)	13.51 (1.11)	13.89 (1.29)	28.38	< 0.001	0.34	AD < MCI AD < HC

† Post hoc comparisons revealed a trend (p < 0.1).

3.7. Alternative segmentation and normalization (FSL)

Detailed results can be found in the Supplementary material (Tables S2, S3 and Fig. S2). In sum, the Pearson correlations of the MD and volume values were moderate to strong (r = 0.54 to 0.85) with the correlation being the lowest in the PVE corrected MD values (Table 5).

The Pearson correlation of the AUCs resulting from the respective processing pipelines (FSL vs. SPM) was r = 0.86, indicating a high overall consistency of the results of the logistic regression.

4. Discussion

In the current study, we evaluated GM MD changes and their diagnostic utility in AD and MCI when considering partial volume effects. In line with our hypothesis, our results showed that partial volume correction led to a decreased size and number of significant clusters in voxel-wise comparisons when compared to uncorrected MD values. Similarly, in the ROI-based analyses, PVE corrected MD was significantly inferior in separating diagnostic groups compared to

uncorrected MD and volume. Thirdly, we showed that the relationship of regional GM MD and diagnosis was significantly mediated by regional volume in AD and HC, and that the correlations between regional MD values and volumes in bilateral hippocampi and PCC were significantly reduced by PVE correction.

The higher correlations of volume with uncorrected MD compared to PVE corrected MD suggest that uncorrected MD values are confounded by atrophy effects resulting in CSF contamination of measured GM MD values. This finding underscores the necessity to correct for PVE if microstructural effects independently of atrophy are the outcome of interest. Our results are in line with a previous study by Jeon et al. (2012), analyzing the regional correlations between cortical thickness and MD in AD, older HC and young HC. Cortical thickness is a measure that is sensitive to changes in gray matter volume, especially cortical thinning (Hutton et al., 2009). Jeon et al. (2012) found significant correlations of MD and cortical thickness in uncorrected data in all subgroups. These correlations decreased significantly in cognitively healthy older and AD participants when applying PVC. We found a similar pattern when examining correlations between regional GM MD

Table 3
Results of the logistic regression analysis, controlled for age, gender, years of education and MRI protocol.

Diagnosis	ROI	PVC	MD		Volume	
			β	AUC [95% CI]	β	AUC [95% CI]
MCI vs. HC	Left hippocampus	With PVC	0.55†	0.67 [0.54, 0.78]	- 0.85**	0.72 [0.60, 0.83]
		Without PVC	0.65*	0.69 [0.56, 0.81]		
	Right hippocampus	With PVC	0.46	0.64 [0.51, 0.76]	- 0.62*	0.68* [0.56, 0.80]
		Without PVC	0.73*	0.71 [0.59, 0.82]		
	Posterior cingulate cortex	With PVC	0.39	0.63 [0.50, 0.75]	- 0.32	0.62 [0.49, 0.74]
		Without PVC	0.39	0.63 [0.50, 0.75]		
AD vs. HC	Left hippocampus	With PVC	1.07**	0.72 [0.61, 0.83]	- 2.64***	0.90** [0.81, 0.96]
		Without PVC	1.89***	0.82** [0.72, 0.90]		
	Right hippocampus	With PVC	0.92**	0.71 [0.59, 0.82]	- 2.22***	0.87** [0.78, 0.94]
		Without PVC	1.65***	0.81*** [0.70, 0.90]		
	Posterior cingulate cortex	With PVC	2.27***	0.81 [0.70, 0.90]	- 2.23***	0.88° [0.80, 0.95]
		Without PVC	2.54***	0.84** [0.74, 0.92]		
AD vs. MCI	Left hippocampus	With PVC	0.45	0.61 [0.48, 0.73]	- 1.40***	0.79** [0.69, 0.89]
		Without PVC	1.07**	0.72*** [0.60, 0.83]		
	Right hippocampus	With PVC	0.41	0.57 [0.44, 0.70]	- 1.29***	0.78** [0.67, 0.88]
		Without PVC	1.06**	0.71*** [0.59, 0.82]		
	Posterior cingulate cortex	With PVC	1.81***	0.75 [0.64, 0.86]	- 1.73***	0.84° [0.75, 0.92]
		Without PVC	2.05***	0.79° [0.68, 0.89]		

Standardized regression weights (β) resulting from logistic regression analysis. The confidence intervals of the area under the receiver operator characteristic curve (AUC) were calculated using bootstrapping.

Asterisks (*) indicate significance of the standardized regression weights (***p < 0.001, **p < 0.01, *p < 0.05, †p < 0.1). Circles (°) indicate significantly larger AUCs compared to the AUC of the MD_{PVC} model, determined using bootstrapping (°p < 0.1, **p < 0.01, ***p < 0.001). None of the other AUCs differed significantly.

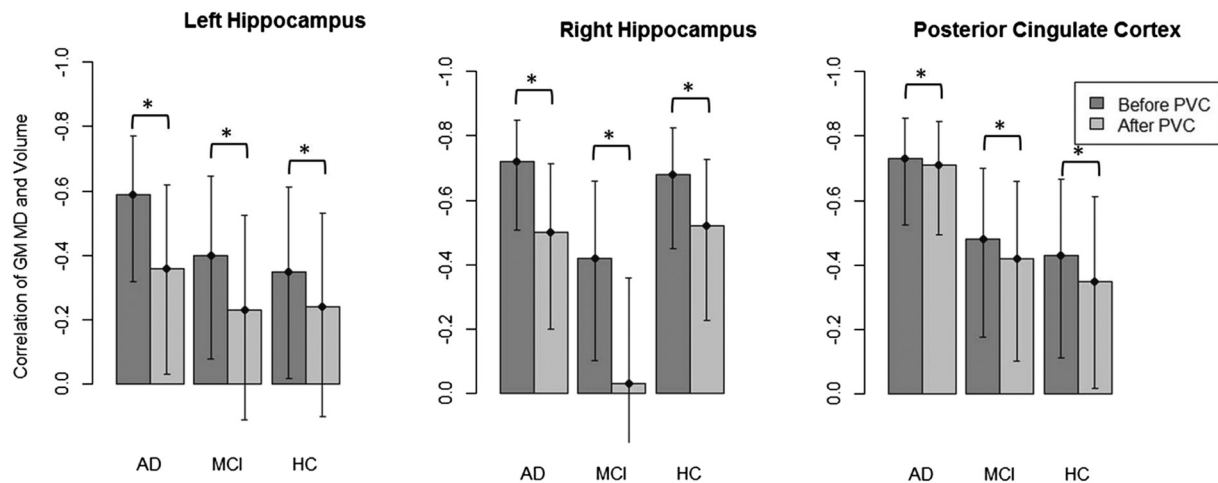


Fig. 3. Partial correlations of uncorrected and corrected GM MD values with GM volume, controlled for MRI protocol, age, gender and years of education. The bars on the left always represent the MD values without PVC, bars on the right represent MD values with PVC. *indicates significant differences of the partial correlations according to the Williams' test, error bars show confidence intervals (95%) of correlation coefficients.

and volume. However, other studies on GM MD in MCI that did not use PVE correction found no significant correlations between regional volumes and MD values (Eustache et al., 2016; Müller et al., 2005). Unfortunately, these studies did not report the correlation sizes because of nonsignificant *p*-values. Also, only zero order correlations were calculated without adjustment for age, gender or education. In our study, in the left hippocampus as well as in the PCC, the correlations between MD and volume were numerically - albeit not significantly - higher in AD subjects than in MCI subjects and higher in MCI than in HC subjects. This is in line with the findings that partial volume effects are more important in brains with stronger atrophy (Jeon et al., 2012). However, the correlations in the right hippocampus were lowest in MCI subjects. As this effect could not be reproduced in the repeated analyses using a different segmentation and normalization pipeline (Fig. S2 in the Supplementary material), we assume that it is an artefact of the segmentation and normalization algorithm (VBM 8).

This is further supported by the fact that in the mediation analysis following the approach recommended by Hayes (2013), volume was a significant mediator in the relationship between MD and diagnosis of AD vs. HC, regardless of PVC. In the left hippocampus, there was also a trend for this mediation on the group separation of MCI vs. HC. The

Table 4
Results of the mediation analysis, controlled for age, gender, years of education and MRI protocol.

Diagnosis	ROI	PVC	Direct Effect (c') β [95% CI]	Indirect effect (a * b) β [95% CI]
MCI vs. HC	Left hippocampus	With PVC	-0.45 [-1.17, 0.28]	-0.23 [†] [-0.73, 0.002]
		Without PVC	-0.53 [-1.44, 0.38]	-0.38 [†] [-1.09, 0.01]
	Right hippocampus	With PVC	-0.46 [-1.24, 0.31]	-0.17 [-0.56, 0.04]
		Without PVC	-0.71 [-1.16, 1.06]	-0.21 [-1.13, 0.31]
	Posterior cingulate cortex	With PVC	-0.32 [-1.43, 0.79]	-0.22 [-0.98, 0.24]
		Without PVC	-0.35 [-1.45, 0.74]	-0.22 [-1.01, 0.33]
AD vs. HC	Left hippocampus	With PVC	-0.47 [-1.47, 0.54]	-1.28* [-2.43, -0.44]
		Without PVC	-0.68 [-1.9, 0.54]	-1.75* [-3.26, -0.63]
	Right hippocampus	With PVC	0.35 [-0.55, 1.24]	-1.44* [-2.64, -0.65]
		Without PVC	-0.06 [-1.22, 1.10]	-1.74* [-3.72, -0.42]
	Posterior cingulate cortex	With PVC	-0.69 [-2.15, 0.76]	-1.83* [-3.32, -0.66]
		Without PVC	-1.04 [-2.52, 0.44]	-1.52* [-2.95, -0.32]
AD vs. MCI	Left hippocampus	With PVC	-0.02 [-0.65, 0.62]	-0.51* [-1.07, -0.14]
		Without PVC	-0.37 [-1.18, 0.44]	-0.74* [-1.41, -0.18]
	Right hippocampus	With PVC	0.05 [-0.52, 0.61]	-0.48* [-0.97, -0.15]
		Without PVC	-0.27 [-1.07, 0.54]	-0.77* [-1.54, -0.1]
	Posterior cingulate cortex	With PVC	-0.23 [-1.52, 1.07]	-1.61* [-3.23, -0.53]
		Without PVC	-0.60 [-1.9, 0.70]	-1.3* [-2.73, -0.33]

* The 95% CI does not include zero.

† the corresponding 90% CI does not include zero.

Table 5

Pearson correlations of the imaging values (MD and volume) resulting from the FSL and SPM processing pipelines.

		Left hippocampus	Right hippocampus	PCC
MD	Without PVC	0.74**	0.77**	0.85**
	With PVC	0.54**	0.59**	0.79**
Volume		0.82**	0.82**	0.53**

** *p* < 0.01.

findings of the mediation analysis, but also the strong correlation of GM MD and GM volume in healthy older subjects and subjects with MCI and AD suggest that changes in uncorrected MD values are largely explained by changes in brain volume in these populations. A strong correlation can be expected in MCI and AD, assuming that alterations of both MD and volume are caused by the same neurodegenerative process. However, the fact that the mediating effect of volume persisted after PVC had been applied could be due to brain atrophy that was already too advanced in these subjects to allow for a sufficient PVE correction. This would mean that MD value estimation would still partly be driven by CSF contamination.

Comparing the diagnostic accuracy of MD and volume, regional MD was not a better indicator for separating AD from HC than regional volume, as indicated by consistently larger AUCs for volume as a predictor. This finding contradicts some previous studies on regional gray matter MD and volume changes in MCI and AD subjects (Fellgiebel et al., 2006; Kantarci et al., 2005; Müller et al., 2007), but is in line with recent findings in an independent sample (Brueggen et al., 2015). As argued there, those studies that found hippocampal MD to be a better predictor for diagnosis or conversion than hippocampal volume did not control for gender, age and education. Also, the sample sizes were smaller than in this recent (Brueggen et al., 2015) and our current study. Considering the different ROIs, we found that the AUCs for both volume and MD of the left hippocampus were greater than those of the right hippocampus and of the PCC. This agrees with a meta-analysis that showed hippocampal asymmetry in MCI and AD cases (Shi et al., 2009).

From our findings, we would conclude that GM MD is not a more useful marker than volume of the hippocampus for detecting prodromal and clinical stages of AD. We showed that the changes in uncorrected GM MD values in MCI and AD patients largely reflect the macrostructural changes in regional brain volume. Although this resulted in a better diagnostic accuracy compared to the PVE corrected MD values, only PVE corrected GM MD values constitute a specific marker of microstructural changes over the course of AD. In contrast to later AD stages, however, one may expect that MD is a useful marker of a neurodegenerative process at very early disease stages that are not yet characterized by overt macrostructural brain atrophy, such as preclinical AD, i.e. cortical amyloidosis without clinical symptoms and limited GM atrophy (Dubois et al., 2014). Also, findings in healthy older people support the assumption that MD may be more sensitive in healthy individuals without significant brain atrophy. Carlesimo et al. (2010) could show that left hippocampal MD was correlated with cognitive performance in cognitively normal older people. Although not correcting for PVE, they observed that hippocampal MD predicted cognitive performance more sensitively than hippocampal volume. Subsequently, it would be interesting to analyze the relationship of PVE corrected hippocampal MD and amyloid accumulation in preclinical AD.

As a limitation, the MCI subjects of our sample were not uniquely amnesic MCI. Thus, our sample is more heterogeneous than the MCI subgroups in other studies (Cherubini et al., 2010; Fellgiebel et al., 2006; Müller et al., 2007; Scola et al., 2010). Second, in our study, we applied the previously validated CSF contamination model by Koo et al. (2009). However, if FLAIR DTI or DTI with more than one non-zero b -value had been available for the sample, DTI maps with suppressed CSF signal could have been obtained so that additional PVC would have been unnecessary (Kantarci et al., 2005; Salminen et al., 2016a). Future studies should assess the impact of different PVC methods on the MD estimates. The approach by Salminen et al. (2016a, 2016b) may be especially promising for clinical practice as it does not require additional scanning time.

In general, as the PVC method by Koo et al. directly uses the segmented maps resulting from VBM8, our results are potentially sensitive to the choice of segmentation and normalization algorithm. However, there was a high consistency of findings in PVE corrected MD data across segmentation/normalization approaches, underscoring the overall relevance of findings irrespective of the normalization or segmentation pipeline. Future studies should systematically evaluate the sensitivity of GM MD estimates to the choice of segmentation and normalization.

5. Conclusion

Evaluating the diagnostic utility of GM MD in AD and MCI, we found that the effects of MD are being overestimated without PVC when using the PVC proposed by Koo et al. (2009). Therefore, when

comparing groups with different levels of atrophy, correction for PVE is indispensable. However, when comparing the diagnostic utility of corrected and uncorrected regional MD to regional volume, regional MD was not superior to volume in separating the potentially prodromal and the clinical stages of AD from matched healthy controls. Future studies need to test whether PVE-corrected MD may be more useful at the preclinical stage of AD.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary Material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nicl.2017.10.005>.

References

- Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. *NeuroImage* 38, 95–113.
- Berlot, R., Metzler-Baddeley, C., Jones, D.K., Sullivan, M.J.O., 2014. NeuroImage CSF contamination contributes to apparent microstructural alterations in mild cognitive impairment. *NeuroImage* 92, 27–35. <http://dx.doi.org/10.1016/j.neuroimage.2014.01.031>.
- Brueggen, K., Dyrba, M., Barkhof, F., Hausner, L., Filippi, M., Nestor, P.J., Hauenstein, K., Klöppel, S., Grothe, M.J., Kasper, E., Teipel, S.J., 2015. Basal forebrain and hippocampus as predictors of conversion to Alzheimer's disease in patients with mild cognitive impairment—a multicenter DTI and volumetry study. *J. Alzheimers Dis.* 48, 197–204. <http://dx.doi.org/10.3233/JAD-150063>.
- Carlesimo, G.A., Cherubini, A., Caltagirone, C., Spalletta, G., 2010. Hippocampal mean diffusivity and memory in healthy elderly individuals A cross-sectional study. *Neurology* 74, 194–200.
- Cherubini, A., Peran, P., Spoleitini, I., Di Paola, M., Di Iulio, F., Hagberg, G.E., Sancsario, G., Gianni, W., Bossu, P., Caltagirone, C., Sabatini, U., Spalletta, G., 2010. Combined volumetry and DTI in subcortical structures of mild cognitive impairment and Alzheimer's disease patients. *J. Alzheimers Dis.* 19, 1273–1282. <http://dx.doi.org/10.3233/JAD-2010-091186>.
- Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31, 968–980.
- Douaud, G., Menke, R.A.L., Gass, A., Monsch, A.U., Rao, A., Whitceter, B., Zamboni, G., Matthews, P.M., Sollberger, M., Smith, S., 2013. Brain microstructure reveals early abnormalities more than two years prior to clinical progression from mild cognitive impairment to Alzheimer's disease. *J. Neurosci.* 33, 2147–2155. <http://dx.doi.org/10.1523/JNEUROSCI.4437-12.2013>.
- Dubois, B., Feldman, H.H., Jacova, C., Hampel, H., Molinuevo, J.L., Blennow, K., DeKosky, S.T., Gauthier, S., Selkoe, D., Bateman, R., et al., 2014. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol.* 13, 614–629.
- Eustache, P., Nemmi, F., Saint-Aubert, L., Pariente, J., Peran, P., 2016. Multimodal magnetic resonance imaging in Alzheimer's disease patients at prodromal stage. *J. Alzheimers Dis.* 50, 1035–1050. <http://dx.doi.org/10.3233/JAD-150353>.
- Fellgiebel, A., Wille, P., Müller, M.J., Winterer, G., Scheurich, A., Vucurevic, G., Schmidt, L.G., Stoeter, P., 2004. Ultrastructural hippocampal and white matter alterations in mild cognitive impairment: a diffusion tensor imaging study. *Dement. Geriatr. Cogn. Disord.* 18, 101–108. <http://dx.doi.org/10.1159/000077817>.
- Fellgiebel, A., Dellani, P.R., Greverus, D., Scheurich, A., Stoeter, P., Müller, M.J., 2006. Predicting conversion to dementia in mild cognitive impairment by volumetric and diffusivity measurements of the hippocampus. *Psychiatry Res.* 146, 283–287. <http://dx.doi.org/10.1016/j.psychres.2006.01.006>.
- Grothe, M., Heinsen, H., Teipel, S., 2013. Longitudinal measures of cholinergic forebrain atrophy in the transition from healthy aging to Alzheimer's disease. *Neurobiol. Aging* 34, 1210–1220. <http://dx.doi.org/10.1016/j.neurobiolaging.2012.10.018>.
- Hayes, A.F., 2013. *Introduction to Mediation, Moderation, and Conditional Process Analysis: a Regression-Based Approach*. Guilford Press.
- Hutton, C., Draganski, B., Ashburner, J., Weiskopf, N., 2009. A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging. *NeuroImage* 48, 371–380. <http://dx.doi.org/10.1016/j.neuroimage.2009.06.043>.
- Jeon, T., Mishra, V., Uh, J., Weiner, M., Hatanpaa, K.J., White 3rd, C.L., Zhao, Y.D., Lu, H., Diaz-Arrastia, R., Huang, H., 2012. Regional changes of cortical mean diffusivities with aging after correction of partial volume effects. *NeuroImage* 62, 1705–1716. <http://dx.doi.org/10.1016/j.neuroimage.2012.05.082>.
- Kantarci, K., CR, J.J., Xu, Y., Campeau, N., O'Brien, P., Smith, G., Ivnik, R., Boeve, B., Kokmen, E., Tangalos, E., Petersen, R., 2001. Mild cognitive impairment and Alzheimer disease: regional diffusivity of water. *Radiology* 101–107. <http://dx.doi.org/10.1148/radiology.219.1.r01ap14101>.
- Kantarci, K., Petersen, R.C., Boeve, B.F., Knopman, D.S., Weigand, S.D., O'Brien, P.C.,

- Shiung, M.M., Smith, G.E., Ivnik, R.J., Tangalos, E.G., et al., 2005. DWI predicts future progression to Alzheimer disease in amnesic mild cognitive impairment. *Neurology* 64, 902–904.
- Kantarci, K., Avula, R., Senjem, M.L., Samikoglu, A.R., Zhang, B., Weigand, S.D., Przybelski, S.A., Edmonson, H.A., Vemuri, P., Knopman, D.S., Ferman, T.J., Boeve, B.F., Petersen, R.C., Jack, C.R., 2010. Dementia with Lewy bodies and Alzheimer disease: neurodegenerative patterns characterized by DTI. *Neurology* 74, 1814–1821. <http://dx.doi.org/10.1212/WNL.0b013e3181e0f7cf>.
- Koo, B.B., Hua, N., Choi, C.H., Ronen, I., Lee, J.M., Kim, D.S., 2009. A framework to analyze partial volume effect on gray matter mean diffusivity measurements. *NeuroImage* 44, 136–144. <http://dx.doi.org/10.1016/j.neuroimage.2008.07.064>.
- Ly, M., Canu, E., Xu, G., Oh, J., McLaren, D.G., Dowling, N.M., Alexander, A.L., Sager, M.A., Johnson, S.C., Bendlin, B.B., 2014. Midlife Measurements of White Matter Microstructure Predict Subsequent Regional White Matter Atrophy in Healthy Adults. 2054. pp. 2044–2054. <http://dx.doi.org/10.1002/hbm.22311>.
- McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morris, J.C., Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub, S., Phelps, C.H., 2011. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dement.* J. Alzheimer's Assoc. 7, 263–269. <http://dx.doi.org/10.1016/j.jalz.2011.03.005>.
- Metzler-Baddeley, C., Sullivan, M.J.O., Bells, S., Pasternak, O., Jones, D.K., 2012. NeuroImage how and how not to correct for CSF-contamination in diffusion MRI. *NeuroImage* 59, 1394–1403. <http://dx.doi.org/10.1016/j.neuroimage.2011.08.043>.
- Morris, J.C., Mohs, R.C., Rogers, H., Fillenbaum, G., Heyman, A., 1987. Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacol. Bull.* 24, 641–652.
- Müller, M.J., Greverus, D., Dellani, P.R., Weibrich, C., Wille, P.R., Scheurich, A., Stoeter, P., Fellgiebel, A., 2005. Functional implications of hippocampal volume and diffusivity in mild cognitive impairment. *NeuroImage* 28, 1033–1042. <http://dx.doi.org/10.1016/j.neuroimage.2005.06.029>.
- Müller, M.J., Greverus, D., Weibrich, C., Dellani, P.R., Scheurich, A., Stoeter, P., Fellgiebel, A., 2007. Diagnostic utility of hippocampal size and mean diffusivity in amnesic MCI. *Neurobiol. Aging* 28, 398–403.
- Pasternak, O., Sochen, N., Gur, Y., Intrator, N., Assaf, Y., 2009. Free water elimination and mapping from diffusion MRI. *Magn. Reson. Med.* 62, 717–730. <http://dx.doi.org/10.1002/mrm.22055>.
- Petersen, R.C., Smith, G.E., Waring, S.C., Ivnik, R.J., Tangalos, E.G., Kokmen, E., 1999. Mild cognitive impairment: clinical characterization and outcome. *Arch. Neurol.* 56, 303–308.
- R Development Core Team, 2008. R: a Language and Environment for Statistical Computing.
- Rose, S.E., Janke, A.L., Chalk, J.B., 2008. Gray and white matter changes in Alzheimer's disease: a diffusion tensor imaging study. *J. Magn. Reson. Imaging* 27, 20–26. <http://dx.doi.org/10.1002/jmri.21231>.
- Salminen, L.E., Conturo, T.E., Bolzenius, J.D., Cabeen, R.P., Akbudak, E., Paul, R.H., 2016a. Reducing CSF Partial Volume Effects to Enhance Diffusion Tensor Imaging Metrics of Brain Microstructure. 18, pp. 5–20.
- Salminen, L.E., Conturo, T.E., Laidlaw, D.H., Cabeen, R.P., Akbudak, E., Lane, E.M., Heaps, J.M., Bolzenius, J.D., Baker, L.M., Cooley, S., Scott, S., Cagle, L.M., Phillips, S., Paul, R.H., 2016b. Regional age differences in gray matter diffusivity among healthy older adults. *Brain Imaging Behav.* 10, 203–211. <http://dx.doi.org/10.1007/s11682-015-9383-7>.
- Scola, E., Bozzali, M., Agosta, F., Magnani, G., Franceschi, M., Sormani, M.P., Cercignani, M., Pagani, E., Falautano, M., Filippi, M., Falini, A., 2010. A diffusion tensor MRI study of patients with MCI and AD with a 2-year clinical follow-up. *J. Neuro. Neurosurg. Psychiatry* 81, 798–805. <http://dx.doi.org/10.1136/jnnp.2009.189639>.
- 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. <http://dx.doi.org/10.1002/hipo.20573>.
- van Uden, I.W.M., Tuladhar, A.M., van der Holst, H.M., van Leijsen, E.M.C., van Norden, A.G.W., de Laat, K.F., Rutten-Jacobs, L.C.A., Norris, D.G., Claassen, J.A.H.R., van Dijk, E.J., Kessels, R.P.C., de Leeuw, F.-E., 2016. Diffusion tensor imaging of the hippocampus predicts the risk of dementia; the RUN DMC study. *Hum. Brain Mapp.* 37, 327–337. <http://dx.doi.org/10.1002/hbm.23029>.
- Uluğ, A.M., Moore, D.F., Bojko, A.S., Zimmerman, R.D., 1999. Clinical use of diffusion-tensor imaging for diseases causing neuronal and axonal damage. *Am. J. Neuroradiol.* 20, 1044–1048.
- Weston, P.S.J., Simpson, I.J.A., Ryan, N.S., Ourselin, S., Fox, N.C., 2015. Diffusion imaging changes in grey matter in Alzheimer's disease: a potential marker of early neurodegeneration. *Alzheimers Res. Ther.* 7, 47. <http://dx.doi.org/10.1186/s13195-015-0132-3>.
- Williams, E.J., 1959. *Regression Analysis*. WILEY, New York.