



## King's Research Portal

DOI:

[10.1007/s10548-015-0455-1](https://doi.org/10.1007/s10548-015-0455-1)

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Falahati, F., Ferreira, D., Soinenen, H., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Lovestone, S., Eriksdotter, M., Wahlund, L. O., Simmons, A., & Westman, E. (2016). The Effect of Age Correction on Multivariate Classification in Alzheimer's Disease, with a Focus on the Characteristics of Incorrectly and Correctly Classified Subjects. *Brain topography*, 29(2), 296-307. <https://doi.org/10.1007/s10548-015-0455-1>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**


Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# The Effect of Age Correction on Multivariate Classification in Alzheimer's Disease, with a Focus on the Characteristics of Incorrectly and Correctly Classified Subjects

Farshad Falahati<sup>1</sup>  · Daniel Ferreira<sup>1</sup> · Hilkka Soininen<sup>2</sup> · Patrizia Mecocci<sup>3</sup> · Bruno Vellas<sup>4</sup> · Magda Tsolaki<sup>5</sup> · Iwona Kłoszewska<sup>6</sup> · Simon Lovestone<sup>7,8</sup> · Maria Eriksson<sup>1</sup> · Lars-Olof Wahlund<sup>1</sup> · Andrew Simmons<sup>7,8</sup> · Eric Westman<sup>1</sup> ·  
for the AddNeuroMed consortium and the Alzheimer's Disease Neuroimaging Initiative

Received: 23 July 2015 / Accepted: 28 September 2015 / Published online: 6 October 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** The similarity of atrophy patterns in Alzheimer's disease (AD) and in normal aging suggests age as a confounding factor in multivariate models that use structural magnetic resonance imaging (MRI) data. To study the effect and compare different age correction approaches on AD diagnosis and prediction of mild cognitive impairment (MCI) progression as well as investigate the characteristics of correctly and incorrectly classified subjects. Data from two multi-center cohorts were included in the study [AD = 297, MCI = 445, controls (CTL) = 340]. 34 cortical thickness and 21 subcortical volumetric measures were extracted from MRI. The age correction approaches

involved: using age as a covariate to MRI-derived measures and linear detrending of age-related changes based on CTL measures. Orthogonal projections to latent structures was used to discriminate between AD and CTL subjects, and to predict MCI progression to AD, up to 36-months follow-up. Both age correction approaches improved models' quality in terms of goodness of fit and goodness of prediction, as well as classification and prediction accuracies. The observed age associations in classification and prediction results were effectively eliminated after age correction. A detailed analysis of correctly and incorrectly classified subjects highlighted age associations in other factors: ApoE genotype, global cognitive impairment and gender. The two methods for age correction gave similar results and show that age can partially mask the influence of other aspects such as cognitive impairment, ApoE-e4 genotype and gender. Age-related brain atrophy may have a more important association with these factors than previously believed.

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

**Electronic supplementary material** The online version of this article (doi:10.1007/s10548-015-0455-1) contains supplementary material, which is available to authorized users.

✉ Farshad Falahati  
farshad.falahati@ki.se

<sup>1</sup> Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Novum, Plan 5, 141 57 Stockholm, Sweden

<sup>2</sup> Department of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland

<sup>3</sup> Institute of Gerontology and Geriatrics, University of Perugia, Perugia, Italy

**Keywords** OPLS · Age correction · Alzheimer's disease · Mild cognitive impairment · MRI · Early diagnosis · Diagnostic misclassification

<sup>4</sup> INSERM U 558, University of Toulouse, Toulouse, France

<sup>5</sup> 3rd Department of Neurology, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece

<sup>6</sup> Medical University of Lodz, Lodz, Poland

<sup>7</sup> Institute of Psychiatry, King's College London, London, UK

<sup>8</sup> NIHR Biomedical Research Centre for Mental Health, London, UK

## Introduction

Alzheimer's disease (AD), the most common form of dementia, is a progressive neurodegenerative disorder that clinically characterizes by gradual loss of cognitive functions. Mild cognitive impairment (MCI), an intermediate condition between normal cognition and dementia, often represents a prodromal form of dementia. MCI patients have a significantly higher risk of converting to AD or other types of dementia. However not all MCI patients develop dementia even after several years. The new criteria for diagnosing "dementia due to AD" and "MCI due to AD" in addition to core clinical criteria, include the use of imaging and other biomarkers to improve the certainty of diagnoses (Albert et al. 2011; McKhann et al. 2011). However, the need of additional work to validate these biomarkers for routine clinical practice is also noted.

Structural magnetic resonance imaging (MRI) is an important marker in clinical practice for dementia diagnosis, particularly in memory clinic settings when younger and rare conditions are examined (Falahati et al. 2014a). MRI has been widely studied for early detection and diagnosis of AD in terms of atrophy of brain structures. In particular, atrophy of medial temporal structures such as hippocampus is demonstrated in AD patients (Scheltens et al. 1992). Due to the complexity of AD, measures of single structures from MRI are probably insufficient for accurate diagnosis. The combination of different structures has proven to be more useful when distinguishing AD from cognitively normal elderly subjects (CTL) (Westman et al. 2011b). With the help of sophisticated image analysis techniques, numerous volumetric and cortical thickness measures can be extracted from structural MRI data.

Machine learning and multivariate data analysis methods provide tools for processing and finding inherent patterns in such data with high complexity and dimensionality. Methods like orthogonal projection to latent structures (OPLS) (Bylesjö et al. 2006; Trygg and Wold 2002) are efficient, robust and validated tools for modeling complex biological data. OPLS was developed with the aim of reducing model complexity and improving model transparency. The improved interpretation property of the OPLS method postures it as a suitable analysis method. OPLS has successfully been applied in research for AD diagnosis and prediction of MCI progression (Aguilar et al. 2014; Spulber et al. 2013).

Confounding factors such as age negatively affect the performance of machine learning and multivariate models. Global and regional brain changes related to increasing age can potentially lead to misclassification of younger AD patients and older CTL subjects. Therefore, there is a need

for developing methods to address this problem. Recently several methods for correcting the age associations are proposed (Dukart et al. 2011; Koikkalainen et al. 2012). The focus of these studies are statistical improvements while their effects on the characteristics of correctly and incorrectly classified subjects were disregarded. Studying the subjects' characteristics is of high importance since it can distinctly reflect the way age correction improves the outcomes. Further, to compare these methods to simply use age as a covariate has not been properly investigated.

In this work two age correction approaches were investigated: simply using age as a variable in the OPLS model and a linear detrending approach that removes age-related effects from each variable based on measures in CTL subjects. The effect of age correction approaches on the classification of AD and CTL subjects, and prediction of progression from MCI to AD was explored. The characteristics of correctly/incorrectly classified and predicted subjects before and after age correction were studied in detail. We hypothesized that age correction would improve the performance of both classification and prediction. Additionally, studying the characteristics of subjects before and after age correction may reveal other clinically relevant aspects.

## Materials and Methods

### Study Setting

Data were obtained from two large multi-center cohorts, the Alzheimer's disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>) and AddNeuroMed. ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations (Mueller et al. 2005). The primary goal of ADNI is to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of MCI and early AD. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California -San Francisco. ADNI subjects were recruited from over 50 sites across the U.S. and Canada. For up-to-date information, see [www.adni-info.org](http://www.adni-info.org). AddNeuroMed, a part of InnoMed (Innovative Medicines in Europe), is an integrated project aimed to develop and validate novel surrogate markers in AD (Lovestone et al. 2007). The neuroimaging part of AddNeuroMed uses MRI collected from six different sites across Europe (<http://www.innomed-addneuromed.com/>).

A total of 1082 subjects were included in the current study (AD = 297, MCI = 445 and CTL = 340). At 12-month follow-up, 85 MCI patients progressed to AD (MCI-p) and 360 remained stable (MCI-s). The demographics of the dataset are given in Table 1. The subjects in the ADNI study have also been followed up at 18, 24 and 36 months after baseline. MCI individuals who progressed to AD were considered as MCI-p and the rest as MCI-s.

### Inclusion and Diagnostic Criteria

Participants' recruitment and eligibility criteria were very similar in both cohorts (Petersen et al. 2010; Simmons et al. 2011). Briefly, AD diagnosis was based on NINCDS-ADRDA and DSM-IV criteria for probable AD, as well as a total clinical dementia rating (CDR) score of 0.5 or above. MCI diagnosis required a MMSE score between 24 and 30; memory complaints; normal activities of daily living; total CDR score of 0.5; and Geriatric Depression Scale (GDS) score of  $\leq 5$ . The inclusion criteria for control participants were a MMSE score between 24 and 30; total CDR score of 0; and GDS score  $\leq 5$ . No significant neurological or psychiatric illness, no significant unstable systemic illness or organ failure, and no history of alcohol or substance abuse or dependence were required for all three groups. MRI information was not used for diagnosis.

## Imaging

### MRI Data Acquisition

In both cohorts, 1.5T MRI data was collected from a variety of MR-systems with protocols optimized for each type of scanner. The MRI protocol included a high-

resolution sagittal 3D T1-weighted MPRAGE volume (voxel size  $1.1 \times 1.1 \times 1.2 \text{ mm}^3$ ) acquired using a custom pulse sequence specifically designed for the ADNI study to ensure compatibility across scanners (Jack et al. 2008). MRI data acquisition in AddNeuroMed was designed to be compatible with the ADNI protocol (Simmons et al. 2011).

### Regional Subcortical Volume Segmentation and Cortical Thickness Parcellation

The FreeSurfer pipeline (version 5.3.0) was applied to the MRI images to produce regional cortical thickness and subcortical volumetric measures. Full details and references of cortical reconstruction and subcortical volumetric segmentation procedure are included in the supplementary material 1. Data was processed through the hive database system (theHiveDB) (Muehlboeck et al. 2014). Visual quality control was performed on all output data. All steps involving brain extraction, automated Talairach transformation, tessellation, surfaces reconstruction, and subcortical segmentation were carefully checked. This segmentation approach has been used for multivariate classification of Alzheimer's disease and healthy controls (Westman et al. 2010), neuropsychological-image analysis (Ferreira et al. 2014) and biomarker discovery (Maioli et al. 2015). In total, 55 MRI measures were used as input variables for OPLS classification, i.e. 34 regional cortical thickness measures and 21 regional subcortical volumes (measures from the left and right sides were averaged). Supplementary material 2 provides a list of measures and their mean and standard deviation in each diagnostic group. All subcortical volumetric and cortical thickness measures were used in their raw form (Westman et al. 2013).

**Table 1** Demographic and clinical characteristics

	CTL	MCI-s	MCI-p <sup>a</sup>	AD
Count	340	360	85	297
Age, years	75.0 $\pm$ 5.7	75.0 $\pm$ 6.9	74.3 $\pm$ 6.5	75.7 $\pm$ 7.0
Gender, Female/Male	172/168	141/219	35/50	165/132
Education, years	14.3 $\pm$ 4.3	13.9 $\pm$ 4.7	13.8 $\pm$ 4.2	12.0 $\pm$ 4.9
MMSE score	29.1 $\pm$ 1.1	27.1 $\pm$ 1.7	26.5 $\pm$ 1.8	22.2 $\pm$ 3.7
CDR	0	0.5	0.5	0.92 $\pm$ 0.43
ApoE-e4, N/P	242/92	182/158	31/52	110/178
Cohort, ADNI/ANM	227/113	260/100	62/23	175/122

Continuous data is represented as mean  $\pm$  SD, CTL control subjects, MCI mild cognitive impairment, MCI-s stable MCI, MCI-p progressive MCI, AD Alzheimer's disease, MMSE mini mental state examination, CDR clinical dementia rating, ApoE apolipoprotein E, N/P negative/positive for at least one e4 allele, ADNI/ANM Alzheimer's Disease Neuroimaging Initiative/AddNeuroMed

<sup>a</sup> MCI patients progressed to AD at month-12 follow-up

## Data Analysis

### Multi and Univariate Data Analysis

Pre-processing was performed using mean-centering and unit variance scaling in order to transform the data into a suitable form for analysis (Eriksson et al. 2013). OPLS (Bylesjö et al. 2006; Trygg and Wold 2002), a supervised multivariate data analysis method, was used to classify AD patients and CTL individuals as well as to predict progression in the MCI patients. The OPLS method is an extension to the projection to latent structures (PLS) method (Wold et al. 1984). PLS has been developed for the purpose of modeling complex data based on the assumption that there are latent variables, which generate the observed data. PLS extracts these latent variables by maximizing the covariance between two sets of data, descriptor and response variables. In OPLS, the systematic variation in descriptor data is separated into two blocks, predictive variation correlated to response data and non-predictive variation orthogonal to response data. This separation improves the model transparency and reduces the model complexity. OPLS and PLS provide the same predictive accuracy, however, particularly for the two-class discriminant problem OPLS has an advantage over PLS that provides only one single predictive component (first component) and the other orthogonal components (if any) are not important in class separation. Accordingly, one single loading vector describes the class discriminating variables.

The performance of an OPLS model is quantified by two parameters, the goodness of fit ( $R^2$ ) and the goodness of prediction ( $Q^2$ ) (Eriksson et al. 2013).  $R^2$  is the fraction of variation of the training data that can be explained by the components of the model.  $R^2$  shows how well the model fits the training data.  $Q^2$  is the fraction of variation of the training data that can be predicted by the model.  $Q^2$  shows how reliable the model predicts new data.  $Q^2$  is used to find the optimal model complexity, which results in the most valid model with a balance between fit and predictive ability. Therefore,  $Q^2$  is more important than  $R^2$  and a model with higher  $Q^2$  is considered as a better model.  $Q^2$  is estimated by cross validation (CV). CV is a practical approach for evaluating learning algorithms that is based on building of a number of parallel models (Wold 1978). In this work, sevenfold CV was used to calculate  $Q^2$ . In addition to  $Q^2$  and  $R^2$  as performance metrics, classification success rates were reported in terms of the accuracy, sensitivity and specificity.

For univariate comparisons of quantitative and qualitative variables, the independent samples  $t$  test and the  $\chi^2$  test were used respectively.

### Age Correction Methods

Two age correction methods were implemented: (1) a simple approach that treats age as a covariate and includes age in the OPLS model as a separate variable along with MRI-derived variables and (2) a linear detrending algorithm based on age-related changes in the CTL group only. The detrending algorithm fits a generalized linear model (GLM) to each MRI-derived variable and age, in the CTL group only, and models the age-related changes as a linear drift. Then, the regression coefficient of the resulted GLM model (linear drift) is used to remove the age-related changes from all individuals (AD, MCI and CTL) and obtain corrected values. The linear model was chosen based on the Good et al. (2001) study where they found an age-related linear decrease in global grey matter volume in healthy individuals. The assumption for age correction method is that the age related changes in the CTL group are due to aging, while the age related changes in the MCI/AD group includes disease-related changes as well. Therefore, the algorithm calculates age-related effects based on the CTL group only, since removing age-related changes based on the AD or MCI group might also remove the disease-related changes. The detrending method was applied prior to further statistical analysis.

### Implementation

In the first step, three OPLS models were created for classification of AD and CTL subjects: (1) a model based on the raw measures (uncorrected model), (2) a model using age as a covariate (covariate model) and (3) a model based on age-detrended measures (detrended model). Subsequently, the resulted classification models were used to predict MCI patients as unseen data.

The output of the OPLS model is a cross-validated score vector where each score corresponds to one subject. A subject with a score close to one displays a pattern similar to AD and a subject with a score close to zero displays a pattern similar to CTL. A fixed cut-off equal to 0.5 was used to assign class membership to the predicted scores and afterwards to calculate accuracy, sensitivity and specificity (Spulber et al. 2013). Similarly, the prediction result for MCI patients is a score vector and by applying an appropriate cut-off (0.5), MCI patients can be predicted as CTL-like/AD-like. These steps were conducted for each age correction method.

All models were created hierarchically, i.e. volumetric and thickness measures were analyzed separately, and the output scores of these base models were used to create the final model. In the simple age correction method, age was included in the model along with base scores.

## Results

Table 2 summarizes the results from the analyses. The OPLS model based on the original MRI variables (uncorrected model) resulted in  $Q^2 = 0.567$  and  $R^2 = 0.568$ . The model using original MRI data and age as a covariate (covariate model) resulted in  $Q^2 = 0.580$  and  $R^2 = 0.586$ . The model based on age corrected MRI data (detrended model) resulted in  $Q^2 = 0.582$  and  $R^2 = 0.583$ . The classification of AD and CTL subjects with uncorrected model resulted in an accuracy = 86.7 %, the covariate model resulted in an accuracy = 87.3 % and the detrended model resulted in an accuracy = 88.2 %. MCI prediction using the uncorrected model resulted in an accuracy = 62.7 %, the covariate model resulted in an accuracy = 62.9 % and finally the accuracy of the detrended model was 65.0 %.

Medial temporal structures including amygdala, entorhinal cortex and hippocampus and the temporal gyrus regions (inferior, middle and superior) were the most important variables for the separation between the AD and CTL group in all three models. Figure 1 shows the thickness values of entorhinal cortex before and after applying age correction in the different groups. The Pearson correlation coefficients between all MRI measures and age, before and after age correction are given in supplementary material 3.

Table 3 shows the prediction results for the MCI subjects from the ADNI cohort at the different follow-up time points. At each time point, the detrended model resulted in the highest accuracy among the three models. In the uncorrected model, by increasing the follow-up duration from 12 months to 36 months, the prediction accuracy was improved from 60.9 to 66.8 %. In the detrended model, the prediction accuracy rose from 63.0 to 70.8 %.

Additional analyses were performed to further investigate the effect of the age correction methods on classification and prediction models. Table 4 shows the comparison between correct and incorrect classified subjects within each diagnostic group (i.e. within AD and CTL group) and the comparison of incorrect classified subjects between AD and CTL subjects. Without age correction, within both the CTL and the AD group, the mean age of

correctly and incorrectly classified subjects was significantly different ( $p < 0.001$  and  $p = 0.006$ , respectively). After accounting by age, both the covariate and detrended models showed no statistically significant age difference. The mean age of the incorrect classified CTL and AD subjects were 79.1 and 73.3 years, respectively ( $p < 0.001$ ) in the uncorrected model. This difference was eliminated in both the covariate and detrended models. The MMSE score of AD subjects was significantly higher in incorrectly classified subjects in all three models. Moreover, the distribution of ApoE-e4 genotype was significantly different between incorrectly classified CTL and AD subjects.

Table 5 shows the comparison between correctly and incorrectly predicted MCI patients within each group (i.e. within MCI-s and MCI-p group) as well as the comparison of incorrectly predicted subjects between MCI-s and MCI-p group. In MCI-s subjects, the mean age of correctly and incorrectly predicted subjects was significantly different without considering age ( $p < 0.001$ ) but not after accounting for age. Using uncorrected data resulted in 3.9 years difference between the mean-age of misclassified MCI-p and MCI-s ( $p = 0.003$ ), where incorrectly predicted MCI-s patients were older than incorrectly classified MCI-p subjects (77.0 and 73.1 years respectively). This difference was eliminated after age correction. Moreover, incorrectly predicted MCI-s patients showed significantly lower MMSE score and lower frequency of male in all three models. In addition, age correction led to a significant difference in ApoE-e4 distribution in MCI-s subjects, showing lower frequency of ApoE-e4 allele in incorrectly predicted subjects.

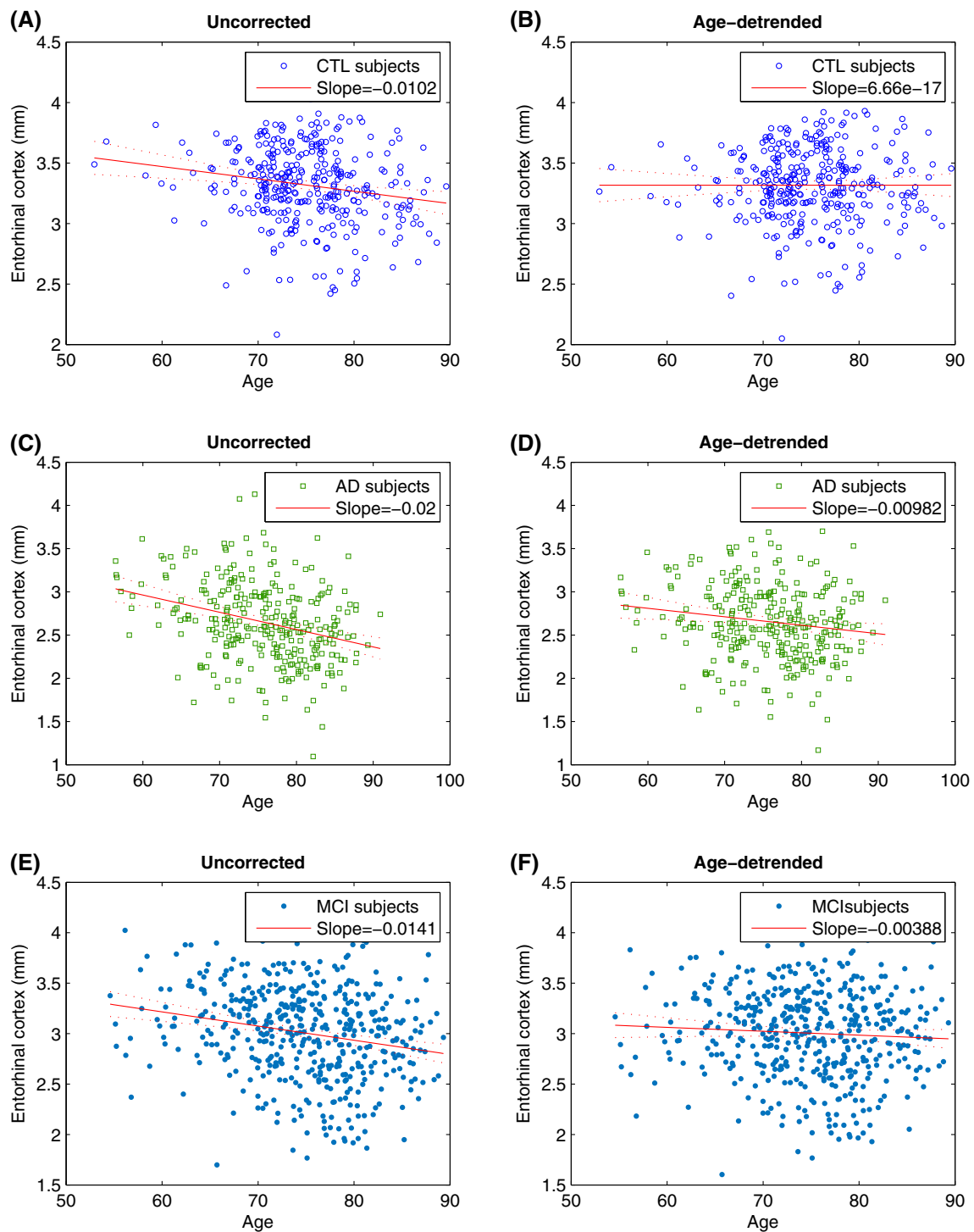
## Discussion

In recent years, there has been an increased interest in using advanced machine learning and multivariate data analysis methods and structural MRI data for early diagnosis of AD. Notably, the discriminative capacity of MRI-derived features and several classifiers for classifying AD patients and CTL individuals and for predicting

**Table 2** AD classification and MCI prediction results

	Model	$Q^2$	$R^2$	Accuracy %	Sensitivity %	Specificity %
AD versus CTL classification (CV)	Uncorrected	0.567	0.568	86.7	81.8	90.9
	Covariate	0.580	0.586	87.3	81.5	92.4
	Detrended	0.582	0.583	88.2	82.8	92.9
MCI prediction	Uncorrected	–	–	62.7	70.2	60.9
	Covariate	–	–	62.9	71.4	60.9
	Detrended	–	–	65.0	75.0	62.6

AD Alzheimer's disease, CTL control subjects, MCI mild cognitive impairment, CV cross-validated,  $Q^2$  goodness of prediction,  $R^2$  goodness of fit



**Fig. 1** The *blue*, *green* and *cyan* markers represent the entorhinal cortex thickness of CTL, AD and MCI subjects before and after age correction with detrending method: **a** CTL subjects before age correction, **b** CTL subjects after age correction, **c** AD patients before

age correction, **d** AD patients after age correction, **e** MCI patients before age correction and **f** MCI patients after age correction. Each marker corresponds to one subject. The *red lines* indicate the age-related drift fitted in the groups (Color figure online)

progression from MCI to AD has been investigated (Liu et al. 2013; Wee et al. 2013; Wolz et al. 2011). The OPLS method in this work resulted in high classification accuracy

and good prediction outcomes. OPLS has previously been used for classification purposes in the same two multi-center cohorts considered here (Westman et al. 2011a).

**Table 3** Prediction results for MCI subjects of ADNI cohort at different follow-up time points

Time point	Model	Accuracy %	Sensitivity %	Specificity %
Month-12 <i>MCI-p</i> = 62, <i>MCI-s</i> = 260	Uncorrected	60.9	72.1	58.2
	Covariate	60.2	70.5	57.9
	Detrended	63.0	75.4	60.2
Month-18 <i>MCI-p</i> = 90, <i>MCI-s</i> = 232	Uncorrected	64.9	72.2	62.1
	Covariate	66.1	74.4	62.9
	Detrended	67.7	75.6	64.7
Month-24 <i>MCI-p</i> = 126, <i>MCI-s</i> = 196	Uncorrected	65.5	66.7	64.8
	Covariate	67.4	69.0	66.3
	Detrended	68.9	69.8	68.4
Month-36 <i>MCI-p</i> = 150, <i>MCI-s</i> = 172	Uncorrected	66.8	65.3	68.0
	Covariate	69.3	68.0	70.3
	Detrended	70.8	68.7	72.7

*MCI* mild cognitive impairment, *MCI-s* stable MCI, *MCI-p* progressive MCI

Despite the input data (dataset subjects) and image processing (FreeSurfer software version) are non-identical in the two studies, the accuracy levels were analogous. It has previously been shown that different advanced classifiers applied to the same data provide similar levels of accuracy (Aguilar et al. 2013). At present, limitations are probably related to input data (quality of data or cohort studied), clinical diagnosis or the confounding effect of some demographic variables such as age, rather than the method used for classification (Falahati et al. 2014b).

Age as a confounding factor can negatively affect the classification and prediction performance. Indeed, the similarity of atrophy patterns in AD patients and in cognitively normal subjects can lead to misclassification of young AD patients and old CTL subjects. Global and regional changes of brain volumes in normal aging have been reported in cross-sectional and longitudinal brain imaging studies (Giorgio et al. 2010; Good et al. 2001; Scahill et al. 2003; Walhovd et al. 2005). Additionally, several studies have reported that brain atrophy accelerates with disease progression in AD and other types of dementia (Fox et al. 1996; Jack et al. 2004; Sabuncu et al. 2011; Tisserand et al. 2004). Particularly, global brain atrophy and reduced volume in the temporal lobe especially in hippocampus and entorhinal cortex have been reported.

Using age as a covariate in statistical models is a common way to deal with this problem. Recently, new approaches such as a data correction method based on linear regression models (Dukart et al. 2011; Koikkalainen et al. 2012), and confounder correcting support vector machine algorithm (Li et al. 2011) have been proposed to more effectively control for the effect of age as a confounding factor. In this study, two approaches i.e. age as a covariate and detrending age-related changes were investigated. Both studied approaches here have pros and cons. In heterogeneous populations containing AD, MCI

and CTL subjects with different patterns and rates of atrophy simply using age as a covariate may not be an optimal approach. However, the OPLS method seems to be able to handle age as a covariate. Detrending age-related changes is challenging since modeling the exact association between age and discriminative features and subsequently remove such associations can be difficult and time consuming. One of the main ideas behind the detrending method was to remove age-related changes while preserving the disease-related changes for each variable separately. The hypothesis was that by detrending the AD and the MCI group based on CTL group, the age-related changes would be omitted and the disease-related changes would be kept. Therefore the control group should be representative of population and equally distributed on the age range. Hence the detrending method may be more effective in a larger dataset. In some variables (e.g. hippocampus) the slopes for MCI and AD were slightly tilted in the opposite direction after age correction, indicating that the algorithm may overcorrect in older AD and MCI cases. A reduction in rates of atrophy in older AD and MCI compared to CTL could be a possible explanation for the observed pattern. A recent study has shown that a pronounced reduction in rates of atrophy can be observed in AD and MCI individuals with increasing age, while for cognitively normal individuals, increasing age leads to increased rates of atrophy (Holland et al. 2012).

The results indicate that when age is included, the quality of models in terms of the goodness of fit and more importantly the goodness of prediction was improved which led to higher classification and prediction accuracies. Although accurate discrimination between AD patients and CTL subjects is of great interest, prediction of progression from MCI to AD is more valuable since it can provide an opportunity for early detection of individuals under risk of developing dementia. Generally, MCI

**Table 4** Subjects' characteristics in AD versus CTL classification: comparison between correctly/incorrectly classified subjects

AD versus CTL Classification	CTL subjects			AD patients			Incorrect-classified AD versus CTL $p$ value
	Correct-classified	Incorrect-classified	$p$ value	Correct-classified	Incorrect-classified	$p$ value	
<i>Uncorrected model</i>							
Count	309	31		243	54		
Age, years	74.5 ± 5.6	79.1 ± 5.2	<b>&lt;0.001<sup>a</sup></b>	76.2 ± 6.7	73.3 ± 7.9	<b>0.006<sup>a</sup></b>	<b>&lt;0.001<sup>a</sup></b>
Education, years	14.3 ± 4.3	14.1 ± 5.5	0.839 <sup>a</sup>	11.9 ± 4.8	12.8 ± 5.1	0.197 <sup>a</sup>	0.294 <sup>a</sup>
MMSE score	29.1 ± 1.1	28.9 ± 1.2	0.326 <sup>a</sup>	21.8 ± 3.7	24.2 ± 2.8	<b>&lt;0.001<sup>a</sup></b>	<b>&lt;0.001<sup>a</sup></b>
Gender, Male/Female	152/157	16/15	0.797 <sup>b</sup>	107/136	25/29	0.762 <sup>b</sup>	0.637 <sup>b</sup>
ApoE-e4, N/P	217/87	26/5	0.138 <sup>b</sup>	86/149	24/29	0.240 <sup>b</sup>	<b>0.001<sup>b</sup></b>
Cohort, ADNI/ANM	207/102	20/11	0.780 <sup>b</sup>	141/102	34/20	0.505 <sup>b</sup>	0.886 <sup>b</sup>
<i>Covariate model</i>							
Count	314	26		242	55		
Age, years	74.8 ± 5.7	77.0 ± 4.3	0.058 <sup>a</sup>	75.7 ± 6.8	75.4 ± 7.7	0.776 <sup>a</sup>	0.237 <sup>a</sup>
Education, years	14.2 ± 4.2	14.0 ± 5.6	0.781 <sup>a</sup>	11.9 ± 4.7	12.5 ± 5.3	0.412 <sup>a</sup>	0.269 <sup>a</sup>
MMSE score	29.1 ± 1.1	28.9 ± 1.1	0.326 <sup>a</sup>	21.7 ± 3.7	24.4 ± 2.7	<b>&lt;0.001<sup>a</sup></b>	<b>&lt;0.001<sup>a</sup></b>
Gender, Male/Female	153/161	15/11	0.380 <sup>b</sup>	103/139	29/26	0.171 <sup>b</sup>	0.675 <sup>b</sup>
ApoE-e4, N/P	223/84	20/5	0.385 <sup>b</sup>	83/151	27/27	<b>0.048<sup>b</sup></b>	<b>0.012<sup>b</sup></b>
Cohort, ADNI/ANM	210/104	17/9	0.876 <sup>b</sup>	143/99	32/23	0.902 <sup>b</sup>	0.536 <sup>b</sup>
<i>Detrended model</i>							
Count	316	24		246	51		
Age, years	75.0 ± 5.8	75.3 ± 4.6	0.756 <sup>a</sup>	75.5 ± 7.0	76.3 ± 7.1	0.493 <sup>a</sup>	0.553 <sup>a</sup>
Education, years	14.4 ± 4.3	12.8 ± 5.6	0.205 <sup>a</sup>	12.0 ± 4.8	12.4 ± 5.0	0.530 <sup>a</sup>	0.756 <sup>a</sup>
MMSE score	29.2 ± 1.1	28.8 ± 1.4	0.168 <sup>a</sup>	21.8 ± 3.7	24.4 ± 2.7	<b>&lt;0.001<sup>a</sup></b>	<b>&lt;0.001<sup>a</sup></b>
Gender, Male/Female	155/161	13/11	0.629 <sup>b</sup>	105/141	27/24	0.180 <sup>b</sup>	0.921 <sup>b</sup>
ApoE-e4, N/P	225/87	18/5	0.524 <sup>b</sup>	86/151	24/27	0.156 <sup>b</sup>	<b>0.012<sup>b</sup></b>
Cohort, ADNI/ANM	213/103	14/10	0.363 <sup>b</sup>	144/102	31/20	0.767 <sup>b</sup>	0.840 <sup>b</sup>

$p$  value in bold indicates statistically significant difference

Continuous data is represented as mean ± standard deviation, AD Alzheimer's disease, CTL control subjects, ApoE apolipoprotein E, N/P negative/positive for at least one e4 allele, MMSE mini mental state examination, ADNI/ANM Alzheimer's disease Neuroimaging Initiative/AddNeuroMed

<sup>a</sup> Independent-samples  $t$  test

<sup>b</sup> Pearson  $\chi^2$

prediction accuracy is not as good as classification of AD and CTL subjects. Reviewing the literature, classification accuracies tend to range between 80 and 90 %, mostly accompanied by lower prediction accuracies for MCI progression (Falahati et al. 2014b). The MCI group is clinically quite heterogeneous. Some MCI subjects progress to AD or even other neurological disorders, some remain stable over time, with a smaller number reverting to a cognitively normal status (Mitchell and Shiri-Feshki 2009). In addition, one-year is a relatively short follow-up time. When the subjects were followed for a longer period, the accuracy increased. Including age in the models also improved prediction accuracies. Although, the covariate

and detrended models performed similarly in terms of model quality values, the detrended model induced the highest accuracies in all settings.

A detailed analysis of correctly classified and incorrectly classified subjects provided valuable information on the age correction performance. As expected, younger AD patients and older CTL subjects were more likely to be misclassified. Similarly, younger MCI-p and older MCI-s patients were prone to misclassification. These findings are in line with previous studies (Dukart et al. 2011).

Incorrectly classified AD subjects had significantly higher MMSE score compared to correctly classified AD subjects. This can be explained by the previous finding that in AD

**Table 5** Subjects' characteristics in MCI prediction: comparison between correctly/incorrectly classified subjects

MCI prediction		MCI-s subjects			MCI-p subjects			Incorrect-predicted MCI-s versus MCI-p <i>p</i> value
		Correct- predicted	Incorrect- predicted	<i>p</i> value	Correct- predicted	Incorrect- predicted	<i>p</i> value	
Uncorrected model	Count	219	141		59	26		
	Age, years	73.6 ± 7.1	77.0 ± 6.1	<b>&lt;0.001<sup>a</sup></b>	75.0 ± 6.9	73.1 ± 5.3	0.220 <sup>a</sup>	<b>0.003<sup>a</sup></b>
	Education, years	13.8 ± 4.6	14.1 ± 4.7	0.582 <sup>a</sup>	13.8 ± 4.4	13.7 ± 3.8	0.820 <sup>a</sup>	0.651 <sup>a</sup>
	MMSE score	27.4 ± 1.7	26.8 ± 1.6	<b>0.001<sup>a</sup></b>	26.5 ± 1.9	26.6 ± 1.6	0.802 <sup>a</sup>	0.682 <sup>a</sup>
	Gender, Male/Female	144/75	75/66	<b>0.017<sup>b</sup></b>	35/24	15/11	0.888 <sup>b</sup>	0.672 <sup>b</sup>
	ApoE-e4, N/P	118/87	65/71	0.077 <sup>b</sup>	22/36	8/16	0.694 <sup>b</sup>	0.190 <sup>b</sup>
	Cohort, ADNI/ANM	151/68	109/32	0.084 <sup>b</sup>	44/15	18/8	0.609 <sup>b</sup>	0.375 <sup>b</sup>
Covariate model	Count	219	141		60	25		
	Age, years	74.7 ± 7.1	75.3 ± 6.6	0.419 <sup>a</sup>	74.3 ± 7.0	74.6 ± 5.0	0.852 <sup>a</sup>	0.635 <sup>a</sup>
	Education, years	13.8 ± 4.6	14.2 ± 4.8	0.394 <sup>a</sup>	13.9 ± 4.4	13.7 ± 3.9	0.853 <sup>a</sup>	0.614 <sup>a</sup>
	MMSE score	27.4 ± 1.7	26.8 ± 1.7	<b>0.001<sup>a</sup></b>	26.6 ± 1.9	26.4 ± 1.6	0.740 <sup>a</sup>	0.383 <sup>a</sup>
	Gender, Male/Female	149/70	70/71	<b>&lt;0.001<sup>b</sup></b>	37/23	13/12	0.409 <sup>b</sup>	0.828 <sup>b</sup>
	ApoE-e4, N/P	122/83	61/75	<b>0.008<sup>b</sup></b>	20/39	10/13	0.418 <sup>b</sup>	0.902 <sup>b</sup>
	Cohort, ADNI/ANM	150/69	110/31	0.050 <sup>b</sup>	43/17	19/6	0.682 <sup>b</sup>	0.824 <sup>b</sup>
Detrended model	Count	225	135		63	22		
	Age, years	75.2 ± 7.1	74.4 ± 6.6	0.273 <sup>a</sup>	74.1 ± 6.9	75.4 ± 4.9	0.408 <sup>a</sup>	0.492 <sup>a</sup>
	Education, years	13.8 ± 4.5	14.2 ± 4.8	0.478 <sup>a</sup>	14.0 ± 4.1	13.2 ± 4.5	0.418 <sup>T</sup>	0.380 <sup>a</sup>
	MMSE score	27.4 ± 1.7	26.7 ± 1.7	<b>0.001<sup>a</sup></b>	26.5 ± 1.9	26.6 ± 1.6	0.775 <sup>a</sup>	0.803 <sup>a</sup>
	Gender, Male/Female	154/71	65/70	<b>&lt;0.001<sup>b</sup></b>	39/24	11/11	0.329 <sup>b</sup>	0.872 <sup>b</sup>
	ApoE-e4, N/P	127/83	56/75	<b>0.001<sup>b</sup></b>	40/22	12/8	0.715 <sup>b</sup>	0.817 <sup>b</sup>
	Cohort, ADNI/ANM	156/69	104/31	0.114 <sup>b</sup>	46/17	16/6	0.979 <sup>b</sup>	0.659 <sup>b</sup>

*p* value in bold indicates statistically significant difference

Continuous data is represented as mean ± SD, *MCI* mild cognitive impairment, *MCI-p* progressive MCI, *MCI-s* stable MCI, *ApoE* apolipoprotein E, *N/P* negative/positive for at least one e4 allele, *MMSE* mini mental state examination, *ADNI/ANM* Alzheimer's disease Neuroimaging Initiative/AddNeuroMed

<sup>a</sup> Independent-samples *t* test

<sup>b</sup> Pearson  $\chi^2$

patients, decreased MMSE score correlates with gray matter reduction (Frisoni et al. 2002) and correlates with aging (Pradier et al. 2014). In MCI-s subjects, the mean MMSE score was slightly higher in correctly classified subjects.

Interestingly, incorrectly classified MCI-s subjects were more frequently ApoE-e4 positive compared to correctly classified MCI-s (Table 5). This is in line with previous studies that reported more regional atrophy in AD patients with presence of genetic risk, especially in the medial temporal cortex (Cherbuin et al. 2007; Ferreira et al. 2015; van der Flier et al. 2011). The incorrectly classified MCI-s subjects may thus have a high risk of developing AD in the future.

The frequency of misclassification was higher in female MCI-s subjects compared to males, indicating that more female subjects had AD-like structural patterns. This is in line with recent findings that the female sex is associated with an increased risk of disease progression (Tifratene

et al. 2015). Although this difference exists even before age correction, after correcting for age, more female subjects were prone to be misclassified which could emphasize the role of age. This can support the age- by sex- related differences in progression rates that proposed in several studies (Mielke et al. 2014; Roberts et al. 2014).

The characteristics of correctly and incorrectly classified/predicted subjects were similar for both correction approaches. In fact, the association between structural brain changes, age, sex, ApoE genotype, cognitive status and other factors are more complicated than pairwise relations. Although the relationship between these factors are explored from several perspectives such as age by sex relations (Fratiglioni et al. 1997), sex by ApoE genotype relation (Altmann et al. 2014), etc., the connection between them is poorly understood. Considering age in multivariate models, regardless of approach, can potentially enhance the outcomes.

## Conclusion

Both age correction approaches (age as a covariate and detrending) could effectively eliminate the age differences in classification and prediction results. Moreover, including age in the models highlighted the role of the other disease-related factors such as cognitive impairment and ApoE-e4 genotype. These results demonstrate that age is partially masking other relevant factors such as ApoE genotype, global cognitive impairment and sex. This is an important finding, suggesting that mechanisms underlying the confounding effect of these factors should be further investigated. At the time being, clinicians are already quite aware about the effect of age when interpreting imaging data for diagnostic purposes. Therefore, the other factors should also be carefully considered when adjusting diagnostic interpretations of imaging data in clinical settings. The exact relationship between normal ageing and AD is far from being fully understood at present and warrants further investigations. Non-linear correction methods and other alternatives for handling confounding factors should be further investigated. Applying correction methods to other confounding factors such as education and sex would be of interest and could potentially improve prediction accuracy of MCI progression further.

**Acknowledgments** Thanks to the Strategic Research Programme in Neuroscience at Karolinska Institutet (StratNeuro), Swedish Brain Power, the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Loo and Hans Osterman foundation for medical research, Stiftelsen för gamla tjänarinnor, the foundation for geriatric diseases at Karolinska Institutet, the research grants from Karolinska Institutet, and the Axel and Signe Lagermans Donation Foundations. Dr. Soininen has received funding from Academy of Finland, EVO Kuopio University Hospital and strategic funding from University of Eastern Finland, UEFBRAIN. This study was supported by InnoMed, (Innovative Medicines in Europe) an Integrated Project funded by the European Union of the Sixth Framework program priority FP6-2004-LIFESCIHEALTH-5, Life Sciences, Genomics and Biotechnology for Health. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support

ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Aguilar C et al (2013) Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. *Psychiatry Res* 212:89–98. doi:[10.1016/j.psychres.2012.11.005](https://doi.org/10.1016/j.psychres.2012.11.005)
- Aguilar C et al (2014) Application of a MRI based index to longitudinal atrophy change in Alzheimer disease, mild cognitive impairment and healthy older individuals in the AddNeuroMed cohort. *Front Aging Neurosci* 6:145. doi:[10.3389/fnagi.2014.00145](https://doi.org/10.3389/fnagi.2014.00145)
- Albert MS et al (2011) The diagnosis of mild cognitive impairment 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 & dementia. *J Alzheimer's Assoc* 7:270–279. doi:[10.1016/j.jalz.2011.03.008](https://doi.org/10.1016/j.jalz.2011.03.008)
- Altmann A, Tian L, Henderson VW, Greicius MD (2014) Alzheimer's Disease Neuroimaging Initiative I sex modifies the APOE-related risk of developing Alzheimer disease. *Ann Neurol* 75:563–573. doi:[10.1002/ana.24135](https://doi.org/10.1002/ana.24135)
- Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J (2006) OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemom* 20:341–351. doi:[10.1002/cem.1006](https://doi.org/10.1002/cem.1006)
- Cherbuin N, Leach LS, Christensen H, Anstey KJ (2007) Neuroimaging and APOE genotype: a systematic qualitative review. *Dement Geriatr Cogn Disord* 24:348–362. doi:[10.1159/000109150](https://doi.org/10.1159/000109150)
- Dukart J, Schroeter ML, Mueller K, Alzheimer's Disease Neuroimaging I (2011) Age correction in dementia—matching to a healthy brain. *PLoS One* 6:e22193. doi:[10.1371/journal.pone.0022193](https://doi.org/10.1371/journal.pone.0022193)
- Eriksson L, Byrne T, Johansson E, Trygg J, Vikström C (2013) Multi- and megavariate data analysis basic principles and applications, Third revised edn. MKS Umetrics AB, Malmö, Sweden
- Falahati F, Fereshtehnejad SM, Religa D, Wahlund LO, Westman E, Eriksdotter M (2014a) The use of MRI CT and lumbar puncture in dementia diagnostics: data from the svedem registry. *Dement Geriatr Cogn Disord* 39:81–91. doi:[10.1159/000366194](https://doi.org/10.1159/000366194)
- Falahati F, Westman E, Simmons A (2014b) Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *J Alzheimers Dis* 41:685–708. doi:[10.3233/JAD-131928](https://doi.org/10.3233/JAD-131928)
- Ferreira D et al (2014) Cognitive decline is mediated by gray matter changes during middle age. *Neurobiol Aging* 35:1086–1094. doi:[10.1016/j.neurobiolaging.2013.10.095](https://doi.org/10.1016/j.neurobiolaging.2013.10.095)
- Ferreira D et al (2015) Practical cut-offs for visual rating scales of medial temporal, frontal and posterior atrophy in Alzheimer's

- disease and mild cognitive impairment. *J Intern Med* 278:277–290. doi:[10.1111/joim.12358](https://doi.org/10.1111/joim.12358)
- Fox NC, Freeborough PA, Rossor MN (1996) Visualisation and quantification of rates of atrophy in Alzheimer's disease. *Lancet* 348:94–97. doi:[10.1016/S0140-6736\(96\)05228-2](https://doi.org/10.1016/S0140-6736(96)05228-2)
- Fratiglioni L, Viitanen M, von Strauss E, Tontodonati V, Herlitz A, Winblad B (1997) Very old women at highest risk of dementia and Alzheimer's disease: incidence data from the Kungsholmen Project. *Stockh Neurol* 48:132–138. doi:[10.1212/wnl.48.1.132](https://doi.org/10.1212/wnl.48.1.132)
- Frisoni GB, Testa C, Zorzan A, Sabatoli F, Beltramello A, Soininen H, Laakso MP (2002) Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry. *J Neurol Neurosurg Psychiatry* 73:657–664. doi:[10.1136/jnnp.73.6.657](https://doi.org/10.1136/jnnp.73.6.657)
- Giorgio A, Santelli L, Tomassini V, Bosnell R, Smith S, De Stefano N, Johansen-Berg H (2010) Age-related changes in grey and white matter structure throughout adulthood. *Neuroimage* 51:943–951. doi:[10.1016/j.neuroimage.2010.03.004](https://doi.org/10.1016/j.neuroimage.2010.03.004)
- Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS (2001) A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* 14:21–36. doi:[10.1006/nimg.2001.0786](https://doi.org/10.1006/nimg.2001.0786)
- Holland D, Desikan RS, Dale AM, McEvoy LK (2012) Rates of decline in Alzheimer disease decrease with age. *PloS One* 7:e42325. doi:[10.1371/journal.pone.0042325](https://doi.org/10.1371/journal.pone.0042325)
- Jack CR Jr et al (2004) Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. *Neurology* 62:591–600. doi:[10.1212/01.wnl.0000110315.26026.ef](https://doi.org/10.1212/01.wnl.0000110315.26026.ef)
- Jack CR Jr et al (2008) The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magne Reson Imaging* 27:685–691. doi:[10.1002/jmri.21049](https://doi.org/10.1002/jmri.21049)
- Koikkalainen J, Polonen H, Mattila J, van Gils M, Soininen H, Lotjonen J, Alzheimer's Disease Neuroimaging I (2012) Improved classification of Alzheimer's disease data via removal of nuisance variability. *PLoS One* 7:e31112. doi:[10.1371/journal.pone.0031112](https://doi.org/10.1371/journal.pone.0031112)
- Li L, Rakitsch B, Borgwardt K (2011) ccSVM: correcting Support Vector Machines for confounding factors in biological data classification. *Bioinformatics* 27:i342–i348. doi:[10.1093/bioinformatics/btr204](https://doi.org/10.1093/bioinformatics/btr204)
- Liu X, Tosun D, Weiner MW, Schuff N, Alzheimer's Disease Neuroimaging I (2013) Locally linear embedding (LLE) for MRI based Alzheimer's disease classification. *Neuroimage* 83:148–157. doi:[10.1016/j.neuroimage.2013.06.033](https://doi.org/10.1016/j.neuroimage.2013.06.033)
- Lovestone S, Francis P, Strandgaard K (2007) Biomarkers for disease modification trials—the innovative medicines initiative and AddNeuroMed. *J Nutr Health Aging* 11:359–361
- Maioli S et al (2015) Alterations in brain leptin signalling in spite of unchanged CSF leptin levels in Alzheimer's disease. *Aging Cell* 14:122–129. doi:[10.1111/ace1.12281](https://doi.org/10.1111/ace1.12281)
- McKhann GM et al (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 & dementia. *J Alzheimer's Assoc* 7:263–269. doi:[10.1016/j.jalz.2011.03.005](https://doi.org/10.1016/j.jalz.2011.03.005)
- Mielke MM, Vemuri P, Rocca WA (2014) Clinical epidemiology of Alzheimer's disease: assessing sex and gender differences. *Clin Epidemiol* 6:37–48. doi:[10.2147/CLEP.S37929](https://doi.org/10.2147/CLEP.S37929)
- Mitchell AJ, Shiri-Feshki M (2009) Rate of progression of mild cognitive impairment to dementia—meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand* 119:252–265. doi:[10.1111/j.1600-0447.2008.01326.x](https://doi.org/10.1111/j.1600-0447.2008.01326.x)
- Muehlboeck J-S, Westman E, Simmons A (2014) TheHiveDB image data management and analysis framework. *Front Neuroinform* 7:49. doi:[10.3389/fninf.2013.00049](https://doi.org/10.3389/fninf.2013.00049)
- Mueller SG et al (2005) Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimer's Dementia* 1:55–66. doi:[10.1016/j.jalz.2005.06.003](https://doi.org/10.1016/j.jalz.2005.06.003)
- Petersen RC et al (2010) Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* 74:201–209. doi:[10.1212/WNL.0b013e3181cb3e25](https://doi.org/10.1212/WNL.0b013e3181cb3e25)
- Pradier C et al (2014) The mini mental state examination at the time of Alzheimer's disease and related disorders diagnosis, according to age, education, gender and place of residence: a cross-sectional study among the French National Alzheimer database. *PLoS One* 9:e103630. doi:[10.1371/journal.pone.0103630](https://doi.org/10.1371/journal.pone.0103630)
- Roberts RO et al (2014) Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal. *Neurology* 82:317–325. doi:[10.1212/WNL.0000000000000055](https://doi.org/10.1212/WNL.0000000000000055)
- Sabuncu MR et al (2011) The dynamics of cortical and hippocampal atrophy in Alzheimer disease. *Arch Neurol* 68:1040–1048. doi:[10.1001/archneurol.2011.167](https://doi.org/10.1001/archneurol.2011.167)
- Scahill RI, Frost C, Jenkins R, Whitwell JL, Rossor MN, Fox NC (2003) A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch Neurol* 60:989–994. doi:[10.1001/archneur.60.7.989](https://doi.org/10.1001/archneur.60.7.989)
- Scheltens P et al (1992) Atrophy of medial temporal lobes on MRI in “probable” Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *J Neurol Neurosurg Psychiatry* 55:967–972. doi:[10.1136/jnnp-2012-302562](https://doi.org/10.1136/jnnp-2012-302562)
- Simmons A et al (2011) The AddNeuroMed framework for multi-centre MRI assessment of Alzheimer's disease: experience from the first 24 months. *Int J Geriatr Psychiatry* 26:75–82. doi:[10.1002/gps.2491](https://doi.org/10.1002/gps.2491)
- Spulber G et al (2013) An MRI-based index to measure the severity of Alzheimer's disease-like structural pattern in subjects with mild cognitive impairment. *J Intern Med* 273:396–409. doi:[10.1111/joim.12028](https://doi.org/10.1111/joim.12028)
- Tifratene K, Robert P, Metelkina A, Pradier C, Dartigues JF (2015) Progression of mild cognitive impairment to dementia due to AD in clinical settings. *Neurology*. doi:[10.1212/WNL.0000000000001788](https://doi.org/10.1212/WNL.0000000000001788)
- Tisserand DJ, van Boxtel MP, Pruessner JC, Hofman P, Evans AC, Jolles J (2004) A voxel-based morphometric study to determine individual differences in gray matter density associated with age and cognitive change over time. *Cereb Cortex* 14:966–973. doi:[10.1093/cercor/bhh057](https://doi.org/10.1093/cercor/bhh057)
- Trygg J, Wold S (2002) Orthogonal projections to latent structures (O-PLS). *J Chemom* 16:119–128. doi:[10.1002/Cem.695](https://doi.org/10.1002/Cem.695)
- van der Flier WM, Pijnenburg YA, Fox NC, Scheltens P (2011) Early-onset versus late-onset Alzheimer's disease: the case of the missing APOE varepsilon4 allele. *Lancet Neurol* 10:280–288. doi:[10.1016/S1474-4422\(10\)70306-9](https://doi.org/10.1016/S1474-4422(10)70306-9)
- Walhovd KB et al (2005) Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiol Aging* 26:1261–1270. doi:[10.1016/j.neurobiolaging.2005.05.020](https://doi.org/10.1016/j.neurobiolaging.2005.05.020) **discussion 1275–1268**
- Wee CY, Yap PT, Shen D (2013) Alzheimer's Disease Neuroimaging I prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns. *Hum Brain Mapp* 34:3411–3425. doi:[10.1002/hbm.22156](https://doi.org/10.1002/hbm.22156)
- Westman E et al (2010) Combining MRI and MRS to distinguish between Alzheimer's disease and healthy controls. *J Alzheimers Dis* 22:171–181. doi:[10.3233/Jad-2010-100168](https://doi.org/10.3233/Jad-2010-100168)
- Westman E et al (2011a) AddNeuroMed and ADNI: similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America. *Neuroimage* 58:818–828. doi:[10.1016/j.neuroimage.2011.06.065](https://doi.org/10.1016/j.neuroimage.2011.06.065)
- Westman E et al (2011b) Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls. *Neuroimage* 54:1178–1187. doi:[10.1016/j.neuroimage.2010.08.044](https://doi.org/10.1016/j.neuroimage.2010.08.044)
- Westman E, Aguilar C, Muehlboeck JS, Simmons A (2013) Regional magnetic resonance imaging measures for multivariate analysis

- in Alzheimer's disease and mild cognitive impairment. *Brain Topogr* 26:9–23. doi:[10.1007/s10548-012-0246-x](https://doi.org/10.1007/s10548-012-0246-x)
- Wold S (1978) Cross-validatory estimation of the number of components in factor and principal components models. *Technometrics* 20:397–405. doi:[10.1080/00401706.1978.10489693](https://doi.org/10.1080/00401706.1978.10489693)
- Wold S, Ruhe A, Wold H, Dunn WJ (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J Sci Stat Comput* 5:735–743. doi:[10.1137/0905052](https://doi.org/10.1137/0905052)
- Wolz R et al (2011) Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. *PLoS One* 6:e25446. doi:[10.1371/journal.pone.0025446](https://doi.org/10.1371/journal.pone.0025446)