Multimodal MRI-based classification of Alzheimer's disease

Vos, F. de

Combining multiple anatomical MRI measures improves Alzheimer's disease classification

Published in Human Brain Mapping, 2016; 37(5), 1920-1929

Frank de Vos, Tijn M. Schouten, Anne Hafkemeijer, Elise G. P. Dopper,
John C. van Swieten, Mark de Rooij, Jeroen van der Grond
& Serge A. R. B. Rombouts

Abstract

Several anatomical MRI markers for Alzheimer's disease (AD) have been identified. Hippocampal volume, cortical thickness and grey matter density have been used successfully to discriminate AD patients from controls. These anatomical MRI measures have so far mainly been used separately. The full potential of anatomical MRI scans for AD diagnosis might thus not yet have been used optimally. In the current study, we therefore combined multiple anatomical MRI measures in order to improve diagnostic classification of AD. For 21 clinically diagnosed AD patients and 21 cognitively normal controls we calculated i) cortical thickness, ii) cortical area, iii) cortical curvature, iv) grey matter density, v) subcortical volumes and vi) hippocampal shape. These six measures were used separately and combined as predictors in an elastic net logistic regression. We made receiver operating characteristic curves and calculated the area under the curve (AUC) to determine classification performance. AUC values for the single measures ranged from 0.67 (cortical thickness) to 0.94 (grey matter density). The combination of all six measures resulted in an AUC of 0.98. Our results demonstrate that the different anatomical MRI measures contain complementary information. A combination of these measures may therefore improve accuracy of AD diagnosis in clinical practice.

Keywords: Alzheimer's disease, Anatomical MRI, Cortical thickness, Cortical area, Cortical curvature, Grey matter density, Subcortical volumes, Hippocampal shape, Classification

2.1 Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder that is characterised by both focal and global grey matter atrophy. Hippocampal atrophy is considered to be the hallmark of AD and it is therefore used as a clinical marker (Morra et al., 2009). Grey matter atrophy in AD patients frequently extends to other brain regions, including subcortical structures and the medial temporal lobe (Seeley et al., 2009; Rombouts et al., 2000; Jack et al., 2004). The location and degree of grey matter atrophy can be accurately visualised with anatomical magnetic resonance imaging (MRI) scans, which is used for the clinical diagnosis of AD (Frisoni et al., 2010). Studies using anatomical MRI scans have showed that AD patients have decreased volumes and altered shapes of the hippocampi compared to cognitively normal elderly (Thompson et al., 2004; Scher et al., 2007) and also the volumes of the putamen and thalamus are reduced in AD patients (de Jong et al., 2008). Moreover, reduced cortical thickness (Lerch et al., 2005) as well as widespread grey matter atrophy have been demonstrated in patients with AD compared with controls (Karas et al., 2003).

However, the group differences found in case control studies are not necessarily useful in a clinical setting. Many AD markers have also been observed in healthy ageing (Salat et al., 1999). AD markers are only helpful in a clinical setting if they can accurately discriminate AD patients from non-affected subjects at the individual level. The focus of research on anatomical MRI biomarkers for AD has therefore shifted from the detection of group differences towards disease classification. Cortical thickness, MRI-based morphometry (VBM) and the volume and shape of the hippocampus have been used to discriminate AD patients from controls with moderate to high classification accuracy. (Cuingnet et al., 2010; Davatzikos et al., 2011; Querbes et al., 2009).

Thus far, these anatomical MRI measures have mainly been used separately to classify AD patients. However, the different measures may possess complementary information, and the combination of these measures could therefore increase AD classification accuracy compared to the separate measures. For example, voxel-based cortical thickness (VBCT) and VBM show different aspects of age-associated decline in grey matter (Hutton et al., 2009). Also, VBM, cortical folding and cortical thickness complement each other in showing neurodegenerative changes related to Parkinson's disease (Pereira et al., 2012). Moreover, the combination of different anatomical MRI measures improves AD classification (Wolz et al., 2011; Westman et al., 2013; Bron et al., 2015).

Furthermore, advances in statistical learning have facilitated the integration of different sources of information into a single predictive model (Zou & Hastie, 2005). This enables the incorporation of many predictors into one predictive model by selecting only the relevant information out of these many predictors. These techniques have already been applied in order to separate AD patients from controls or to separate MCI converters from MCI non-converters. (Cui et al., 2011; Dyrba et al., 2015a; Dyrba et al.; 2015b; Schouten et al., 2016; Trzepacz et al., 2014; Teipel et al., 2015; Wee et al., 2013; Zhang et al., 2013). It would therefore make sense to combine all anatomical MRI measures that have been shown to be discriminative for AD into a single model to improve sensitivity and specificity.

In this study we will use anatomical MRI scans from a group of AD patients and a group of cognitively normal controls and calculate several commonly used measures that are informative for AD. These measures are: i) cortical thickness, ii) cortical surface area, iii) cortical curvature, iv) grey matter density, v) the volume of the subcortical structures and vi) the shape of the hippocampus. We will combine all these measures into a single predictive model and calculate its classification performance. We hypothesise that the combination will outperform the separate measures.

2.2 Methods

Participants

Anatomical MRI scans were obtained from 21 probable AD patients (10 females) between ages 50 and 87 (M = 71.7, SD = 9.3) and 21 cognitively healthy controls (10 females) between ages 57 and 80 (M = 68.0, SD = 7.5). The AD patients had an average score on the mini-mental state examination (MMSE) of 23 (SD = 2.4) and the cognitively healthy controls had an average score of 28 (SD = 1.5). All AD patients underwent a standardised dementia screening that included their medical history, informant-based history, physical and neurological examination, and an extensive neuropsychological assessment including the MMSE. Diagnoses were made in a multidisciplinary consensus meeting according to the core clinical criteria of the National Institute on Aging and the Alzheimer's Association workgroup for probable AD (Mckhann et al., 1984; McKhann, 2011). As control subjects, we included cognitively healthy elderly volunteers. This study was approved by the medical ethical committee of the Leiden University Medical Center, and all participants provided written informed consent.

MR image acquisition

All participants were scanned on a Philips 3 Tesla Achieva MRI scanner in the Leiden University Medical Center. Three-dimensional T1-weighted structural scans were acquired with the following parameters: TR = 9.8 ms, TE = 4.6 ms, flip angle = 8, 140 slices, voxel size = 0.88 x 0.88 x 1.20 mm.

Cortical thickness, area, and curvature

To calculate cortical thickness, cortical area, and cortical curvature we processed the T1-weighted images using the Freesurfer software package version 5.3.0 (Dale et al., 1999; Fisch et al., 1999). First, intensity normalisation was applied to the brain-extracted image to create an image with relatively high contrast to noise ratio. This image was used to locate the boundary between grey and white matter. A triangular mesh was then constructed around the white matter surface. This triangular mesh consists of over 160,000 vertices for each hemisphere. To create the grey matter surface, the mesh was deformed outwards so that it closely followed the boundary between grey matter and cerebrospinal fluid (CSF). Cortical thickness was calculated as the distance between the white matter surface and the grey matter surface for each vertex. The image was then registered to the Freesurfer common template, using the image's cortical folding pattern. The neocortex was parcellated into the 68 neocortical regions (34 regions for each hemisphere) of the Desikan-Killiany atlas (Desikan et al., 2006). The thickness of each parcellation unit was calculated as the mean thickness of all the vertices within that parcellation. This yielded 68 cortical thickness features per subject. To calculate cortical surface area we summed the areas of the grey matter mesh triangles for each parcellation, which yielded 68 cortical area features per subject. Cortical curvature was calculated as the mean of the curvature values in the two principal directions of the surface. The curvature of a vertex in these directions was calculated as the inverse of the length of the radius of the osculating circles in these directions (Ronan et al., 2011). For each of the parcellations we averaged the curvature values of the vertices, which yielded 68 cortical curvature features per subject.

Grey matter density of the cortical structures

We calculated grey matter density using VBM in the FMRIB Software Library (FSL version 5.0.7; Ashburner et al., 2000; Smith et al., 2004). First, we segmented the brain-extracted images into grey matter, white matter, and cerebrospinal fluid. Next, we created a study specific grey matter template in two steps. In a first run we affine-registered the grey matter images to the ICBM-152 grey matter template and we averaged the resulting images to create a first-pass template. In a second run we non-

linearly registered the grey matter images to the first-pass template and we averaged these images to obtain the final template at 2x2x2 mm^3 resolution in standard space. Finally, we non-linearly registered the grey matter images to the final template and smoothed these images with a Gaussian kernel with sigma = 3 mm. The voxel values in these images range between 0 and 1, representing the percentage of a voxel being grey matter tissue. We averaged the voxel wise values within the 48 regions of the probabilistic Harvard-Oxford cortical atlas. We calculated the weighted averages of the regions, with voxels contributing to the average of a region based on their probability of being part of that region. This yielded 48 grey matter density values per subject.

Subcortical volumes

We calculated the volumes of the subcortical structures using the FMRIB's Integrated Registration and Segmentation Tool (FIRST) in FSL (Patenaude et al., 2011). First, the whole-head images were affine registered to the non-linear MNI-152 template. In a second stage, initialised by the result of the first stage, we used a subcortical mask to achieve a more accurate and robust affine registration. The shapes of the subcortical structures were then modelled by deformable meshes and the boundary voxels were classified as being part of the subcortical structure using structural segmentation (Zhang et al., 2001). Finally, we corrected the subcortical volumes for intracranial volume as obtained by FSL. This yielded 14 subcortical volume features per subject (thalamus, caudate, putamen, pallidum, hippocampus, amygdala, and accumbens for both hemispheres).

Hippocampal shape

To calculate hippocampal shape, we used the vertex analysis in FSL (Patenaude et al., 2011). The shape values represent the distance of a vertex on the hippocampal mesh of a specific subject to the mean location of that vertex within the whole sample. A negative value represents a decrease of the size of the hippocampus on that specific location for a subject relative to the mean sample. Vice versa, a positive value represents a relative increase of the size of the hippocampus on that location. In the absence of a sufficiently detailed brain atlas of the human hippocampus we used a data-driven method to reduce the number of features. We ran a principal component analysis on the vertex shape values of both hippocampi and extracted only the first ten components, because these alone explained 86 percent of the variance in hippocampal shape values. This yielded ten hippocampal shape features per subject.

Reference measures: whole brain atrophy and hippocampal volume

To provide a reference for the classification performance of the anatomical MRI measures, we also calculated two simple measures that are commonly used for clinical diagnosis of AD, whole brain atrophy and hippocampal volume (Frisoni et al., 2010). We used Freesurfer to calculate the ratio of total brain volume to intracranial volume as a measure of whole brain atrophy. We used FSL FIRST to calculate the volumes of the left and the right hippocampus.

Statistical analyses

The features of the six anatomical MRI measures were used in an elastic net logistic regression to classify the subjects as either AD or control. Elastic net regression uses penalties to hinder the features from entering the regression model (Zou & Hastie, 2005; Friedman et al., 2012). Thus, only the most relevant predictors will enter the regression model, which is helpful if the number of features outnumbers the number of subjects. Elastic net regression uses a combination of an L1 (LASSO) (Tibshirani, 1996) and L2 (Ridge) (Hoerl and Kennard, 1970) penalty. Therefore, two hyperparameters should be set: the α parameter determines the relative weight of the two different penalties and λ determines the size of those penalties. Elastic net logistic regression has been used for AD classification by Schouten et al. (2016), Trzepacz et al. (2014) and Teipel et al. (2015).

We used cross-validation to ensure that we are not overfitting the prediction models. In our case there are two potential sources of overfitting. We could either include too many predictors in our logistic regression model or we could overestimate the classification accuracy by looping over all the values of the hyperparameters and only pick the best result. To ascertain that we are not subject to any of these two sources of overfitting we used a nested cross-validation approach (Krstajic et al., 2014). We used the inner loop of the nested cross-validation to fit the logistic regression model and the outer loop to tune the hyperparameters. For both the inner and outer loop we used 10-fold cross-validation, thus using 90 percent of the subjects in the training set and 10 percent in the test set, and repeating this 10 times such that all subjects were part of the test set once.

We plotted receiver operating characteristic (ROC) curves and calculated the area under the curve (AUC) as a measure of classification accuracy. We repeated the cross-validation procedure 50 times to get a more reliable cross-validation error (Krstajic et al., 2014). We extracted the median AUC value instead of the mean, because we expect that the distribution of AUC values is skewed to the left due to a ceiling effect.

First, we calculated the classification accuracy for each measure separately. Then we calculated the classification accuracy for the combination of all measures and for all pairs of two measures. We used two different methods to combine the features of different measures. In the first method we concatenated the features and we used the concatenated feature set for classification similarly as the single measures. Concatenation is commonly used to combine different sets of features for prediction (De Magalhães Oliveira et al., 2010; Westman et al., 2012). It might however not be the most optimal method for combination, because all sets of features are then weighted equally whereas some sets might be more predictive than others. Therefore, in the second method we used a weighted average of the single measure predictions in order to increase classification performance (Wolpert, 1992; Breiman, 1994). Measures that achieved a higher accuracy were given a larger weight. We used the inverse of the binomial deviance as a measure of accuracy. Binomial deviance is a continuous measure of prediction error and is therefore sensitive for small accuracy differences. We determined the measures' weights for each subject individually. To avoid overfitting we used the binomial deviances within the training set to determine the contributions of the measures for the left out subjects. The weighted average was calculated for all the 50 cross-validation repetitions and the median result was presented.

2.3 Results

Figure 1 shows ROC curves for the classification of AD vs. cognitively normal controls for the six anatomical MRI measures separately, and for the combination of all six measures. The accompanying AUCs, representing a measure of classification accuracy, are shown in Table 1. Table 1 also presents the AUC values for the two reference measures: hippocampal volume and whole brain atrophy. The two reference measures perform reasonably well. Especially, the hippocampal volumes can discriminate well between AD patients and controls. Yet, the grey matter density of the cortical structures and the volumes of subcortical structures discriminate better than the reference measures. Cortical thickness, cortical area, cortical curvature, and hippocampal shape cannot improve over the reference measures. Most importantly however, the combination of all six measures outperforms the separate measures. The weighted combination of all measures discriminates somewhat better than the concatenated combination

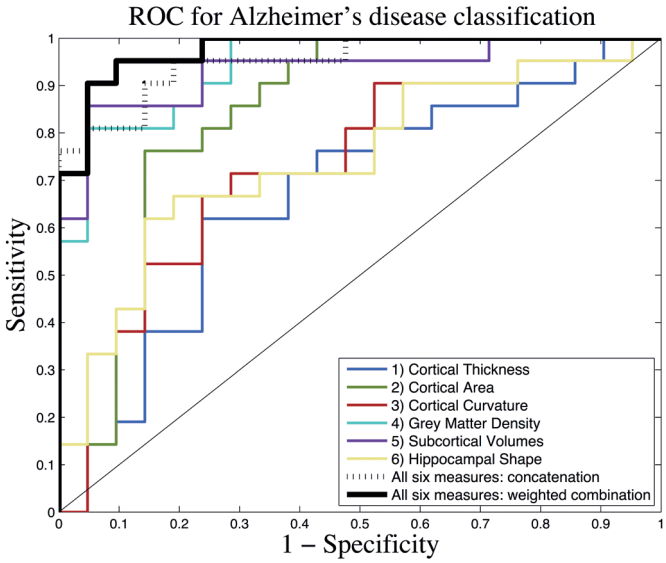

Figure 1. ROC curves for discriminating AD patients from cognitively healthy controls. ROC curves are plotted for the six anatomical measures separately and for the two types of combinations of all six measures.

Table 1. AUC values discriminating AD patients from cognitively healthy controls for the different anatomical MRI measures. Hippocampal volumes and whole brain atrophy are added for comparison, because these are fairly simple measures that are commonly used by clinicians to diagnose AD. The other six measures are used separately and combined to discriminate the two groups. We used two methods for the combination. Both of them outperform the separate measures, and the weighted combination works best.

Anatomical MRI measures	AUC
Reference: hippocampal volumes	0.87
Reference: whole brain atrophy	0.77
1) Cortical thickness	0.67
2) Cortical area	0.85
3) Cortical curvature	0.73
4) Grey matter density	0.94
5) Subcortical volumes	0.93
6) Hippocampal shape	0.74
All six measures: concatenation	0.95
All six measures: weighted combination	0.98

To further investigate the additive value of combining different measures, we calculated the AUCs for all pairs of two measures, using both measure concatenation (Figure 2, left side) and weighed combinations (Figure 2, right side). The tall bars represent the single measure AUCs, like in Table 1. The short bars represent the additive value of a second measure. Note that the additive value can also be negative, when the addition of a second measure worsens the classification accuracy. Both for concatenation and weighted combinations, the AUC of a single measure often improves when a second measure is added. When using concatenation, the highest AUC values are obtained by combining grey matter density with either subcortical volumes or cortical thickness (AUC = 0.98), which is even higher than the concatenation of all six modalities. Using weighted combinations, the highest AUC is obtained by combining grey matter density with subcortical volumes (AUC = 0.98), which is equal to the weighted combination of all six modalities.

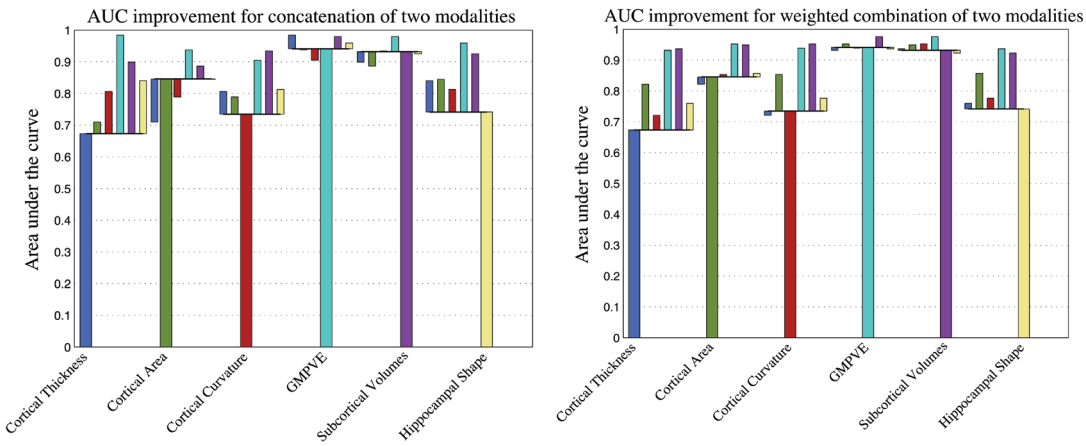

Figure 2. AUC values for all the possible combinations of two measures using feature concatenation (left) and weighted combinations (right). The tall bars represent the single measure AUCs, which are the same as those in Table 1. The short bars represent the additive value of a second measure. The additive values are mostly positive. For example, when cortical area is concatenated with cortical thickness, the AUC increases from 0.67 to 0.71. However, sometimes the additive value a second measure is negative. For example, cortical area is on its self a better predictor (0.85) than concatenated with cortical thickness (0.71).

These statistical models are primarily meant to predict class membership, and not to make claims about which features were most important to distinguish AD patients from controls. If a statistical model is built for prediction, rather than for explanation, one should be careful when interpreting the explanatory part of that model (Shmueli, 2010). We elaborate more on this in the discussion section. Yet, to illustrate the content of the classification models we show the standardised beta values (averaged over the 50 cross-validation repetitions) for the predictors of the grey matter density measure (Figure 3) and the subcortical volumes measure (Figure 4). We used these two measures for illustration, because they discriminate best between AD patients and cognitively normal controls. For regions with a negative beta value, low grey matter density values or low subcortical volumes increase the odds for AD. For regions with a positive beta value, high grey matter density values or high subcortical volumes increase the odds for AD. This might seem contradictionary, because we do not expect to see increased grey matter density values or increased subcortical volumes in AD patients. The interpretation of these effects could be something like: if that region is relatively unaffected (high grey matter density, or large subcortical volume) while some other regions are more affected, this is evidence in favour of AD. We plotted the regions in colour coding for the grey matter density measure (Figure 5) and the subcortical volumes measure (Figure 6).

The grey matter density classification model was mostly driven by decreased grey matter density within the cortical areas in the medial temporal lobes, and to a lesser extent in the occipital and frontal lobes. The regions with large weights include the orbitofrontal cortex, the subcallosal cortex, the insular cortex, and the inferior temporal gyrus. Most regions contribute to the classification model to a certain extent, suggesting that a global pattern of atrophy is predictive for AD. The subcortical volumes classification model was mostly driven by decreased sizes of the hippocampus, putamen and thalamus, and to a lesser extent by decreased sizes of the accumbens and the amygdala.

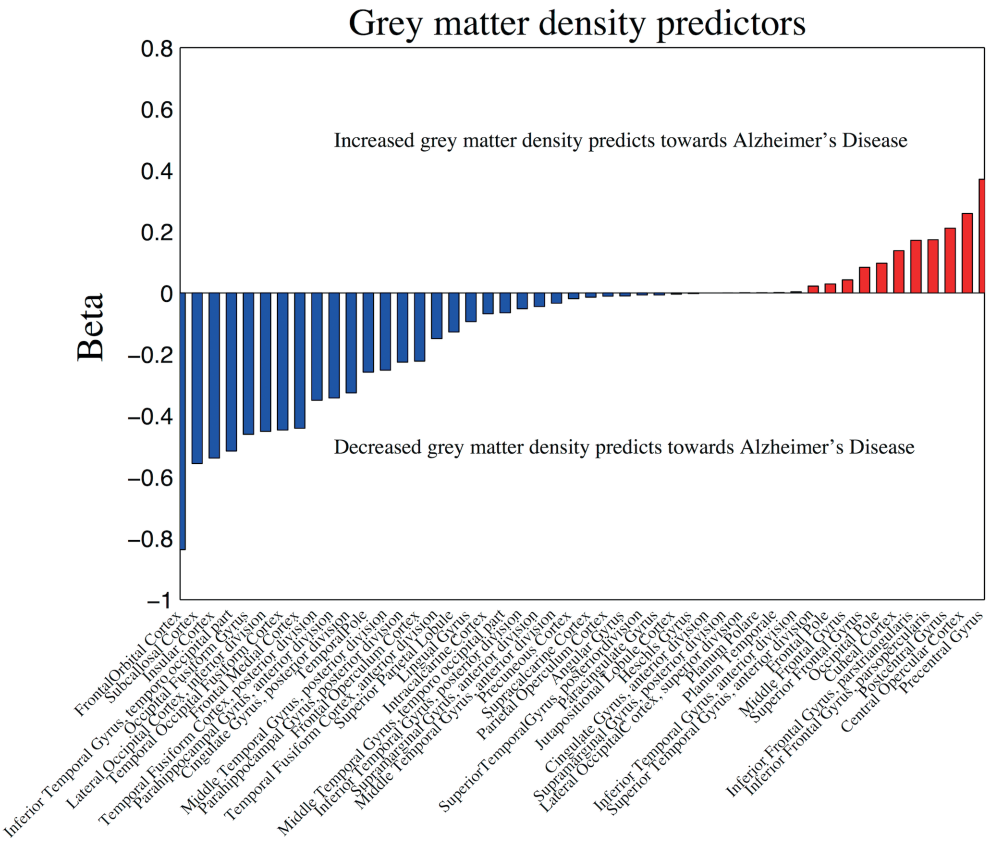

Figure 3. Beta values for the cortical grey matter density estimates. The beta values represent the mean beta values over all the cross-validation folds of all the cross-validation repetitions. The beta values are ordered according to size. Negative beta values are coloured blue, and positive beta values are coloured red. Most beta values are negative, and the meaning of those is that low grey matter density predicts toward AD. Some beta values are positive, which is counterintuitive, but could mean something like: if that region is relatively unaffected (high grey matter density) while some other regions are more affected, this is evidence in favour of AD. Note that grey matter density was only calculated for the cortical regions. The results for the subcortical volumes are presented in Figure 4.

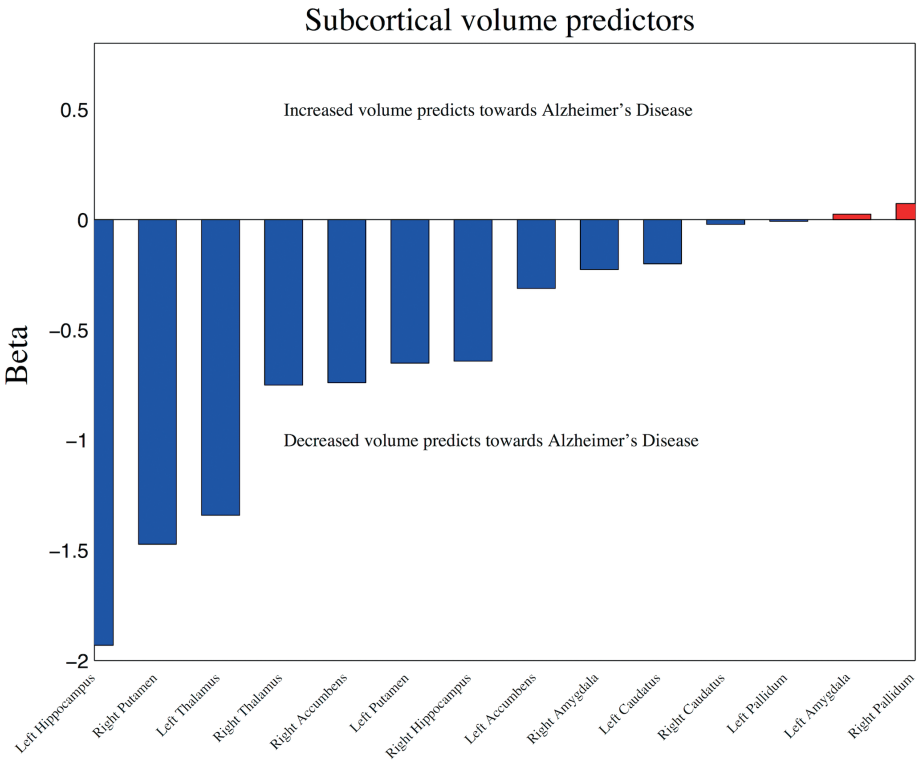

Figure 4. Beta values for the subcortical volumes. The beta values represent the mean beta values over all the cross-validation folds of all the cross-validation repetitions. The beta values are ordered according to size. Negative beta values are coloured blue, and positive beta values are coloured red. Most beta values are negative, and the meaning of those is that a small volume predicts toward AD. Some beta values are positive, which is counterintuitive, but could mean something like: if that region is relatively unaffected (large volume) while some other regions are more affected, this is evidence in favour of AD.

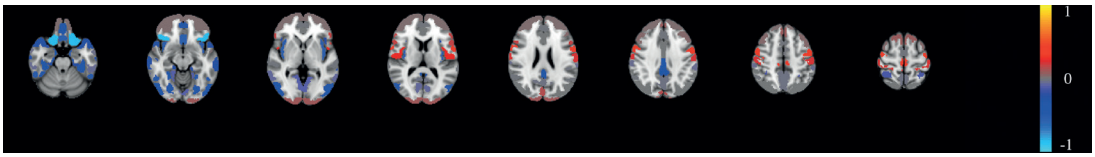

Figure 5. The beta values from Figure 4 are presented here in colour coding. The 'cool' regions correspond with the blue bars and the meaning of those is that low grey matter density predicts toward AD. The 'hot' regions correspond with the red bars and the meaning of those is that high grey matter density predicts toward AD. This is counterintuitive, but could mean something like: if that region is relatively unaffected (high grey matter density) while some other regions are more affected, this is evidence in favour of AD.

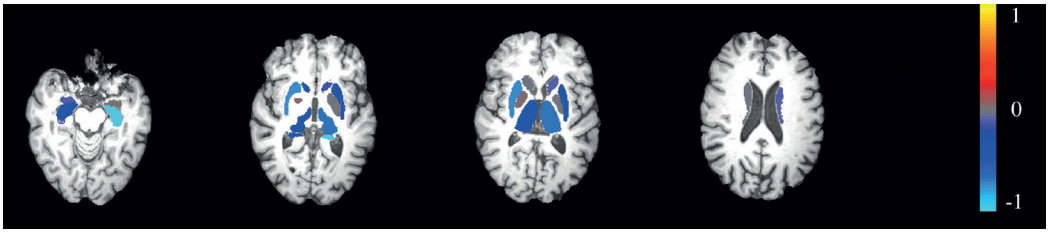

Figure 6. The beta values from Figure 5 are presented here in colour coding. The 'cool' regions correspond with the blue bars and the meaning of those is that a small volume predicts toward AD. The 'hot' regions correspond with the red bars and the meaning of those is that a large volume predicts toward AD. This is counterintuitive, but could mean something like: if that region is relatively unaffected (large volume) while some other regions are more affected, this is evidence in favour of AD.

2.4 Discussion

In this study we used different anatomical MRI measures to separate AD patients from controls. Our main finding is that the combination of the different anatomical MRI measures improves AD classification accuracy over each single measure. The combination of different anatomical MRI measures thus captures more information on grey matter loss than each of the measures separately, and AD classification benefits from this extra information through an automated classification algorithm.

When used separately, all measures were sensitive to the detection of AD. We replicated early findings in which grey matter density values (Cuingnet et al., 2010) and the volumes of the subcortical structures (De Magalhães Oliveira et al., 2010) are highly discriminative for AD. However, the high classification accuracy of cortical surface area compared to cortical thickness is relatively surprising. Previous studies found that cortical grey matter atrophy is primarily reflected in cortical thinning rather than in a decrease of cortical area (Dickerson et al., 2009; Westman et al., 2012).

We have used two different methods to combine the measures. The weighted combination of the measures showed higher accuracy than the concatenation of the measures. Due to a relatively small sample size, it is unclear whether this difference will generalise to other data sets, but it does demonstrate that the method that is used for the combination of the information can influence diagnostic accuracy.

The classification model for the grey matter density estimates (which were only calculated for the cortical regions) was mostly driven by decreased grey matter density within the medial temporal lobes, and to a lesser extent in the occipital and frontal lobes. These findings are in line with other observations of AD atrophy (Karas et al., 2004; Frisoni et al., 2010; Risacher et al., 2010). The subcortical volumes classification model was mostly driven by decreased sizes of the hippocampus, putamen and thalamus, which is also in accordance to previous findings (Leung et al., 2010; de Jong et al., 2008).

It should be noted that we have built predictive models, without the explicit goal to make explanatory models. We should therefore be careful when interpreting the explanatory part of our models (Shmueli, 2010). For example, we have forced sparsity in our models, which hinders potentially relevant features from entering the model. Furthermore, the effect of a feature is conditional on the effect of all the other features in

the model, and can therefore not be interpreted independently. Also, multicollinearity is not an issue for prediction models, but it can be problematic for the explanatory part of a model. These constraints should thus be taken into consideration when interpreting the content of the prediction models. For these reasons the effects do not denote one to one relationships between the feature and the probability of being classified as an AD patient, nor do they reflect mean differences between the groups.

We did not use non-linear effects or interaction-effects in our classification algorithm, which improves the reproducibility of our results. More complex classification models might further enhance the classification accuracy. For example, longitudinal AD studies have found that atrophy in some areas in the brain follows a non-linear trend (Chan et al., 2003; Fotenos et al., 2005). On the contrary, more complex classification algorithms can cause overfitting, resulting in a poorer classification performance (Hastie et al., 2009). Neither did we use prior feature selection, because feature selection generally does not improve AD classification results (Chu et al., 2012) and avoiding this extra step eases the reproducibility of our results.

We have used the most up to date versions of Freesurfer and FSL to calculate the structural measures. Freesurfer and FSL have both been validated for the analysis of anatomical MRI scans for both healthy subjects and AD patients. Freesurfer is sensitive to detect cortical thinning in AD patients compared to controls (Redolfi et al., 2015). and it has good scan-rescan reproducibility (Tustison et al., 2014). The results of Freesurfer analyses do however differ between different Freesurfer versions (Groenschild et al., 2012), but the most recent Freesurfer versions correspond better with the results of manual outlining (Clerx et al., 2015). The shape model that is used by FSL First to segment the subcortical structures has been trained on both healthy subjects and pathological brains (including AD patients) to reduce bias towards healthy brains (Patenaude et al., 2011). Furthermore, FSL First calculation of the subcortical volumes is roughly similar to the results of manual outlining for AD patients, MCI patients and controls from the ADNI data set (Mulder et al., 2014). FSL voxel-based morphometry (VBM) is relatively accurate because it makes use of a non-linear registration tool (Callaert et al., 2014). Furthermore, FSL VBM accurately detects hippocampal atrophy, but is biased towards detecting medial temporal lobe atrophy in AD patients (Diaz-de Grenu et al., 2014).

In conclusion, we demonstrated that the combination of i) cortical thickness, ii) cortical area, iii) cortical curvature, iv) grey matter density, v) subcortical volumes and vi) hippocampal shape improves AD diagnosis. The added value of combining different anatomical MRI measures should be considered in AD scanning protocols. It is still common practice to only use the size of the hippocampus or a single measure of whole brain atrophy for AD diagnosis. Our results demonstrate that clinical AD diagnosis could benefit from calculating multiple measures from an anatomical MRI scan and incorporate these all in an automated analysis. Our results further suggest that the grey matter density of the cortical structures and the volumes of the subcortical structures are sufficient for optimal AD classification based on an anatomical MRI scan. These results might also be relevant to studies of early AD diagnosis and other neurodegenerative diseases studies.

2.5 Acknowledgments

This study is supported by grant no. 016130677 of the Netherlands Organisation for Scientific Research (NWO). A. Hafkemeijer is sponsored by NWO grant no. 05613010. E. G. Dopper is sponsored by grants from the Dioraphte Foundation grant no. 09-02-03-00, the association for Frontotemporal Dementia research and NWO grant no. 05613018.