Improving prediction of Alzheimer's disease using patterns of cortical thinning and homogenizing images according to disease stage

Simon Eskildsen, Pierrick Coupé, Daniel García-Lorenzo, Vladimir Fonov, Jens Pruessner, Louis Collins

Improving prediction of Alzheimer's disease using patterns of cortical thinning and homogenizing images according to disease stage

Simon F. Eskildsen[1,2], Pierrick Coupé[2,3], Daniel García-Lorenzo[2,4], Vladimir Fonov[2], Jens C. Pruessner[5], D. Louis Collins[2] and the Alzheimer's Disease Neuroimaging Initiative

[1]Center of Functionally Integrative Neuroscience, Aarhus University, Aarhus, Denmark. [2]McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada. [3]Laboratoire Bordelais de Recherche en Informatique, Unité Mixte de Recherche CNRS (UMR 5800), Bordeaux, France. [4]ICM, UPMC/INSERM UMR_975 CNRS UMR 7225, Hôpital de la Salpêtrière, Paris, France. [5]Departments of Psychiatry, Neurology and Neurosurgery, McGill University, Montreal, Canada

Abstract. Predicting Alzheimer's disease (AD) in individuals with some symptoms of cognitive decline may have great influence on treatment choice and guide subject selection in trials on disease modifying drugs. Structural MRI has the potential of revealing early signs of neurodegeneration in the human brain and may thus aid in predicting and diagnosing AD. Surface-based cortical thickness measurements from T1-weighted MRI have demonstrated high sensitivity to cortical gray matter changes. In this study, we investigated the possibility of using patterns of cortical thickness measurements for predicting AD in subjects with mild cognitive impairment (MCI). Specific patterns of atrophy were identified at four time periods before diagnosis of probable AD and features were selected as regions of interest within these patterns. The selected regions were used for cortical thickness measurements and applied in a classifier for testing the ability to predict AD at the four stages. The accuracy of the prediction improved as the time to conversion from MCI to AD decreased, from 70% at 3 years before the clinical criteria for AD was met, to 76% at 6 months before AD. These results show that prediction accuracies of conversion from MCI to AD can be improved by learning the atrophy patterns that are specific to the different stages of disease progression. This has the potential to guide the further development of imaging biomarkers in AD.

1 Introduction

The ability to diagnose and predict Alzheimer's disease (AD) at an early or even preclinical stage may have great impact on the possibility for improving treatment choices for AD patients. This may lead to reduced costs associated with long-term care. In addition, accurate prediction may also reduce costs associated with selecting subjects for pharmaceutical trials when performing large scale tests on disease modifying drugs, since false positives can be excluded in the initial stage.

AD is characterized by accumulation of amyloid-β (Aβ) and hyperphosphorylated tau in the brain, eventually leading to neurodegeneration. To support an early diagnosis of AD, various biomarkers are currently being investigated. Even though the accumulation of Aβ can be detected in the cerebrospinal fluid (CSF), or by using positron emission tomography (PET) years before structural changes can be detected, structural imaging markers based on magnetic resonance imaging (MRI) are considered more sensitive to change after the first symptoms appear [1]. Signs of atrophy in the medial temporal lobes may aid in differentiating AD from other pathologies as MRI examinations often are part of the clinical assessment standard of care in patients with mild cognitive impairment (MCI). While studies investigating the usefulness of medial temporal lobe atrophy in the diagnosis of AD are ongoing [2-4], the assessment of patterns of cortical thinning across the cerebrum may aid to increase the specificity of the diagnosis for the disease [5, 6].

The high tissue contrast offered by T1-weighted (T1w) MRI enables accurate structural neuroimaging analysis, which may be used as a possible surrogate biomarker for diagnosing and predicting AD [7]. Measurements of cortical thickness based on MRI are highly sensitive to small structural changes across the cortex. However, results from previous studies seem to have suggested that cortical thickness measurements do not perform better than other techniques when trying to predict AD in subjects with MCI, yielding accuracies from 56% to 68% depending on the technique [3, 4]. Cortical thickness is usually measured at a very high resolution (tens of thousands of points on the cerebral cortex). Using such high numbers of measurements in prediction may lead to over-fitting in a discriminatory model. The dimensionality can be reduced by defining regions of interests (ROI) in which measurements are averaged. This reduces the inherent noise of high resolution data and reduces the risk of over-fitting. Usually, such ROIs are predefined from a structural or functional perspective. However, the pattern of neurodegeneration may not follow standard definitions for anatomical or functional regions; thus, such ROIs may lead to loss of discriminative information. Therefore, data driven approaches to select discriminative cortical thickness ROIs, independent of any predefined parcellation, may lead to better prediction results.

Another factor preventing high predictive power is the heterogeneity of images used when training a classifier. Usually, a classifier is trained with images from individuals who convert to AD at some future time point and images from individuals who do not convert to AD in the follow-up period. Inherently, there is variability in the degree of neurodegeneration simply due to the variability of when the converters actually convert. This variability affects how well the coefficients of the classifier can be fitted. By homogenizing the images with respect to "time to conversion" and thereby the pattern of atrophy, the performance of the classifier may be improved.

In this study, we investigated the possibilities of improving the prediction accuracy by *i)* automatically selecting the most discriminative ROIs, and *ii)* homogenizing the training data by time to conversion.

2 Methods

2.1 Subjects and acquisition

All data used in the preparation of this article were obtained from the ADNI database (www.loni.ucla.edu/ADNI). The ADNI database contains 1.5T and 3.0T T1w MRI scans for AD, MCI, and cognitively normal controls (CN) at several time points. The aim was to follow and collect scans from MCI patients at baseline, 6 months, 12 months, 18 months, 24 months, 36 months and 48 months. The number of participating subjects was reduced over the course of the study due to drop outs. At each time point a clinical diagnosis was made to identify MCI subjects who converted to probable AD according to the NINCDS-ADRDA Alzheimer's Criteria [8].

Several studies have used ADNI data to predict which MCI patients would convert to probable AD using a single MRI scan [3, 4, 9-15]. All of these studies have used baseline data for the analysis, which rendered the group of progressive MCI (pMCI) heterogeneous with respect to "time to conversion", since the pMCI patients would convert anytime over the course of 6 months to 4 years followup. Such heterogeneity may conceal the specific neurodegenerative processes that may be attributed to the different sub-stages of disease progression. For example, the pattern of atrophy may be different in patients one year before diagnosis compared to the pattern two years earlier. In this study, we therefore utilized the full ADNI database and selected scans at various intervals prior to diagnosis. We selected pMCI scans six months, 12 months, 24 months and 36 months prior to AD diagnosis and grouped these into time-homogeneous groups of pMCI. A total of 163 different pMCI subjects were included and the majority of these were represented in more than one pMCI subgroup. To identify characteristic traits for disease progression in the pMCI groups and determine if these could be used as markers for prediction, we compared with the group of stable MCI (sMCI) patients who did not have a change of diagnosis over the course of the ADNI study. Our sMCI group only included those MCI patients who were followed for at least three years. Thus for the sMCI group, we selected scans at baseline. Table 1 lists the selected MCI groups after removing subjects due to scanner acquisition artefacts or image processing problems. No differences in age were found between the groups, while all pMCI groups had significantly (t-test, $p<0.05$) smaller MMSE scores than the sMCI group. All scans used in the study were acquired at 1.5T MRI scanner, with T1w imaging modality.

2.2 Image processing

Images were denoised [16] using an estimated standard deviation of noise [17], bias field corrected [18], registered to MNI space [19] and skull stripped [20]. Cortical thickness was calculated using FACE (fast accurate cortex extraction) [21] and mapped to the cortical surface of a population-specific average non-linear anatomical template [22]. Cortical segmentations were manually checked for errors by an expert and subjects were excluded if errors were found in one of the image processing steps mentioned above. The quality control removed only 2.7% of the scans processed dis-

tributed on six different subjects, resulting in the group sizes listed in Table 1. In general, the scans were excluded due to image artefacts (n=2), insufficient stereotaxic registration (n=2, same subject) and insufficient skull stripping (n=9, three different subjects).

Table 1. Demographics and proportion of subjects used in one subgroup only (unique rate). Significant differences (t-test, $p<0.05$) for the respective progressive MCI (pMCI) groups compared to the stable MCI (sMCI) group are in bold.

Group	N (females)	Age±sd	MMSE±sd	Unique rate
sMCI	134 (40)	75.0±7.5	27.6±1.8	100%
pMCI6	122 (47)	75.3±7.6	**25.3±2.6**	20%
pMCI12	**128 (54)**	75.6±7.1	**26.0±2.3**	5%
pMCI24	61 (24)	74.4±7.2	**26.7±1.8**	2%
pMCI36	29 (13)	75.8±6.4	**26.9±1.7**	10%

2.3 Validation strategy

In the experiments below, we use a leave-one-out (LOO) validation strategy where for each comparison (e.g., pMCI12 vs sMCI), all the subjects (the ensemble of all pMCI12 and sMCI) except one are used to select features and generate a classification model and the one subject left out is then used for testing. This procedure is repeated for every subject in the two groups compared, thus validating the method with every subject. Since the test subject is not used in the selection of features, nor in building the classifier, we avoid any bias or "double dipping" in our efforts of predicting converters. It is important to note that we obtain a unique set of features and classifier for each LOO test; a total of 876 feature sets.

2.4 Feature generation and selection

To explore the patterns of atrophy at the different stages of the progression from MCI to AD, we devised a data-driven feature selection method for classification. Using the training sets (i.e., the groups to be compared, less the subject to be tested), statistical parametric maps of differences in cortical thickness between the sMCI group and the pMCI groups were constructed by one-sided t-tests per vertex of the template surface (from a total of 162,582 vertices). Our aim was to generate a compact set of features with high discriminating power. Thus, candidate features were restricted to a proportion, φ, of the cortical area with the largest magnitude t-values. Within this thresholded t-map (Fig. 1a), local maxima were detected and used as seed points for a constrained region growing algorithm. For each maxima, region growing were performed downhill only in a circular fashion constraining the area to a maximum of 300 mm^2 and a maximum accumulated t-value of δ, where δ is defined as

$$\delta = \frac{\sum_{t \in \varphi} t}{c} \tag{1}$$

where c is the number of seed points. Only regions that reached an accumulated t-value of δ were kept. These regions constituted the candidate ROIs (Fig. 1b). The proportion, φ, of the cortical surface is used to guide the region growing algorithm and limiting the number of seed points. However, restricting the features to a too small proportion of the surface leads to very small patches of cortical thickness which are more affected by noise. In our experiments we found a proportion in the range of 10% - 15% to be a suitable trade-off.

Cortical features were determined as the mean cortical thickness for each ROI. To remove redundant features and to keep discriminant features, we used the minimal-redundancy-maximal-relevance (mRMR) criterion [23] as feature selection. We used the mutual information difference metric and discretized the features to three states (-1,0,1) by thresholding at μ±σ. This was done to ease the mutual information computation as described in [23]. Based on experiments, we chose to keep the 10 best features according to the mRMR criterion (Fig. 1c).

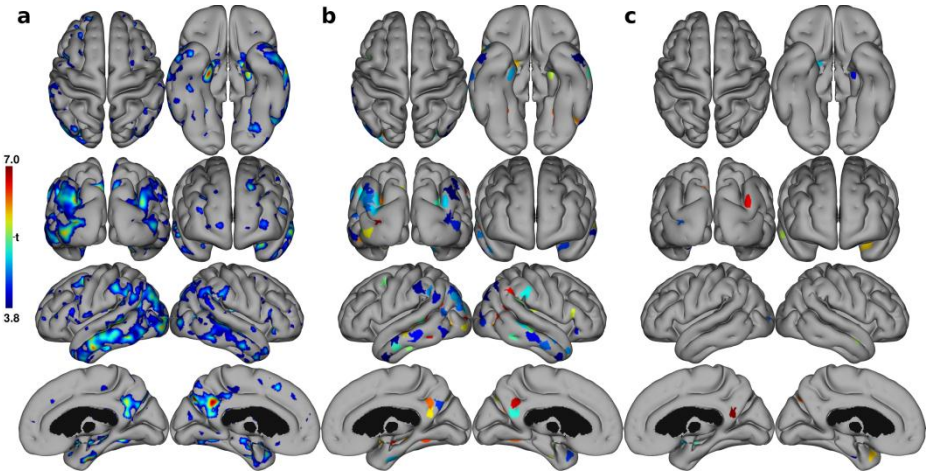

Fig. 1. A single instance of the feature generation and selection process for the pMCI12 classification showing a) thresholded t-map, b) ROIs after constrained region growing from seed points, and c) ROIs left after feature selection using mRMR criterion.

2.5 Classification

Linear discriminant analysis (LDA) was used for the classification. As described above, for each subject to classify a separate classifier was trained with the remaining subjects using the subject-specific feature set. The correct classification rate, the sensitivity, the specificity, and the area under the receiver operating characteristic curve (AUC) were calculated from the results. Furthermore, McNemar's chi-square test was used to assess whether the classification performed better than a random classifier. This was done by performing a random classification and calculating the p-value of

McNemar's test with significance level 0.05 that our classification was better than the random classification. This validation process was repeated 1000 times for each of the 876 LOO experiments and the median p-value was reported.

3 Results

Medial temporal lobe structures were automatically selected as relevant features in all the studied stages. However, the hippocampus, which is considered to be affected in the incipient stages of AD, was rarely chosen in the pMCI36 classification problem. Fig. 2 shows the selected ROIs color coded by the frequency they were selected. ROIs for short-term prediction (<12 months) were selected more consistent than for long-term prediction (>12 months).

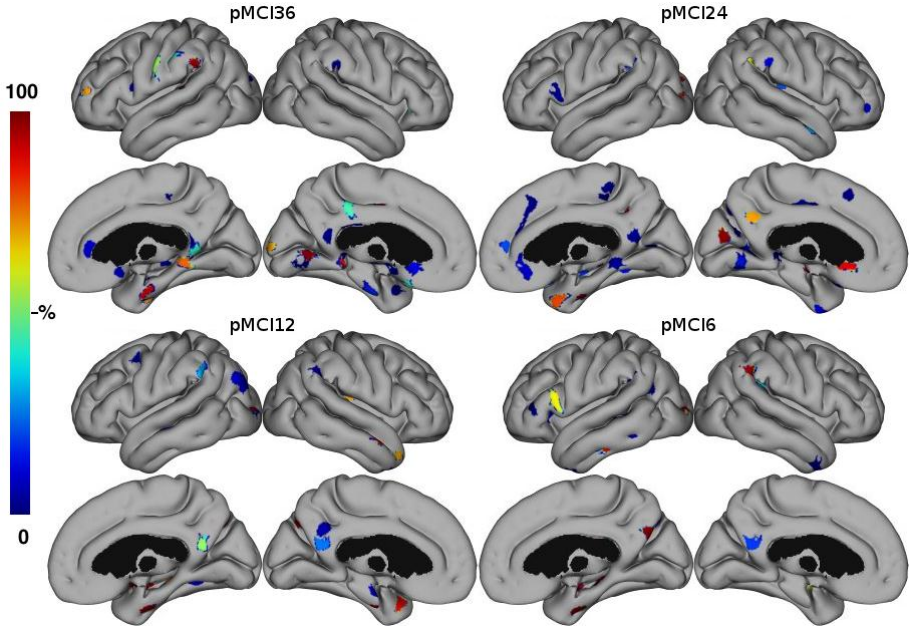

Fig. 2. ROIs color coded by frequency of selection for each classification problem.

Classifier performances for the studied classification problems are shown in Table 2. For comparison, we performed similar classifications for the combined group of all pMCI using the baseline scans. Since age is an important associated risk factor in AD, Table 2 also lists classification results after including age in the LDA. Using the age, the AUC is improved in all classification problems, while the accuracy is improved only for the long term prediction. Also, by including age, all classifiers performed significantly better than a random classifier. Receiver operating characteristic (ROC) curves for the classifications are shown in Fig. 3.

Table 2. Classification results using feature sets based on cortical thickness features only and including age in the classifier. Acc=Accuracy, Sen=Sensitivity, Spe=Specificity, AUC=Area under the ROC curve, MN=McNemar's test.

	Cortical thickness only					Cortical thickness and age				
Classifica-tion	Acc (%)	Sen (%)	Spe (%)	AUC (%)	MN (p-val)	Acc (%)	Sen (%)	Spe (%)	AUC (%)	MN (p-val)
pMCI36	69.9	55.2	73.1	63.5	0.198	72.4	48.3	77.6	63.7	0.001
pMCI24	66.7	59.0	70.2	67.3	0.062	67.2	55.7	72.4	70.7	0.001
pMCI12	72.9	75.8	70.2	76.2	0.001	70.6	72.7	68.7	76.3	0.001
pMCI6	75.8	75.4	76.1	80.9	0.001	74.6	72.1	76.9	81.1	0.001
pMCIall	69.0	68.7	69.4	71.2	0.001	69.4	70.1	68.7	73.7	0.001

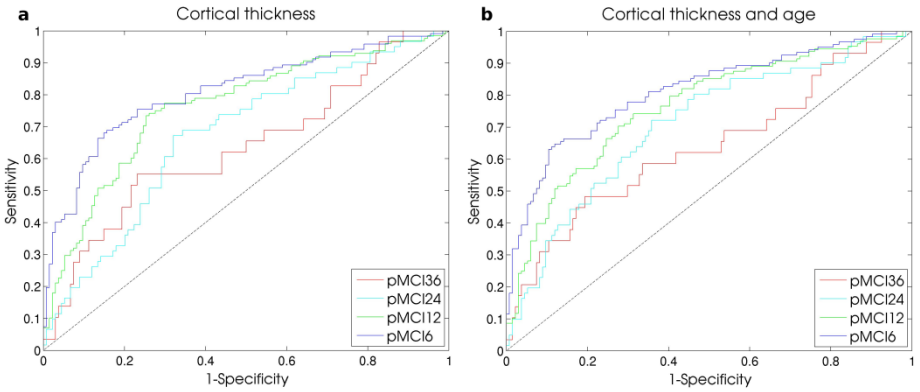

Fig. 3. ROC curves for the four classification problems using a) cortical thickness only and b) cortical thickness and age.

4 Discussion

Using cortical thickness features, we obtained prediction accuracies in the range of 67% - 76% (Table 2). Surprisingly, the accuracy for the pMCI36 prediction is nominally higher than the pMCI24 prediction. However, the sensitivity and the AUC are lower for the pMCI36 prediction. The higher specificity at 36 months compared to 24 and 12 months prior to conversion suggests that additional confounding factors appear later in the process at 24 months. It is unclear at this point what those factors could be, perhaps synaptic and network changes in the CNS in response to the pathological processes associated with the prodromal period of AD. The inflated accuracy of the pMCI36 prediction may also be caused by the low number (n=29) of pMCI individuals in this group combined with a relatively high specificity. Better balanced groups might yield a more confident estimate of the prediction accuracy for MCI subjects three years before AD diagnosis. Whatever the case, clearly, future research is needed in this specific area.

As expected, the sensitivity and the AUC increase as the time to AD diagnosis is reduced. According to the McNemar's test, only the predictions of pMCI less than 12 months prior to diagnosis (short-term prediction) are significantly better than a random classifier when not using the age. However, when adding the age information all predictions become significantly better than random even though the sensitivity is slightly reduced.

Compared to previously published studies, which use cortical thickness for predicting pMCI in the ADNI cohort [3, 4, 9, 14], our prediction accuracies are overlapping. We obtain higher accuracies for short-term prediction (<12 months) than previous studies. In contrast, our results for long-term prediction (>12 months) are similar to the results from previously published studies. Our efforts of homogenising data by "time to conversion" did not have the expected effect on the long-term prediction. It seems that the increased predictive power at the short-term may be attributed to the progression of the disease, which yields a more consistent pattern of cortical neurodegeneration. Another perhaps more important factor is the relatively small samples available to evaluate the long-term predictions. The statistical maps, which are the basis for the feature generation, are more affected by noise in the case of small sample sizes. With the continuation of ADNI, the number of long-term converters will increase and more consistent feature patterns can be generated for these groups.

All our prediction accuracies are similar or better than previously published results, except for the results in [13], where only 27 pMCI subjects were tested. These results highlight the competitive prediction accuracy obtained by the proposed method.

Compared to the conventional grouping of pMCI and sMCI as published in Wolz et al. [4], only the pMCI24 prediction performed slightly worse (66.7% vs. 67.3%), however, this difference is not significant. This is an indication of the improved accuracy by homogenizing data. In addition, our prediction accuracy when combining all pMCI subjects is better than the accuracy obtained by Wolz et al.

With these new methods, superior prediction accuracies can be obtained. The specificity was higher than the sensitivity for the long-term prediction. From an economic perspective a high specificity is very important in clinical trials when recruiting subjects. Reducing the number of false positives in trials may save time and reduce the associated costs. On the other hand, the sensitivity for the long-term prediction lags behind. Here, clearly more research needs to be done to improve the technique, and to be able to benefit from eventual neuroprotective therapies.

4.1 Regions selected for prediction model

The ROIs most often chosen by the feature selection for prediction indicate which anatomical regions are involved at different times prior to the progression from MCI to AD (Fig. 2). The shorter time to conversion the more consistent ROI selection is seen. This may reflect a better defined structural pattern of degeneration at the later stages of the disease or simply an effect of the larger sample sizes yielding a more consistent pattern in the LOO experiment.

Several regions seem to be consistently chosen in all classification problems. As shown in Fig. 2, medial temporal lobe structures are selected in all prediction prob-

lems. Hippocampus is frequently included in all predictions except the pMCI36, while the right parahippocampal gyrus is used in all cases. The para-hippocampal gyrus has previously been found to be highly discriminative for separating AD patients from healthy controls using cortical thickness and multivariate analysis, however, in the absence of hippocampal segmentation [24]. In addition to medial temporal lobe structures, ROIs were mostly selected from the cingulate gyrus. The cingulate gyrus is well known to be affected in early AD.

4.2 Image processing

In this study, we used a combination of highly consistent and robust image processing methods to measure the cortical thickness. Previous studies using cortical thickness have suffered from high exclusion rates due weak links in the image processing pipeline [3, 4, 25]. In our experience, the main factor for excluding scans due to processing is the skull stripping step. By using BEaST, a recent robust brain extraction algorithm [20], we were able to effectively reduce the exclusion rate. In our study, we excluded only 2.7% of all scans, which, to the best of our knowledge, is the lowest exclusion rate of any published cortical thickness analysis on ADNI data. The low exclusion rate enabled us to construct relatively large samples of pMCI subjects homogenous with respect to time to conversion. Furthermore, our strategy of denoising the images before cortical surface extraction has been shown to provide more accurate results than processing unfiltered images [26].

4.3 Feature generation

The feature generation approach used in our study is similar in spirit to the approach suggested by Fan et al. [27], who used a watershed algorithm on a VBM based statistical parametric map to select ROIs. Using a watershed algorithm on the statistical maps generated by cortical thickness would yield less compact ROIs, as the regions with high t-values often are elongated following a sulcus or gyrus across several anatomical regions. In the proposed method, we sought more focal features and therefore applied the constrained region growing as described above. We acknowledge that our method for feature generation is just one among many possibilities using the statistical map. Nevertheless, we expect the tendency of increased sensitivity by homogenizing the data by time to conversion will remain the same irrespective of the feature generation method applied.

5 Conclusion

Using patterns of characteristic cortical thinning in disease stages of progressive MCI compared to MCI patients who remained stable for three years demonstrated promising results for the prediction of patients with prodromal AD progressing to probable AD. We obtained a more accurate and unbiased estimate of the predictive power of cortical thickness measurements than published to date. The prediction ac-

curacies obtained by subgrouping progressive MCI patients with respect to “time to conversion” were better than previously published results on the same cohort. The improved accuracies are likely caused by this homogenization and improvements in the image processing pipeline in terms of robustness and accuracy.

The accuracy for the short-term prediction (<12 months) was relatively high (73%-76%), while the sensitivity for long-term prediction (>12 months) was relatively low (55%-59%). To be clinically applicable the sensitivity for the long-term prediction needs to be improved to be able to benefit from eventual neuroprotective therapies. The relatively high specificity for the long-term prediction holds promises of reduced costs associated with recruiting subjects for clinical trials; a reduction in false positives may save both time and money.

Acknowledgements

The research was supported by the MINDLab UNIK initiative at Aarhus University, funded by the Danish Ministry of Science, Technology and Innovation, grant agreement number 09-065250. We acknowledge funding from Canadian Institutes of Health Research (MOP-111169) and les Fonds de Recherche Santé Québec.

References

1. Frisoni, G.B., Fox, N.C., Jack, C.R., Jr., Scheltens, P., Thompson, P.M.: The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 6, 67-77 (2010)
2. Coupe, P., Eskildsen, S.F., Manjon, J.V., Fonov, V.S., Collins, D.L.: Simultaneous segmentation and grading of anatomical structures for patient's classification: application to Alzheimer's disease. NeuroImage 59, 3736-3747 (2012)
3. Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehericy, S., Habert, M.O., Chupin, M., Benali, H., Colliot, O.: Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. NeuroImage 56, 766-781 (2011)
4. Wolz, R., Julkunen, V., Koikkalainen, J., Niskanen, E., Zhang, D.P., Rueckert, D., Soininen, H., Lotjonen, J.: Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. PLoS One 6, e25446 (2011)
5. Apostolova, L.G., Thompson, P.M.: Mapping progressive brain structural changes in early Alzheimer's disease and mild cognitive impairment. Neuropsychologia 46, 1597-1612 (2008)
6. Vemuri, P., Whitwell, J.L., Kantarci, K., Josephs, K.A., Parisi, J.E., Shiung, M.S., Knopman, D.S., Boeve, B.F., Petersen, R.C., Dickson, D.W., Jack, C.R., Jr.: Antemortem MRI based STructural Abnormality iNDex (STAND)-scores correlate with postmortem Braak neurofibrillary tangle stage. NeuroImage 42, 559-567 (2008)
7. Ritchie, K., Lovestone, S.: The dementias. Lancet 360, 1759-1766 (2002)
8. McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M.: Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the

auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34, 939-944 (1984)
9. Cho, Y., Seong, J.K., Jeong, Y., Shin, S.Y.: Individual subject classification for Alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. NeuroImage 59, 2217-2230 (2012)
10. Chupin, M., Gerardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehericy, S., Benali, H., Garnero, L., Colliot, O.: Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. Hippocampus 19, 579-587 (2009)
11. Davatzikos, C., Bhatt, P., Shaw, L.M., Batmanghelich, K.N., Trojanowski, J.Q.: Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. Neurobiol Aging 32, 2322 e2319-2327 (2011)
12. Koikkalainen, J., Lotjonen, J., Thurfjell, L., Rueckert, D., Waldemar, G., Soininen, H.: Multi-template tensor-based morphometry: application to analysis of Alzheimer's disease. NeuroImage 56, 1134-1144 (2011)
13. Misra, C., Fan, Y., Davatzikos, C.: Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. NeuroImage 44, 1415-1422 (2009)
14. Querbes, O., Aubry, F., Pariente, J., Lotterie, J.A., Demonet, J.F., Duret, V., Puel, M., Berry, I., Fort, J.C., Celsis, P.: Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve. Brain 132, 2036-2047 (2009)
15. Westman, E., Simmons, A., Muehlboeck, J.S., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., Soininen, H., Weiner, M.W., Lovestone, S., Spenger, C., Wahlund, L.O.: AddNeuroMed and ADNI: similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America. NeuroImage 58, 818-828 (2011)
16. Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C.: An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE Trans Med Imaging 27, 425-441 (2008)
17. Coupe, P., Manjon, J.V., Gedamu, E., Arnold, D., Robles, M., Collins, D.L.: Robust Rician noise estimation for MR images. Med Image Anal 14, 483-493 (2010)
18. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging 17, 87-97 (1998)
19. Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C.: Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J Comput Assist Tomogr 18, 192-205 (1994)
20. Eskildsen, S.F., Coupe, P., Fonov, V., Manjon, J.V., Leung, K.K., Guizard, N., Wassef, S.N., Ostergaard, L.R., Collins, D.L.: BEaST: Brain extraction based on nonlocal segmentation technique. NeuroImage (2011)
21. Eskildsen, S.F., Ostergaard, L.R.: Active surface approach for extraction of the human cerebral cortex from MRI. Medical image computing and computer-assisted intervention : MICCAI - International Conference on Medical Image Computing and Computer-Assisted Intervention 9, 823-830 (2006)
22. Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L.: Unbiased average age-appropriate atlases for pediatric studies. NeuroImage 54, 313-327 (2011)

23. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 27, 1226-1238 (2005)
24. Lerch, J.P., Pruessner, J., Zijdenbos, A.P., Collins, D.L., Teipel, S.J., Hampel, H., Evans, A.C.: Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. Neurobiol Aging 29, 23-30 (2008)
25. Eskildsen, S.F., Fonov, V., Coupé, P., Østergaard, L.R., Collins, D.L.: Prediction of Alzheimer's Disease in Subjects with Mild Cognitive Impairment Using Structural Patterns of Cortical Thinning*. In: Dremstrup, K., Rees, S., Jensen, M.Ø. (eds.) 15th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC 2011), vol. 34, pp. 156-159. Springer Berlin Heidelberg (2011)
26. Eskildsen, S.F., Coupe, P., Fonov, V., Østergaard, L.R., Collins, D.L.: Effect of non-local means denoising on cortical segmentation accuracy with FACE. In: 17th Annual Meeting of the Organization on Human Brain Mapping, pp. 673. (Year)
27. Fan, Y., Shen, D., Gur, R.C., Gur, R.E., Davatzikos, C.: COMPARE: classification of morphological patterns using adaptive regional elements. IEEE Trans Med Imaging 26, 93-105 (2007)