



Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features

Yang Li, Yaping Wang, Guorong Wu, Feng Shi, Luping Zhou, Weili Lin, Dinggang Shen*

The Alzheimer's Disease Neuroimaging Initiative, Department of Radiology and BRIC, University of North Carolina at Chapel Hill, 27599, USA

Received 22 August, 2010; received in revised form 22 October 2010; accepted 10 November 2010

Note: proofreading version, may contain errors/tipos

Abstract

Neuroimage measures from magnetic resonance (MR) imaging, such as cortical thickness, have been playing an increasingly important role in searching for biomarkers of Alzheimer's disease (AD). Recent studies show that, AD, mild cognitive impairment (MCI) and normal control (NC) can be distinguished with relatively high accuracy using the baseline cortical thickness. With the increasing availability of large longitudinal datasets, it also becomes possible to study the longitudinal changes of cortical thickness and their correlation with the development of pathology in AD. In this study, the longitudinal cortical thickness changes of 152 subjects from 4 clinical groups (AD, NC, Progressive-MCI and Stable-MCI) selected from Alzheimer's Disease Neuroimaging Initiative (ADNI) are measured by our recently developed 4 D (spatial+temporal) thickness measuring algorithm. It is found that the 4 clinical groups demonstrate very similar spatial distribution of grey matter (GM) loss on cortex. To fully utilize the longitudinal information and better discriminate the subjects from 4 groups, especially between Stable-MCI and Progressive-MCI, 3 different categories of features are extracted for each subject, i.e., (1) static cortical thickness measures computed from the baseline and endline, (2) cortex thinning dynamics, such as the thinning speed (mm/year) and the thinning ratio (endline/baseline), and (3) network features computed from the brain network constructed based on the correlation between the longitudinal thickness changes of different regions of interest (ROIs). By combining the complementary information provided by features from the 3 categories, 2 classifiers are trained to diagnose AD and to predict the conversion to AD in MCI subjects, respectively. In the leave-one-out cross-validation, the proposed method can distinguish AD patients from NC at an accuracy of 96.1%, and can detect 81.7% (AUC = 0.875) of the MCI converters 6 months ahead of their conversions to AD. Also, by analyzing the brain network built via longitudinal cortical thickness changes, a significant decrease ($p < 0.02$) of the network clustering coefficient (associated with the development of AD pathology) is found in the Progressive-MCI group, which indicates the degenerated wiring efficiency of the brain network due to AD. More interestingly, the decreasing of network clustering coefficient of the olfactory cortex region was also found in the AD patients, which suggests olfactory dysfunction. Although the smell identification test is not performed in ADNI, this finding is consistent with other AD-related olfactory studies.

© 2010 Published by Elsevier Inc.

Keywords: Cortical thickness; Longitudinal analysis; Classification; Alzheimer's disease; Mild cognitive impairment; Dynamics; Brain network

1. Introduction

Alzheimer's disease (AD) is the most common cause of dementia. With the incidence rate doubling every 5 years after the age of 65, AD poses significant medical, social and socioeconomic challenges. Due to the fact that the cortical atrophy associated with AD can be detected *in vivo* using

MRI, neuroimaging measures have been playing an increasingly important role in searching for bio-markers of AD that can be used for early diagnosis, progression monitoring and also therapy responses measure. Recent studies have focused on individuals with mild cognitive impairment (MCI), a concept used to describe a high-risk "pre-dementia" state (the prodromal stage of Alzheimer's disease (Dubois et al., 2007)), in which a deterioration of cognitive skills can be measured while the ability to manage tasks of daily living remains intact (Petersen et al., 1999; Petersen,

* Corresponding author. Tel.: +1 919 843 8464; fax: +1 919 843 2641.
E-mail address: dgshen@med.unc.edu (D. Shen).

2004). According to epidemiological studies, each year 6~25% of subjects with MCI eventually convert to clinical AD, which is a markedly higher annual rate of conversion than that observed in normal elderly control population, which is around 1% (Petersen, 2001). However, MCI is a heterogeneous group. There have been approaches to more precisely define different MCI clinical presentations (Petersen, 2004). When individuals have impairments in domains other than memory it is classified as non-amnesic single- or multiple-domain MCI and these individuals are believed to be more likely to convert to other dementias. Additionally, when memory loss is the predominant symptom, it is termed “amnesic” MCI. Subjects in this subtype show greatest risk for conversion to AD in 6 years. These MCI subjects finally develop probable AD, which is referred to as Progressive-MCI (P-MCI). In contrast, some other MCI subjects do not convert to AD even after several years of follow-up, which is referred to as Stable-MCI (S-MCI). Therefore, more and more emphasis has been placed on identifying those P-MCI individuals from the MCI group. Successful identification of such individuals at an early-stage before the onset of clinical symptoms may lead to effective intervention of pharmacological treatments for AD as they become available.

It has been shown that many P-MCIs present evolving AD-like patterns of brain atrophy. Therefore, to find out how AD-induced grey matter (GM) atrophy spreads spatially and temporally in the brain, many longitudinal studies of normal aging or AD-degenerated cortical thinning have been conducted, in which GM volume/density are measured and analyzed within regions of interest (ROIs) or voxel-wise (Resnick et al., 2003; Thompson et al., 2003; Toga and Thompson, 2003; Gogtay et al., 2004; Thompson et al., 2004; Chetelat et al., 2005; Davatzikos et al., 2009; Desikan et al., 2009; Dickerson et al., 2009). Recently, cortical thickness is also proposed as a more stable parameter for AD diagnosis than volume/density measures, because it is a more direct measure of GM atrophy due to the cytoarchitectural feature of the GM (Regeur, 2000; Singh et al., 2006). Hence, as an alternative to volumetric measures, the assessment of thickness of the cerebral cortex has been recently proposed as a more sophisticated way to measure brain atrophy resulting from the AD neuropathological changes (Lerch et al., 2005). This measure has been proven both precise and sensitive in detecting alterations in cortical morphology (Lerch and Evans, 2005). Cortical thickness analysis has been successfully used in various studies as a marker to separate AD patients from healthy controls and MCI subjects (Chetelat et al., 2005; Singh et al., 2006; Du et al., 2007; Desikan et al., 2009; Hutton et al., 2009). In addition, more recently, attempts to distinguish P-MCI from S-MCI by analyzing the baseline cortical thickness have also been reported (Querbes et al., 2009).

So far, although there are plenty of longitudinal thickness studies focused on the comparative analysis between differ-

ent clinical groups (Toga and Thompson, 2003; Thompson et al., 2004; Chetelat et al., 2005; Lerch et al., 2005; Singh et al., 2006), research work on classification or AD conversion prediction using the longitudinal cortical thickness changes still remains scarce. Compared to the volumetric measure, the obstacles in incorporating longitudinal thickness change information into the classification largely come from the difficulties in measuring the subtle longitudinal thickness changes accurately and reliably. Comparing the thickness of cortical structures (1.2~4.5 mm) with the resolution of MR images (~1 mm), cortical structures are only a few voxels thick in the images and the possible errors in the measuring process could be considerably large. This situation becomes even worse when the thickness at different time-points are measured independently, because the expected change in GM thickness during the early stage of AD is less than 1 mm in most brain regions (Lerch et al., 2005; Singh et al., 2006), which can be easily overwhelmed by the noises. In our recently developed 4-D cortical thickness measuring method (Li et al., 2010), the accuracy and the robustness of thickness measurement for longitudinal images have been significantly improved by incorporating the information from all time-points as a constraint or guidance to each other. This leads to a much higher correlation detected between the thickness measured by our method and the MMSE (Mini-Mental State Examination score (Folstein et al., 1975)) or CDR-SOB (Clinical Dementia Rating-Sum of Boxes (Morris, 1993)) scores, when compared with the common 3D measuring method.

With this 4-D cortical thickness measuring method, in this study, experiments are designed to compare the thickness change patterns among four different clinical groups (AD, NC, S-MCI and P-MCI). Besides comparing the static features (such as thickness measured at each time-point), the temporal dynamic measures are also extracted, such as the thinning speed (mm/year) and thinning ratio (endline/baseline). In addition, to capture the system-level information, a brain network is first constructed based on the correlation of longitudinal cortical thickness changes between different ROIs, and then the network clustering coefficient (Rubinov and Sporns, 2009) which represents the wiring efficiency of the network is also extracted as feature in this study. Finally, we combine these three different categories of features to find 4D patterns that can be used to effectively diagnose AD or predict the MCI converters (the P-MCI patients) in MCI group.

2. Methods

2.1. Participants and inclusion criteria

The longitudinal dataset used in the preparation of this article were obtained from the ADNI database (www.adni-info.org). The 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

Table 1
Descriptive demographic information of the subjects in the study

Group	NC	S-MCI	P-MCI	AD	P-value
No. of subjects	40	36	39	37	
No. and percentage of males	23 (58%)	25 (69%)	25 (64%)	23 (62%)	NS
Baseline age, mean (SD)	73.7 (6.3)	75.3 (6.9)	75.6 (7.4)	74.8 (8.0)	NS
MMSE at first visit, mean (SD)	29.4 (0.7)	26.8 (1.6)	26.5 (1.7)	23.0 (1.9)	$P < 0.001^*$
MMSE at last visit, mean (SD)	28.8 (1.3)	25.8 (2.8)	21.1 (4.1)	13.8 (4.7)	$P < 0.001$
Significant decrease in MMSE	$P < 0.05$	$P < 0.001$	$P < 0.001$	$P < 0.001$	
No. and percentage of ApoE4 carriers	9 (22.5%)	21 (58.3%)	28 (71.8%)	27 (72.8%)	

NS = Not significant

* Significant difference in MMSE scores, except between S-MCI and P-MCI.

Administration (FDA), private pharmaceutical companies and non-profit organizations as a US\$60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early Alzheimer's disease progression is intended to aid researchers and clinicians in the development of new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. The image acquisition parameters have been described in www.adniinfo.org. The ADNI protocol included a sagittal volumetric 3D MPRAGE with 1.25×1.25 mm in-plane spatial resolution and 1.2 mm thick sagittal slices (8° flip angle). The TR and TE values of the ADNI protocol were somewhat variable, but the target values were TE ~ 3.9 ms and TR ~ 8.9 ms. In ADNI, other than 1.5T, for some subjects and time-points, MR scans of 3.0-T strength field are also available. However, to be consistent between scans of different subjects and different time-points, 1.5T scans are used in this study, since they are available for more subjects and time-points.

According to ADNI clinical procedures, diagnosis of Alzheimer's disease was made if the subject had a MMSE score between 20 and 26, a CDR score of 0.5 or 1, and met NINCDS/ADRDA (McKhann et al., 1984) criteria for probable Alzheimer's disease. The diagnosis of single-domain amnesic MCI was made according to the revised MCI criteria (Petersen, 2004): (i) MMSE scores between 24–30; (ii) reported memory complaint; (iii) having objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II (D, 1987); (iv) a CDR of 0.5; (v) absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and absence of dementia.

In this study, subjects from ADNI are included in the P-MCI group, if they satisfy the following criteria: (i) every subject was re-scanned and re-evaluated every six months for a period of at least 24 months. Therefore, including baseline, each has at least 5 time-points (ii). Each subject developed probable AD after at least 12 months from the baseline scan. Thus, there are at least 3 time-points before conversion. This is the minimum number of time-points

needed for the calculation of dynamic features. As of December 2009, when the data was downloaded, 39 subjects from ADNI were found to meet the above criteria and are thus included in the P-MCI group. For the other three clinical groups, subjects are selected to match the demographic information in the P-MCI group and to have the same gender and age. Specifically, for each subject in P-MCI group, a subject with the same gender and age is sought in the other group. If this is not possible, the gender constraint is relaxed first, and then the age constraint. However, under no circumstances are subjects with an age difference greater than 3 years included. Due to the similar reason for the calculation of dynamic features, the subjects in the other clinical groups are required to have at least 4 time-points. The demographic and clinical information of all the selected subjects in the four groups are summarized in Table 1.

2.2. Image preprocessing

The preprocessing step performed in our method is a standard procedure as used in general brain image analysis, aiming to reduce image noise and remove non-brain tissues. For each image, intensity inhomogeneity is firstly corrected using the N3 algorithm (Sled et al., 1998). Many algorithms have been developed to remove non-brain tissues, such as the skull and the extra-cranial tissues, including Brain Surface Extractor (BSE) (Shattuck and Leahy, 2001) and Brain Extraction Tool (BET) (Smith, 2002). However, it is still hard to achieve good skull stripping results on every subject through automatic algorithms. For better results, we combine BSE and BET to take advantage of both algorithms, followed by a manual editing to ensure that skull removal is accurate. In order to maintain the consistency between different time-points, the above skull stripping is first performed on baseline scan and then applied to all rigidly-aligned follow-up scans. We note here that the common method of registering all follow-ups to the baseline may introduce bias. The main reason is that only the follow-up images are interpolated and, therefore, smoothed a little bit after registration (Yushkevich et al., 2010). In our implementation, the baseline image is first rigidly (instead of affine which will affect the accuracy of thickness measures

due to the shearing transformation) registered to a standard template to introduce the similar interpolation/smoothing effect on the baseline image. After this, all follow-up images are rigidly registered to the warped baseline image.

2.3. Temporal consistent longitudinal cortical thickness measurement

To avoid getting jittery longitudinal thickness measures, in our recently developed method (Li et al., 2010), the information from all time-points is incorporated as a constraint or guidance when measuring the cortical thickness. Specifically, after obtaining the 4-D segmentation results using CLASSIC (Xue et al., 2006), an unbiased group-wise registration algorithm (Wu et al., 2010) is adopted to register each time-point image to a common space by iteratively constructing the group-mean image (via averaging the images from all time-points of a subject) and registering each time-point to the latest estimated group-mean image. In this way, the spatial correspondence among the scans of the same subject is established. The reason for not using the baseline as the template for registration is to avoid the possible bias which generally tends to exaggerate the longitudinal change (as reported in Thomas et al., 2009; Yushkevich et al., 2010). After performing the group-wise registration, the Jacobian determinants of the resultant deformation fields are computed to quantify the expansion or contraction caused by the deformation on each voxel. Then, in the group-mean space, the warped GM map from each time-point is multiplied by the Jacobian determinant voxel-wise to preserve the GM volume before warping. Due to partial volume effects, limited resolution and noise in imaging, “buried sulci” (part of the cortical surface hidden within the deep and narrow sulci) is often observed. This usually leads to the over estimation of the cortical thickness. To overcome this problem, a method proposed by Hutton et al. (2008) is adopted to recover the buried sulci. In this method, starting from the initial white matter (WM) map, GM layers are successively recovered to surround the WM, until the 2 sides meet in the middle of the buried sulci. Thus, the set of central voxels can be relabeled as the restored CSF.

In the final step, since these GM maps have been all aligned in a standardized stereotaxic space (i.e., group-mean space), for each GM voxel in this space, the cortical thickness of each time-point can be computed as the line-integral (Aganj et al., 2009) of the underlying GM density map along a direction on which the summation of the integrals in all time-points is minimized (Li et al., 2010). In this way, the information from different time-points is fully utilized as a constraint or guidance for finding the measuring direction (which is the key factor in consistent measurement of longitudinal thickness). Due to the fact that the obtained thickness measuring direction is optimal and consistent for all time-points in the longitudinal image dataset, this 4-D measurement of thickness is much more temporally consistent, robust and accurate than the conventional 3-D mea-

surement that estimates the thickness of each time-point image independently. For more details of this 4-D thickness measuring method, we refer the readers to Li et al. (2010).

After the longitudinal thickness measurement, for each subject, the obtained thickness maps of different time-points need to be mapped to a standard space (Tzourio-Mazoyer et al., 2002), in order to compare the thickness across different subjects at the corresponding location. Since all images at different time-points of each subject have been registered to the group-mean image, we just need to register the group-mean image to the template space, and then bring the obtained thickness maps of all time-points to this space by using the estimated deformation field. Finally, in order to suppress the registration errors between subjects, the mapped thickness in the template space is smoothed by a 6 mm full-width-at-half-magnitude (FWHM) Gaussian filter before across-subject comparison. In voxel-based morphometry (VBM) (Good et al., 2001), following spatial normalization, data are often smoothed using FWHM filter to improve the validity of statistical inferences and to reduce inter-individual variation or registration residual. The size of the smoothing kernel is generally chosen empirically. An often quoted “rule of thumb” rule of selecting the size of the smoothing kernel is that a kernel should be at least 2–3 times the voxel dimension (Jones et al., 2005). Therefore, in our study, starting from $1.5 \times 3 = 4.5$ mm to 9.0 mm, we tested the statistical inferences of different kernel size in intervals of 0.5 mm, and finally selected the kernel size as 6 mm, which gives the highest correlation between thickness measures and clinical scores (CDR).

2.4. Feature extraction

After estimating the longitudinal cortical thickness changes using the 4-D measuring method, the static, dynamic and network features can be extracted as detailed below.

2.4.1. Static and dynamic features

In this study, we call the thickness measure computed from a single time-point as the static feature, since it describes a static state in the progression of GM loss or disease development. Generally, only the baseline measurements are used for cortical thickness based classification in the existing works. However, in our study, the endline (last time-point) is also used to provide the most recent information of the pathological development. Moreover, with the increasing availability of longitudinal image datasets, dynamic features that directly describe the temporal changes of cortical thickness can also be calculated to provide more information. For instance, compared with static measures, the thinning ratio defined as $(1 - \text{endline/baseline}) \times 100\%$ may be a more accurate measurement of the severity of the disease. Meanwhile, a linear regression on thickness values of all time-points can also be performed voxel-wise to find the thinning speed (mm/year). This measure gives the kinetic of the GM loss and could be more helpful in prediction

of the disease development. Therefore, besides the static features (i.e., baseline and endline measures), we also calculate 2 dynamic features, i.e., thinning ratio and thinning speed in this study.

2.4.2. Adaptive ROI parcellation

In order to improve the robustness and reduce the dimensionality of the feature set, usually, the mean feature value computed within each ROI is used in the classification, instead of using the thickness measures on every single voxel directly. Therefore, the partition of the whole cortex into different ROIs is a very important step in the ROI-based feature extraction. Conventionally, the cortical parcellation can be achieved by mapping the anatomical labels of an atlas to each subject. However, due to the fact that the spatial distribution of thickness changes does not necessarily coincide with the anatomical partitions, the thickness changes on different voxels within a ROI could be heterogeneous. Therefore, directly averaging the feature values within each predefined anatomical ROI may cancel or at-

tenuate the information contained by the features. To better reflect the spatial distribution of cortical thickness changes, we apply data-driven method (Fan et al., 2007; Davatzikos et al., 2009) to automatically parcellate the brain cortex into different ROIs.

Here, we use the ROI generation for the baseline thickness as an example to explain the idea of adaptive ROI definition. First, on each voxel, across all subjects, the Pearson's correlation between the baseline thickness and the classification labels (for instance, '1' for AD and '0' for normal control (NC)) is computed. One axial slice of the resultant correlation map is shown in Fig. 1(a) with high correlation indicating high discriminative power. As the figure shows, the voxels with relatively high correlations form many different clusters. Using the "watershed" segmentation algorithm, these clusters can be segmented and defined as ROIs shown in Fig. 1(b). In this process, the small clusters with relatively low correlation are excluded in order to limit the number of ROIs, because of their low and

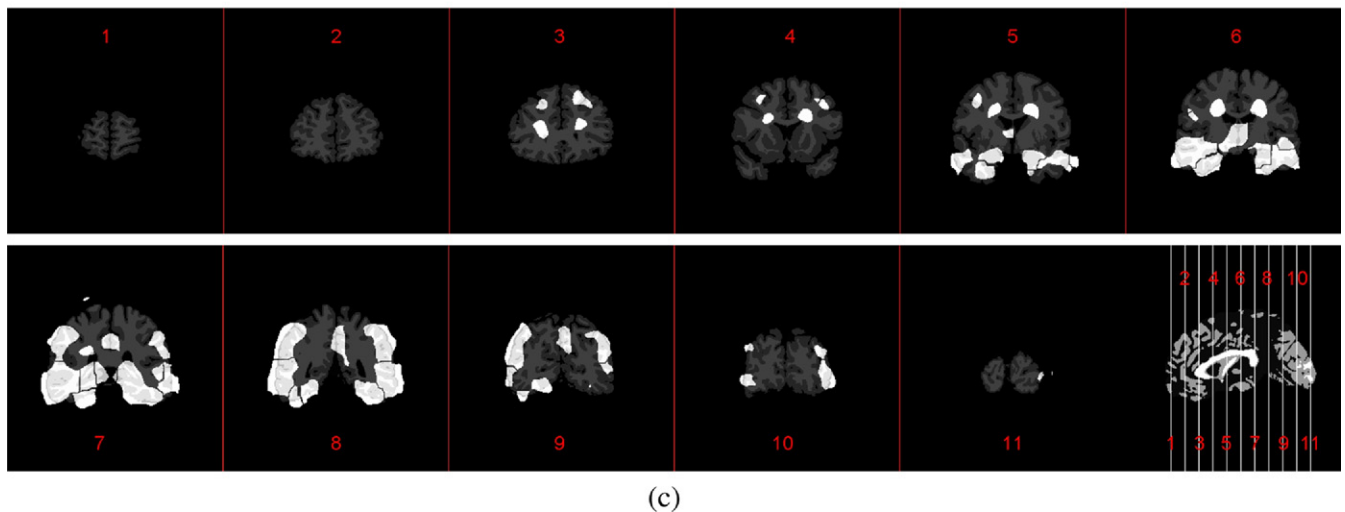
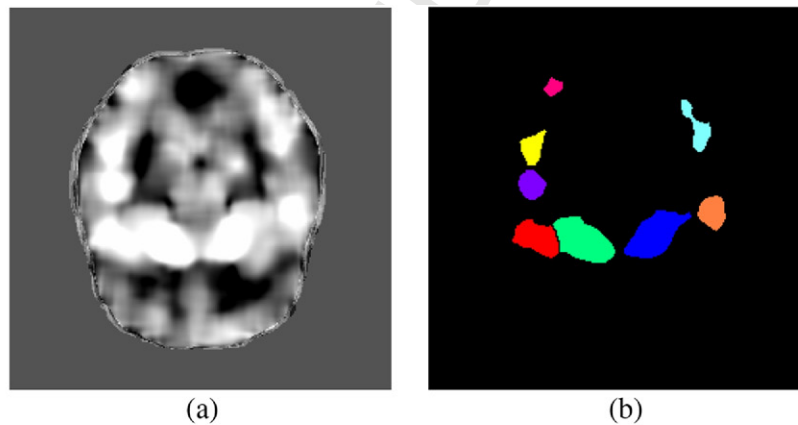


Fig. 1. An example of generating the spatial-adaptive ROIs for baseline thickness. (a) Correlation map between the baseline thickness and the classification labels; (b) The ROIs obtained via watershed segmentation corresponding to (a); (c) The ROIs generated for baseline thickness shown in different coronal slices.

less-reliable discriminative ability for the classification. This means that, for the purpose of classification, the obtained ROIs do not necessarily cover all the cortical regions. As an example, the final ROI partition results on baseline thickness are shown in Fig. 1(c). As can be seen, most of the ROIs exist on the temporal and parahippocampal cortex area. This is consistent with the findings in (Desikan et al., 2009; Holland et al., 2009), in which these regions are also reported as the most affected by AD. In the similar manner, ROIs for endline thickness, thinning speed and thinning ratio can also be generated, respectively. After that, within each ROI, the mean value of each feature can then be computed.

2.4.3. Network features

An important problem in neuroscience is how to characterize the underlying architectures of complex brain networks. Electroencephalography (EEG) (Micheloyannis et al., 2006; Stam et al., 2007) and fMRI (Eg  luz et al., 2005; Achard et al., 2006) have been used to explore the connections in brain. The brain network constructed based on these functional measures are called the functional network. More recently, structural measurements, such as cortical thickness, were also used to build the anatomical network (He et al., 2007). In this network, 2 areas are considered anatomically connected if they show statistically significant correlations ($p < 0.01$) on cortical thickness across different subjects (as shown in Fig. 2(a)). The small-worldness property was also observed in this thickness-revealed network, which is consistent with the findings from the networks constructed with functional measures. Since the network constructed in this way (He et al., 2007) represents the "average" connection information in a population, we refer it as the population network.

As a source of higher level information, the thickness-based network information is also included in this study. However, different from the network information in a population, we need network information at individual level for the classification task. Therefore, in our study, we use a

slightly different method to construct the network. Firstly, the whole brain cortex is partitioned into 90 ROIs according to the AAL atlas (Tzourio-Mazoyer et al., 2002). It should be noticed that here we do not use the spatial-adaptive ROIs (as described in Section 2.4.2) for network construction. This is because the adaptively defined ROIs may not cover all the cortical regions as shown in Fig. 1. Therefore, to make the constructed network complete, the ROI partitions used for network construction are generated according to an anatomical atlas (Tzourio-Mazoyer et al., 2002). Next, for each time-point, the mean cortical thickness of each ROI is computed. After that, a pairwise ROI connection is constructed, if the longitudinal cortical thickness changes in two ROIs are statistically correlated (as shown in Fig. 2(b)). Lastly, the brain network is constructed based on such connections. In this way, the constructed network represents the brain connection information of each subject. Thus, measures of this subject-specific network can be used for classification purposes. Corresponding to the population network in (He et al., 2007), we refer to our proposed network as an individual network in the rest of this paper.

With the proposed individual network, in this study, we calculate the clustering coefficient (Rubinov and Sporns, 2009) on every ROI as the network feature, which is defined as the ratio of the number of connections between one node's neighbors divided by the maximum possible connections between its neighbors. Intuitively, the clustering coefficient measures the local wiring efficiency of the brain network and is closely correlated to the small-worldness property of the underlying brain-network.

To summarize, for each ROI of a subject, the following five groups of features (as shown in Table 2) are extracted in our study: Baseline Thickness (BT, mm), Endline Thickness (ET, mm), TR (%), Thinning Speed (TS, mm/year) and Clustering Coefficient (CC) of brain network. Consequently, each subject is represented by 262 longitudinal image features.

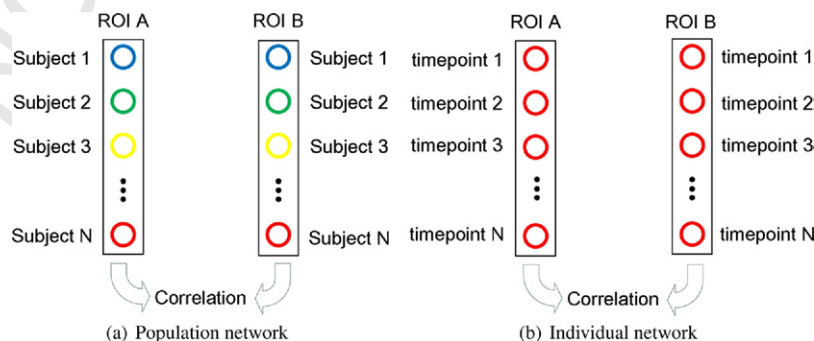


Fig. 2. Two different methods for thickness-based brain network construction. (a) In the population network, the pairwise connection between two ROIs is established, if the thickness measures from different subjects show significant correlation between these two ROIs; (b) In the individual network, the pairwise connection between two ROIs is established, if the thickness measures from different time-points of the same subject show significant correlation between these two ROIs.

Table 2
Features defined and extracted in the study

Static			Dynamic		Network
Features	Baseline thickness (BT)	Endline thickness (ET)	Thinning ratio (TR)	Thinning speed (TS)	Clustering coefficient (CC)
Number	26	29	56	61	90

2.5. Feature selection and classifier training

In this study, as shown in Table 2, each subject has 262 features from 3 categories (static, dynamic and network), which give different aspects of information about a subject. Among these features, some information may be complementary and some may be redundant or overlapped. Therefore, for the best classification performance, the optimal set of the features from 3 different categories should contain the most comprehensive information with the minimum redundancy. To achieve this, we use the mRMR (minimum redundancy maximum relevance) feature selection algorithm (Peng et al., 2005) to select the best feature subset. This filtering feature selection approach considers all features jointly and identifies a minimal set of features which jointly maximally differentiate the classes. With this feature selection method, we first train an SVM (support vector machine) nonlinear classifier to recognize the spatial-temporal thinning and the brain network changing patterns that distinguish AD from NC individuals. For the discrimination between P-MCI and S-MCI, a similar approach is adopted except that only the thickness values measured 6 months ahead of AD conversion are used. Thus, the classification results actually give a short-term prediction of the conversion from MCI to AD.

3. Results

3.1. Improved accuracy in measuring the longitudinal thickness changes

To show the improved accuracy in our temporal-consistent 4-D thickness measurement, correlation between the thickness measurements and the CDR-SOB scores are computed on each subject and then averaged in the P-MCI group. The reason to choose P-MCI group is because subjects in this group have relatively large change of thickness and clinical scores. The results of the conventional 3-D method (Aganj et al., 2009) and the 4-D method (Li et al., 2010) are compared in Fig. 3, in which much higher correlations are found between the clinical score and the 4-D thickness measures. To show the significance of improvement, the *T*-score map indicating the places where 4-D measures give statistically higher correlations is also provided. To control the multiple comparison errors, False Discovery Rate (FDR, defined as the expected proportion of false positives among all significant tests) correction is adopted. As an indirect validation for thickness measurement (Lerch and Evans, 2005; Aganj et al., 2009), the higher correlation with the clinical score indicates that the new 4-D thickness measuring method provides more sensi-

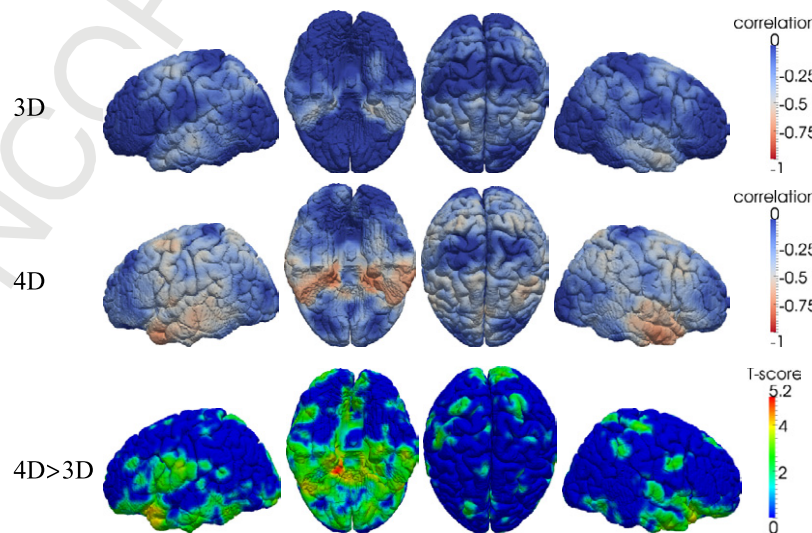


Fig. 3. Average correlation between thickness and CDR-SOB scores within the P-MCI subjects (each has 5 time-points). The higher correlation implies the improved accuracy using the 4-D thickness measurement (Li et al., 2010). From left to right: the left, inferior, superior and right views, respectively. The bottom shows the *T*-score ($p < 0.05$, FDR corrected) where the 4-D method gives significantly higher correlation than the 3-D method.

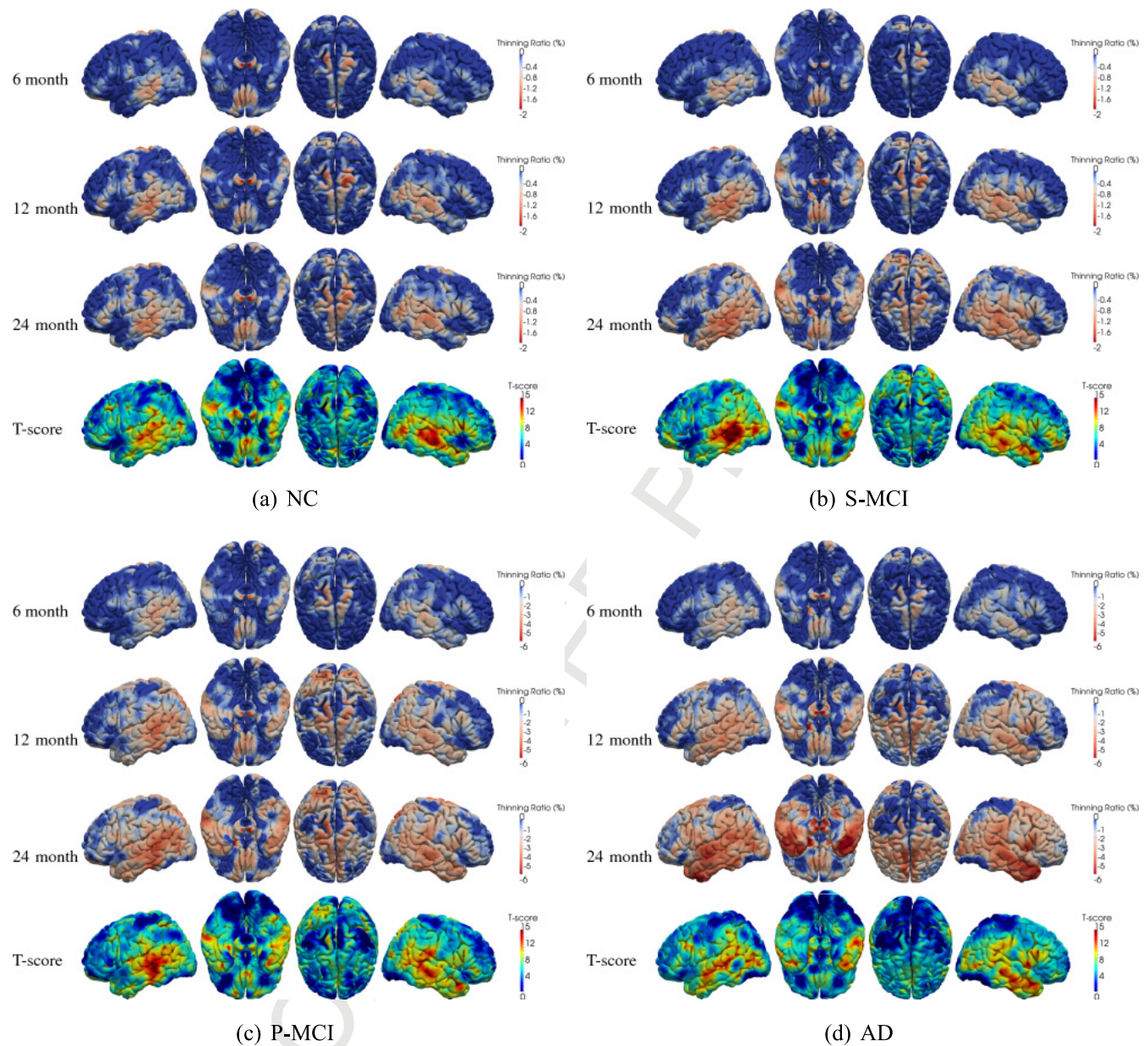


Fig. 4. Cortical thinning patterns in different clinical groups. The fourth row shows the location where significant ($p < 0.05$, FDR corrected) cortex thinning (by comparing baseline and endline) is detected.

tive and accurate measurement of the longitudinal thickness changes.

3.2. 4-D Pattern of cortex thinning in different groups

With the longitudinal thickness measured for each subject at all time-points, 4-D maps of the thickness thinning ratio within each clinical group can be computed. First, in the template space (Tzourio-Mazoyer et al., 2002), for each subject the thickness thinning ratio is computed by comparing the thickness at each follow-up time-point with that at baseline. Then, by averaging the thinning ratio maps of different subjects at the corresponding time-point, the mean thinning ratio map in different clinical groups is computed.

The results for NC, S-MCI, P-MCI and AD group are shown in Fig. 4(a)-4(d), respectively. As the figures show, on average, the 4 groups have very different cortical thinning ratios (NC: $< 1.5\%$, S-MCI: $< 2\%$, P-MCI: $1 \sim 4\%$ and AD: $2 \sim 6\%$, for 24 months).

The similar spatial distribution of cortical thinning in all 4 groups also indicates that, instead of distinguishing between the 4 groups only by specific spatial patterns based on static features, adding dynamic features may give valuable complementary information for the purpose of classification, especially between P-MCI and S-MCI. Therefore, the thinning speeds of the subjects in the P-MCI and S-MCI group are compared voxel-wise by adding age and gender as

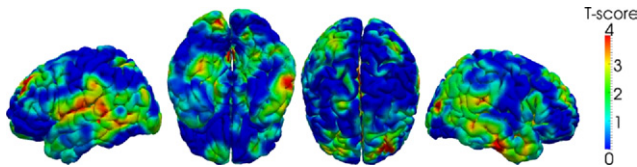


Fig. 5. Cortical areas with higher thinning speed in P-MCI group than in S-MCI group. From left to right, T-scores ($p < 0.05$, FDR-corrected) are shown in left, inferior, superior and right views, respectively.

confounding variables. The results are shown in Fig. 5. As we can see, significant difference exists between the thinning speed of the two groups. Specifically, the cortical thinning on the anterior prefrontal cortex, temporal gyrus and parahippocampal area is much faster in the P-MCI group than in the S-MCI group. This indicates that adding the dynamic features may help differentiate the P-MCI subjects from S-MCI subjects.

3.3. Discriminative power of different features

To show the discriminative power of different features, the correlation between each of the 5 proposed features and the classification labels is calculated for different classification tasks, i.e., AD vs. NC and P-MCI vs. S-MCI, respectively. The results are shown in protocols 6 and 7, in which higher correlation suggests higher discriminative ability in the classification tasks. In both figures, baseline and endline features show quite similar spatial patterns, although higher correlation is observed at endline. This is because these two static features carry similar information of the disease development. Since the cortical thinning process continues during the study, the GM loss becomes more severe and much easier to detect at the endline. However, when compared with the other two dynamic features, i.e., thinning ratio and thinning speed, very different spatial distributions can be observed. This is because the dynamic features carry different information about the pathology: the thinning ratio indicates the severity of the disease and the thinning speed represents the trend of degeneration. This information is not available in the static features and thus is complementary to the static features.

As high-level information, the clustering coefficient of the thickness-change-revealed brain network also shows relatively high correlations with the classification labels, as indicated by the bottom figures in protocols 6 and 7. Not surprisingly, parahippocampal cortex, temporal lobe and supramarginal gyrus are found among the regions with high correlations, since relatively high GM loss was also detected in these regions.

It may be noticed that dynamic and network features both show relatively low correlation to the classification labels, when compared with static features. One possible reason is that these features rely on multiple time-points and need more complex computation, therefore they are more sensitive to noises. However, as these new features provide

complementary information to the static features, by adding them, the discriminative ability of the selected features as a whole set will be improved. In our classification experiments, a similar phenomenon was also observed. The performance of the classifiers has been successfully improved by adding the dynamic and network features, as we will see in details in the next section.

3.4. Thickness-based classification results

To validate the effectiveness of the proposed static, dynamic and network features for AD diagnosis and conversion prediction, 3 different combinations of features, i.e., static, static+dynamic and static+dynamic +network, are tested in the leave-one-out cross-validation.

The accuracies of the 2 classification tasks, i.e., AD vs. NC and P-MCI vs. S-MCI, with respect to 3 different combinations of features are summarized in Table 3. As we can see, as more complementary information is added (by adding more features), a higher classification rate can be achieved. Among these three combinations of features, the static+dynamic +network combination gives the best results. For AD and NC classification, the accuracy reaches as high as 96.1%, with addition of dynamic features contributing most of the improvement. More importantly, for classification between P-MCI and S-MCI, the classification rate is improved from 76.7% to 81.7%. This means that we can achieve a higher accuracy in the prediction of MCI-to-AD conversion. For more information about the performance of the classifiers, the Receiver Operation Curves (ROC) of different classifiers with different feature combinations are also provided in Fig. 8.

3.5. Discriminative features selected for classification

By adding the complementary information (provided by dynamic and network features) into the thickness-based AD diagnosis and prediction of the MCI to AD conversion, the classification accuracy is improved significantly. Therefore, it is interesting to analyze the optimal features selected by the classifier. Table 4 lists the optimal features selected to best discriminate between AD and NC, and also between P-MCI and S-MCI. The "order" in the table does not represent the importance of a feature individually, because the mRMR feature selection algorithm (Peng et al., 2005) con-

Table 3
Comparisons of the classification accuracy of three different combinations of features

Group	Static	Static + dynamic	Static + dynamic + Network
AD vs. NORMAL	89.6%	94.8% ($P < 0.0001$)	96.1% ($P < 0.002$)
P-MCI vs. S-MCI	76.7%	80.3% ($P < 0.0001$)	81.7% ($P < 0.003$)

The P -value shows the significance of improvement by comparing with the static method.

NC vs. AD

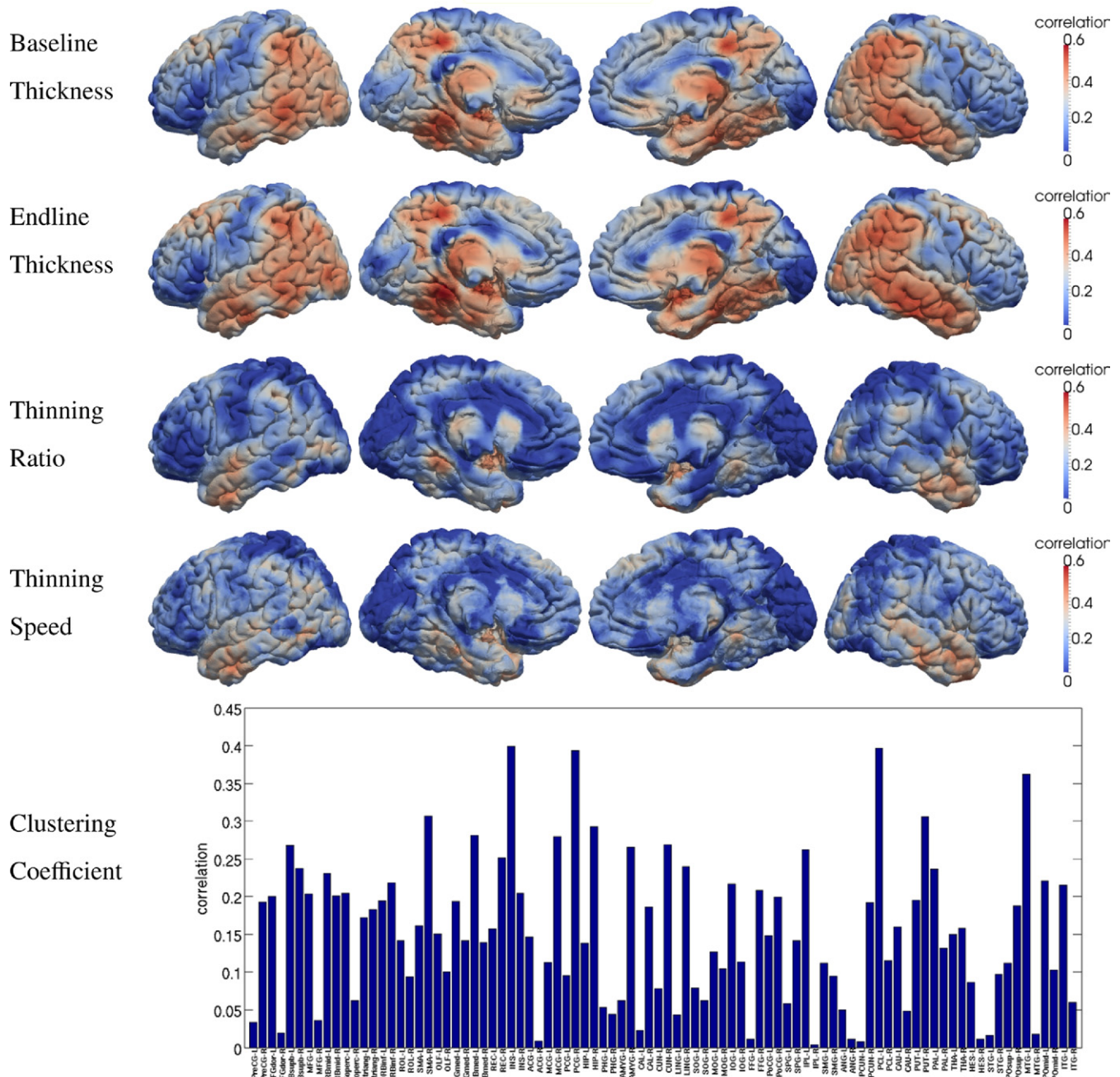


Fig. 6. Discriminative power of five different types of features in distinguishing NC and AD. High correlation indicates high discrimination power.

siders the relationship between all the features to favor the complementary and avoid redundancy. Intuitively, these orders only mean that, if only one feature is used for classification, the feature with order “2.5 cm should be selected. Then, if 2 features are considered, the order “5.1 cm feature should be included, because the combination of these 2 features has the best performance, compared with any other two-feature combinations. In other words, every newly added feature in the table contains some complementary

information to the existing ones, until no more complementary information can be found. For instance, for distinguishing AD and NC, it is found that the combination of 7 features (listed in Table 4(a)) gives the best accuracy and adding more features will not help. As we can see, features from all 3 categories (static, dynamic and network) are selected, because they provide complementary information. As expected, more dynamic features (4 out of 6) are selected for the discrimination of P-MCI and S-MCI. In contrast, for

P-MCI vs. S-MCI

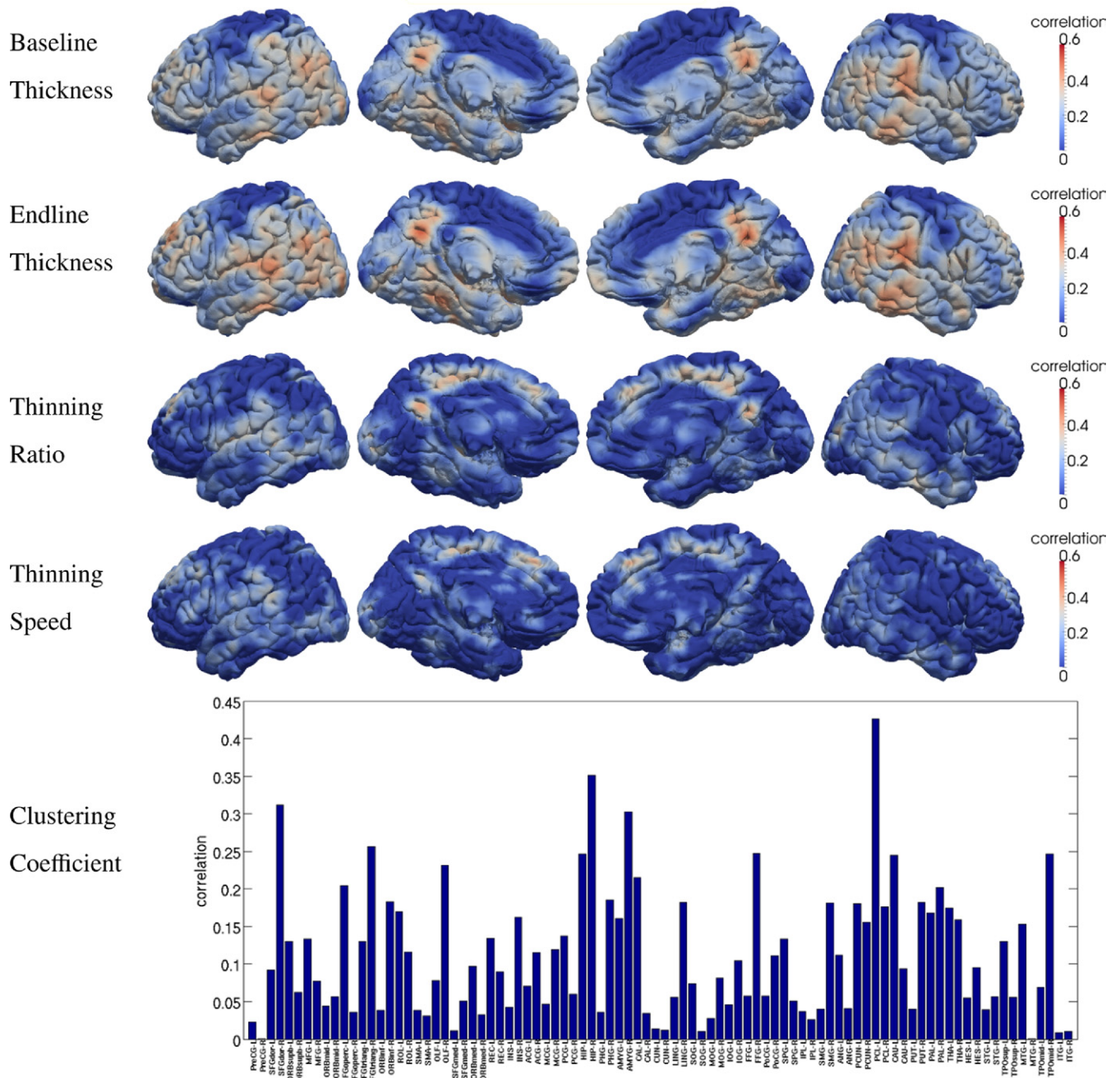


Fig. 7. Discriminative power of five different types of features in distinguishing P-MCI and S-MCI. High correlation indicates high discrimination power.

the discrimination of AD and NC, more static features (3 out of 7) are selected, whereas only 2 dynamic features are chosen. The selection of network features shows very interesting results. For the diagnosis of AD, the clustering coefficient on olfactory is selected as the second feature. This result is consistent with the findings in other nonimaging-based studies focusing on the association between AD and olfactory dysfunction (Koss et al., 1988; Talamo et

al., 1989; Murphy et al., 1990; McCaffrey et al., 2000; Wesson et al., 2010b).

3.6. Decreased network clustering efficiency in P-MCI group

As described in Section 2.4.3, for the purpose of classification, the subject-specific brain network is constructed

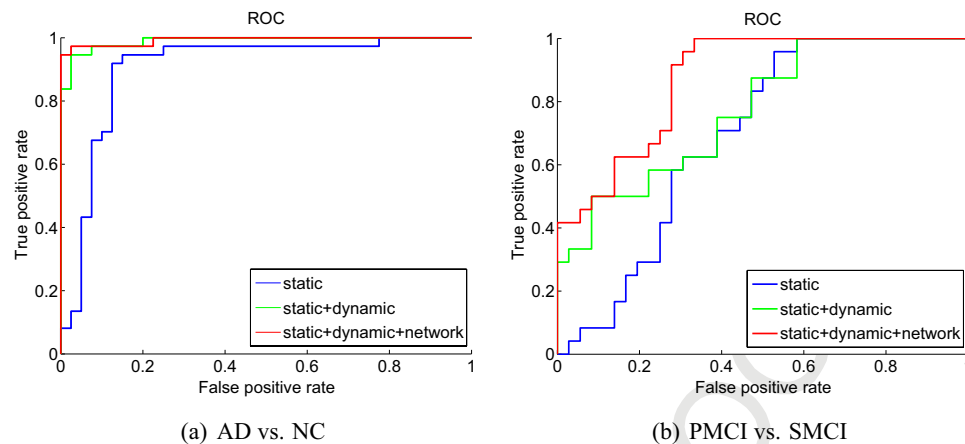


Fig. 8. Receiver Operation Curve (ROC) of different feature combinations.

for providing additional information for classification. On the other hand, it will also be interesting to investigate the general longitudinal network changes in a population. For this purpose, we also construct such population network (He et al., 2007) on each time-point within the 4 clinical groups. Especially, by comparing the networks at different time-points in the P-MCI group, we can find the longitudinal changes of brain network associated with the progression of AD. In our study, significant decrease of the clustering coefficient of the brain network is observed in the P-MCI group (Fig. 9(c)), which shows the degenerated small-worldness in the brain network caused by AD. This indicates that, with the development of AD, not only the localized GM loss can be observed, but also on a larger scale the brain connections between different regions are degenerated. Similar decreases of the clustering coefficient are also observed in S-MCI and AD group, but not significant. In the NC group, with a much smaller scale, such decrease is also observed.

4. Discussion

4.1. Similar spatial patterns of cortical thinning in 4 clinical groups

Based on the measured longitudinal thickness changes within 4 different clinical groups, we first explore the possibility of finding different spatial patterns of cortical thinning to differentiate the 4 clinical groups. The similar idea has been used to differentiate AD and frontotemporal dementia (al Duet, 2007). However, in our study, by comparing the spatial pattern of cortical thinning in different groups (shown in Fig. 4), it is found that the spatial pattern of GM loss associated with the AD-related pathology is similar to the pattern observed in normal aging. Specifically, parahippocampal cortex, middle/inferior temporal gyrus, supramarginal gyrus, angular gyrus and superior frontal gyrus (consistent with the findings in (Sowell et al., 2004; Desikan et al., 2009; Holland et al., 2009)) are the regions in which

relatively higher thinning ratios are detected in all 4 groups. This indicates that, regarding the cortex atrophy, the pathological development of AD can be viewed as an “accelerated” process of normal aging, because the AD-related GM loss has the similar spatial distribution as that in normal aging, but at a much higher speed. The answer to the question regarding why this thinning process stops (at least temporarily) in the S-MCI group could be very helpful in the effort to find effective pharmacological treatments for AD. Although the thinning ratio and thinning speed are different in all 4 groups, the anatomical areas where the cortical thinning is observed are mainly temporal lobe and around parahippocampal region. As expected, the thinning speed is relatively higher in P-MCI group than in S-MCI and NC group. The highest thinning speed is observed in the

Table 4

The selected optimal feature set for two classification tasks

Order	Feature	Regions
(a) AD vs. NC		
1	Endline Thickness	Parahippocampal Left (PHG-L)
		Fusiform Left (FFG-L)
2	Clustering Coefficient	Olfactory Left (OLF-L)
3	Thinning Speed	Temporal Inferior Right (ITG-R)
4	Endline Thickness	Precentral Right (PreCG-R) Frontal
		Middle Right (MFG-R)
5	Clustering Coefficient	Cuneus Right (CUN-R)
6	Endline Thickness	Supra Marginal Left (SMG-L)
		Temporal Supra Left (STG-L)
7	Thinning Speed	Temporal Middle Right (MTG-R)
		Temporal Inferior Right (ITG-R)
(b) P-MCI vs. S-MCI		
1	Baseline Thickness	Supra Marginal Left (SMG-L)
		Angular Left (ANG-L) Parietal
		Inferior Left (IPL-L)
2	Thinning Speed	Postcentral Left (PoCG-L)
3	Thinning Ratio	Temporal Middle Left (MTG-L)
4	Clustering Coefficient	Fusiform Left (FFG-L)
5	Thinning Speed	Paracentral Lobule Left (PCL-L)
6	Thinning Speed	Temporal Middle Right (MTG-R)
		Temporal Supra Right (STG-R)

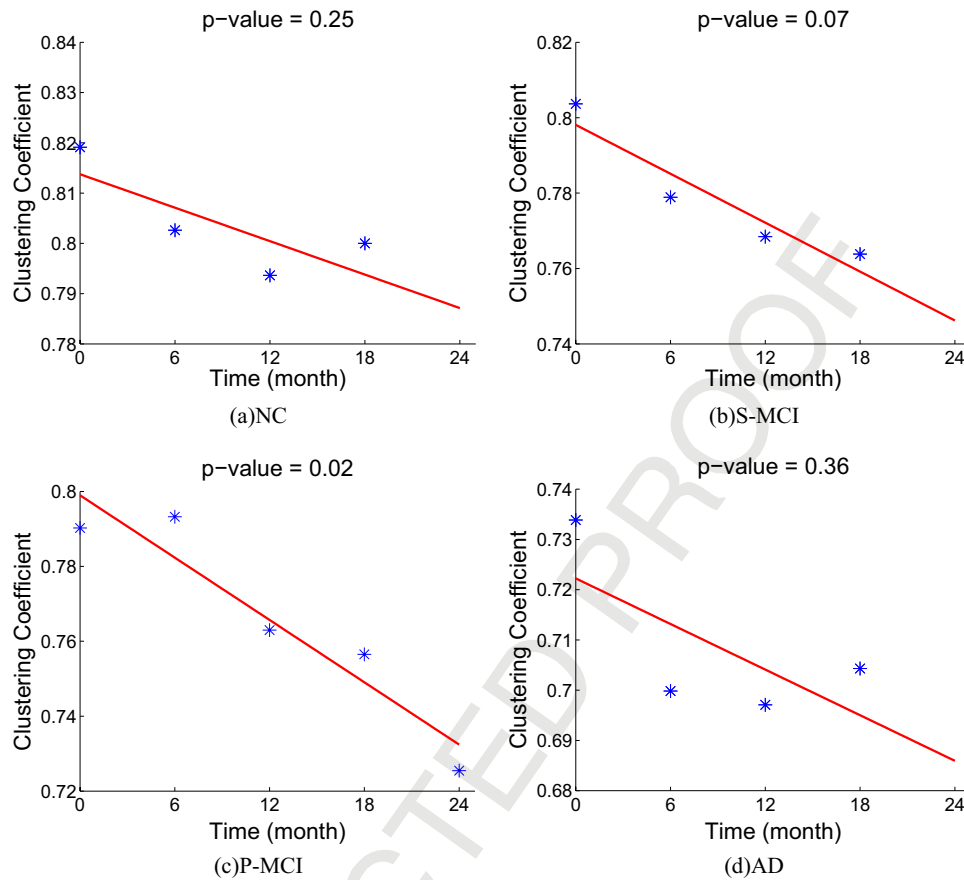


Fig. 9. Decrease of the average clustering coefficient (on 90 nodes) observed in 4 clinical groups.

AD group. To summarize the observed atrophy patterns in the sequence of time-points, the first area in which the thinning is observed is the supramarginal gyrus, angular gyrus and posterior superior/middle temporal gyrus. Then, the cortex thinning spreads, within the temporal lobe, from posterior to anterior, to fusiform gyrus and temporal pole. The highest thinning ratio is observed on parahippocampal cortex. The thinning speed becomes even higher in subjects with probable AD (the AD group). These findings are consistent with the results in (Thompson et al., 2003; Gogtay et al., 2004; Thompson et al., 2004; Lerch et al., 2005; Singh et al., 2006; Julkunen et al., 2009).

Our finding that 4 groups share a common spatial pattern of thickness change is also supported by the results from other 2 related studies (Davatzikos et al., 2009; Querbes et al., 2009). In both studies, a classifier was first trained to distinguish AD and NC by learning the different pattern of changes in either GM density (Davatzikos et al., 2009) or cortical thickness (Querbes et al., 2009). After that, in both studies, the trained classifier was applied to the subjects in MCI group. The output score of the classifier, which is continuous before thresholding, was used as a measure of severity of the AD-like brain change. The effectiveness of this AD measure was then validated by the experiments, in which high correlations were observed between MMSE

scores and classifier-generated scores (Davatzikos et al., 2009). In addition, in the NC group, the AD vs. NC classification score was also found to be highly correlated with age. Both studies show that the brain change patterns are similar in different groups but with different magnitudes.

4.2. Complementary information provided by the dynamic and network features

In order to find the biomarker which can differentiate P-MCI from S-MCI for the prediction of the conversion from MCI to AD, 2 dynamic measures, i.e., the thinning ratio (related to the severity of AD) and the thinning speed (related to AD's development speed), are proposed in our study. In Fig. 5, the thinning speed is found to be significantly higher in the P-MCI group than in the S-MCI group, mainly on supramarginal gyrus, inferior parietal lobule and right inferior temporal gyrus. To demonstrate the ability of these features in the discrimination of P-MCI and S-MCI, the correlations between these features and the classification labels are shown in Fig. 7. Relatively high correlations are observed on paracentral lobule, precuneus, lingual, inferior/middle temporal, parahippocampal and entorhinal cortex, respectively. The proposed network feature (clustering coefficient) represents high-level information revealed by lon-

gitudinal thickness changes. Among the anatomical regions found with high correlation between network features and classification labels (as shown in protocols 6 and 7), the GM loss is also observed in some of these regions, such as fusiform and mid temporal pole. However, no significant GM loss was found in other regions, such as supplementary motor area, superior frontal and paracentral lobule. This can be explained as that, since network features measure the wiring efficiency of the brain network, the degeneration of 1 cortical region may also affect the network property of other regions (which connect to the degenerated regions). Therefore, the network features are more global than the local GM loss features and thus give additional information.

4.3. Network degeneration around olfactory gyrus

It has been noticed that Alzheimer's disease often results in impaired olfactory perceptual acuity (Koss et al., 1988; Talamo et al., 1989; Murphy et al., 1990; Doty, 2001). Due to this fact, the olfactory deficit was used to discriminate AD from major depression (McCaffrey et al., 2000) and even proposed as a potential biomarker for early-stage AD diagnosis (Wesson et al., 2010b). More recent pathological study also suggests the correlation between olfactory dysfunction and AD (Wesson et al., 2010a). Recently, MR image study (Frisoni et al., 2009) also reported reduced density in the orbitofrontal olfactory cortex in the earliest stages of the AD. In our study, we did not detect significant atrophy in olfactory gyrus by directly using the thickness information provided by either static or dynamic features. However, by adding the high-level network information, the clustering coefficient from the olfactory gyrus is selected as the second feature in Table 4 for discriminating AD from NC. One explanation for this finding is that the olfactory deficit observed in AD probably can be caused by either the degeneration of the olfactory gyrus itself, or by the dysfunction of information transmission channel (connecting to the olfactory gyrus). This finding also indicates the effectiveness of the network features in differentiating AD from NC.

5. Conclusions

By using our longitudinal cortical thickness measurement method, we first compared the spatial pattern of cortical thinning in 4 clinical groups. It is found that a similar pattern of GM loss is shared among all clinical groups. At the same time, thinning dynamics, such as the thinning speed and the thinning ratio, give complementary information in distinguishing the 4 groups, especially between S-MCI and P-MCI. In addition to the features directly based on the thickness measures, brain network is also constructed to extract high-level connectivity information, by analyzing the correlation of longitudinal thickness changes between different ROIs. Combining the complementary information provided by features from all 3 different categories (static, dynamic and network), the classification accuracy between

AD and NC, as well as the accuracy in predicting the MCI-to-AD conversion, are both improved. The analysis of brain network changes also shows significant decrease of the clustering coefficient associated with AD development in the P-MCI group. In addition, the decreased network wiring efficiency around the olfactory cortex may suggest the olfactory dysfunction in AD which is consistent with other AD-related olfactory research with smell identification test.

In addition to these interesting findings with substantial clinical implications, this study also has some limitations. First, we used linear cortex thinning model in the study. Within a relative short period of time (24 months) and a limited number of scans (at most 5 time-points), constant thinning speed may be a reasonable simplification. In future studies, if longitudinal information of more time-points becomes available, a more sophisticated thinning model could be investigated. Second, for the prediction of AD-conversion in the MCI group, the validation was made according to the diagnosis obtained in the following 6 months. However, after 6 months, it is possible that some S-MCI subjects may convert to probable AD. Therefore, for more accurate evaluation of the cortical thickness based prediction, the results should be re-visited when more longitudinal information becomes available. Third, the clustering coefficient, a network measure used in this study, is defined as "the ratio of the number of connections between one node's neighbors divided by the maximum possible connections between its neighbors", which means only the neighbors are included. Recently, another network measure, "communicability", was proposed by (Crofts et al., 2010). In the definition of this measure, not only the local direct connections but also the long distance indirect connections are also taken into account. In the future study, such more advanced network measures could be applied. Lastly, due to the fact that the ADNI study is primarily concerned with developing imaging measures and biomarkers for incipient AD, the appropriate subtype of MCI (amnestic-MCI) which likely represents prodromal AD is emphasized. Because of this, the MCI population of ADNI is very amnestic and may not reflect the true distributions of different subtypes of MCI in the whole population.

Disclosure statement

Data used in this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) (www.loni.ucla.edu/ADNI). The ADNI data were previously collected across 50 research sites. Study subjects gave written informed consent at the time of enrollment for imaging and genetic sample collection and completed questionnaires approved by each participating sites' Institutional Review Board (IRB). None of the authors have any conflicts of interest. Fig. 6.

Acknowledgments

This work was supported by the National Institute of Health (NIH), Bethesda, MD 20 892, USA. Data used in the preparation of this manuscript were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

Uncited references

Davatzikos et al., 2008.

References

- Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E., 2006. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* 26, 63–72.
- Aganj, I., Sapiro, G., Parikshak, N., Madsen, S.K., Thompson, P.M., 2009. Measurement of cortical thickness from MRI by minimum line integrals on soft-classified tissue. *Hum. Brain Mapp.* 30, 188–199.
- Chetelat, G., Landeau, B., Eustache, F., Mezenge, F., Viader, F., de la Sayette, V., Desgranges, B., Baron, J.-C., 2005. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: A longitudinal MRI study. *Neuroimage* 27, 934–946.
- Crofts, J., Higham, D., Bosnell, R., Jbabdi, S., Matthews, P., Behrens, T., Johansen-Berg, H., 2010. Network analysis detects changes in the contralateral hemisphere following stroke. *Neuroimage* 54, 161–169.
- D. W., 1987. Wechsler Memory Scale-Revised Manual. Psychological Corporation, San Antonio, TX.
- Davatzikos, C., Fan, Y., Wu, X., Shen, D., Resnick, S. M., 2008. Detection of prodromal Alzheimer's disease via pattern classification of MRI. *Neurobiol Aging* 29, 514–523.
- Davatzikos, C., Xu, F., An, Y., Fan, Y., Resnick, S.M., 2009. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. *Brain* 132, 2026–2035.
- Desikan, R.S., Cabral, H.J., Hess, C.P., Dillon, W.P., Glastonbury, C.M., Weiner, M.W., Schmansky, N.J., Greve, D.N., Salat, D.H., Buckner, R.L., Fischl, B., 2009. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. *Brain* 132, 2048–2057.
- Dickerson, B.C., Feczko, E., Augustinack, J.C., Pacheco, J., Morris, J.C., Fischl, B., Buckner, R.L., 2009. Differential effects of aging and Alzheimer's disease on medial temporal lobe cortical thickness and surface area. *Neurobiol. Aging* 30, 432–440.
- Doty, R.L., 2001. Olfactory Deficit in Alzheimer's Disease? *Am. J. Psychiatry* 158, 1533–1534.
- Du, A.-T., Schuff, N., Kramer, J. H., Rosen, H. J., Gorno-Tempini, M. L., Rankin, K., Miller, B. L., W. Weiner, M., 2007. Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. *Brain* 130, 1159–1166.
- Dubois, B., Feldman, H.H., Jacova, C., DeKosky, S.T., Barberger-Gateau, P., Cummings, J., 2007. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol.* 6, 734–746.
- Egúiluz, V.M., Chialvo, D.R., Cecchi, G.A., Baliki, M., Apkarian, A.V., 2005. Scale-free brain functional networks. *Phys. Review Lett.* 94, 018102.
- Fan, Y., Shen, D., Gur, R.C., Gur, R.E., Davatzikos, C., 2007. Compare: Classification of morphological patterns using adaptive regional elements. *IEEE Trans Med Imaging* 26, 93–105.
- Folstein MF, Folstein SE, McHugh PR., 1975. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198.
- Frisoni, G.B., Prestia, A., Rasser, P.E., Bonetti, M., Thompson, P.M., 2009. In vivo mapping of incremental cortical atrophy from incipient to overt Alzheimer's disease. *J. Neurol.* 256, 16–924.
- Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., Thompson, P.M., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. USA* 101, 8174–8179.
- Good, C.D., Johnsrude, I., Ashburner, J., Henson, R.N.A., Friston, K.J., Frackowiak, R.S.J., 2001. Cerebral asymmetry and the effects of sex and handedness on brain structure: A voxel-based morphometric analysis of 465 normal adult human brains. *Neuroimage* 14, 685–700.
- He, Y., Chen, Z.J., Evans, A.C., 2007. Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. *Cereb. Cortex* 17, 2407–2419.
- Holland, D., Brewer, J.B., Hagler, D.J., Fenema-Notestine, C., Dale, A.M., 2009. Subregional neuroanatomical change as a biomarker for Alzheimer's disease. *Proc. Natl. Acad. Sci. USA* 106, 20954–20959.
- Hutton, C., De Vita, E., Ashburner, J., Deichmann, R., Turner, R., 2008. Voxel-based cortical thickness measurements in MRI. *Neuroimage* 40, 1701–10.
- Hutton, C., Draganski, B., Ashburner, J., Weiskopf, N., 2009. A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging. *Neuroimage* 48, 371–380.
- Jones, D., Symms, M., Cercignani, M., Howard, R., Jun, 2005. The effect of filter size on VBM analyses of DT-MRI data. *Neuroimage* 26, 546–554.
- Julkunen, V., Niskanen, E., Muehlboeck, S., Pihlajamäki, M., Mervikönnö, Hallikainen, M., Kivipelto, M., 2009. Cortical thickness analysis to detect progressive mild cognitive impairment: A reference to Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders* 28, 404–412.
- Koss, E., Weiffenbach, J.M., Haxby, J.V., Friedland, R.P., 1988. Olfactory detection and identification performance are dissociated in early Alzheimer's disease. *Neurology* 38, 1228–1232.
- Lerch, J., Pruessner, J., Zijdenbos, A., Hampel, H., Teipel, S., Evans, A., 2005. Focal decline of cortical thickness in Alzheimer's disease identified by computational neuroanatomy. *Cereb. Cortex* 15, 995–1001.
- Lerch, J.P., Evans, A.C., 2005. Cortical thickness analysis examined through power analysis and a population simulation. *Neuroimage* 24, 163–173.
- Li, Y., Wang, Y., Xue, Z., Shi, F., Lin, W., Shen, D., 2010. Consistent 4d Cortical Thickness Measurement for Longitudinal Neuroimaging Study. MICCAI, Beijing.
- McCaffrey, R.J., Duff, K., Solomon, G.S., 2000. Olfactory Dysfunction Discriminates Probable Alzheimer's Dementia From Major Depression: A Cross-Validation and Extension. *J. Neuropsychiatry Clin. Neurosci.* 12, 29–33.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E., 1984. 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.
- Michelyonannis, S., Pachou, E., Stam, C.J., Vourkas, M., Erimaki, S., Tsirka, V., 2006. Using graph theoretical analysis of multi channel EEG to evaluate the neural efficiency hypothesis. *Neurosci. Lett.* 402, 273–277.
- Morris, J., 1993. The clinical dementia rating (CDR): Current version and scoring rules. *Neurology* 42, 2412–2414.
- Murphy, C., Gilmore, M.M., Seery, C.S., Salmon, D.P., Lasker, B.R., 1990. Olfactory thresholds are associated with degree of dementia in Alzheimer's disease. *Neurobiol. Aging* 11, 465–469.
- Peng, H., Long, F., Ding, C., 2005. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transact. Pattern Anal. Machine Intell.* 27, 1226–1238.

- Petersen, R., 2001. Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 56, 1133–1142.
- Petersen, R., 2004. Mild cognitive impairment as a diagnostic entity. *J. Intern. Med.* 256, 183–194.
- Petersen, R., Smith, G., Waring, S., Ivnik, R., Tangalos, E., Kokmen, E., 1999. Mild cognitive impairment: clinical characterization and outcome. *Arch. Neurol.* 56, 303–308.
- Querbes, O., Aubry, F., Pariente, J., Lotterie, J.-A., Démonet, J.-F., Duret, V., Puel, M., Berry, I., 2009. Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve. *Brain* 132, 2036–2047.
- Regeur, L., 2000. Increasing loss of brain tissue with increasing dementia: a stereological study of post-mortem brains from elderly females. *Eur. J. Neurol. Off. J. Eur. Fed. Neurol. Soc.* 7, 47–54.
- Resnick, S.M., Pham, D.L., Kraut, M.A., Zonderman, A.B., Davatzikos, C., 2003. Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain. *J. Neurosci.* 23, 3295–3301.
- Rubinov, M., Sporns, O., 2009. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* 52, 1059–69.
- Shattuck, D., Leahy, R., 2001. Automated graph-based analysis and correction of cortical volume topology. *IEEE Trans. Med. Imaging* 20, 1167–1177.
- Singh, V., Chertkow, H., Lerch, J., Evans, A., Dorr, A., Kabani, N., 2006. Spatial patterns of cortical thinning in mild cognitive impairment and Alzheimer's disease. *Brain* 129, 2885–2893.
- Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* 17, 87–97.
- Smith, S., 2002. Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155.
- Sowell, E.R., Thompson, P.M., Leonard, C.M., Welcome, S.E., Kan, E., Toga, A.W., 2004. Longitudinal mapping of cortical thickness and brain growth in normal children. *J. Neurosci.* 24, 8223–8231.
- Stam, C.J., Jones, B.F., Nolte, G., Breakspear, M., Scheltens, P., 2007. Small-world networks and functional connectivity in Alzheimer's disease. *Cereb. Cortex* 19, 92–99.
- Talamo, B.R., Rudel, R., Kosik, K.S., Lee, V.M., Neff, S., Adelman, L., Kauer, J.S., 1989. Pathological changes in olfactory neurons in patients with Alzheimer's disease. *Nature* 337, 736–739.
- Thomas, A.G., Marrett, S., Saad, Z.S., Ruff, D.A., Martin, A., Bandettini, P.A., 2009. Functional but Not Structural Changes Associated with Learning: an Exploration of Longitudinal voxel-Based Morphometry VBM. *Neuroimage* 48, 117–125.
- Thompson, P.M., Hayashi, K.M., de Zubicaray, G., Janke, A.L., Rose, S.E., Semple, J., 2003. Dynamics of gray matter loss in Alzheimer's disease. *J. Neurosci.* 23, 994–1005.
- Thompson, P.M., Hayashi, K.M., Sowell, E.R., Gogtay, N., Giedd, J.N., Rapoport, J.L., de Zubicaray, G.I., Janke, A.L., Rose, S.E., Semple, J., 2004. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. *Neuroimage* 23, S2–S18.
- Toga, A.W., Thompson, P.M., 2003. Temporal dynamics of brain anatomy. *Annu. Rev. Biomed. Eng.* 5, 119–145.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single subject brain. *Neuroimage* 15, 273–289.
- Wesson, D.W., Levy, E., Nixon, R.A., Wilson, D.A., 2010a. Olfactory dysfunction correlates with amyloid-beta burden in an Alzheimer's disease mouse model. *J. Neurosci. Off. J. Soc. Neurosci.* 30, 505–514.
- Wesson, D.W., Wilson, D.A., Nixon, R.A., 2010b. Should olfactory dysfunction be used as a biomarker of Alzheimer's disease? *Expert Review Neurother.* 10, 633–635.
- Wu, G., Jia, H., Wang, Q., Shen, D., 2010. Groupwise Registration With Sharp Mean. *MICCAI, Beijing*.
- Xue, Z., Shen, D., Davatzikos, C., 2006. CLASSIC: Consistent longitudinal alignment and segmentation for serial image computing. *Neuroimage* 30, 388–399.
- Yushkevich, P.A., Avants, B.B., Das, S.R., Pluta, J., Altainay, M., Craige, C., 2010. Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: An illustration in ADNI 3T MRI data. *Neuroimage* 50, 434–445.