Brain Imaging and Behavior
DOI 10.1007/s11682-015-9448-7

ORIGINAL RESEARCH

Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease

Ali Khazaee[1] · **Ata Ebrahimzadeh**[1] · **Abbas Babajani-Feremi**[2,3,4]

© Springer Science+Business Media New York 2015

Abstract The study of brain networks by resting-state functional magnetic resonance imaging (rs-fMRI) is a promising method for identifying patients with dementia from healthy controls (HC). Using graph theory, different aspects of the brain network can be efficiently characterized by calculating measures of integration and segregation. In this study, we combined a graph theoretical approach with advanced machine learning methods to study the brain network in 89 patients with mild cognitive impairment (MCI), 34 patients with Alzheimer's disease (AD), and 45 age-matched HC. The rs-fMRI connectivity matrix was constructed using a brain parcellation based on a 264 putative functional areas. Using the optimal features extracted from the graph measures, we were able to accurately classify three groups (i.e., HC, MCI, and AD) with accuracy of 88.4 %. We also investigated performance of our proposed method for a binary classification of a group (e.g., MCI) from two other groups (e.g., HC and AD). The classification accuracies for identifying HC from AD and MCI, AD from HC and MCI, and MCI from HC and AD, were 87.3, 97.5, and 72.0 %, respectively. In addition, results based on the parcellation of 264 regions were compared to that of the automated anatomical labeling atlas (AAL), consisted of 90 regions. The accuracy of classification of three groups using AAL was degraded to 83.2 %. Our results show that combining the graph measures with the machine learning approach, on the basis of the rs-fMRI connectivity analysis, may assist in diagnosis of AD and MCI.

Keywords Resting-state functional magnetic resonance imaging (rs-fMRI) · Alzheimer's disease (AD) · Mild cognitive impairment (MCI) · Graph theory · Machine learning · Support vector machine (SVM)

Introduction

Alzheimer's disease (AD) is the most commonly diagnosed type of dementia in elderly patients (Reitz et al. 2011). AD is a progressive neurodegenerative disease characterized by memory dysfunction and cognitive decline, likely caused by aberrant neural circuitry. Prior to the onset of dementia, patients may develop an intermediate stage of dysfunction known as mild cognitive impairment (MCI). Patients with MCI have a higher risk of progressing to AD (Petersen et al. 1999). Indeed, it has been shown that patients with MCI progress to AD at a rate of 10–15 % per year, whereas healthy age-matched individuals with no evidence of MCI develop dementia at a rate of 1–2 % per year (Petersen et al. 2001). Developing accurate and automated methods with the ability to identify individuals with MCI and AD, and differentiate the disease state from normal aging, would be a powerful clinical tool to identify patients for early treatment which may possibly delay disease progression. To date, several studies have investigated AD or MCI induced alterations of the brain network using resting-state functional magnetic resonance imaging (rs-fMRI) data (Brier et al. 2014; Buckner et al. 2009; Jie et al. 2014; Mccarthy et al. 2014; Sanz-Arigita et al. 2010;

✉ Ali Khazaee
khazaee.a@gmail.com

[1] Department of Electrical and Computer Engineering, Babol University of Technology, Babol, Iran

[2] Department of Pediatrics, University of Tennessee Health Science Center, Memphis, TN, USA

[3] Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, TN, USA

[4] Neuroscience Institute, Le Bonheur Children's Hospital, Memphis, TN, USA

Published online: 12 September 2015

Supekar et al. 2008; Toussaint et al. 2014; Zhao et al. 2012). Although rs-fMRI based brain network analysis has been used as a promising technique for identification of AD and MCI, an accurate and reliable predictive biomarker based on this technique for these diseases has not yet been established. Importantly, several studies have shown that the topological organization of brain networks are disrupted in neurodegenerative diseases (Anderson and Cohen 2013; Dey et al. 2012; dos Santos Siqueira et al. 2014; Fekete et al. 2013; Jie et al. 2013; Li et al. 2013; Rish et al. 2012; Wang et al. 2013; Wee et al. 2012a), possibly providing a quantifiable biological phenomena to indicate disease. Recent studies suggested that topological organization of the human whole-brain networks, constructed using rs-fMRI data, can be characterized using various graph metrics (Bassett et al. 2009; Bullmore and Sporns 2009).

Machine learning approach has been used in diagnosis of MCI and AD in order to automatically classify healthy controls (HC) from patients with MCI or AD (Challis et al. 2015; Davatzikos et al. 2011). Selection of the most discriminative features for use in machine learning methods is very important. Graph measures are appropriate features in automatic diagnosis of diseases, since they provide information about the entire brain network. Other local features such as regional homogeneity and amplitude of low frequency fluctuations may also be used for diagnosis of neurodegenerative diseases, however, these features only consider specific local changes and are less efficient than the graph measures corresponding to the entire brain network (Zhang et al. 2012). A few studies investigated automated diagnosis of AD or MCI using the topological properties of the brain network (Jie et al. 2013; Khazaee et al. 2015; Li et al. 2013; Wang et al. 2013; Wee et al. 2012a). However, to the best of our knowledge, there is no study to use topological properties of the brain network for automatic identification of three groups (i.e., AD, MCI and HC).

This study investigated abnormality of connectivity between different brain areas in patients with AD and MCI, and developed and evaluated an accurate and automated method for diagnosis of these diseases using machine learning techniques. To this end, we calculated the connectivity matrices using the rs-fMRI data collected from three groups (i.e., AD, MCI, and HC) and then performed two different analyses on these matrices. In the first analysis, raw connectivity matrices were analyzed using the network based statistics (NBS) method (Zalesky et al. 2010) for finding elements of the connectivity matrices that differ significantly between three groups. The second analysis was based on the calculated local and global graph measures from the connectivity matrices. First, graph measures were analyzed using analysis of variance (ANOVA) to find those measures that differed significantly between three groups. Then a sophisticated machine learning approach was developed for automatic identification of AD, MCI, and HC using the extracted graph measures. The machine learning approach employed the extracted graph measures as discriminating features and performed forward sequential feature selection (FSFS) algorithm on them to find the most discriminative feature vector. Finally, a support vector machine (SVM) was trained, based on the most discriminative feature vector, to automatically distinguish three groups.

The main goal of our study was to demonstrate an automated and accurate method for identification of AD, MCI, and HC based on the graph theoretical approach. In addition to achieving this, we found: (1) a connected component within the connectivity matrix which was significantly different in three groups (found from the NBS method); (2) graph measures which were significantly different in three groups (found from the ANOVA on graph measures); and (3) graph measures with most discriminating ability between groups (found from the machine learning approach). Another novelty of our method is the use of a functional brain atlas for defining brain network nodes. Previous AD or MCI studies have used structural atlases such as the automated anatomical labeling (AAL) atlas, which consists of 90 regions within the cerebral cortex and provides a coarse parcellation of the brain based on the non-overlapping structural units. A coarse parcellation of the brain based on the structural features may not appropriately represent the functional brain network. In this study, we used a parcellation based on 264 putative functional areas, proposed in (Power et al. 2011), and obtained superior results using this atlas compared to the AAL atlas.

Methods

Subjects

Data from 34 patients with AD (average age 72.5 years, 18 female), 89 patients with MCI (average age 71.8 years, 46 female), and 45 age-matched HCs (average age 75.9 years, 26 female) from the Alzheimer's disease neuroimaging initiative (ADNI) database (http://adni.loni.ucla.edu) were analyzed in this study.

1) The patients with AD had a Mini-Mental State Examination (MMSE) score of 14–26, a Clinical Dementia Rating (CDR) of 0.5 or 1.0 and met the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for probable AD.
2) The patients with MCI had MMSE scores between 24 and 30, a memory complaint, objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, a CDR of 0.5, absence

of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living and an absence of dementia.

3) The HC subjects were non-depressed, non-MCI, non-demented and had an MMSE score of 24–30 and a CDR of 0. Demographic information of subjects is summarized in Table 1. Data for this study were selected based on availability of resting-state fMRI datasets for age-matched control subjects and patients with AD and MCI. Two subjects had MMSE scores of 24 (one subject achieved MMSE scores of 30 and 28 in next months) and one subject had an MMSE score of 25. These three subjects had CDRs of zero and were classified as HC in the ADNI database. Other subjects had MMSE scores greater than 27.

Data acquisition and preprocessing

The functional and structural MRI images were collected using 3-T Philips scanners. Acquisitions were performed according to the ADNI acquisition protocol (Jack et al. 2008). A total of 140 functional volumes (TR/TE 3000/30 msec, flip angle=80°, 3.313 mm slice thickness, 48 slices) were obtained. For each subject, the first few volumes (7 volumes) of the functional images were discarded for signal equilibrium and to allow the participant's adaptation to the circumstances, leaving rest volumes for next steps. Image preprocessing was carried out using the Data Processing Assistant for Resting-State fMRI (DPARSF) toolbox (Chao-Gan and Yu-Feng 2010) and the SPM5 package (http://www.fil.ion.ucl.ac.uk/spm).

Standard pre-processing steps were performed on the rs-fMRI datasets. Slice-timing correction to the last slice was performed. fMRI time-series realigned using a six-parameter rigid-body spatial transformation to compensate for head movement effects (Friston et al. 1995). Four subjects (2 HC, 1 MCI, and 1 AD) who exhibited head motions more than 2.5 mm of displacement and 2.5° of rotation in any direction were excluded. All images were then normalized into the Montreal Neurological Institute (MNI) space and resampled to 3-mm isotropic voxels. Resulted images then detrended, smoothed using a Gaussian filter with FWHM=4 mm and band-pass filtered (0.01–0.08 Hz). To reduce the effect of the physiological artifacts, the whole-brain signal was removed by a multiple linear regression analysis (Fox et al. 2005; Fransson 2005; Greicius et al. 2003). In addition to the global mean signal, six head motion parameters, the cerebrospinal fluid (CSF), and the white matter signals were also removed as nuisance covariates to reduce the effects of motion and non-neuronal BOLD fluctuations (Fox et al. 2005; Kelly et al. 2008).

Brain network analysis

The 264 putative functional areas atlas parcellated the whole brain into 264 distinct regions (Power et al. 2011). These areas were formed using two sets of region of interests (ROIs) resulted from two different methods: meta-analytic and functional connectivity (fc)-mapping. The first method of identifying putative functional areas searched a large fMRI datasets (n>300) for brain regions that showed significant activity when certain tasks (e.g., button-pressing) were performed or certain signal types (e.g., error-related activity) were expected. This method identified 151 ROIs across the cortex, subcortical structures, and cerebellum. The fc-mapping method was applied to the rs-fMRI data from 40 healthy young adults. This method generated 193 ROIs across the cortex. The interested reader is referred to (Cohen et al. 2008) and (Nelson et al. 2010) for more information about the conceptual and technical details of fc-mapping. The ROIs from meta-analytic and fc-mapping methods were merged to form a maximally-spanning collection of 264 ROIs (Power et al. 2011). The preference was given to the meta-analytic ROIs and non-overlapping fc-mapping ROIs were then added.

Table 1 Demographics of healthy control subjects, patients with MCI, and patients with AD

	Healthy Control	Patients with MCI	Patients with AD
Number	45	89	34
Male/Female	19/26	43/46	16/18
Age (mean±SD)	75.9±6.79	71.77±7.78	72.54±7.02
MMSE (mean±SD)+	28.95±1.56	27.56±2.2	21. 24±3.37
CDR (mean±SD)+	0.07±0.21	0.49±0.17	0.92±0.31
FAQ (mean±SD)+	0.61±2.69	3.73±5.16	17.52±6.55
GD scale (mean±SD)*	1±1.16	2.18±2.51	1.67±1.52

MMSE Mini-Mental State Examination, *CDR* Clinical Dementia Rating, *FAQ* Functional Activities Questionnaire, *GD* Geriatric Depression

+ p-value<0.01 (one-way analysis of variance)

* p-value=0.02 (one-way analysis of variance)

We employed the identified 264 putative functional areas as graph nodes to construct the brain network. The representative signal of each node was obtained by averaging the time series of all voxels within the corresponding region. Edges of the brain network were defined as functional connectivity of all pairs of 264 putative functional areas using the Pearson's correlation coefficient. The resulting weighted undirected functional connectivity matrix was a dense network. We thresholded the connectivity matrix by preserving a proportion of the strong weights (PSW) which is defined as the number of the retained strong weights divided by the total number of weights. All diagonal weights (self-connections) were set to zero. By assigning a large value to PSW and keeping the weaker edges corresponding to less significant and noisy correlations, the resulted binary graph will be a denser graph with more edges. On the other hand, setting a small value to PSW removes many edges from the graph and generates a disconnected graph in which the graph measures may not be calculated. To find the optimal threshold value, we performed a graph theory-based analysis (Bassett et al. 2009; Bullmore and Sporns 2009; Khazaee et al. 2015). This analysis sought a threshold value for PSW that maximizes the global cost efficiency (GCE). The GCE is defined as the global efficiency E minus the proportional threshold PSW (i.e., $GCE=E - PSW$). The global efficiency E is computed as:

$$E = \frac{1}{n}\sum_{i\in N} E_i = \frac{1}{n}\sum_{i\in N} \frac{\sum_{j\in N, j\neq i} d_{ij}^{-1}}{n-1} \quad (1)$$

Where E_i is the efficiency of node i, N is the set of all nodes in the network, n is the number of nodes and d_{ij} is the shortest path length (distance) between nodes i and j.

The GCE of an economical small-world network has a positive maximum value at some proportional threshold $PSW_{\max}$. Small-world networks have a dense network of local connections and a sparse network of long-range connections, and consequently high levels of local and global efficiencies. The proportional threshold PSW in the GCE quantifies the required energy expenditure for the network to maintain its efficiency. Therefore, the proportional threshold PSW can be considered as the cost of the network. The main benefit of finding the optimal value of the threshold based on maximizing GCE is that this method defines a data-driven threshold value and does not need user-defined thresholds. In fact, this method searches over a wide range of thresholds and selects the one that maximizes the GCE function. Adoption of the measures of network efficiency was for directly estimating the economic performance of small-world brain functional networks. Global efficiency has several advantages over the other classical measures of small-worldness such as average shortest path length (characteristic path length). First, global efficiency is computed as the average inverse shortest path length and, unlike the characteristic path length, can be computed in disconnected graphs. During the search for finding optimum threshold value, application of high threshold values may result in disconnecting some nodes. In this case, the path between disconnected nodes will be infinite and correspondingly zero efficiency. Second, path length measures the capacity of network for serial transfer of information between nodes while global efficiency measures the capacity of network for parallel transfer of information between nodes via multiple series of edges. Because the information processing in the brain is massively parallel, it seems reasonable to use the measure of global efficiency.

Computation of graph measures

Once the graphs were constructed, we computed five global and eleven local graph measures and used them as features to train and test the classifier. Global measures have only one value for each graph whereas local measures have n values (n is the number of nodes in the graph) for each graph. Three measures of functional segregation (i.e., clustering coefficient (Watts and Strogatz 1998), local efficiency (Latora and Marchiori 2001), and ratio of local to global efficiency) were computed to characterize the ability of the brain for specialized processing within densely interconnected groups of regions. Two functional integration measures (i.e., characteristic path length (Watts and Strogatz 1998) and global efficiency (Latora and Marchiori 2001)) were used to assess ability of the brain in rapidly combining specialized information from distributed regions. Nine local nodal measures (i.e., degree, node strength, participation coefficient (Guimera et al. 2007), diversity coefficient (Rubinov and Sporns 2011), betweenness centrality (Brandes 2001), K-coreness centrality (Hagmann et al. 2008), subgraph centrality (Estrada and Higham 2010), eigenvector centrality (Newman 2008), and PageRank centrality (Boldi et al. 2009)), one measure of resilience (i.e., assortativity (Foster et al. 2010)) and one measure of network small-worldness (Humphries and Gurney 2008; Watts and Strogatz 1998) were computed to investigate properties of 264 putative functional areas of brain. The graph measures were calculated based on the binary adjacency matrices, except the node strength which was computed based on weighted networks, using the *Brain Connectivity Toolbox* (www.brain-connectivity-toolbox.net) (Rubinov and Sporns 2010). These measures together construct the final feature vector for each HC, MCI, and AD subject. The size of final feature vector for each subject was 2909 ($11\times264=2904$ local and 5 global features).

Statistical analysis

Two types of statistical analysis were performed: statistical analysis on raw connectivity matrices and statistical analysis on graph measures. To determine the altered connectivity

networks in AD and MCI patients in the first statistical analysis, we analyzed raw connectivity matrices using network based statistics (NBS) method implemented in NBS toolbox (Zalesky et al. 2010). Firstly, ANOVA F-statistic was calculated independently for each connection in the network to test the null hypothesis of equality in the mean value of that connection (across intra-group subjects) between groups. Connections with an F-statistic value exceeding a primary threshold (F-threshold) comprise supra-threshold connections set. The connectivity matrices were searched for any connected graph components defined by the set of supra-threshold connections. The size of any observed component was computed as the number of connections it comprised. Then permutation testing was used to assign a family-wise error rate (FWER) corrected p-value for each component. In each permutation, subjects were randomly assigned to a group (either HC, MCI, or AD) and the above steps were repeated on the permuted data. The size of the largest connected component was then recorded in each permutation. After 10,000 permutations, an empirical null distribution for the size of the largest connected component was obtained. FWER-corrected p-value for a given component size in the original data was calculated as the proportion of permutations for which the largest component size was equal or greater. The value of the primary threshold (i.e., F-threshold) is chosen experimentally. However, the FWER is controlled irrespective of the choice of the F-threshold. The choice of F-threshold can only affect the sensitivity of the NBS method. Significant networks can lose their significance or include fewer/more connections. To obtain reliable results by the NBS method, we experimented with a range of F-thresholds and reported consistent networks across F-thresholds.

In the second type of statistical analysis, we used ANOVA to determine between-group significant differences in both global and local graph measures. Most of the graph measures (described in Section Methods–Discussion) require the underlying network to be sparse. Different threshold values for PSW ranging from 0.1 to 0.5 by a step of 0.01 were examined. Consequently, the binary graphs were constructed by keeping a proportion of PSW of the strong edges, i.e., the elements of the connectivity matrix excluding the diagonal elements. As explained in Section Methods–Discussion, 16 functional segregation, functional integration, and regional nodal graph measures were calculated for the thresholded connectivity matrix. Comparisons were performed over 41 different threshold values (from 0.1 to 0.5 by a step of 0.01) and ANOVA on graph measures resulted in F-statistics and corresponding p-values. Considering a significance level of $p<0.05$, we identified the global and local graph measures which had significantly different values in three groups of HC, MCI, and AD.

Feature selection

An efficient feature selection algorithm is an essential part of a machine learning approach in the case of high dimensional datasets, such as the graph measures extracted from the connectivity matrix based on rs-fMRI data. In the feature selection stage, an optimal subset of features are selected from the original feature set. Feature selection facilitates data understanding, reduces the storage requirements and training–testing times, and improves accuracy of classification. Feature selection algorithms roughly divide into two categories: filter and wrapper methods. The filter methods select a subset of features according to the general characteristics of data, independently of chosen classifier. However, the wrapper methods require a predetermined classifier and evaluate features according to their performances in discrimination of classes. The Fisher algorithm and the sequential feature selection algorithm are the most popular filter and wrapper methods, respectively. These algorithms are explained briefly here.

Fisher score is a univariate filter method that is commonly employed to determine the discriminatory power of individual features between two classes of equal probability (Duda et al. 2012). It is independent of the class distribution. Fisher score for each feature in a two class problem is defined as:

$$FS = \frac{n_1(m_1 - m)^2 + n_2(m_2 - m)^2}{\left(n_1\sigma_1^2 + n_2\sigma_2^2\right)} \tag{2}$$

where m is the mean value of the feature, m_1 and m_2 are feature mean values on each class, σ_1^2 and σ_2^2 are respective variances, and n_1 and n_2 are number of samples in two classes.

Forward sequential feature selection (FSFS), starting from the empty set of features, sequentially adds features that result in the highest objective function when combined with the features that have already been selected. Backward sequential feature selection works in the opposite direction of FSFS and starting from the full set, sequentially removes the features that result in the smallest decrease in the value of the objective function.

As mentioned above, our goal was to reduce the dimension of the data by finding a small set of important features which can give good classification performance. We have used a two-level feature selection by combining filter and wrapper methods. Filter methods rely on general characteristics of the data to evaluate and to select the feature subsets without involving the chosen learning algorithm. Wrapper methods use the performance of the chosen learning algorithm to evaluate each candidate feature subset. Wrapper methods search for features better fit for the chosen learning algorithm, but they can be significantly slower than filter methods if the learning algorithm takes a long time to run. Thus, applying the FSFS algorithm on the whole features is very time consuming. To

overcome this problem, we have used filter methods as preprocessing step before wrapper methods. The wrapper methods were applied to the first half of the sorted features by the filter methods. This will reduce the classification time while maintaining high performance of the algorithm by eliminating the half of the worst features. The interested reader is referred to (Kohavi and John 1997) for more information about the concepts of filters and wrappers.

Classification and performance metrics

In this study, supervised machine learning methods were used to construct the classifier. A set of input data, namely the training dataset, was used to train the supervised machine learning algorithm and produce desired output. We used the support vector machine (SVM) as the supervised machine learning algorithm (Noble 2006). The SVM is a supervised machine learning method that was originally developed for binary classification problems (Vapnik 1998). The classification algorithm was implemented in MATLAB (The Math Works, Natwick, MA) using LIBSVM software package (www.csie.ntu.edu.tw/~cjlin/libsvm/).

Holdout cross validation was used to train and test the classifier. To reduce the variance of the results, the holdout cross validation was repeated several times. Because our data was unbalanced, we hypothesized that using of the other cross validation methods can cause the classifier to be overfitted to the samples of a specific class. The holdout method randomly selects K subjects from each group for training, leaving the others for testing. In this study, the value of K was set to 23. Therefore, number of training samples was 69 (23 from each group) and number of testing samples was 95 (20 HC, 65 MCI, and 10 AD). Four performance metrics including accuracy, sensitivity, specificity, and positive predictivity were used to assess the performance of the classifier in discriminating HC, MCI, and AD groups. Eqs. (3) to (6) define these metrics.

$$Accuracy = \frac{total\,number\,of\,subjects - number\,of\,misclassifications}{total\,number\,of\,subjects} \times 100 \quad (3)$$

$$Sensitivity = \frac{true\,positives}{true\,positives + false\,negatives} \times 100 \quad (4)$$

$$Specificity = \frac{true\,negatives}{true\,negatives + false\,positives} \times 100 \quad (5)$$

$$Positive\,predictivity = \frac{true\,positives}{true\,positives + false\,positives} \times 100 \quad (6)$$

Results

Statistical analysis on raw connectivity matrix

The data processing and classification procedures used in this study are shown in Fig. 1. After pre-processing of the rs-fMRI data, the 264 putative functional areas were employed to parcellate the brain (Fig. 2). Signals of all voxels within each region were then averaged to produce 264 signals for each subject. Pearson's correlation was used to construct a 264-by-264 connectivity matrix of the brain network graph. Using the raw connectivity matrices of all subjects, we performed the NBS analysis based on 10,000 random permutations to find disrupted connectivity patterns in patients with AD and MCI. A range of primary threshold values (F-threshold) were examined. At low F-thresholds there were significant networks with a lot of connections and at high F-thresholds there were no significant network. By using a primary threshold value (F-threshold) of 10, a single network of 15 connections (p=0.001, Bonferroni corrected) was revealed to show disrupted functional connectivity pattern between groups (Fig. 3). By increasing the primary threshold (F-threshold) to 11, the network split into two independently connected sub-networks of 3 (p=0.036, corrected) and 6 (p=0.004, corrected) connections (Fig. 4). These results remained unchanged when we performed analysis using the false discovery rate (FDR) method (an alternative method to correct for multiple comparisons) in 10,000 permutations.

Statistical analysis on graph measures

There was no significant difference in values of global efficiency (F<1.483, p>0.230, the lowest p was at PSW=0.5), characteristic path length (F<1.483, p>0.230, the lowest p was at PSW=0.5), clustering coefficient (F<0.322, p>0.726, the lowest p was at PSW=0.5), and assortativity (F<2.427, p>0.092, the lowest p was at PSW=0.18) between groups. Post hoc analysis of assortativity between HC and AD groups using two-sample t-test, revealed significant differences at some PSW threshold values: t=0.043, p=0.049 at PSW=0.14; t=0.047, p=0.042 at PSW=0.16; t=0.048, p=0.043 at PSW=0.17; and t=0.048 , p=0.041 at PSW=0.18.

All subjects showed an organization of small-world functional network. However, the value of small-world index was significantly different between groups over the PSW threshold values from 0.16 to 0.25. Figure 5 shows variation of the F-values and the p-values by changing the threshold PSW. Figure 6 shows the group mean and standard error values for the small-world index.

Local network measures have 264 values for each subject (one measure for each node). We identified regions (nodes) with corresponding local measures which were significantly different between three groups (i.e., HC, MCI, and AD) in

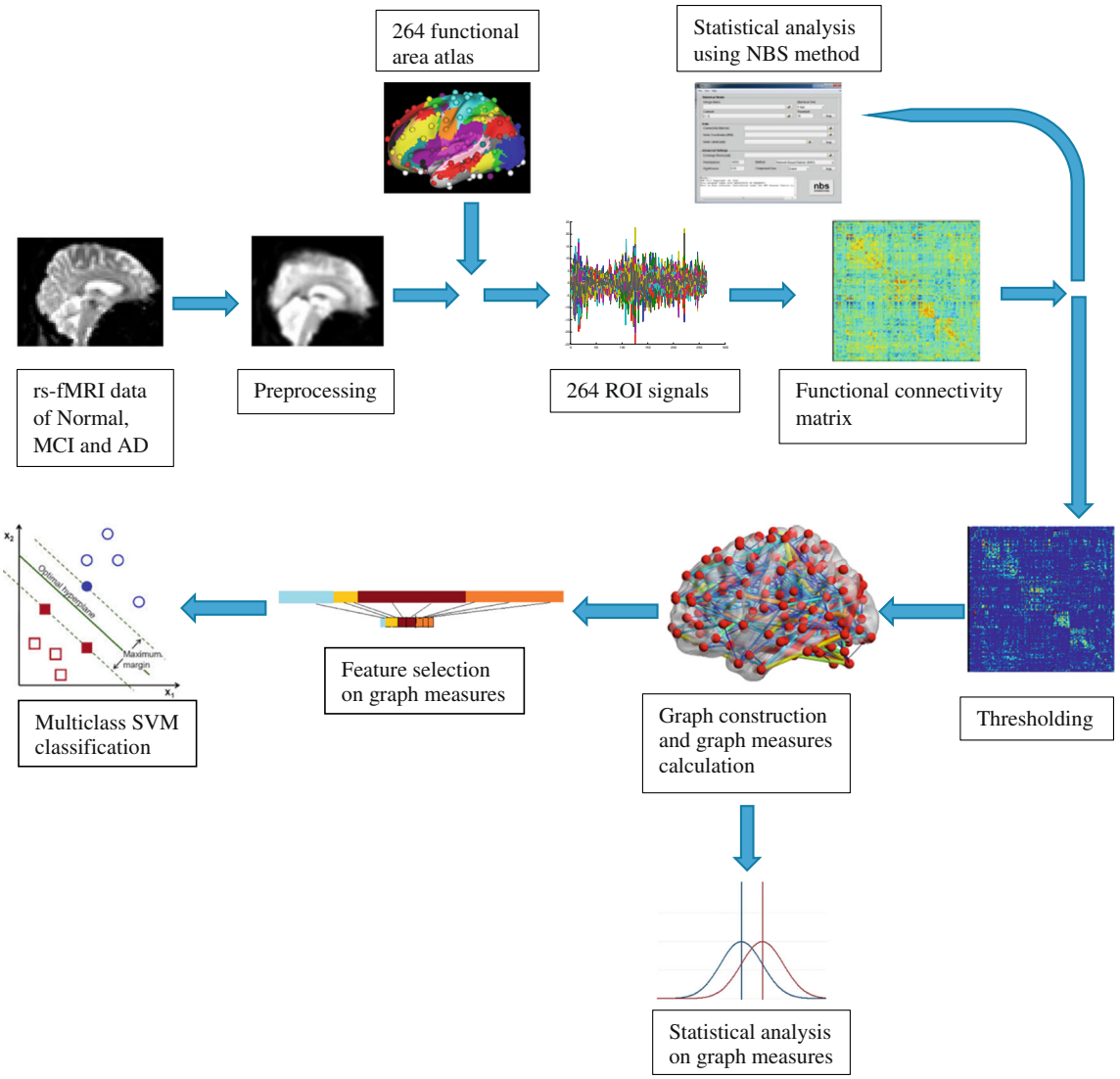

Fig. 1 Procedure of the proposed method in this study

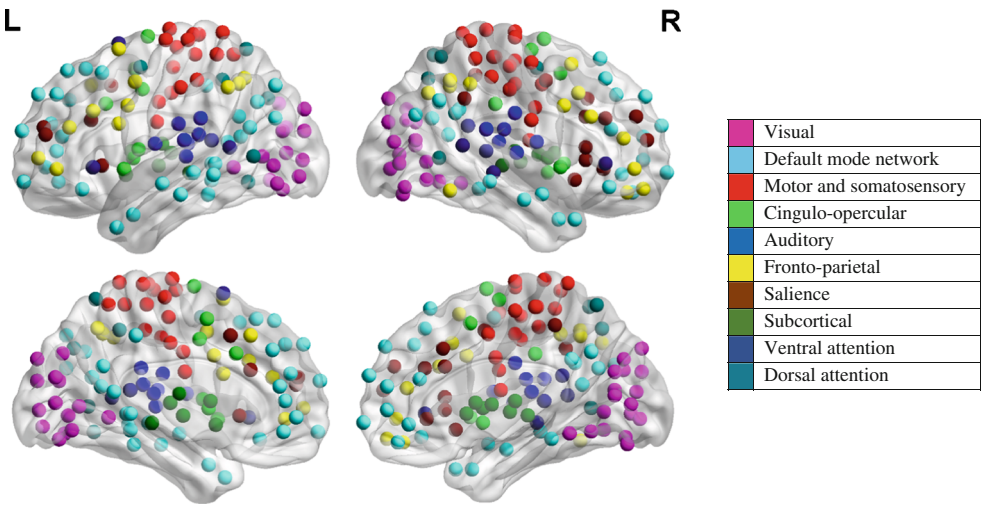

Fig. 2 Nodes of the graph defined by 264 putative functional areas (Power et al. 2011). Ten major resting-state networks are indicated on the right. Plots of this figure were created using the BrainNet Viewer software package (http://nitrc.org/projects/bnv/)

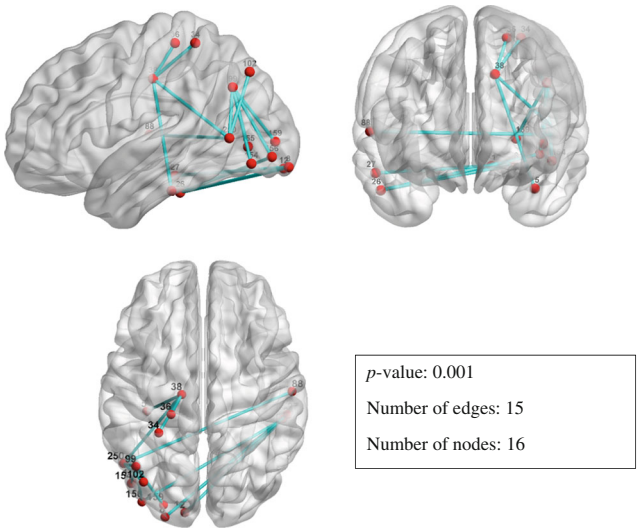

Fig. 3 Significant differences in the inter-regional connections (edges) between three groups (HC, MCI, and AD) at threshold value (*F*-threshold) of 10. The statistical analysis using the network based statistics (NBS) method identified one significant altered network in patients with AD and MCI

more than 20 threshold PSW values (over the threshold range of 0.1 to 0.5, step of 0.01). Those regions are listed in Table 2. It is notable that there was no region with either the subgraph centrality, K-coreness centrality, or participation coefficient values significant in more than 20 threshold PSW values.

Application of machine learning approaches to graph measures

After the statistical analysis, we investigated whether graph measures can differentiate patients with AD and MCI from

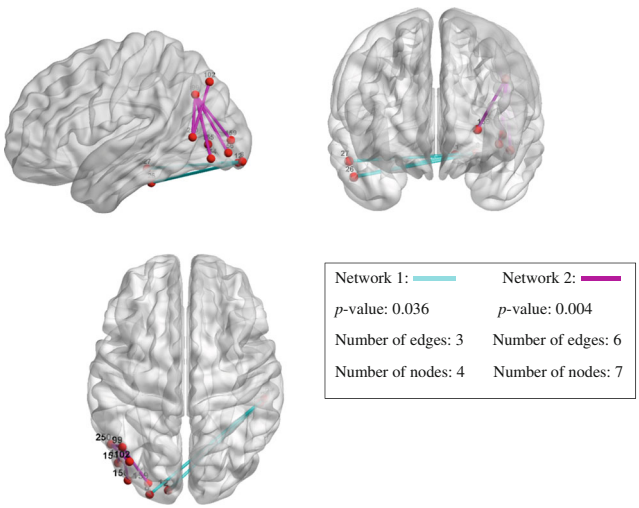

Fig. 4 Significant differences in the inter-regional connections (edges) between three groups (HC, MCI, and AD) at threshold value (*F*-threshold) of 11. The statistical analysis using the network based statistics (NBS) identified two significant altered network in patients with AD and MCI

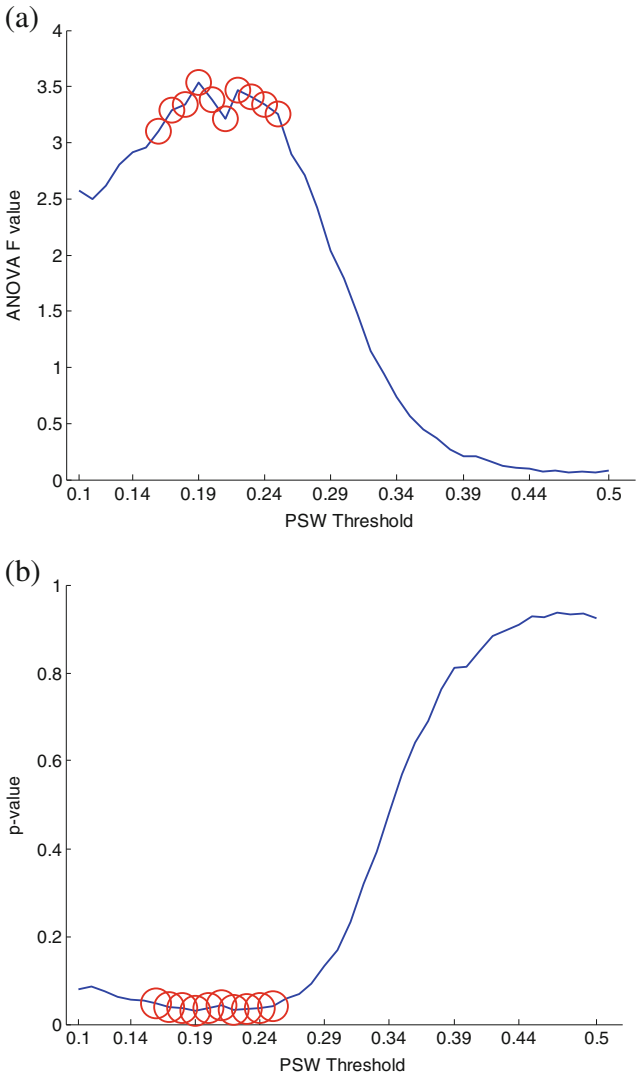

Fig. 5 The *F*-statistic (**a**) and *p*-value (**b**) of comparing between group differences (between three groups HC, MCI, and AD) in mean value of small-world index over different PSW thresholds. Red circles show significant values corresponding to $p<0.05$

healthy subjects. To this end, we used the proposed method in Section Methods–Results and the graph-based thresholding method was applied to the average fully connected graphs of HC subjects and patients with AD and MCI. The value of *GCE* for different PSW proportional thresholds (costs) was positive and was higher in low cost PSW thresholds because economical small-world properties were generally most salient in conservatively thresholded networks. As an example, Fig. 7 shows a plot of global cost efficiency as a function of cost for the group of HC subjects. Threshold values of 0.19 for HC and MCI groups (Fig. 7) and 0.18 for patients with AD were obtained. To ensure all subjects had the same number of edges after thresholding, we thresholded the fully connected graphs of all subjects at a PSW threshold value of 0.19 (i.e., preserving 19 % of the strongest weights). The graph measures were then calculated for sparse brain networks. The

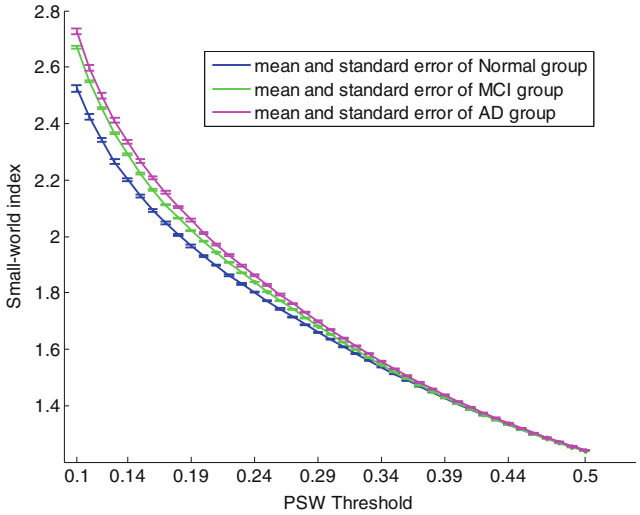

Fig. 6 Measure of small-worldness of brain network with respect to the PSW threshold for three groups

calculated measures were used as discriminative features to classify three classes: HC subjects, patients with MCI, and patients with AD. The number of graph nodes in this study was 264 and thus we calculated 264 features for each of the following measures: degree, node strength, participation coefficient, betweenness centrality, K-coreness centrality, subgraph centrality, eigenvector centrality, PageRank centrality, diversity coefficient, local efficiency, and ratio of local to global efficiency. In addition, we calculated one feature for the following global measures: average characteristic path length, average clustering coefficient, global efficiency, assortativity, and small-worldness. These features were combined to construct the final feature vector for each subject. The size of final feature vector was 2909 (11×264=2904 local and 5 global features). Performing the classification using this high dimensional feature space is time consuming and usually results in poor performance due to the existence of redundant and irrelevant features. In the current study, feature selection was performed on the extracted 2909 features to select the most discriminative features. Filter feature selection algorithms were used as a pre-processing step since they are simple and fast. We found that the Fisher algorithm provided the best results in the filtering stage.

The filter feature selection algorithm (i.e., Fisher algorithm in this study) sorts the features based on their discrimination ability. The best feature is on the top of the list with maximum discrimination. We first trained and tested the classifier (SVM with the radius basis function (RBF) kernel) using the best feature and obtained the corresponding classification accuracy. Then we trained and tested the classifier using the best two features and obtained the classification accuracy. Number of used features increased in each stage of the experiment until all features used to train and test the classifier. As shown in Fig. 8, the maximum achieved accuracy is less than 70 % which is not high enough. Therefore, we added the FSFS wrapper stage after the filter stage. The FSFS was applied on the first half of the selected features from the filter stage to find the final features. Since the typical goal of classification is to maximize the accuracy, the FSFS feature selection procedure performs a sequential search using the accuracy of the learning algorithm on each candidate feature subset. The training set was used to select the features and to fit the model, and the test set was used to evaluate the performance of the final selected features.

Comparison against other learning algorithms

Several learning algorithms (classifiers) have been used in neuroscience literature. To find the best fit for the present study, we examined some of these algorithms and the results were summarized in Tables 3 and 4. Results in Table 3 were obtained by utilizing the same algorithm used in the Table 4, except that the feature selection stage in Table 3 was ignored and classification was performed using full set of features. As shown by the results in Table 4, the SVM with the RBF kernel achieved the best performance in terms of the classification accuracy, sensitivity, specificity, and positive predictivity. Therefore, we used SVM-RBF learning algorithm in the subsequent analyses.

Performing classification using individual graph measures

As explained in Section Methods–Discussion, various graph measures were extracted and used as discriminative features. To show the efficiency of individual measures in discriminating classes, we performed classification of three groups (i.e., HC, MCI and AD) using each individual graph measure. Local graph measures have 264 feature values (each value for one node in 264 putative functional areas) and the global graph measures have only one feature value. The accuracies of classifiers corresponding to individual features and also all features combined are depicted in Table 5. The classification accuracies of the local graph measures were more than that of global measures. While the maximum classification accuracy of the global measures was less than 54 %, the local PageRank centrality measure was able to classify three groups with an accuracy of 80 %. Moreover, the maximum classification accuracy of 88.42 % was obtained when an optimal 44 features were used.

Another strategy for finding the optimal threshold value of the best classification performance

To find the appropriate threshold value, we performed a graph theory-based analysis on the mean connectivity

Table 2 A subset of 264 brain regions which have local graph measures that significantly differ between three groups (HC, MCI, and AD). For each column, regions in which the corresponding local graph measure showed significant differences in three groups of more than 20 PSW threshold values (from 0.1 to 0.5 in step of 0.01) were identified. The right seven columns are corresponding to graph measures. The numbers in these columns represent the lowest *p*-value of the corresponding region across thresholds and the numbers in parenthesis represent the PSW threshold in which the lowest *p*-value occurred

Area	MNI coordinates	AAL	RSN	EC	LE	Deg	BC	Div	PC	S
16	[17–91 −14]	Lingual_R	U		0.01280 (0.28)					
23	[35–67 −34]	Cerebelum_Crus1_R	U	0.00040 (0.18)		0.02472 (0.34)				
26	[52–34 −27]	Temporal_Inf_R	U	0.00230 (0.11)		0.00114 (0.12)		0.00064 (0.49)	0.00114 (0.12)	0.00113 (0.12)
27	[55–31 −17]	Temporal_Inf_R	U	0.00416 (0.11)	0.02067 (0.46)	0.00635 (0.3)			0.00798 (0.3)	0.00878 (0.3)
30	[−45 −32 47]	Postcentral_L	SMH	0.00046 (0.1)		0.00031 (0.24)			0.00028 (0.24)	0.00447 (0.24)
34	[−29 −43 61]	Postcentral_L	SMH	0.00048 (0.14)	0.00136 (0.43)	0.00085 (0.14)	0.00056 (0.43)	0.00211 (0.23)	0.00224 (0.14)	0.00366 (0.14)
42	[−7 −33 72]	Paracentral_Lobule_L	SMH			0.00631 (0.31)			0.00547 (0.31)	0.03552 (0.31)
50	[20–29 60]	–	SMH				0.00947 (0.35)			
53	[29–39 59]	Postcentral_R	SMH				0.00903 (0.49)			
63	[66–8 25]	Postcentral_R	SMM							0.02540 (0.1)
110	[−13 −40 1]	Lingual_L	DMN	0.00014 (0.42)		0.00114 (0.35)			0.00184 (0.46)	0.03266 (0.35)
131	[11–54 17]	Precuneus_R	DMN					0.00350 (0.38)		
138	[27–37 −13]	Fusiform_R	DMN	0.01257 (0.22)						
146	[52 7–30]	Temporal_Pole_Mid_R	DMN				0.00372 (0.31)			
149	[−7 −71 42]	Precuneus_L	MR		0.01097 (0.41)			0.00172 (0.32)		
151	[2–24 30]	–	MR	0.01466 (0.37)						
153	[11–66 42]	Precuneus_R	MR			0.01353 (0.39)			0.00773 (0.39)	0.02067 (0.38)
158	[−28 −79 19]	Occipital_Mid_L	V	0.01200 (0.42)		0.01197 (0.42)				
203	[40 18 40]	Frontal_Mid_R	FPTC	0.00317 (0.14)						
220	[31 33 26]	Frontal_Mid_R	SA	0.00375 (0.1)		0.00970 (0.11)			0.01081 (0.11)	0.00464 (0.11)
228	[−31 −11 0]	Putamen_L	SUB		0.00116 (0.22)					
230	[−15 4 8]	–	SUB	0.00231 (0.34)						
231	[−10 −18 7]	Thalamus_L	SUB	0.01255 (0.49)						
235	[9–4 6]	Thalamus_R	SUB	0.00514 (0.2)						
236	[12–17 8]	Thalamus_R	SUB	0.00038 (0.38)		0.00040 (0.5)			0.00036 (0.5)	
237	[15 5 7]	–	SUB	0.00659 (0.25)			0.00113 (0.32)			
240	[31–14 2]	Putamen_R	SUB				0.00207 (0.46)			
250	[−52 −63 5]	Temporal_Mid_L	DA	0.00621 (0.49)						
264	[22–58 −23]	Cerebelum_6_R	CER	0.01828 (0.31)						

AAL the automated anatomical labeling atlas, *BC* betweenness centrality, *CER* Cerebellum, *DA* Dorsal attention, *Deg* degree, *Div* diversity, *DMN* Default mode network, *EC* eigenvector centrality, *FPTC* Fronto-parietal Task Control, *MNI* Montreal Neurological Institute space, *MR* Memory retrieval, *LE* local efficiency, *PC* PageRank centrality, *RSN* Resting-state network, *S* strength, *SA* Salience, *SMH* Sensory/somatomotor Hand, *SMM* Sensory/somatomotor Mouth, *SUB* Subcortical, *U* Uncertain, *V* Visual

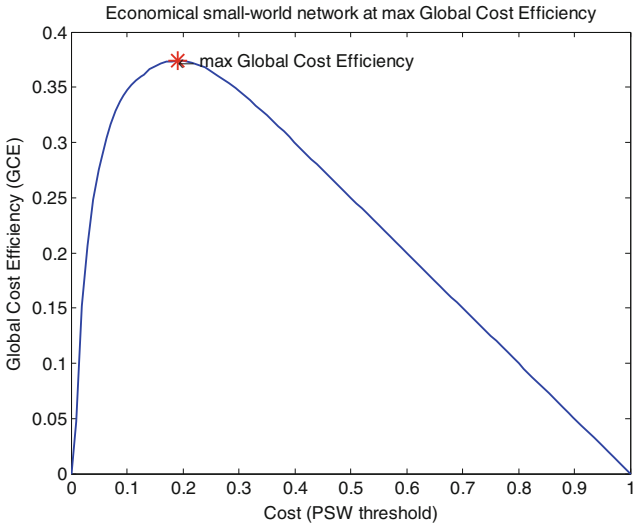

Fig. 7 Global cost efficiency as a function of cost (PSW threshold). Plot of global cost efficiency versus cost averaged across the entire set of HC subjects. The cost (or PSW threshold) is defined as the ratio of existing connections, after a given PSW threshold, divided by the total number of possible pairwise connections in a network. The global cost efficiency is defined as the global efficiency *E*, at a given cost PSW, minus the cost (i.e., *E* - *PSW*)

matrices of subjects and obtained the optimal PSW threshold value of 0.19, i.e., we preserved 19 % of the strongest weights in the connectivity matrix of any subject. All of the diagonal elements of the connectivity matrix were set to zero. To show the optimality of the selected threshold value, all stages including feature extraction based on the graph measures, feature selection, and classification were performed using different threshold values. Figure 9 shows the results. The highest recognition accuracy (88.42 %) was obtained at PSW threshold value of 0.19.

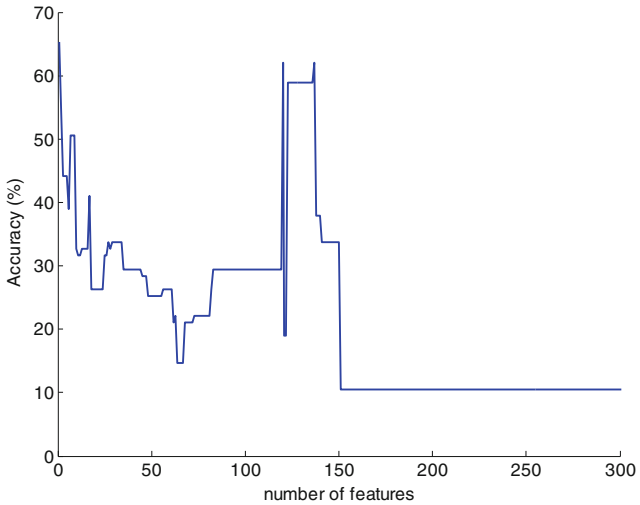

Fig. 8 Performance of the Fisher score feature selection algorithm in various number of selected features

Binary classification of three groups

We investigated performance of our proposed method for a binary classification of a group from two other groups (e.g., HC vs. MCI and AD). To this end, 23 subjects from each group (e.g., 23 healthy controls and 23 patients with AD or MCI for HC vs. AD and MCI classification) were used for training and remaining subjects (i.e., 164–46=118 subjects) were used for testing the classifier. The results are depicted in Table 6. The classification accuracies for distinguishing HC from AD&MCI, AD from HC&MCI, and MCI from HC&AD, were 87.29, 97.46, and 72.03 %, respectively.

Using different atlas to define graph nodes of the brain network

Most of the previous studies have used AAL template (with 90 nodes) to parcellate brain and define the nodes and corresponding graph. Performance of classification can be affected by using different parcellations of the brain. We compared performance of the classification using AAL template with that using the 264 putative functional areas. Accuracy of classification using the 264 putative functional areas was 87.29 %, which was improvement over the accuracy of classification using the AAL atlas, which was found to be 83.16 %. Comparison of the classification performance of two parcellation approaches are shown in Fig. 10.

Discussion

Comparison of the dense connectivity matrix of the brain network of three groups (HC, MCI, and AD) using the network based statistics (NBS) revealed significant altered networks in patients with MCI and AD (Figs. 3 and 4). At the lower threshold value for *F*-statistic, an altered network with 16 nodes and 15 edges was found (Fig. 3). These nodes correspond to the following AAL regions: left fusiform gyrus, left inferior occipital gyrus, left lingual gyrus, right inferior temporal gyrus, left postcentral gyrus, right superior temporal gyrus, left angular gyrus, left inferior parietal, supramarginal and angular gyri, left middle occipital gyrus, and left middle temporal gyrus. With reference to the modular structure in Fig. 2, the following observations can be inferred about the altered connections in three groups. Four connections were related to the altered connectivity between dorsal attention network and other three resting-state networks (RSNs) (i.e., sensory/somatomotor hand, auditory, and two regions in the default mode network (DMN)). Four altered connections were observed between the visual system and DMN. Two altered connections were related to the intra-connections in sensory/somatomotor hand. The remaining connections including two inter-regional and three intra-regional connections was related

Table 3 Comparison of classification performance of some common classifiers using full set of features (without feature selection) in identification of three groups of HC, MCI, and AD

Classifier	Accuracy (%)	Sensitivity (%)			Specificity (%)			Positive predictivity (%)		
		N	MCI	AD	N	MCI	AD	N	MCI	AD
SVM (RBF)	40	25	43.08	50	70.67	66.67	70.59	18.52	73.68	16.67
SVM (poly)	36.84	30	33.85	70	65.33	63.33	72.94	18.75	66.67	23.33
SVM (linear)	40	40	40	40	66.67	76.67	70.59	24.24	78.79	13.79
KNN	31.58	25	32.31	40	57.33	76.67	69.41	13.51	75	13.33
Naïve bayes	55.79	45	73.85	40	97.33	33.33	76.47	33.33	70.59	16.67
Fisher linear	34.74	45	29.23	50	64	80	65.88	25	76	14.71
Linear discriminant (ldc)	66.32	15	92.31	0	93.33	10	100	37.5	68.97	–
Quadratic classifier (qdc)	21.05	100	0	0	0	100	100	21.05	–	–
Decision tree	46.32	35	53.85	20	70.67	63.33	78.83	24.14	76.09	10

SVM Support vector machine, *RBF* radius basis function, *KNN* K-nearest neighbor

to the regions in the brain with unknown RSN (which are labeled uncertain in Table 2). By increasing the threshold of *F*-statistic, the network split in two independent sub-networks (Fig. 4). These results are in agreement with previous studies that reported reduction of the strength of connections between the temporal lobe and the parietal and occipital regions in AD patients (Sanz-Arigita et al. 2010). In addition, these results suggest alteration of memory, attentional, and default mode networks in patients with AD, as these networks are essential for successful memory performance (Celone et al. 2006; Daselaar et al. 2004; Grady et al. 2006; Miller et al. 2008; Pihlajamäki et al. 2008; Sperling et al. 2010; Sperling et al. 2009).

It has been shown that the brain functional network has the property of small-worldness (Bassett et al. 2008; Latora and Marchiori 2001). A loss of small-world characteristics in patients with AD has also been reported (He et al. 2008; Stam et al. 2007). As shown in Fig. 6, all three groups showed a small-world property in different PSW thresholds. By increasing the PSW threshold value, the small-worldness was reduced. Furthermore, a decrease in small-world value was observed from HC to MCI, and from MCI to AD groups in all PSW threshold values. Our results are in agreement with previous studies that reported reduction of the small-worldness in patients with AD (Zhao et al. 2012).

At low PSW threshold values, a significant difference in mean values of small-world index was observed between HC, MCI, and AD groups (Fig. 5). However, we did not find any significant difference in three groups in values of other global measures (i.e., clustering coefficient, characteristic path length, global efficiency, and assortativity). Some previous studies reported no significant difference in values of

Table 4 Comparison of classification performance of some common classifiers using full set of features (with feature selection) in identification of three groups of HC, MCI, and AD

Classifier	Accuracy (%)	Sensitivity (%)			Specificity (%)			Positive predictivity (%)		
		N	MCI	AD	N	MCI	AD	N	MCI	AD
SVM (RBF)	88.42	70	96.92	70	100	73.33	96.47	100	88.73	70
SVM (poly)	71.58	15	100	0	97.33	16.67	100	60	72.22	–
SVM (linear)	75.79	15	98.46	50	100	33.33	96.47	100	76.19	62.5
KNN	66.32	75	69.23	30	84	73.33	85.88	55	84.91	20
Naïve bayes	80	65	84.62	80	94.67	76.67	90.59	76.47	88.71	50
Fisher linear	75.79	85	75.38	60	82.67	86.67	92.94	56.67	92.45	50
Linear discriminant (ldc)	71.58	20	98.46	0	98.67	13.33	100	80	71.11	–
Quadratic classifier (qdc)	72.63	25	98.46	0	98.67	16.67	100	83.33	71.91	–
Decision tree	66.32	35	80	40	89.33	63.33	84.71	46.67	82.54	23.53

SVM Support vector machine, *RBF* radius basis function, *KNN* K-nearest neighbor

Table 5 Performance of classification of three groups (i.e., HC, MCI, and AD) using individual groups of features

Local measures (each group has 264 features)	No of selected features (of the original 264 features)	Accuracy (%)
Degree	39	76.84
Participation coefficient	64	77.89
Diversity coefficient	26	74.74
Betweenness centrality	24	76.84
Local efficiency	86	76.84
Local/Global efficiency	6	71.58
K-coreness centrality	41	78.95
Subgraph centrality	1	10.53
Eigenvector centrality	102	77.89
Pagerank centrality	63	80
Node strength	28	74.74
Global measures (each group has a single feature)		
Characteristic path length	–	53.68
Clustering coefficient	–	23.16
Global efficiency	–	33.68
Small-world index	–	37.89
Assortativity	–	18.95
All features together	44 (of the original 2909 features)	88.42

clustering coefficient (Lo et al. 2010; Sanz-Arigita et al. 2010; Stam et al. 2007) and global efficiency (Sanz-Arigita et al. 2010) between HC and AD groups. Table 2 lists regions showing significant difference of some local graph measures in three groups. Interestingly, those regions were located in sensory/somatomotor (6 regions), DMN (4 regions), memory retrieval (3 regions), dorsal attention (1 region), and visual (1 region) areas. These observations are in agreement with previous results reported that regions in DMN, attention, motor, and visual modules show impaired functioning in patients with AD (Bai et al. 2011; Wang et al. 2012; Wee et al. 2012b). Seven subcortical regions also showed significant alteration in some graph measures that is also in agreement with the reported results in (Bai et al. 2011; Wang et al. 2013). Local efficiency of left putamen and betweenness centrality of right putamen in this study have significant differences between groups. Some studies reported that putamen show gray matter atrophy (Madsen et al. 2010) and metabolic disruption (Koivunen et al. 2011) in patients with MCI. In addition, abnormal functional connectivity was observed in this region (Bai et al. 2011; Wang et al. 2013). The putamen is an important region in the neostriatum which has a key role in transmitting connections from the cortex to the basal ganglia nuclei and then projecting back to the cortex via thalamus (Postuma and Dagher 2006). The eigenvector centrality of thalamus (left and right) is another significantly different measure between three groups. Previous studies reported reduced connectivity between thalamus and other regions in AD (Celone et al. 2006; Wang et al. 2012).

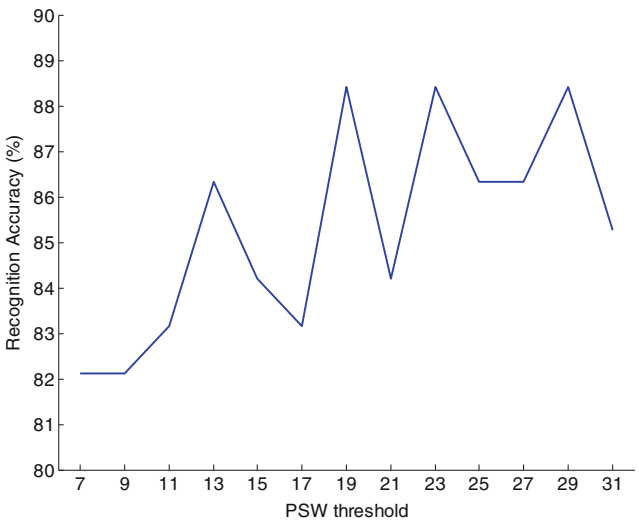

Fig. 9 Performance of classification in various PSW threshold values

MMSE test is a behavioral assessment of AD severity. We calculated the correlation coefficient between MMSE scores of patients with AD and each of the 2909 graph measures to assess the association between them. 272 out of 2909 features had significant ($p<0.05$) positive or negative correlation with the MMSE scores. Figure 11 shows the scatter plot of the features versus the MMSE scores for six of the most correlative features. Of 33 patients with AD, the MMSE scores of 8 patients were not included in the ADNI database and those patients were excluded in the correlation analysis.

Table 6 Performance of classification of HC from AD&MCI, AD from HC&MCI, and MCI from HC&AD using full set of features with feature selection

Classes	Accuracy	Sensitivity		Specificity		Positive predictivity	
		Class 1	Class 2	Class 1	Class 2	Class 1	Class 2
Normal vs. MCI&AD	87.29	35	97.96	97.96	35	77.78	88.07
AD vs. Normal&MCI	97.46	100	70	70	100	97.3	100
MCI vs. Normal&AD	72.03	84.91	61.54	61.54	84.91	64.29	83.33

We investigated whether graph theoretical measures derived from the brain network were able to accurately discriminate between patients with AD, patients with MCI, and HC subjects. Selecting an appropriate learning algorithm is critical for an accurate classification of three groups (i.e., AD, MCI, and HC). Various classifiers have been used in bioinformatics literature, including *k*-nearest neighbor, naïve Bayes, Fisher linear, linear discriminant, quadratic classifier, and decision tree, as well as SVM with linear, polynomial, and RBF kernels. We examined these classifiers to find the best one for our study. As presented in Table 4, SVM with RBF kernel outperformed other classifiers in terms of accuracy, sensitivity, specificity, and positive predictivity. The SVM-RBF classifier achieved the accuracy of 88.42 % for classification of three groups of HC, MCI, and AD. This accuracy was achieved using only 44 optimal features of the original 2909 features. These optimal features were selected by the feature selection stage. It is noteworthy that, of 44 selected optimal features, 15 features were related to nodes in the default mode network.

This result is consistent with the previously reported results that DMN is affected during the process of AD. The consistently reported parts of DMN are posterior cingulate cortex, precuneus, medial prefrontal cortex, inferior parietal lobules, lateral temporal cortices, and hippocampus (Buckner et al. 2008; Raichle et al. 2001). These regions are among the earliest to show abnormal amyloid deposition (Mintun et al. 2006) and, therefore, they should have an important role in AD and MCI. Of 15 features that were related to the DMN in the current study, two features were related to the posterior parts of DMN, i.e., posterior cingulate cortex and precuneus, and three features were related to the anterior parts of DMN, i.e., ventral medial prefrontal cortex and dorsal medial prefrontal cortex. Since these features were related to the centrality measure, these results may be suggestive of a loss of hubs (Buckner et al. 2009; Drzezga et al. 2011; Tijms et al. 2013). In addition, these findings support the anterior-posterior disconnection syndrome in patients with MCI and AD (Bai et al. 2008; Koch et al. 2012; Sheline and Raichle 2013; Sperling et al. 2010; Wee et al. 2012b) and show decreased functional connectivity from the posterior to anterior portions of the DMN.

Moreover, most of the 44 selected features were measures of PageRank centrality of the graph. This result was confirmed in Table 5, where each group of graph measures was used individually as features and classification was performed utilizing these features. It should be noted that, all of the classification steps including feature selection were performed on individual groups. It can be seen from Table 5 that, if we include only PageRank centrality measures as discriminating features and exclude all of the other features, the accuracy was 80 % (the highest accuracy in feature groups). These observations indicate importance of the PageRank centrality measure in discriminating HC, MCI and AD. The PageRank centrality algorithm has been used by Google and favors nodes that connect to other highly central nodes (van den Heuvel and Sporns 2013). Nodes with large values of the PageRank centrality are network hubs that makes strong contribution to global network function. The importance of the PageRank centrality in identification of patients with AD and MCI in our results may indicate that these patients have a problem in connectivity of their main brain hubs. Our observation is in agreement with a recent study which reported that the eigenvector centrality, which is a variant of the PageRank centrality, was an efficient graph measure in identification of AD (Binnewijzend et al. 2014).

Setting the threshold value to convert fully connected brain network to sparse network is critical for accurate classification of three groups. A small threshold value eliminates more graph edges and may result in missing some informative edges. On the other hand, a large threshold value incorporates weaker edges and may introduce noise in the graph. We found an optimal PSW threshold value of 0.19 (corresponding to retain only 19 % of strong connections) using a graph-based method. To confirm optimality of the calculated threshold value, we performed classification using proportional threshold values in the range from 0.07 to 0.31, and found an optimal value of 0.19 for this parameter.

Figure 10 compares classification performance of two parcellation approaches (i.e., AAL template and 264 putative functional areas) for defining nodes of the graph, in terms of sensitivity, specificity, and positive predictivity. The accuracies of classification using the AAL atlas and the 264 putative functional areas atlas were 83.16 and 88.42 %, respectively. In addition to the reduction of the accuracy, other performance metrics were degraded when using AAL atlas compared to the 264 putative functional areas atlas. In particular, sensitivity of AD detection (Fig. 10a) using AAL atlas was significantly smaller than that using 264 regions. Given that a large

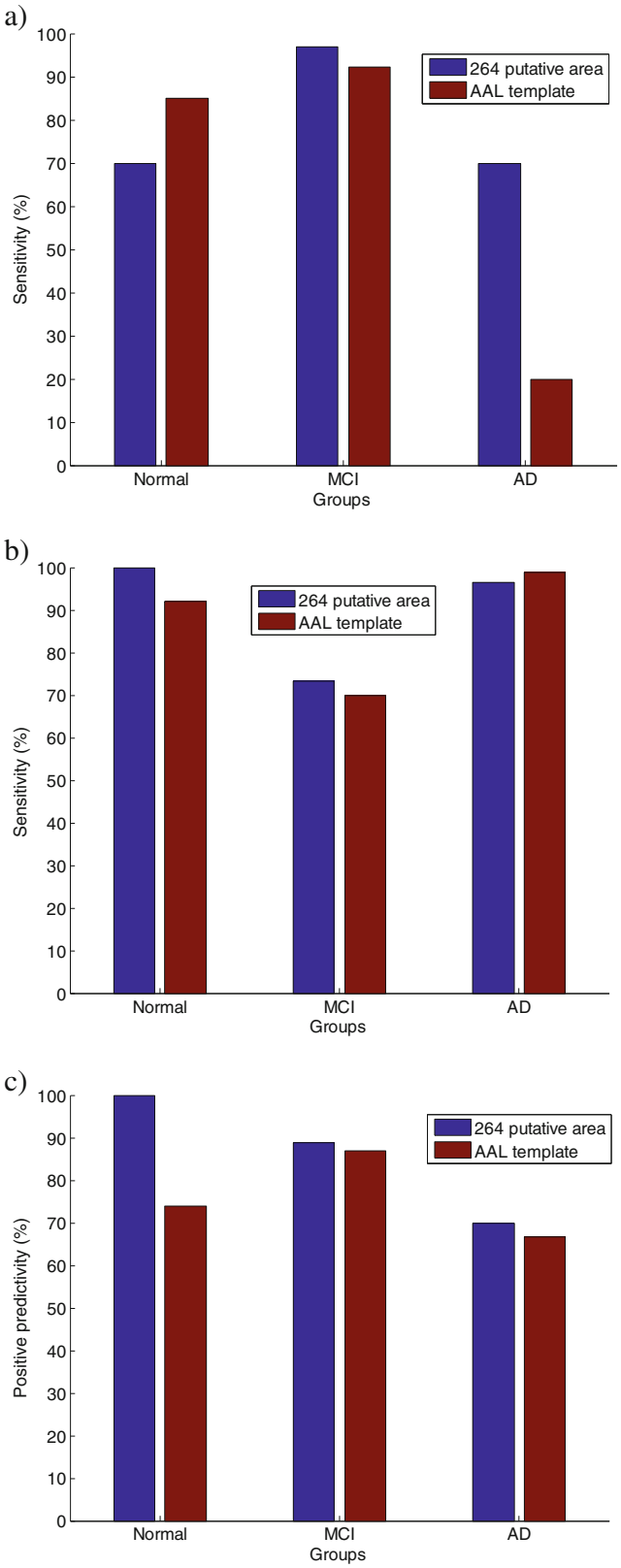

Fig. 10 Comparison of performance of classification using AAL atlas and 264 functional areas atals, in terms of sensitivity, specificity, and positive predictivity

sensitivity of an automatic diagnostic method is crucial for its reliability, parcellation using 264 regions will be more appropriate in likely real clinical applications. On the other hand, positive predictivity of the healthy control group using AAL atlas was degraded significantly (Fig. 10c) which means more AD and MCI patients were incorrectly classified as healthy subjects. Classifying patients as healthy subject is a crucial limitation of an automatic diagnostic algorithm in clinical applications. Degradation in performance metrics may originate from the fact that AAL atlas parcels the brain based on the anatomical trait, which may not correspond to the functional brain organization. In contrast, the 264 putative functional areas are based on functional brain network and, thus, may be more sensitive to the brain functional organization. Previous studies have also shown that functionally defined ROIs better suited to fMRI data than the structurally defined ROIs and provide better classification accuracy (Shirer et al. 2012). Beyond the nature of the ROIs, number of them can also affect classification accuracy. However, there is not a common finding about the optimal number of ROIs to achieve maximum classification accuracy. Shirer et al. compared classification of 4 cognitive states using 90 functionally defined ROIs and 112 structurally defined ROIs from AAL atlas. They reported that the functional ROIs consistently and significantly outperformed the structural ROIs. They strongly suggested that the functionally defined ROIs provides superior results compared to the structurally defined ROIs in whole-brain connectivity analyses. The authors in (Ota et al. 2014) compared performance of MCI prediction using three different atlases, i.e., 116 ROIs from AAL atlas, 82 ROIs from Brodmann's Areas (BA) and 56 ROIs from the LONI Probabilistic Brain Atlas (LPBA40). The overall performance of classification using LPBA40 was better than that of AAL and BA. The authors in (Mesrob et al. 2008) used the 90 ROIs from AAL atlas and refined 487 ROIs from the original 90 AAL regions for identification of AD. Despite the larger number of ROIs in the second set of ROIs, the 90 ROIs from AAL atlas achieved better performance.

Overall, the proposed algorithm in this study, using the rs-fMRI data and graph theory, provided promising results for classification of HC, MCI and AD. Recently, some authors suggested to use a multi-modality approach for binary classification of HC, MCI and AD (Cheng et al. 2015; Suk et al. 2015a, b). They also developed their method for binary classification of MCI converters (MCI-C) from MCI non-converters (MCI-NC). The authors in (Cheng et al. 2015) obtained an accuracy of 79.4 % in classification of MCI-C patients from MCI-NC patients using a novel domain transfer learning method. Although classifying MCI-C and MCI-NC is important in diagnostic of AD, but it is beyond the scope of our current study and we plan to perform this classification in future.

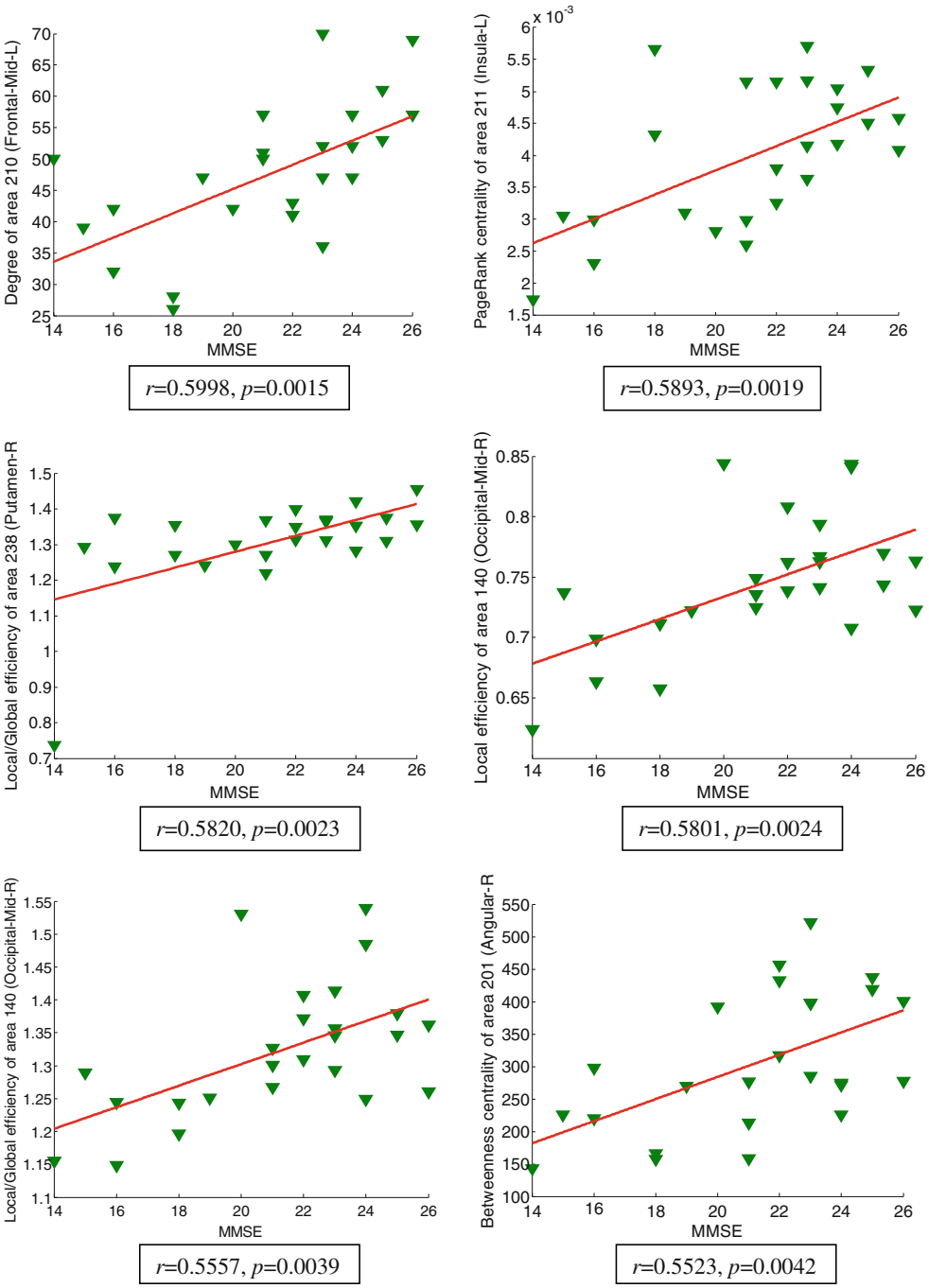

Fig. 11 The scatter plot of features (i.e., graph measures) versus the MMSE scores for six of the most correlative features. The red line shows the linear regression. The correlation (r) and p-value (p) for each feature were added to the corresponding subplot

The presented method in this study, is an improved version of previous method in (Khazaee et al. 2015). There are fundamental differences between two methods. The previous study was differentiated HC from AD based on 40 subjects, whereas we differentiated three groups i.e., HC, MCI, and AD based on 168 subjects. This study has more statistical analysis than the previous study. In the present study, extensive statistical analysis was performed on raw connectivity matrices and on graph measures. Another main advantage of present study is that we have used a functional atlas i.e., 264 putative functional area instead of AAL structural atlas. In the previous study, only filter feature selection algorithms were employed which provided poor results in this study (Fig. 8). An improved feature selection was used in this study by using both filter and wrapper feature selection algorithms. Fisher algorithm in combination with FSFS provided the best performance. In addition, features that were used to discriminate classes, were different than the previous study. Here, we used

binary graph measures instead of weighted ones in the previous study. In addition, we extended the types of graph measures and included more graph measures.

Conclusions

We applied the graph theoretical approach and the pattern recognition method on the resting state fMRI data to classify three groups: HC, patients with AD, and patients with MCI. The proposed method identified specific sets of the brain graph measures that are informative to distinguish three groups (i.e., HC, MCI and AD). Informative graph measures were related to the brain cortical regions using edges and nodes and provided information about disrupted brain functional nodes. When applied to functional data from groups of HC subjects and MCI and AD patients, we were able to identify AD and MCI induced causes to the brain network. The proposed method in this study classified patients with AD, MCI, and healthy control subjects with a high accuracy.

Acknowledgments Data used in this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://ADNI.loni. usc.edu). The investigators within the ADNI, who can be found at http://ADNI.loni.usc.edu/study-design/ongoing-investigations, contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this article. This study was supported by the Children's Foundation Research Institute, Le Bonheur Children's Hospital, Memphis, TN.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

We thank Dr. Amanda Preston for her assistance with manuscript preparation.

Compliance with Ethical Standards

Funding This study was funded by the Children's Foundation Research Institute, Le Bonheur Children's Hospital, Memphis, TN.

Conflict of Interest The authors declared that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Anderson, A., & Cohen, M.S. (2013). Decreased small-world functional network connectivity and clustering across resting state networks in schizophrenia: an fMRI classification tutorial. *Frontiers in Human Neuroscience, 7*.

Bai, F., Zhang, Z., Yu, H., Shi, Y., Yuan, Y., Zhu, W., & Qian, Y. (2008). Default-mode network activity distinguishes amnestic type mild cognitive impairment from healthy aging: a combined structural and resting-state functional MRI study. *Neuroscience Letters, 438*(1), 111–115.

Bai, F., Liao, W., Watson, D. R., Shi, Y., Wang, Y., Yue, C., & Jia, J. (2011). Abnormal whole-brain functional connection in amnestic mild cognitive impairment patients. *Behavioural Brain Research, 216*(2), 666–672.

Bassett, D. S., Bullmore, E., Verchinski, B. A., Mattay, V. S., Weinberger, D. R., & Meyer-Lindenberg, A. (2008). Hierarchical organization of human cortical networks in health and schizophrenia. *The Journal of Neuroscience, 28*(37), 9239–9248. doi:10.1523/JNEUROSCI.1929-08.2008.

Bassett, D. S., Bullmore, E. T., Meyer-Lindenberg, A., Apud, J. A., Weinberger, D. R., & Coppola, R. (2009). Cognitive fitness of cost-efficient brain functional networks. *Proceedings of the National Academy of Sciences of the United States of America, 106*(28), 11747–11752. doi:10.1073/pnas.0903641106.

Binnewijzend, M. A., Adriaanse, S. M., Van der Flier, W. M., Teunissen, C. E., de Munck, J. C., Stam, C. J., & Wink, A. M. (2014). Brain network alterations in Alzheimer's disease measured by eigenvector centrality in fMRI are related to cognition and CSF biomarkers. *Human Brain Mapping, 35*(5), 2383–2393. doi:10.1002/hbm.22335.

Boldi, P., Santini, M., & Vigna, S. (2009). PageRank: functional dependencies. *ACM Transactions on Information Systems (TOIS), 27*(4), 19.

Brandes, U. (2001). A faster algorithm for betweenness centrality*. *Journal of Mathematical Sociology, 25*(2), 163–177.

Brier, M. R., Thomas, J. B., Fagan, A. M., Hassenstab, J., Holtzman, D. M., Benzinger, T. L., & Ances, B. M. (2014). Functional connectivity and graph theory in preclinical Alzheimer's disease. *Neurobiology of Aging, 35*(4), 757–768.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences, 1124*(1), 1–38.

Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., & Johnson, K. A. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *The Journal of Neuroscience, 29*(6), 1860–1873.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience, 10*(3), 186–198.

Celone, K. A., Calhoun, V. D., Dickerson, B. C., Atri, A., Chua, E. F., Miller, S. L., & Blacker, D. (2006). Alterations in memory networks in mild cognitive impairment and Alzheimer's disease: an independent component analysis. *The Journal of Neuroscience, 26*(40), 10222–10231.

Challis, E., Hurley, P., Serra, L., Bozzali, M., Oliver, S., & Cercignani, M. (2015). Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. *NeuroImage, 112*, 232–243.

Chao-Gan, Y., & Yu-Feng, Z. (2010). DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Frontiers in Systems Neuroscience, 4*, 13. doi:10.3389/fnsys.2010.00013.

Cheng, B., Liu, M., Zhang, D., Munsell, B. C., & Shen, D. (2015). Domain transfer learning for MCI conversion prediction. *Biomedical Engineering, IEEE Transactions on, 62*(7), 1805–1817. doi:10.1109/TBME.2015.2404809.

Cohen, A. L., Fair, D. A., Dosenbach, N. U., Miezin, F. M., Dierker, D., Van Essen, D. C., & Petersen, S. E. (2008). Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage, 41*(1), 45–57.

Daselaar, S., Prince, S., & Cabeza, R. (2004). When less means more: deactivations during encoding that predict subsequent memory. *NeuroImage, 23*(3), 921–927.

Davatzikos, C., Bhatt, P., Shaw, L. M., Batmanghelich, K. N., & Trojanowski, J. Q. (2011). Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiology of Aging, 32*(12), 2322 e2319–2327. doi:10.1016/j.neurobiolaging.2010.05.023.

Dey, S., Rao, A.R., & Shah, M. (2012). Exploiting the brain's network structure in identifying ADHD subjects. *Frontiers in Systems Neuroscience, 6*.

dos Santos Siqueira, A., Biazoli Junior, C.E., Comfort, W.E., Rohde, L.A., & Sato, J.R. (2014). Abnormal functional resting-state networks in ADHD: graph theory and pattern recognition analysis of fMRI Data. *BioMed Research International, 2014*.

Drzezga, A., Becker, J. A., Van Dijk, K. R., Sreenivasan, A., Talukdar, T., Sullivan, C., & Greve, D. (2011). Neuronal dysfunction and disconnection of cortical hubs in non-demented subjects with elevated amyloid burden. *Brain, 134*(6), 1635–1646.

Duda, R.O., Hart, P.E., & Stork, D.G. (2012). *Pattern classification*. Wiley.

Estrada, E., & Higham, D. J. (2010). Network properties revealed through matrix functions. *SIAM Review, 52*(4), 696–714.

Fekete, T., Wilf, M., Rubin, D., Edelman, S., Malach, R., & Mujica-Parodi, L. R. (2013). Combining classification with fMRI-derived complex network measures for potential neurodiagnostics. *PloS One, 8*(5), e62867.

Foster, J. G., Foster, D. V., Grassberger, P., & Paczuski, M. (2010). Edge direction and the structure of networks. *Proceedings of the National Academy of Sciences, 107*(24), 10815–10820.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America, 102*(27), 9673–9678.

Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping, 26*(1), 15–29.

Friston, K. J., Frith, C. D., Frackowiak, R. S., & Turner, R. (1995). Characterizing dynamic brain responses with fMRI: a multivariate approach. *NeuroImage, 2*(2PA), 166–172.

Grady, C., Springer, M., Hongwanishkul, D., McIntosh, A., & Winocur, G. (2006). Age-related changes in brain activity across the adult lifespan. *Journal of Cognitive Neuroscience, 18*(2), 227–241.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences, 100*(1), 253–258.

Guimera, R., Sales-Pardo, M., & Amaral, L. A. (2007). Classes of complex networks defined by role-to-role connectivity profiles. *Nature Physics, 3*(1), 63–69.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology, 6*(7), e159.

He, Y., Chen, Z., & Evans, A. (2008). Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. *The Journal of Neuroscience, 28*(18), 4756–4766.

Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PloS One, 3*(4), e0002051.

Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., & Ward, C. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging, 27*(4), 685–691.

Jie, B., Zhang, D., Suk, H.-I., Wee, C.-Y., & Shen, D. (2013). Integrating Multiple Network Properties for MCI Identification. In G. Wu, D. Zhang, D. Shen, P. Yan, K. Suzuki & F. Wang (Eds.), *Machine Learning in Medical Imaging* (Vol. 8184, pp. 9–16). Springer International Publishing.

Jie, B., Zhang, D., Wee, C. Y., & Shen, D. (2014). Topological graph kernel on multiple thresholded functional connectivity networks for mild cognitive impairment classification. *Human Brain Mapping, 35*(7), 2876–2897. doi:10.1002/hbm.22353.

Kelly, A., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2008). Competition between functional brain networks mediates behavioral variability. *NeuroImage, 39*(1), 527–537.

Khazaee, A., Ebrahimzadeh, A., & Babajani-Feremi, A. (2015). Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory. *Clinical Neurophysiology*. doi:10.1016/j.clinph.2015.02.060.

Koch, W., Teipel, S., Mueller, S., Benninghoff, J., Wagner, M., Bokde, A. L., & Meindl, T. (2012). Diagnostic power of default mode network resting state fMRI in the detection of Alzheimer's disease. *Neurobiology of Aging, 33*(3), 466–478.

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence, 97*(1), 273–324.

Koivunen, J., Scheinin, N., Virta, J., Aalto, S., Vahlberg, T., Någren, K., & Rinne, J. (2011). Amyloid PET imaging in patients with mild cognitive impairment A 2-year follow-up study. *Neurology, 76*(12), 1085–1090.

Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters, 87*(19), 198701.

Li, Y., Qin, Y., Chen, X., & Li, W. (2013). Exploring the functional brain network of Alzheimer's disease: based on the computational experiment. *PloS One, 8*(9), e73186. doi:10.1371/journal.pone.0073186.

Lo, C.-Y., Wang, P.-N., Chou, K.-H., Wang, J., He, Y., & Lin, C.-P. (2010). Diffusion tensor tractography reveals abnormal topological organization in structural cortical networks in Alzheimer's disease. *The Journal of Neuroscience, 30*(50), 16876–16885.

Madsen, S. K., Ho, A. J., Hua, X., Saharan, P. S., Toga, A. W., Jack, C. R., Jr., & Initiative, A. s. D. N. (2010). 3D maps localize caudate nucleus atrophy in 400 Alzheimer's disease, mild cognitive impairment, and healthy elderly subjects. *Neurobiology of Aging, 31*(8), 1312–1325.

Mccarthy, P., Benuskova, L., & Franz, E.A. (2014). The age-related posterior-anterior shift as revealed by voxelwise analysis of functional brain networks. *Frontiers in Aging Neuroscience, 6*. doi: 10.3389/fnagi.2014.00301.

Mesrob, L., Magnin, B., Colliot, O., Sarazin, M., Hahn-Barma, V., Dubois, B., & Benali, H. (2008). Identification of atrophy patterns in Alzheimer's disease based on SVM feature selection and anatomical parcellation. In T. Dohi, I. Sakuma, & H. Liao (Eds.), *Medical*

imaging and augmented reality (Vol. 5128, pp. 124–132). Berlin Heidelberg: Springer.

Miller, S. L., Celone, K., DePeau, K., Diamond, E., Dickerson, B. C., Rentz, D., & Sperling, R. A. (2008). Age-related memory impairment associated with loss of parietal deactivation but preserved hippocampal activation. *Proceedings of the National Academy of Sciences, 105*(6), 2181–2186.

Mintun, M., Larossa, G., Sheline, Y., Dence, C., Lee, S. Y., Mach, R., & Morris, J. (2006). [11C] PIB in a nondemented population potential antecedent marker of Alzheimer disease. *Neurology, 67*(3), 446–452.

Nelson, S. M., Cohen, A. L., Power, J. D., Wig, G. S., Miezin, F. M., Wheeler, M. E., & Schlaggar, B. L. (2010). A parcellation scheme for human left lateral parietal cortex. *Neuron, 67*(1), 156–170.

Newman, M. E. J. (2008). mathematics of networks. In S. N. Durlauf & L. E. Blume (Eds.), *The new palgrave dictionary of economics*. Palgrave Macmillan: Basingstoke.

Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology, 24*(12), 1565–1567.

Ota, K., Oishi, N., Ito, K., Fukuyama, H., & Group, S.-J. S. (2014). A comparison of three brain atlases for MCI prediction. *Journal of Neuroscience Methods, 221*, 139–150.

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology, 56*(3), 303–308.

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., & Winblad, B. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology, 58*(12), 1985–1992.

Pihlajamäki, M., DePeau, K. M., Blacker, D., & Sperling, R. A. (2008). Impaired medial temporal repetition suppression is related to failure of parietal deactivation in Alzheimer disease. *The American Journal of Geriatric Psychiatry, 16*(4), 283–292.

Postuma, R. B., & Dagher, A. (2006). Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. *Cerebral Cortex, 16*(10), 1508–1521.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., & Schlaggar, B. L. (2011). Functional network organization of the human brain. *Neuron, 72*(4), 665–678.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, 98*(2), 676–682.

Reitz, C., Brayne, C., & Mayeux, R. (2011). Epidemiology of Alzheimer disease. *Nature Reviews Neurology, 7*(3), 137–152.

Rish, I., Cecchi, G.A., & Heuton, K. (2012). *Schizophrenia classification using functional network features*. Paper presented at the SPIE Medical Imaging.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage, 52*(3), 1059–1069.

Rubinov, M., & Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *NeuroImage, 56*(4), 2068–2079.

Sanz-Arigita, E. J., Schoonheim, M. M., Damoiseaux, J. S., Rombouts, S. A. R. B., Maris, E., Barkhof, F., & Stam, C. J. (2010). Loss of 'small-world' networks in Alzheimer's disease: graph analysis of fMRI resting-state functional connectivity. *PloS One, 5*(11), e13788. doi:10.1371/journal.pone.0013788.

Sheline, Y. I., & Raichle, M. E. (2013). Resting state functional connectivity in preclinical Alzheimer's disease. *Biological Psychiatry, 74*(5), 340–347.

Shirer, W., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex, 22*(1), 158–165.

Sperling, R. A., LaViolette, P. S., O'Keefe, K., O'Brien, J., Rentz, D. M., Pihlajamaki, M., & Hedden, T. (2009). Amyloid deposition is associated with impaired default network function in older persons without dementia. *Neuron, 63*(2), 178–188.

Sperling, R. A., Dickerson, B. C., Pihlajamaki, M., Vannini, P., LaViolette, P. S., Vitolo, O. V., & Selkoe, D. J. (2010). Functional alterations in memory networks in early Alzheimer's disease. *Neuromolecular Medicine, 12*(1), 27–43.

Stam, C. J., Jones, B. F., Nolte, G., Breakspear, M., & Scheltens, P. (2007). Small-world networks and functional connectivity in Alzheimer's disease. *Cerebral Cortex, 17*(1), 92–99. doi:10.1093/cercor/bhj127.

Suk, H.-I., Lee, S.-W., & Shen, D. (2015a). Deep sparse multi-task learning for feature selection in Alzheimer's disease diagnosis. *Brain Structure and Function*, 1–19. doi: 10.1007/s00429-015-1059-y.

Suk, H.-I., Lee, S.-W., & Shen, D. (2015b). Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. *Brain Structure and Function, 220*(2), 841–859. doi:10.1007/s00429-013-0687-3.

Supekar, K., Menon, V., Rubin, D., Musen, M., & Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Computational Biology, 4*(6), e1000100. doi:10.1371/journal.pcbi.1000100.

Tijms, B. M., Wink, A. M., de Haan, W., van der Flier, W. M., Stam, C. J., Scheltens, P., & Barkhof, F. (2013). Alzheimer's disease: connecting findings from graph theoretical studies of brain networks. *Neurobiology of Aging, 34*(8), 2023–2036.

Toussaint, P.-J., Maiz, S., Coynel, D., Doyon, J., Messé, A., de Souza, L. C., & Benali, H. (2014). Characteristics of the default mode functional connectivity in normal ageing and Alzheimer's disease using resting state fMRI with a combined approach of entropy-based and graph theoretical measurements. *NeuroImage, 101*, 778–786. doi:10.1016/j.neuroimage.2014.08.003.

van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. *Trends in Cognitive Sciences, 17*(12), 683–696. doi:10.1016/j.tics.2013.09.012.

Vapnik, V. (1998). *Statistical learning theory*. New York: Wiley.

Wang, Z., Jia, X., Liang, P., Qi, Z., Yang, Y., Zhou, W., & Li, K. (2012). Changes in thalamus connectivity in mild cognitive impairment: evidence from resting state fMRI. *European Journal of Radiology, 81*(2), 277–285.

Wang, J., Zuo, X., Dai, Z., Xia, M., Zhao, Z., Zhao, X., & He, Y. (2013). Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. *Biological Psychiatry, 73*(5), 472–481.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *Nature, 393*(6684), 440–442.

Wee, C.-Y., Yap, P.-T., Denny, K., Browndyke, J. N., Potter, G. G., Welsh-Bohmer, K. A., & Shen, D. (2012a). Resting-state multi-spectrum functional connectivity networks for identification of MCI patients. *PloS One, 7*(5), e37828.

Wee, C.-Y., Yap, P.-T., Zhang, D., Denny, K., Browndyke, J. N., Potter, G. G., & Shen, D. (2012b). Identification of MCI individuals using structural and functional connectivity networks. *NeuroImage, 59*(3), 2045–2056.

Zalesky, A., Fornito, A., & Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *NeuroImage, 53*(4), 1197–1207.

Zhang, Z., Liu, Y., Jiang, T., Zhou, B., An, N., Dai, H., & Zhang, X. (2012). Altered spontaneous activity in Alzheimer's disease and mild cognitive impairment revealed by Regional Homogeneity. *NeuroImage, 59*(2), 1429–1440. doi:10.1016/j.neuroimage.2011.08.049.

Zhao, X., Liu, Y., Wang, X., Liu, B., Xi, Q., Guo, Q., & Wang, P. (2012). Disrupted small-world brain networks in moderate Alzheimer's disease: a resting-state fMRI study. *PloS One, 7*(3), e33540. doi:10.1371/journal.pone.0033540.