Journal of Neuroscience Methods 282 (2017) 69–80

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Research paper

Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM

Seyed Hani Hojjati[a], Ata Ebrahimzadeh[a], Ali Khazaee[b], Abbas Babajani-Feremi[c,d,e,*], for the Alzheimer's Disease Neuroimaging Initiative[1]

[a] Department of Electrical Engineering, Babol University of Technology, Babol, Iran
[b] Department of Electrical Engineering, University of Bojnord, Bojnord, Iran
[c] Department of Pediatrics, Division of Clinical Neurosciences, University of Tennessee Health Science Center, Memphis, TN, USA
[d] Neuroscience Institute and Children's Foundation Research Institute, Le Bonheur Children's Hospital, Memphis, TN, USA
[e] Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, TN, USA

HIGHLIGHTS

- We investigated identifying MCI converter (MCI-C) from MCI non-converter (MCI-NC) using rs-fMRI.
- Graph theory and machine learning approach were used to classify MCI-C from MCI-NC (accuracy = 91.4%).
- We demonstrated potential of the proposed approach based on rs-fMRI for early AD diagnosis.

ARTICLE INFO

Article history:
Received 8 December 2016
Received in revised form 13 February 2017
Accepted 7 March 2017
Available online 9 March 2017

Keywords:
Alzheimer's disease (AD)
Mild cognitive impairment (MCI)
MCI converter
MCI non-converter
Resting-state fMRI
Graph theory
Machine learning approach

ABSTRACT

Background: We investigated identifying patients with mild cognitive impairment (MCI) who progress to Alzheimer's disease (AD), MCI converter (MCI-C), from those with MCI who do not progress to AD, MCI non-converter (MCI-NC), based on resting-state fMRI (rs-fMRI).
New method: Graph theory and machine learning approach were utilized to predict progress of patients with MCI to AD using rs-fMRI. Eighteen MCI converts (average age 73.6 years; 11 male) and 62 age-matched MCI non-converters (average age 73.0 years, 28 male) were included in this study. We trained and tested a support vector machine (SVM) to classify MCI-C from MCI-NC using features constructed based on the local and global graph measures. A novel feature selection algorithm was developed and utilized to select an optimal subset of features.
Results: Using subset of optimal features in SVM, we classified MCI-C from MCI-NC with an accuracy, sensitivity, specificity, and the area under the receiver operating characteristic (ROC) curve of 91.4%, 83.24%, 90.1%, and 0.95, respectively. Furthermore, results of our statistical analyses were used to identify the affected brain regions in AD.
Comparison with existing method(s): To the best of our knowledge, this is the first study that combines the graph measures (constructed based on rs-fMRI) with machine learning approach and accurately classify MCI-C from MCI-NC.
Conclusion: Results of this study demonstrate potential of the proposed approach for early AD diagnosis and demonstrate capability of rs-fMRI to predict conversion from MCI to AD by identifying affected brain regions underlying this conversion.

* Corresponding author at: Department of Pediatrics, Division of Clinical Neurosciences, The University of Tennessee Health Science Center, and Neuroscience Institute at Le Bonheur Children's Hospital, 848 Adams Ave, Suite L445A, Memphis, TN 38103, USA.
E-mail address: ababajan@uthsc.edu (A. Babajani-Feremi).

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

http://dx.doi.org/10.1016/j.jneumeth.2017.03.006
0165-0270/

1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease that is clinically characterized by the decline in memory and other cognitive functions (Albert et al., 2011). Due to the alteration in functional and structural connectivity of the brain networks in patients with AD, this disease has been named as a disconnection syndrome (Bokde et al., 2006). Mild cognitive impairment (MCI) refers to a transitional stage between age-related cognitive decline and AD. The cognitive decline in MCI is abnormal for an individual's age and educational level but does not meet criteria for AD (Petersen, 2000). Prevalence of MCI is approximately 15% in adults older than 65 years and more than half of patients with MCI progress to dementia within 5 years (Farlow, 2009). The amnestic subtype of MCI has a high risk of progression to AD constituting a prodromal stage of AD (Gauthier et al., 2006). Earlier diagnosis of AD is important in individuals with MCI, who present a conversion rate of approximately 15% per year (Allison et al., 2014). Identifying patients at an early stage of AD can influence its course using disease-modifying treatments.

It has been shown that AD pathology can be detected using neuroimaging techniques (Markesbery, 2010). A variety of approaches have been proposed to derive magnetic resonance imaging (MRI) biomarkers for early diagnostic of AD (Misra et al., 2009; Wolz et al., 2011). In a recent study, Eskildsen et al. used patterns of cortical thickness and identified cortical regions potentially discriminative for separating MCI converter (MCI-C) patients from MCI patients who remained stable for three years (Eskildsen et al., 2013). They demonstrated promising results for the prediction of patients with prodromal AD progressing to probable AD. Although utility and efficacy of MRI-based biomarkers in classifying MCI-C from MCI-NC have been demonstrated in previous studies, there is no study, to the best of our knowledge, to investigate utility of resting-state functional MRI (rs-fMRI) for differentiating these two groups.

The rs-fMRI has emerged as a powerful and efficient tool for mapping and assessing the functional architecture of the brain networks (van den Heuvel and Hulshoff Pol, 2010). To date, several studies have used rs-fMRI to identify AD induced alterations of the brain networks (Khazaee et al., 2017; Khazaee et al., 2015a; Khazaee et al., 2016; Khazaee et al., 2015b; Zhao et al., 2012). While previous studies demonstrated ability of rs-fMRI in identification of patients with AD from healthy controls or patients with MCI (Khazaee et al., 2015a; Khazaee et al., 2016; Khazaee et al., 2015b; Stam et al., 2007), utility of rs-fMRI in identification of MCI converter (MCI-C) and MCI non-converter (MCI-NC) has not been explored yet. Identifying reliable biomarkers for classifying MCI-C from MCI-NC is crucial for early diagnosis and treatment of AD. Furthermore, the neural mechanism of conversion to AD has been investigated in some studies (Gomar et al., 2016), though the neural correlates underlying this mechanism remain poorly understood. To this end, main aim of this study was to develop and evaluate a method based on rs-fMRI and predict conversion of MCI to AD.

It was reported that patients with AD compared to healthy controls had significantly lower regional connectivity and showed disrupted global functional organization (Supekar et al., 2008). Moreover, cognitive decline in patients with AD was associated with disrupted functional connectivity in the entire brain (Supekar et al., 2008). The graph theory has been shown to be an efficient approach in identifying alteration of the brain network in psychiatric and neurological diseases (Armstrong et al., 2016). It has also been shown that combination of the graph theory with machine learning approach on the basis of rs-fMRI can accurately classify patients with MCI, patients with AD, and normal subjects (Khazaee et al., 2015a; Khazaee et al., 2016; Khazaee et al., 2015b; Stam et al., 2007).

In the current study, we developed and evaluated a novel method based on the graph theory, a machine learning approach, and baseline rs-fMRI data to accurately classify MCI-C from MCI-NC. While previous studies have investigated utility of the structural MRI for classification of MCI-C from MCI-NC, this is the first study that used rs-fMRI for this classification. MCI-NC patients in our study did not convert to AD in 36 months after the baseline rs-fMRI, though MCI-C patients converted to AD 6-36 months after the baseline rs-fMRI. We did not use the "time to conversion" in MCI-C patients in the proposed method to test performance of our method in a challenging condition where this information is unknown in real clinical application. We developed a novel feature selection algorithm that investigated a wide range of local and global graph measures to identify a subset of optimal measures and classify MCI-C from MCI-NC accurately using a minimum number of features. We also investigated the neural mechanism of conversion to AD by comparing the brain networks of MCI-C with that of MCI-NC and found brain regions involved in this conversion. To this end, results of two statistical approaches were used to identify the brain regions underlying the early stages of AD.

2. Methods

2.1. Overall procedure

The overall procedure of this study is shown in Fig. 1. After data acquisition (a), pre-processing of rs-fMRI data (b), and applying the automated anatomical labeling (AAL) parcellation (c), 90 time series were extracted from ROIs (d) and connectivity matrix was calculated for each subject (e). The connectivity matrix was then thresholded by preserving 19% of the strongest weights (f), and 10 local and 13 global graph measures were computed that resulted in 913 features (g). Combinations of filter and wrapper feature selection algorithms were used to select an optimal subset of features from the original 913 features. To this end, we used five different feature selection algorithms, e.g. Multivariate Minimal Redundancy Maximal Relevance (MRMR), to sort the features based on their individual discrimination ability (h). Then our proposed sequential features collection (SFC) algorithm was applied to the sorted features and find an optimal subset of features (i). We used SVM (with linear kernel) and the optimal subset of features to classify MCI-C from MCI-NC (j). The k-fold ($k=9$) cross-validation was employed to evaluate performance of the SVM, separately based on each feature selection algorithm (e.g. MRMR), for classifying MCI-C from MCI-NC (m). Finally, brain areas with significantly different features in two groups were identified based on the optimal local graph measure (n).

We performed a statistical analysis on two global graph measures (i.e. smallworldness and transitivity) by calculating the Pearson's correlation coefficient between these measures and following behavior data: Mini Mental State Examination (MMSE), Clinical Dementia Rating (CDR), and Functional Activities Questionnaire (FAQ). We compared values of the optimal features (identified in Section 2.7) in MCI-C and MCI-NC to find features that were significantly different in two groups (l). We also performed the network based statistics (NBS) analysis on raw connectivity matrices (k).

2.2. Patients

Data for this study were selected from publicly available Alzheimer's disease neuroimaging initiative (ADNI) database.[1] In

[1] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to

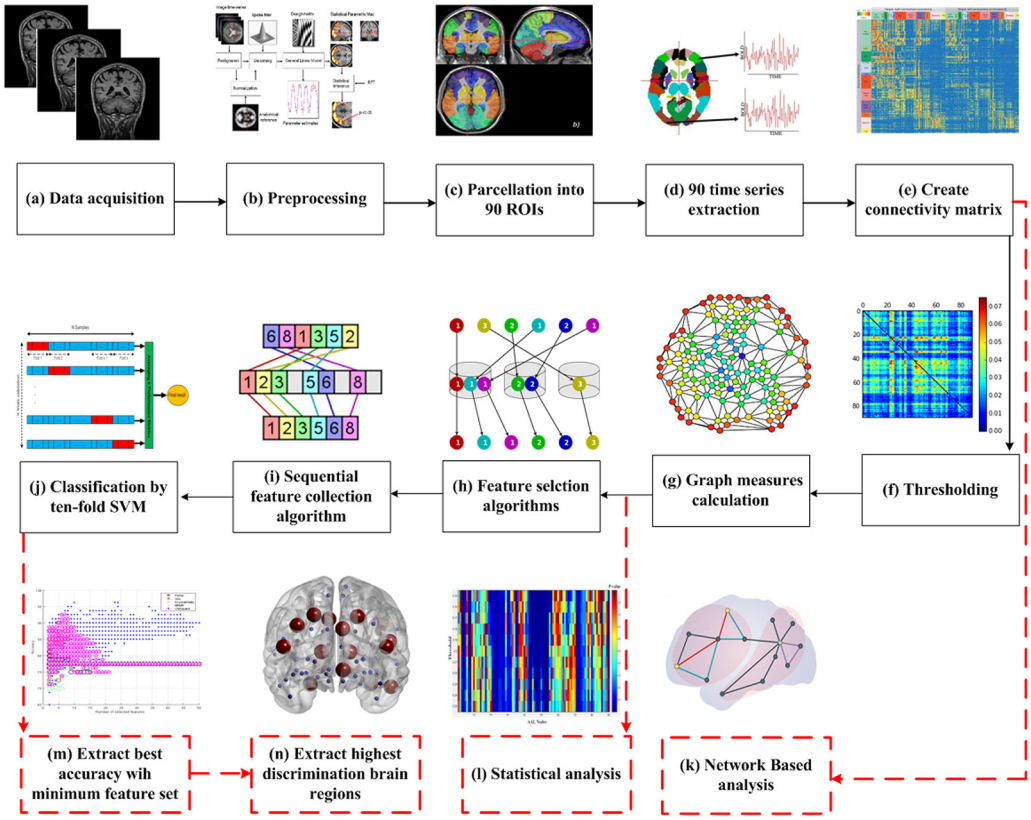

Fig 1. The overall procedure of this study. See text for details.

this study we included MCI patients who had a complete session of rs-fMRI data. The diagnosis of MCI patients was based on cognitive tests like MMSE score, instrumental activities of daily living, and CDR score. General inclusion/exclusion criteria for MCI patients were: MMSE scores between 24 and 30, a CDR of 0.5, objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, an absence of dementia, and an absence of significant levels of impairment in other cognitive domains that preserved activities of daily living.

We used the diagnostic summary file in ADNI assessment files of ADNIGO 2 subjects for identifying MCI-C and MCI-NC patients. To this end, we utilized the diagnosis variable "DXCHANGE" which had following values MCI patients: 2 = Stable in MCI; 5 = Conversion from MCI to AD; 7 = Revision from MCI to Normal; 8 = Revision from AD to MCI. MCI patients with DXCHANGE values of 2 and 5 were identified as MCI-NC and MCI-C, respectively. We exclude MCI patients with DXCHANGE values of 7 or 8 in this study. Eighteen MCI-C (average and standard deviation age 73.6 and 15.7 years, respectively; 11 male) and 62 age-matched MCI-NC (average and standard deviation age 73 and 16.3 years, respectively; 28 male) were included in this study. It is noteworthy that we included all MCI-C and MCI-NC subjects, listed in ADNI database and had a complete set of rs-fMRI data, in the current study. The MCI-C patients were converted to AD after 6–36 months. The MCI-C had a Mini Mental State Examination (MMSE) score of 21–29, a Clinical Dementia Rating (CDR) of 0.5 or 1.0, and a Functional Activities Questionnaire (FAQ) of 1–22. The MCI-NC patients did not converted to AD after 36 months of follow up. The MCI-NC subjects had a MMSE score of 19–30, a CDR of 0.5 or 1.0, and a FAQ of 0–18. All patients had no significant impairment in other cognitive domains, preserved activities of daily living, and an absence of dementia. Demographic information of subjects are summarized in Table 1.

measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org.

2.3. Data acquisition and preprocessing

The functional and structural MRI images were collected using 3-T Philips scanner. Acquisitions were performed according to the ADNI acquisition protocol (Jack et al., 2008).A total of 140 functional volumes (TR/TE 3000/30 ms, flip angle = 80°, 3.313 mm slice thickness, 48 slices) were obtained. For each subject, the first few volumes (7 vols) 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 (Khazaee et al., 2015a). Standard preprocessing was applied on rs-fMRI dataset of all patients using Data Processing Assistant for Resting State fMRI (DPARSF) toolbox (Chao-Gan and Yu-Feng, 2010) and SPM12 package (http://www.fil.ion.ucl.ac.uk/spm). 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. The average and standard deviation of maximum movement parameters across subjects in two groups are listed in Table 1. The movement parameters were not significantly different in two groups (Table 1). Then all images were 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 bandpass 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. 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 (Greicius et al., 2003).

Table 1
Demographic, clinical, and fMRI movement information of MCI-C and MCI-NC patients.

			MCI converter (MCI-C)	MCI non-converter (MCI-NC)	P-value
Demographic information	Number		18	62	–
	Male/Female		11/7	28/34	0.233
	EMCI/LMCI		6/12	38/24	0.035
	Age		73.6 ± 15.7	73.0 ± 16.3	0.256
	MMSE score		26.0 ± 2.0	27 ± 3.0	0.056
	CDR score		0.5 ± 0.0	0.5 ± 0.0	0.177
	FAQ score		11.5 ± 10.5	9.0 ± 9.0	0.049
movement parameters	Translation (mm)	X	0.362 ± 0.232	0.291 ± 0.130	0.228
		Y	0.647 ± 0.285	0.764 ± 0.466	0.198
		Z	0.715 ± 0.421	0.778 ± 0.486	0.590
	Rotation (radian)	X	0.013 ± 0.008	0.015 ± 0.012	0.345
		Y	0.006 ± 0.005	0.005 ± 0.002	0.688
		Z	0.008 ± 0.008	0.006 ± 0.006	0.527

MCI: mild cognitive impairment; EMCI: early MCI; LMCI: late MCI; MMSE: Mini-Mental State Examination; CDR: Clinical Dementia Rating. FAQ: Functional Activities Questionnaire.

2.4. Computation of graph measures

The brain was parcellated using the AAL atlas (Tzourio-Mazoyer et al., 2002). Edges of the brain network were defined as connectivity matrix. A connectivity matrix was constructed using correlation between time series of the fMRI signals of all pairs of the brain regions using the Pearson's correlation coefficient. We converted the undirected graphs to sparse ones by thresholding the graph at an optimal threshold value by preserving 19% of the strongest weights (Khazaee et al., 2015a).

Graph theory was used to calculate different measures of integration and segregation, We utilized the GraphVar (version 6.2) software for computing graph measures (Kruschwitz et al., 2015). Classification of MCI-C from MCI-NC is a difficult task, because of the similarity between their brain networks. To overcome this problem, we used a wide variety of graph measures. We computed 10 local and 13 global graph measures based on the binary undirected graphs. The local graph measures were: betweenness centrality, clustering coefficient, characteristic path, community structure Newman, community structure Louvain, eccentricity, eigenvector centrality, rich club coefficient, sub graph centrality, and participation coefficient (Rubinov and Sporns, 2010). The global graph measures were: assortativity, clustering coefficient, characteristic path, community structure Newman output, community structure Louvain output, cost efficiency (relative threshold), cost efficiency (absolute threshold), density, efficiency, graph radius, graph diameter, transitivity, and small-worldness (Rubinov and Sporns, 2010). Integration of the local and global graph measures resulted in a feature vector with 913 elements (913 = 10 (local measures) × 90 (AAL areas) + 13 (global measures)).

2.5. Statistical analysis

Two types of statistical analysis were performed: statistical analysis on graph measures and statistical analysis on raw connectivity matrices. In the first statistical analysis, we investigated correlation between the graph measures and patients' behavioral data (e.g., MMSE score). We utilized the GraphVar software to perform this statistical analysis. To this end, the connectivity matrix was thresholded by preserving 19% of the strongest weights, and then graph measures were calculated. Consequently, the Pearson correlation between two global graph measures (i.e. smallworldness and transitivity) and patients' behavioral data was calculated. Finally, we identified the global graph measures which had a significant correlation ($p < 0.05$) with the behavioral data.

In the second statistical analysis, we applied the NBS on raw connectivity matrices (Zalesky et al., 2010). To this end, first, a *t*-test was performed independently for each connection in the network to test the null hypothesis for equality of the mean value of that connection in two groups (MCI-C and MCI-NC). Connections with a *t*-value above a primary threshold contain supra-threshold connection set. The connectivity matrix was searched for any connected graph components defined by the set of supra-threshold connections, and size of those connected component was calculated. Then permutation testing (n = 10,000), by assigning subjects randomly to a group (either MCI-C or MCI-NC), was performed to find an empirical null distribution for the size of the largest connected component. Finally, the family-wise error rate corrected *p*-value for a given component size was calculated, and used to identify components that were significantly different in two groups (MCI-C and MCI-NC). In a separate statistical analysis, we also compared values of the optimal features (identified in Section 2.7) in MCI-C and MCI-NC to find features that were significantly different in two groups. It is noteworthy that in contrast with the statistical analysis on graph measure, the NBS method exclusively focuses on multiple hypotheses testing, where one instead tests the hypothesis of interest at each network connection, thereby introducing more localizing power at the cost of a massive number of multiple comparisons. The main contribution of NBS method is to present a potentially more powerful method to control the family-wise error rate (FWER) when performing this kind of analysis (Genovese et al., 2002).

2.6. Feature selection and classification

The extracted feature vector for each subject was a large vector with 913 elements. Because of the large number of features, feature selection algorithm was utilized to select an optimal subset of features from the original feature set. In addition, feature selection reduces the training-testing times and improves classification performance. Here, five different feature selection methods, namely Multivariate Minimal Redundancy Maximal Relevance (MRMR), Fisher score, Chi-square score, Gini score, Kruskal-Wallis test, were applied to the feature vector. We utilized MATLAB R2014b software for feature selection.

The Fisher score is a univariate filter that is applied to independently compute the discriminatory power of individual features between two classes of equal probability (Gu et al., 2012). Fisher score for each feature in a two class problem is defined as:

$$FS = \frac{n_1(m_1 - m)^2 + n_2(m_2 - m)^2}{\left(n_1\sigma_1^2 + n_2\sigma_2^2\right)} \quad (1)$$

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

The multivariate minimal redundancy maximal relevance (MRMR) feature selection algorithm selects features that have maximal statistical dependency based on mutual information (Yang, 2013). The algorithm considers relevant features and redundant features simultaneously. The relevance of a feature set S for the class C is defined by the average value of all mutual information values between the individual feature f_i and the class C, and the redundancy of all features in the set f_jS is the average value of all mutual information values between the features f_i and

The MRMR is defined as:

$$MRMR = MAX_S \left\{ \frac{1}{|S|} \sum_{f_i \varepsilon S} I(f_i; c) - \frac{1}{|S|^2} \sum_{f_i \varepsilon S} I\left(f_i; f_j\right) \right\} \quad (2)$$

The chi-square score algorithm is a common statistical test that is used to measure divergence from the distribution expected by the assumption of independence of the class value from feature occurrence (Geng et al., 2007). The chi-square score for a feature with k different values and C classes is calculated as:

$$\tau^2 = \sum_{i=1}^{k} \sum_{j=1}^{C} \frac{\left(n_{ij} - m_{ij}\right)^2}{m_{ij}} \quad (3)$$

where n_{ij} is the number for samples with i^{th} feature value in class j and $m_{ij} = (n_{i*} n_{*j}/n)$ where n_{i*} is the number of samples with i^{th} value for a particular feature, n_{*j} is the number of samples in class j, n is the number for samples.

The Gini index quantifies ability of a feature to discriminate classes (Shang et al., 2007). The Gini Index of a feature f assuming C classes is defined as:

$$GI(f) = 1 - \sum_{i=1}^{C} [p(i|f)]^2 \quad (4)$$

where p is the probability operator. The maximum value of the Gini index for binary classification is 0.5. Features with the smallest value are selected as the most relevant features.

The Kruskal-Wallis test is a non-parametric version of one-way analysis of variance (ANOVA) that compares medians of samples from two or more groups and estimates the p-value for the null hypothesis that all samples are produced from the same population (Breslow, 1970). After ranking all data points across all groups together, a value can be calculated as follow:

$$K = \frac{12}{N(N+1)} \sum_{i=1}^{8} n_i (\bar{r}_i)^2 - 3(N+1) \quad (5)$$

where N is the total number of samples across all groups, n_i is number of samples in group 'i', r_{ij} is rank of sample 'j' in the group 'i', and $.\bar{r}_i = (\sum_{j=1}^{n_i} r_{ij})/n_i$

We developed a sequential features collection (SFC) algorithm to find an optimal subset of features (with a small number of features) and accurately classify MCI-C from MCR-NC. To this end, we sorted all featured using five feature selection algorithms, i.e. MRMR, Fisher score, Chi-square score, Gini score, and Kruskal-Wallis test. Output of each feature selection algorithm was a sorted feature vector. The first and the last features in the sorted feature vector had maximum and minimum discrimination ability, respectively, to classify MCI-C from MCR-NC. The sorted feature vectors were inputs of the SFC algorithm.

Suppose that the sorted feature vector F had m features. The SFC algorithm started with first two elements of the sorted feature vector (feature length = 2) and computed performance of SVM classifier using 9-fold cross validation based on these two features. Then length of the feature vector increased to the first three feature (feature length = 3) and computed performance of SVM classifier using these three features. This procedure was continued until feature length was equal to m. In the next step of the SFC algorithm, we eliminated the first element of the sorted feature vector and the SFC algorithm was run using $m-1$ elements of this vector. Similar to the pervious step, performance of SVM classifier was computed based on $m-1$ elements of the sorted feature vector and this step was continued until the last element of the sorted feature vector was incorporated, i.e. feature length = $m-1$. These steps were continued by eliminating top features of the sorted feature vector one by one. At the end, we found an optimal subset of features by comparing performances of SVM classifier using all subsets of features. The following pseudo-code shows the algorithm of SFC.

Pseudo-code for SFC algorithm

Suppose that the sorted feature vector F had M features.

For $i = 1 : m$

  For $j = i + 1 : m$

    Collected Features $= F[\, i : j \,]$

    Feature Length $= j - i + 1$

    Computed performance of SVM classifier for Collected Features:

    a_{ij} = Accuracy of SVM classifier using Collected Features

  End

End

Output:

  Calculate maximum of accuracy and corresponding values of j_{max} and i_{max}

  Optimal subset of features: Collected Features $= F[i_{max} : j_{max}]$

  If there are more than two subset of features with equal accuracy

    Choose subset of features with minimum length of features

  End

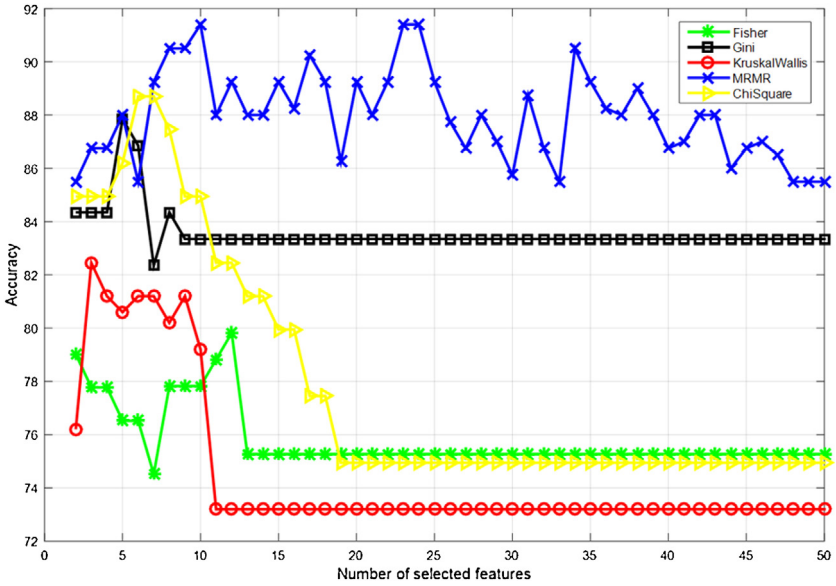

Fig. 2. Illustration of performances of five feature selection algorithms for classifying MCI-C from MCI-NC. The x-axis represents the number of collected features by the sequential features collection (SFC) algorithm. Only the top 50 features out of 913 original features were considered in this figure.

We used support vector machine (SVM) with linear kernel in this study to classify MCI-C from MCI-NC based on an optimal subset of features extracted from the rs-fMRI graph measures. We tested performances of SVM with linear and non-linear kernels and observed that SVM with linear kernel outperformed SVM with non-linear kernel. Since the number of features was larger than the number of observation in the current study, the nonlinear kernel was not expected to outperform the linear kernel (Hsu et al., 2003). In fact, the linear kernel was good enough in this study by searching for only one parameter, i.e. parameter C, and solving the optimization problem for the linear kernel was much faster than that for non-linear kernel (Smola and Schölkopf, 2004a). The SVM was implemented in MATLAB (The Math Works, Natwick, MA) and LIBSVM (http://www.csie.ntu.edu.tw/cjlin/libsvm/). Neuroimaging studies are usually based on small number of subjects. Therefore, various cross-validation strategies were suggested to overcome the loss of generalization due to the small training and testing sample size in the neuroimaging applications.

Two popular cross validation (CV) approached are Leave One Out Cross Validation (LOOCV) and *k*-fold Cross-Validation. In LOOCV, one data-point in the available dataset is held-out and the model is trained with respect to the rest and then test with that data-point. The bias of LOOCV error is expected to be small since almost entire dataset are used for training and the trained model is pretty close to the real one. However, variance of LOOCV error is expected to be large since there is only one data point being used for testing every time. Moreover, LOOCV is computationally very intensive. The *k*-fold Cross-Validation was introduced as an alternative to the computationally expensive LOOCV (Sylvain and Alain, 2010; Byun and Lee, 2002; Smola and Schölkopf, 2004b). The *k*-fold Cross-Validation approach provides a tradeoff between bias and variance of error (Sylvain and Alain, 2010). Choosing an appropriate value for '*k*' ensures that the testing procedure provides a good estimate of how well the model works. By choosing a small value for *k*, the Cross-Validation can result in large bias for error. The *k*-fold Cross-Validation with a large value for *k* is similar to LOOCV which has a large error variance and is computationally intensive. A value for *k* between 5–10 is usually used in practical machine learning scenarios. We used $k = 9$ for *k*-fold cross-validation in this study.

SVM classification with unequal group sizes (in this study 18 and 62 samples) can lead to biased accuracy of classification. To address this issue, we randomly selected 18 out of 62 MCI-NC subjects and then evaluated performance of the SVM using those 18 selected MCI-NC subjects and 18 MCI-C subjects based on *k*-fold CV approach. The total number of subjects in each random iteration was 36 (18 MCI-C and 18 MCI-NC). Since we had 36 subjects in each random iteration, we used 9-fold CV, instead of 10-fold CV which is the most commonly used *k*-fold CV approach, in order to have 4 subjects in each of 9 folds. 9-fold cross-validation was repeated 1000 times by randomly selecting 18 out of 62 MCI-NC subjects and the average accuracy, sensitivity, specificity, positive predictively, and the area under the receiver operating characteristic (ROC) curve (AUC) across this repetition were calculated. We also performed a permutation test to estimate the empirical distribution of the classifier accuracy under the null hypothesis. To this end, we labeled each patient to two groups randomly, trained and tested the SVM classifier with this random labeling, and calculated the accuracy of SVM classifier. The permutation test was repeated 1000 times.

3. Results

Results of the current study can be divided into three parts. In the first part (Sections 3.1), we present results of the proposed algorithm for predicting the early stages of AD. Using our proposed algorithm, we identified an optimal subset of features which extracted from the undirected binary graph measures. Then the optimum subset of feature was used to classify MCI-C from MCI-NC. In the second part (Sections 3.2), the results of statistical analysis for identification of significant alterations of the graph measures in MCI-C and MCI-NC are presented. Then the results of NBS analysis on the raw connectivity matrices are presented.

3.1. Predicting early stages of AD

Performances of the feature selection algorithms in different number of collected features by SFC are compared in Fig. 2. Only the top 50 features were considered in this figure. As shown in Fig. 2, the MRMR feature selection algorithm outperformed other algorithms. The accuracy, sensitivity, specificity, positive predictive value (PPV) of SVM in classifying MCI-C from MCI-NC using the optimal subset of features identified by SFC and five feature

Table 2
Accuracy, sensitivity, specificity, positive predictive value (PPV) of SVM classifier in discriminating MCI-C from MCI-NC using a subset of optimal features identified by five feature selection algorithms.

Feature selection methods	No. of selected features	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)
MRMR	10	91.40	83.24	90.10	84.22
Chisquare	6	88.70	81.68	68.44	91.89
Gini	5	86.85	87.1	72.81	75.34
Krushalwails	3	82.40	73.72	67.39	71.94
Fisher	2	79.00	61.12	53.78	68.14

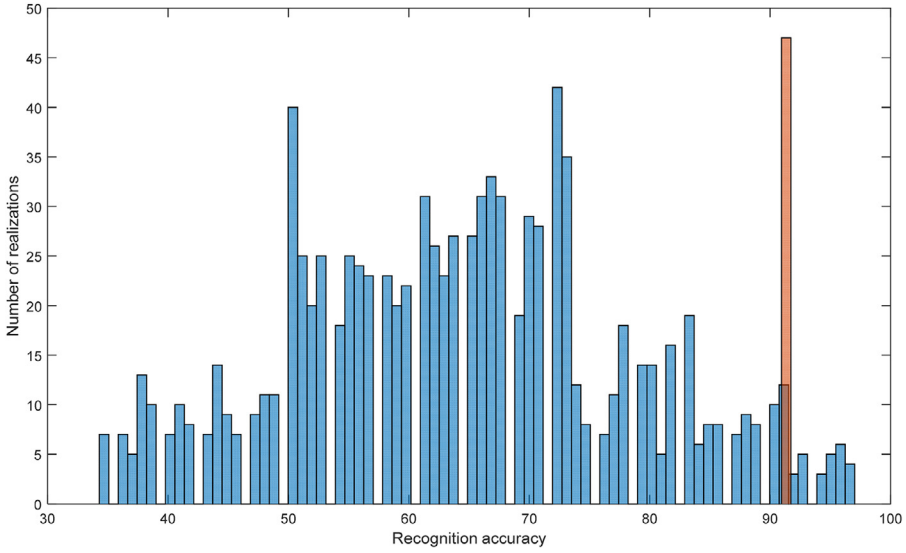

Fig. 3. Results of the permutation test to estimate the empirical distribution of the accuracy of the SVM classifier based on the MRMR feature selection algorithm. The vertical bar in brown color represents accuracy of the MRMR feature selection algorithm using the original data without permutation.

Table 3
Selected features and the corresponding brain regions using five feature selection algorithms.

Feature selection algorithms	Graph measurements	Type of measurements	Brain regions
MRMR	Community structure Newman	Local	Angular_L
	Community structure Newman	Local	ParaHippocampal_R
	Community structure Louvain	Local	Frontal_Inf_Tri_R
	Betweenness centrality	Local	Frontal_Inf_Oper_L
	Eigenvector centrality	Local	Rolandic_Oper_L
	Subgraph centrality	Local	Lingual_R
	Eccentricity	Local	Fusiform_L
	Eigenvector centrality	Local	Lingual_L
	Clustering coefficient	Local	Frontal_Inf_Tri_L
	Transitivity	Global	–
Chisquare	Rich club coefficient	Local	Frontal_Sup_R
	Rich club coefficient	Local	Frontal_Sup_L
	Rich club coefficient	Local	Frontal_Mid_R
	Subgraph centrality	Local	Lingual_L
	Participation coefficient	Local	Cingulum_Mid_R
	Smallworldness	Global	–
Gini	Rich club coefficient	Local	Frontal_Sup_R
	Rich club coefficient	Local	Frontal_Sup_L
	Rich club coefficient	Local	Frontal_Mid_R
	Subgraph centrality	Local	Lingual_L
	Smallworldness	Global	–
Krushalwails	Betweenness centrality	Local	Temporal_Inf_L
	Subgraph centrality	Local	Caudate_L
	Subgraph centrality	Local	Caudate_R
Fisher	Betweenness centrality	Local	Frontal_Sup_Medial_L
	Betweenness centrality	Local	Frontal_Med_Orb_L

Angular: Angular gyrus; Frontal_Inf_Tri: Triangular part of inferior frontal gyrus; Frontal_Inf_Oper: opercular part of inferior frontal gyrus; Rolandic_Oper: Rolandic operculum; Lingual: Lingual gyrus; Fusiform: Fusiform gyrus; ParaHippocampal: Parahippocampal gyrus; Frontal_Sup: dorsolateral superior frontal gyrus; Frontal_Mid: Middle frontal gyrus; Cingulum_Mid: Middle cingulate and paracingulate gyri; Temporal_Inf: Inferior temporal gyrus; Caudate: Caudate nucleus; Frontal_Sup_Medial: Medial superior frontal gyrus; Frontal_Med_Orb: Medial orbital part of superior frontal gyrus; _R: Right hemisphere; _L: Left hemisphere.

selection algorithms are listed in Table 2. The MRMR algorithm outperformed other feature selection algorithms by providing a classification accuracy, specificity, sensitivity, and PPV of 91.40%, 83.24%, 90.10%, and 84.22%, respectively. Fig. 3 shows the empirical distribution of the classification accuracy using random permutation where only 29 out of 1000 random permutations resulted accuracies greater than 91.4%. The ROC curve was plotted in Fig. 4. The area under ROC curve was 0.95. Graph measures and cortical areas corresponding to the optimal subset of features for five feature selection algorithms are listed in Table 3. Since the graph measures and corresponding cortical areas listed in Table 3 had the most discrimination ability in identifying MCI-C patients from MCI-NC patients, these areas might be affected in the early stages of AD. Locations of these areas are shown in Fig. 5.

3.2. Statistical analysis

Statistical analysis was performed on all local graph measures listed in Table 3 to investigate significant change in those features in two groups (i.e. MCI-C and MCI-NC), and the results are listed in Table 4. Twelve local graph measures were significantly different ($p < 0.05$) in two groups, and the most significant difference was observed in the Betweenness centrality in the left inferior temporal gyrus ($p < 1.0e^{-11}$).

The results of correlation between the behavior data (FAQ, CDR, and MMSE) and two global graph measures (i.e. Smallworldness and Transitivity) are listed in Table 5. These two global graph measures were selected since these measures were the only ones included in the optimal subset of features in Table 3. As listed in Table 5, FAQ and CDR were significantly correlated with Small-worldness and Transitivity ($p < 0.003$), though MMSE did not show a significant correlation with these global measures.

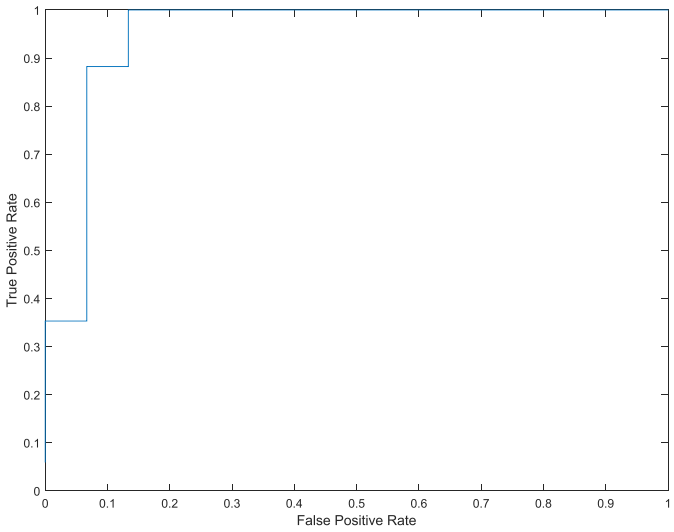

Fig. 4. ROC curves for binary classification of MCI-C *vs.* MCI-NC for optimal feature selected by MRMR feature selection method. Area under ROC curve (AUC) is 0.94902.

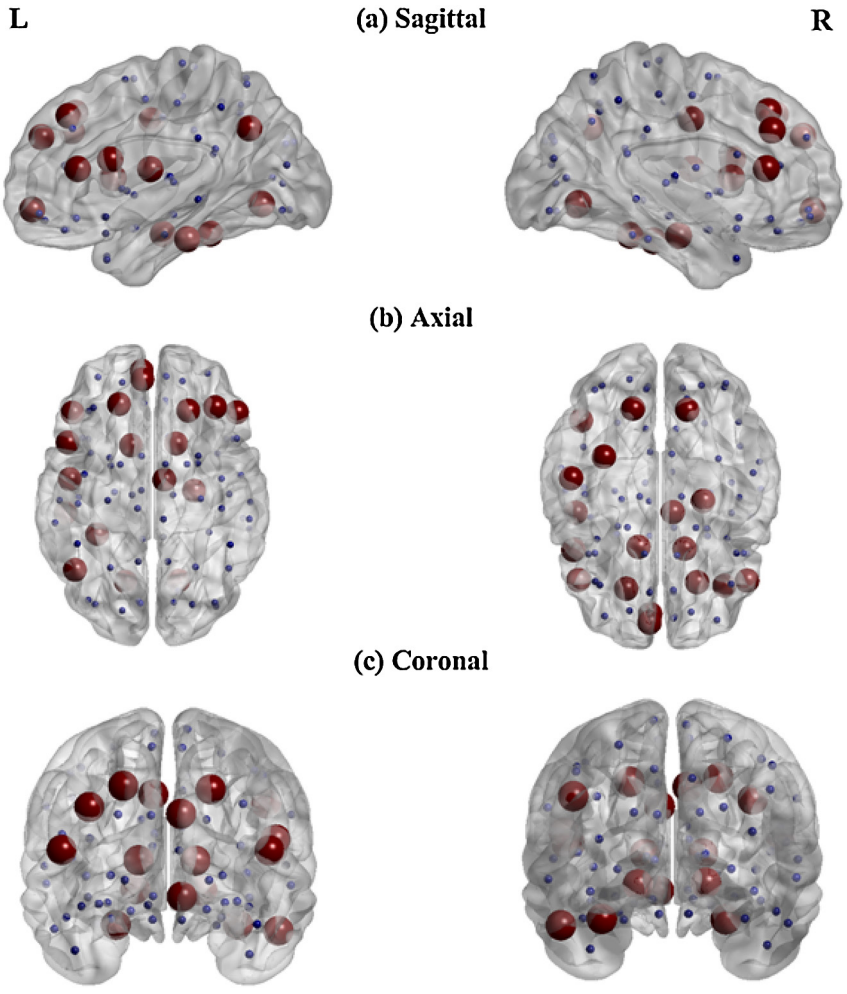

Fig. 5. Locations of the cortical areas, listed in Table 3, which might be affected in early stage of AD in (a) sagittal (b) axial (c) coronal views.

Table 4
Comparison of local graph measures, listed in Table 3, in MCI-C and MCI-NC.

Graph measurements (Brain regions)	MCI-C (mean ± standard deviation)	MCI-NC (mean± standard deviation)	P-value
Betweenness centrality (Temporal_Inf_L)	1.0340 ± 1.0267	3.7512 ± 5.9779	8.34e-12
Rich club coefficient (Frontal_Mid_R)	1.0190 ± 0.0129	1.0042 ± 0.0056	2.91e-10
Subgraph centrality (Lingual_L)	8.2893 ± 8.0760	1.0585 ± 1.3134	1.34e-9
Betweenness centrality (Frontal_Med_Orb_L)	0.6892 ± 0.5643	1.5855 ± 1.3099	0.72e-8
Local Clustering coefficient (Frontal_Inf_Tri_L)	1.6138 ± 0.8747	2.5069 ± 0.9724	2.00e-6
Rich club coefficient (Frontal_Sup_R)	1.0052 ± 0.0064	1.0006 ± 0.0018	2.46e-6
Betweenness centrality (Frontal_Inf_Oper_L)	19.7342 ± 66.8236	3.4695 ± 6.7705	4.5e-6
Rich club coefficient (Frontal_Sup_L)	1.0190 ± 0.0129	1.0042 ± 0.0056	7.64e-6
Eigenvector centrality (Lingual_L)	0.8815 ± 0.3549	0.4568 ± 0.3643	4.3e-5
Subgraph centrality (Lingual_R)	10.3497 ± 14.2786	1.4814 ± 3.4714	0.0001
Eccentricity (Fusiform_L)	1.3935 ± 0.2678	1.1304 ± 0.1971	0.0002
Eigenvector centrality (Rolandic_Oper_L)	0.6525 ± 0.3983	0.9774 ± 0.3203	0.0053
Subgraph centrality (Caudate_L)	1.2800 ± 1.6798	6.7549 ± 10.2424	0.0255
Participation coefficient (Cingulum_Mid_R)	0.7931 ± 0.2109	0.5982 ± 0.3266	0.0261
Subgraph centrality (Caudate_R)	1.3533 ± 2.1263	8.4012 ± 14.4462	0.0264
Community structure Louvain(Frontal_Inf_Tri_R)	2.1324 ± 2.6279	0.9017 ± 0.8178	0.1262
Community structure Newman (ParaHippocampal_R)	0.7750 ± 0.5688	1.4247 ± 1.0162	0.1669
Community structure Newman (Angular_L)	1.8426 ± 1.1736	1.2925 ± 0.9180	0.3397
Betweenness centrality (Frontal_Sup_Medial_L)	3.4488 ± 6.0589	3.6806 ± 10.8855	0.7649

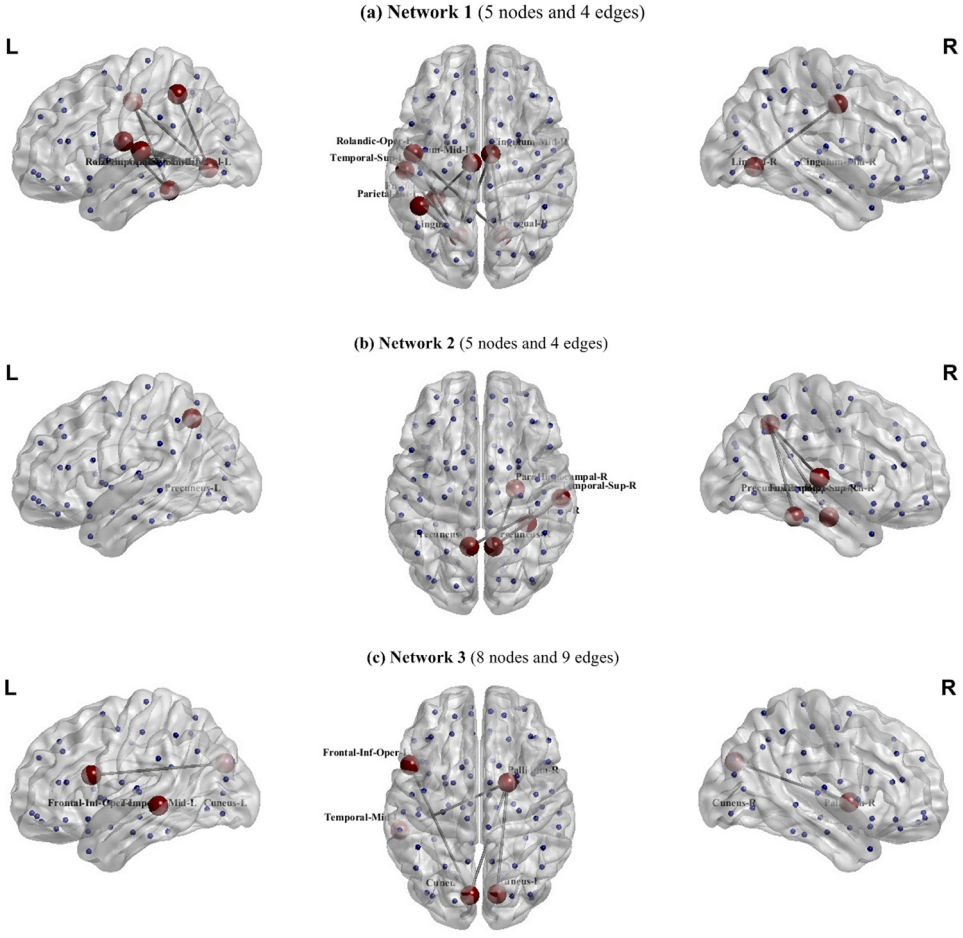

Fig. 6. Three brain networks that were significantly different in MCI-C and MCI-NC. These networks were calculated using the network based statistics (NBS) on the raw connectivity matrices in patients with MCI-C and MCI-NC. List of cortical areas are as follows.
(1) Network 1: opercular part of the left inferior frontal gyrus (Frontal-Inf-Oper-L), left middle temporal gyrus (Temporal-Mid-L), left and right cuneus, and right pallidum.
(2) Network 2: Left and right precuneus (Precuneus-L and Precuneus-R), right fusiform gyrus (Fusiform-R), right superior temporal gyrus (Temporal-Sup-R), and right parahippocampal gyrus (Parahippocampal-R).
(3) Network 3: left rolandic operculum (Rolandic-oper-L), left superior temporal gyrus (Temporal-Sup-L), left inferior parietal gyrus, excluding supramarginal and angular gyri (Parietal-Inf-L), left fusiform gyrus (Fusiform-L), left lingual gyrus (Lingual-L), left Middle cingulate and paracingulate gyri (Cingulate-Mid-L), right Middle cingulate and paracingulate gyri (Cingulate-Mid-R), and right Lingual gyrus (Lingual-R).

Table 5
Correlation between the behavior data and two global graph measures, listed in Table 3.

Measures		FAQ	CDR	MMSE score
Smallworldness	Correlation	0.425	0.363	0.060
	p-value	0.0001	0.0008	0.6067
Transitivity	Correlation	0.339	0.332	0.115
	p-value	0.0021	0.0025	0.3243

By applying NBS to the raw connectivity matrices, we identified three networks that were significantly different in two groups (i.e. MCI-C and MCI-NC, Fig. 6). These networks were identified using a threshold at $t = 9.18$ and a p-value < 0.001. The first network comprised four edges and five nodes in bilateral visual cortex (i.e. cuneus) and left language circuit (i.e. opercular part of inferior frontal gyrus and middle temporal gyrus in the left hemisphere). Second network comprised four edges and five nodes, located bilaterally in precuneus as well as in the parahippocampal, fusiform, and superior temporal gyri in the right hemisphere. It is noteworthy that this network comprises cortical areas involved in memory function. Third network comprised nine edges and eight nodes, located mostly in the left hemisphere.

4. Discussion

In this study, we used rs-fMRI and combined the graph theory with a machine learning approach to accurately classify patients with MCI who progress to AD (MCI-C) from those who do not progress to AD (MCI-NC). The major findings of this study are as follows: 1) The proposed method classified MCI-C from MCI-NC with a high accuracy of 91.4%. It is noteworthy that achieving high accuracy in classifying MCI-C from MCI-NC using rs-fMRI is a challenging problem since alteration of the resting-state brain network in an early stage of AD is expected to be small. 2) We achieved high accuracy in classifying MCI-C from MCI-NC using a small number of features (i.e. equal or less than 10 features) that caused our proposed algorithm to be less complex and more general. 3) We identified alteration of the brain network and cortical regions involved in this alteration in the early stage of AD.

As listed in Table 2, we predicted conversion to AD accurately with a small number of features, and achieved a maximum accuracy of 91.4% using MRMR feature selection algorithm and 10 out of 913 features. It can be seen from this table that MRMR method outperformed slightly other methods in sensitivity, specificity, and positive predictive value. The proposed method provided a superior accuracy compared to previous studies in early AD diagnosis (Misra et al., 2009; Wolz et al., 2011; Eskildsen et al., 2013; Zalesky et al., 2010; Moradi et al., 2015). Results in Table 2 revealed that accuracy and sensitivity of MRMR and Chi-square are generally comparable, though Chi-square, compared to MRMR, has a worse specificity but a better PPV. The different degree of concordance between MRMR and Chi-square methods for accuracy, sensitivity, specificity, and PPV was mainly related to the numbers of true positives (TP) and true negatives (TN) of these methods. The order of difference between TP of MRMR and Chi-square was higher than the order of difference between their TN. Since different features and cortical areas were used in Chi-square and MRMR methods, the calculated confusion matrices were dissimilar for these methods that resulted in different TP and TN.

Performances of five feature selection algorithms are compared in Fig. 2. In Fig. 3, we used the permutations to estimate the empirical cumulative distribution of the classifier accuracy under the null hypothesis for MRMR feature selection method. Based on the permutation results, our classification results were significant. In Fig. 5 the value of AUC for MCI-C vs. MCI-NC classification was 0.95. Results of this figure showed our proposed method has a high performance for predicting AD. These results suggest that the proposed machine learning approach based on the directed graph measures may serve as a biomarker for predicting AD.

The MRMR feature selection algorithm outperformed other algorithms, though we achieved an accuracy of 79.0% using only 2 features in the Fisher feature selection algorithm. As shown in this figure, the MRMR feature selection algorithm provided a maximum accuracy using 10, 23, or 24 features and the SFC algorithm selected minimum number of features, i.e. 10 features.

Results in Table 3 show that features based on 23 out of 913 graph measures in 18 cortical areas were able to accurately classify MCI-C from MCI-NC. Eskildsen et al. considered "time to conversion" and separated patients in different groups based on the time to conversion from MCI to AD, and reported 70%-76% accuracy for predicting conversion to AD depending on the time to conversion. In another recent study, six anatomical MRI markers, e.g. cortical thickness and grey matter density, were used for AD diagnosis (de Vos et al., 2016). In this study, the area under the receiver operating curve (AUC) was used to determine the classification performance. It was reported that combination of all six measures resulted in an AUC of 0.98 while the AUC values for the single measures ranged from 0.67 (cortical thickness) to 0.94 (grey matter density). In another study, a MRI-based method was proposed to predict conversion from MCI to AD from one to three years before the clinical diagnosis (Moradi et al., 2015), and potential of MRI biomarkers for early diagnosis of AD was demonstrated. Cheng et al. develop a domain transfer learning method to predict MCI conversion using MRI, FDG-PET, and cerebrospinal fluid (CSF) data from normal subjects and patients with MCI and AD. They classified MCI-C from MCI-NC with an accuracy of 79.4% (Cheng et al., 2015).

Fig. 5 shows cortical areas listed in Table 3. Referring to discrimination ability of graph measures in these areas in classifying MCI-C from MCI-NC, these areas may be affected in the early stage of AD. Our findings for association of the caudate nucleus and frontal cortex with the early stage of AD are in agreement with previous MRI studies demonstrating atrophy in these regions in patients with AD (Busatto et al., 2003; Frisoni et al., 2002). Studies on functional and structural brain networks in patients with AD have demonstrated that cognitive function deficits in these patients may relate to the abnormalities in the connectivity between different brain areas in temporal lobe (Stam et al., 2007; He et al., 2008). Previous studies reported that the inferior temporal gyrus (ITG) is affected during the prodromal stage of AD (Scheff et al., 2011). In fact, the ITG plays an important role in verbal fluency, a cognitive function affected in early AD, and disruption in function of this area may underlie some of the early AD-related clinical dysfunctions. Our results revealed that the betweenness centrality in the left ITG was significantly smaller ($p < 1e^{-11}$) in MCI-C compared to that in MCI-NC (Table 4). The betweenness centrality in a node of a network represents influence of that node on the transfer of information to other nodes. Since the ITG is an essential area in the verbal fluency circuit, our results might be indicative of disruption in communication between this area and other regions involved in this cognitive function in early stage of AD.

Previous PET studies demonstrated significant abnormalities in the lingual gyrus in patients with AD (Eustache et al., 2004). In agreement with these studies, we found that the left lingual gyrus may be affected during the early stage of AD and the subgraph centrality in this region was significantly larger ($p < 2e^{-9}$) in MCI-C than in MCI-NC (Table 4). Our results revealed that the rich club coefficient in the superior frontal gyrus (SFG) was significantly larger ($p < 3e^{-6}$) in MCI-C than in MCI-NC (Table 4). Our findings for involvement of the SFG in early stage of AD are in agreement with rs-fMRI studies showing a significant decrease in connectivity

between this region and the posterior cingulate (PCC) in early AD patients (Wang et al., 2007).

Cortical thickness of the parahippocampal gyrus has been reported to be discriminative for separating patients with AD from healthy controls (Lerch et al., 2008). However, it was reported that the hippocampus is not a discriminative factor at 36 months prior to diagnosis of AD in the MRI based analysis (Eskildsen et al., 2013). In agreement with these studies, we found that the community structure Newman in the right parahippocampal gyrus was a discriminative feature for separating MCI-C from MCI-NC (Table 3), though we failed to find a significantly difference in this feature between two groups ($p > 0.16$, Table 4).

We found that only two out of 13 global graph measures, i.e. transitivity and smallwordness, were discriminative for separating MCI-C from MCI-NC (Table 3). We also found a significant correlation ($p < 0.003$) between these global graph measures and behavioral measures (i.e. FAQ and CDR) in patients with MCI, though the correlations of the behavioral measures and the smallwordness were more significant than that and the transitivity (Table 5). The fact that there are many short-range anatomical connections and relatively few long-range connections in the human brain network indicates that this network has small world properties (He et al., 2007a). It has been reported that the small world architecture of the brain network was disrupted in patients with AD (Zhao et al., 2012; Stam et al., 2007; Supekar et al., 2008; Wang et al., 2013), and this disruption is more prominent in patients with AD compared to patients with MCI (Yao et al., 2010). In agreement with these studies, we found that the smallwordness is highly discriminative for separating MCI-C from MCI-NC.

We identified three brain networks that were significantly different in MCI-C and MCI-NC (Fig. 6). The first network comprised areas in the language circuit (i.e. left inferior frontal gyrus (IFG) and middle temporal gyrus (MTG)) and bilateral visual cortex (Fig. 6). Several previous studies have shown involvement of the IFG and MTG in AD (Agosta et al., 2012), though rule of these regions in the early stage of AD remains unclear. Our results revealed that the betweenness centrality in the opercular part of IFG (IFG-operc) in the left hemisphere was significantly different in MCI-C and MCI-NC (Table 4). Since the left IFG-operc is an essential area in the expressive language circuit, our results for dysfunction of the betweenness centrality in this region might be indicative of disruption in communication between this area and other language specific cortices in early stage of AD. It has been shown that the MTG is involved in accessing lexical and semantic information (Hickok and Poeppel, 2007), and our results might indicate involvement of this region in disruption of these functions in early stage of AD.

We identified a right-lateralized network, namely Network 2 in Fig. 6, which was significantly different in MCI-C and MCI-NC. This network comprised the parahippocampal gyrus (PHG), superior temporal gyrus (STG), and fusiform in the right hemisphere and precuneus bilaterally. It was demonstrated that volume of the PHG was significantly smaller in patients with AD than in healthy control subjects, and atrophy of this region in AD was related to distinct aspects of the patients' memory impairments (Kohler et al., 1998). Furthermore, volume of the PHG was shown to be a biomarker for the early phase of AD (Echavarri et al., 2011). Consistent with these structural MRI studies, our results revealed that the functional connectivity of right PHG in Network 2 was disrupted in the early stage of AD. Previous studies have shown that patients with AD exhibited significant reduction in volume of the gray and white matters in the STG (Guo et al., 2010). Wang et al. reported that the rs-fMRI functional connectivity between the PCC and STG was significantly decreased in early AD patients (Wang et al., 2007). Consistent with these studies, we found that the functional connectivity of right STG in Network 2 was disrupted in the early stage of AD.

We identified another network, namely Network 3 in Fig. 6, which was significantly different in MCI-C and MCI-NC (Fig. 5). This network comprised eight areas in the left STG, fusiform gyrus, inferior parietal gyrus, and rolandic operculum as well as bilateral lingual gyrus and middle cingulate and paracingulate gyri. We already discusses involvement of STG, fusiform gyrus, and lingual gyrus in early stage of AD. The rolandic operculum has previously been found to be highly discriminative for separating patients with AD from healthy controls, and the cingulate gyrus was well known to be affected in early stage of AD (Eskildsen et al., 2013; Khazaee et al., 2015a). The inferior parietal gyrus has also been associated with the early stage of AD in previous studies (He et al., 2007b; Reed et al., 2009).

5. Conclusions

We combined the graph theory with a machine learning approach to use rs-fMRI and accurately classify MCI-C from MCI-NC. We developed a sequential features collection (SFC) algorithm to find an optimal subset of features that were highly discriminative in classifying MCI-C from MCI-NC. The proposed algorithm in this study was able to accurately classify two groups with an accuracy of 91.4%. In addition, we identified 23 graph measures in 18 cortical areas and three brain networks that were significantly different in MCI-C and MCI-NC. Results of this study show that combining the graph theory with an efficient machine learning approach, based on the rs-fMRI connectivity analysis, might be useful for the diagnosis of early AD.

Disclosure

None of the authors has any conflict of interest to disclose.

Acknowledgments

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

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. 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 Insti-

tute 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.

References

Agosta, F., et al., 2012. Resting state fMRI in Alzheimer's disease: beyond the default mode network. Neurobiol. Aging 33 (8), 1564–1578.

Albert, M.S., et al., 2011. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 7 (3), 270–279.

Allison, J.R., et al., 2014. A relationship between the transient structure in the monomeric state and the aggregation propensities of alpha-synuclein and beta-synuclein. Biochemistry 53 (46), 7170–7183.

Armstrong, C.C., et al., 2016. Graph-theoretical analysis of resting-state fMRI in pediatric obsessive-compulsive disorder. J. Affect. Disord. 193, 175–184.

Bokde, A.L., et al., 2006. Functional connectivity of the fusiform gyrus during a face-matching task in subjects with mild cognitive impairment. Brain 129 (Pt 5), 1113–1124.

Breslow, N., 1970. A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship. Biometrika 57 (3), 579–594.

Busatto, G.F., et al., 2003. A voxel-based morphometry study of temporal lobe gray matter reductions in Alzheimer's disease. Neurobiol. Aging 24 (2), 221–231.

Byun, H., Lee, S.W., 2002. Applications of support vector machines for pattern recognition: a survey. In: Pattern Recognition with Support Vector Machines. Springer, Berlin Heidelberg, pp. 213–236.

Chao-Gan, Y., Yu-Feng, Z., 2010. DPARSF: a MATLAB toolbox for pipeline data analysis of resting-state fMRI. Front. Syst. Neurosci. 4, 13.

Cheng, B., et al., 2015. Domain transfer learning for MCI conversion prediction. IEEE Trans. Biomed. Eng. 62 (7), 1805–1817.

de Vos, F., et al., 2016. Combining multiple anatomical MRI measures improves Alzheimer's disease classification. Hum. Brain Mapp. 37 (5), 1920–1929.

Echavarri, C., et al., 2011. Atrophy in the parahippocampal gyrus as an early biomarker of Alzheimer's disease. Brain Struct. Funct. 215 (3–4), 265–271.

Eskildsen, S.F., et al., 2013. Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. Neuroimage 65, 511–521.

Eustache, F., et al., 2004. 'In the course of time': a PET study of the cerebral substrates of autobiographical amnesia in Alzheimer's disease. Brain 127 (Pt 7), 1549–1560.

Farlow, M.R., 2009. Treatment of mild cognitive impairment (MCI). Curr. Alzheimer Res. 6 (4), 362–367.

Frisoni, G.B., et al., 2002. Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry. J. Neurol. Neurosurg. Psychiatry 73 (6), 657–664.

Gauthier, S., et al., 2006. Mild cognitive impairment. Lancet 367 (9518), 1262–1270.

Geng, X., et al., 2007. Feature selection for ranking. Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 407–414.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15 (4), 870–878.

Gomar, J.J., et al., 2016. Lack of neural compensatory mechanisms of BDNF val66 met met carriers and APOE E4 carriers in healthy aging, mild cognitive impairment, and Alzheimer's disease. Neurobiol. Aging 39, 165–173.

Greicius, M.D., et al., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc. Natl. Acad. Sci. U. S. A. 100 (1), 253–258.

Gu, Q., Li, Z., Z., Han, J., J., 2012 2012. Generalized fisher score for feature selection. arXiv preprint. p. arXiv:1202.3725.

Guo, X., et al., 2010. Voxel-based assessment of gray and white matter volumes in Alzheimer's disease. Neurosci. Lett. 468 (2), 146–150.

He, Y., Chen, Z.J., Evans, A.C., 2007a. Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. Cereb. Cortex 17 (10), 2407–2419.

He, Y., et al., 2007b. Regional coherence changes in the early stages of Alzheimer's disease: a combined structural and resting-state functional MRI study. Neuroimage 35 (2), 488–500.

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

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8 (5), 393–402.

Hsu, C.-W., Chang, C.-C., Lin, C.-J., 2003. A Practical Guide to Support Vector Classification.

Jack Jr., C.R., et al., 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27 (4), 685–691.

Khazaee, A., Ebrahimzadeh, A., Babajani-Feremi, A., 2015a. Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory. Clin. Neurophysiol. 126 (11), 2132–2141.

Khazaee, A., Ebrahimzadeh, A., Babajani-Feremi, A., 2015b. Application of pattern recognition and graph theoretical approaches to analysis of brain network in Alzheimer's disease. J. Med. Imaging Health Inf. 5 (6), 1145–1155.

Khazaee, A., Ebrahimzadeh, A., Babajani-Feremi, A., 2016. Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease. Brain Imaging Behav. 10 (3), 799–817.

Khazaee, A., Ebrahimzadeh, A., Babajani-Feremi, A., 2017. Classification of patients with MCI and AD from healthy controls using directed graph measures of resting-state fMRI. Behav. Brain Res., 339–350, 322(B).

Kohler, S., et al., 1998. Memory impairments associated with hippocampal versus parahippocampal-gyrus atrophy: an MR volumetry study in Alzheimer's disease. Neuropsychologia 36 (9), 901–914.

Kruschwitz, J.D., et al., 2015. GraphVar: a user-friendly toolbox for comprehensive graph analyses of functional brain connectivity. J. Neurosci. Methods 245, 107–115.

Lerch, J.P., et al., 2008. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. Neurobiol. Aging 29 (1), 23–30.

Markesbery, W.R., 2010. Neuropathologic alterations in mild cognitive impairment: a review. J. Alzheimers Dis. 19 (1), 221–228.

Misra, C., Fan, Y., Davatzikos, C., 2009. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. Neuroimage 44 (4), 1415–1422.

Moradi, E., et al., 2015. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. Neuroimage 104, 398–412.

Petersen, R.C., 2000. Mild cognitive impairment: transition between aging and Alzheimer's disease. Neurologia 15 (3), 93–101.

Reed, T.T., et al., 2009. Proteomic identification of nitrated brain proteins in early Alzheimer's disease inferior parietal lobule. J. Cell. Mol. Med. 13 (8b), 2019–2029.

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

Scheff, S.W., et al., 2011. Synaptic loss in the inferior temporal gyrus in mild cognitive impairment and Alzheimer's disease. J. Alzheimers Dis. 24 (3), 547–557.

Shang, W., et al., 2007. A novel feature selection algorithm for text categorization. Expert Syst. Appl. 33 (1), 1–5.

Smola, A.J., Schölkopf, B., 2004a. A tutorial on support vector regression. Stat. Comput. 14 (3), 199–222.

Smola, A.J., Schölkopf, B., 2004b. A tutorial on support vector regression. Stat. Comput. 14 (3), 199–222.

Stam, C.J., et al., 2007. Small-world networks and functional connectivity in Alzheimer's disease. Cereb. Cortex 17 (1), 92–99.

Supekar, K., et al., 2008. Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. PLoS Comput. Biol. 4 (6), e1000100.

Sylvain, A., Alain, C., 2010. A survey of cross-validation procedures for model selection. Stat. Surv. 4, 40–79.

Tzourio-Mazoyer, N., et al., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15 (1), 273–289.

van den Heuvel, M.P., Hulshoff Pol, H.E., 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur. Neuropsychopharmacol. 20 (8), 519–534.

Wang, K., et al., 2007. Altered functional connectivity in early Alzheimer's disease: a resting-state fMRI study. Hum. Brain Mapp. 28 (10), 967–978.

Wang, J., et al., 2013. Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. Biol. Psychiatry 73 (5), 472–481.

Wolz, R., et al., 2011. Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. PLoS One 6 (10), e25446.

Yang, J., et al., 2013. Minimal-redundancy-maximal-relevance feature selection using different relevance measures for omics data classification. In: IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), April 2013, pp. 246–251.

Yao, Z., et al., 2010. Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease. PLoS Comput. Biol. 6 (11), e1001006.

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

Zhao, X., et al., 2012. Disrupted small-world brain networks in moderate Alzheimer's disease: a resting-state FMRI study. PLoS One 7 (3), e33540.