CrossMark

ORIGINAL RESEARCH

Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification

Chong-Yaw Wee[1] · Sen Yang[3] · Pew-Thian Yap[1] · Dinggang Shen[1,2] ·
for the Alzheimer's Disease Neuroimaging Initiative

Abstract In conventional resting-state functional MRI (R-fMRI) analysis, functional connectivity is assumed to be temporally stationary, overlooking neural activities or interactions that may happen within the scan duration. Dynamic changes of neural interactions can be reflected by variations of topology and correlation strength in temporally correlated functional connectivity networks. These connectivity networks may potentially capture subtle yet short neural connectivity disruptions induced by disease pathologies. Accordingly, we are motivated to utilize disrupted temporal network properties for improving control-patient classification performance. Specifically, a sliding window approach is firstly employed to generate a sequence of overlapping R-fMRI sub-series. Based on these sub-series, sliding window correlations, which characterize the neural interactions between brain regions, are then computed to construct a series of temporal networks. Individual estimation of these temporal networks using conventional network construction approaches fails to take into consideration intrinsic temporal smoothness among successive overlapping R-fMRI sub-series. To preserve temporal smoothness of R-fMRI sub-series, we suggest to jointly estimate the temporal networks by maximizing a penalized log likelihood using a fused sparse learning algorithm. This sparse learning algorithm encourages temporally correlated networks to have similar network topology and correlation strengths. We design a disease identification framework based on the estimated temporal networks, and group level network property differences and classification results demonstrate the importance of including temporally dynamic R-fMRI scan information to improve diagnosis accuracy of mild cognitive impairment patients.

Keywords Mild Cognitive Impairment (MCI) · Resting-state functional MRI (R-fMRI) · Sliding window correlation · Temporal dynamics · Temporal smoothness · Sparse temporal networks

✉ Dinggang Shen
dgshen@med.unc.edu

1 Image Display, Enhancement, and Analysis (IDEA) Laboratory, Biomedical Research Imaging Center (BRIC) and Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

2 Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea

3 Department of Computer Science and Engineering, Arizona State University, Tempe, AZ, USA

Introduction

There is ample literature that suggests the pathological manifestation of Alzheimer's disease (AD) begins many years or decades before any clinical symptom can be observed (Johnson et al. 2006; Thompson and Apostolova 2007; Whitwell et al. 2007). When AD symptoms are observed, significant neurodegeneration has already occurred in the brain, either anatomically or functionally (Li et al. 2012;

Wee et al. 2013; Zhang et al. 2012). Furthermore, if memory loss is the predominant symptom (Dubois and Albert 2004; Economou et al. 2007; Han et al. 2012b; Tabert et al. 2006), patients with Mild Cognitive Impairment (MCI), an intermediate stage between the expected cognitive decline of normal aging and AD, have an increased risk of developing AD. This risk can be reduced if appropriate interventions and treatments are provided during the early stages of MCI. However, for MCI patients changes in the brain are very subtle (Johnson et al. 2006; Li et al. 2011; McEvoy et al. 2009; Smith et al. 2007), therefore the early detection of MCI in neuroimaging data (e.g., MRI and fMRI) can be very challenging.

Neuroimaging data has been widely used in AD and MCI analysis to explore brain structure atrophies and brain function disruptions to enhance our knowledge of the biological underpinnings of AD/MCI. Recently, resting-state functional MRI (R-fMRI) (Smith et al. 2011; Sporns 2011; Stam et al. 2007; Supekar et al. 2008; Wang et al. 2010), which characterizes hemodynamic response related to neural activity, has been widely used because of its ability to detect patients with MCI or AD before the clinical symptoms (Sheline and Raichle 2013). This early detection is achieved by identifying disrupted functional connectivity, which is related to neural activity, using a graph theoretic technique. Through this technique, disruptions in functional connectivity are primarily reflected in the disruptions of network topology and connection strengths, which propels the consideration of using R-fMRI as a potential biomarker to effectively detect early MCI (eMCI).

In conventional network-based analysis, it is assumed that correlations between different brain regions in a typical R-fMRI scanning session of approximately $5 \sim 10$ minutes are not changing over time (i.e., *temporally stationary*) with a repetition time (TR) in the range of $2 \sim 3$ seconds. Because of this assumption, brain region correlations are computed over the entire duration of R-fMRI scan to characterize network connection strengths (Achard et al. 2006; Dosenbach et al. 2007; Fransson and Marrelec 2008; Greicius et al. 2004; Rombouts et al. 2005; Suk et al. 2013; Suk et al. 2014; Wee et al. 2012a; Wee et al. 2012b; Wee et al. 2014). However, recent studies suggest that wealth of information is contained within the temporal features of spontaneous BOLD functional connectivity (Allen et al. 2014; Chang and Glover 2010; Chang et al. 2013; Handwerker et al. 2012; Hutchison et al. 2013a; Jia et al. 2013; Smith et al. 2012; Tomasi et al. 2014) and functional connectivity metrics are able to index changes in temporal patterns of neural activity either in rest or task conditions (Hutchison et al. 2013a; Tomasi et al. 2014). Given the increasing evidence of dynamic functional connectivity during rest and its important for characterizing brain's intrinsic functional organization, multiple network

estimation technique that can characterize fluctuations of whole-brain functional connectivity becomes necessary to better reveal connectivity patterns with *coherent dynamics*. In this paper, we suggest to use our recently proposed sparse learning-based method to jointly estimate multiple temporally correlated networks, which simultaneously preserves the sparsity and coherent dynamics (temporal smoothness) among networks.

Recently, disease-related alterations in the dynamic properties of functional connectivity have been reported (Jones et al. 2012; Saköglu et al. 2010), further suggesting a neural origin and raising the intriguing possibility that temporal features of functional connectivity could serve as a disease biomarker. Inspiring by the advantages of the previous mentioned findings, we hypothesize that the dynamic patterns of whole-brain functional connectivity of R-fMRI and thus their network properties derived based on graph theory may be disrupted by disease pathologies. Motivated by this hypothesis, we design a novel framework that utilizes a series of temporal functional connectivity networks derived from an R-fMRI scan to possibly improve disease identification performance between eMCI patients and healthy controls. The novelty is three-fold: 1) a temporal series of sparse functional connectivity networks (*sparse temporal networks* for short) are constructed to estimate the time varying dynamic properties of functional connectivity over the duration of R-fMRI scan, 2) sparse temporal networks are estimated simultaneously to preserve temporal smoothness, i.e. temporally adjoint sparse networks have similar topology and connection strengths, 3) a network measure-based framework is proposed to improve patient-control identification performance.

The rest of the paper is organized as follows: Section Material and methods provides detailed descriptions on the proposed temporal network-based disease identification framework and the dataset used in this study. Performance of the proposed framework is evaluated extensively in Section Experimental results. Findings, methodological issues, and limitations of the proposed framework are discussed in Section Discussion. Section Conclusions concludes this study.

Material and methods

Participants and data acquisition

Data used in this study were obtained from the ADNI dataset.[1] The primary goal of ADNI was to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological

---

[1] http://www.adni-info.org

assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. A more detailed description of the ADNI dataset is provided in Appendix A.

Twenty nine eMCI subjects (13F/16M) and 30 normal controls (NCs) (17F/13M) were obtained from ADNI 2 dataset. Subjects from both groups are age-matched ($p = 0.6174$), with mean age in terms of year for eMCI and NC groups as 73.6 ± 4.8 and 74.3 ± 5.7, respectively. All subjects were scanned at different centers using 3.0 T Philips Achieva scanners with the following parameters: TR/TE = 3000/30 mm, flip angle = 80°, imaging matrix = 64 × 64, 48 slices, 140 volumes, and voxel thickness = 3.3 mm. Standard R-fMRI preprocessing procedure was performed using the SPM8 software package.[2] The first 10 acquired R-fMRI volumes of each subject were initially discarded before any further processing to ensure magnetization equilibrium. The remaining 130 volumes were then corrected for the staggered order of slice acquisition that was used during echo-planar scanning. The correction ensures the data on each slice correspond to the same point in time. The interpolated time point was chosen as the TR/2 time to minimize relative errors across each TR in the study. After acquisition time delay correction, the slice timing corrected R-fMRI time-series of each subject were realigned using a least squares approach and a rigid body spatial transformation. The first volume of each subject was used as the reference to which all subsequent volumes were realigned. This step removed the head-motion artifacts in the R-fMRI time-series. There were no significant group differences in head-motion for all participants used in the study. After realignment, the volumes were resliced such that they match the first volume voxel-by-voxel. R-fMRI images were then normalized to the MNI space with resolution of $3 \times 3 \times 3$ mm³.

To further reduce the effects of nuisance signals before inferring functional connectivity, regression of ventricle and WM signals as well as six head-motion profiles was performed (Van Dijk et al. 2010). Given the controversy of removing the global signal in the post-processing of R-fMRI data (Fox et al. 2009; Murphy et al. 2009), we did not regress the global signal out (Supekar et al. 2008; Achard et al. 2006; Lynall et al. 2010). The regressed R-fMRI images were parcellated into 116 regions-of-interest (ROIs) according to the Automated Anatomical Labeling (AAL) template (Tzourio-Mazoyer et al. 2002). Prior to functional connectivity estimation, mean R-fMRI time series of each ROI was band-pass filtered ($0.01 \leq f \leq 0.08$ Hz).

Proposed framework

The proposed sparse temporal network-based disease identification framework is graphically shown in Fig. 1. Specifically, for each ROI entire mean time series of an R-fMRI scan is decomposed into multiple overlapping sub-series using a sliding window (Allen et al. 2014; Chang and Glover 2010; Handwerker et al. 2012; Smith et al. 2012) where its length shorter than the duration of the R-fMRI scan. For each R-fMRI sub-series, correlation coefficients between different brain regions are computed to generate one functional connectivity network. Sparse temporal networks are estimated simultaneously from these sub-series using fused multiple group LASSO (FMGL) to preserve the temporal smoothness between adjoining networks. Network measure, i.e., local clustering coefficients, are calculated from each sparse temporal network. Local clustering coefficients of all temporal networks are then concatenated into a long feature vector to train a Support Vector Machine (SVM) classifier.

Functional connectivity estimation

Generation of multiple R-fMRI sub-series using the sliding window approach is graphically illustrated in Fig. 2. In particular, given an R-fMRI time series with M temporal image volumes, $K = \lfloor (M - N)/s \rfloor + 1$ is the number of sub-series that can be generated, where N denotes the length of sliding window and s denotes the translation step size. The k-th sub-series is represented in the form of a matrix $\mathbf{X}^{(k)} = \{\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, \ldots, \mathbf{x}_P^{(k)}\} \in \mathbb{R}^{N \times P}, k = 1, \ldots, K$, where $P = 116$ is the total number of regions-of-interest (ROIs), and $\mathbf{x}^{(k)} = [x^{(k)}(1), x^{(k)}(2), \ldots, x^{(k)}(N)]^T$ is the k-th sub-series of N image volumes for a particular ROI. A symmetric connectivity matrix $\mathbf{C}^{(k)} = [C_{ij}^{(k)}] \in \mathbb{R}^{P \times P}$ can be constructed using $\mathbf{X}^{(k)}$, where each element in the connectivity matrix defines the correlation strength between two different ROIs. Specifically, the correlation strength is defined as

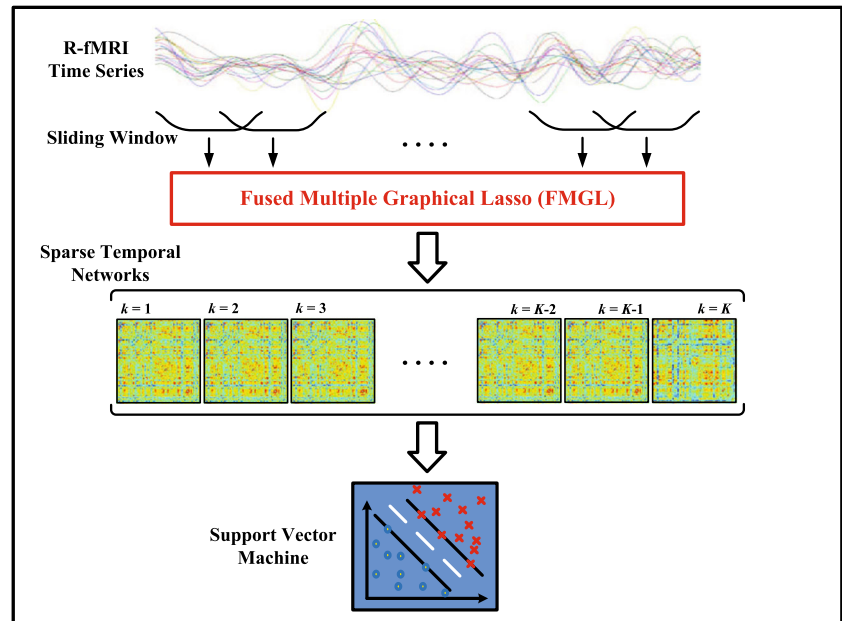$$C_{ij}^{(k)} = corr\left(\mathbf{x}_i^{(k)}, \mathbf{x}_j^{(k)}\right), \tag{1}$$

where i and j are two different brain regions, and $corr(\cdot)$ computes the correlation between region i and region j. For instance, if $i = 1$ and $j = 2$ (1) would compute the correlation strength between regions 1 and 2 for sub-series k. It is noteworthy that the conventional stationary-based functional connectivity matrix can be estimated by replacing $corr(\cdot)$ in Eq. 1 with Pearson or partial correlations and $K = 1$.

Graphical model

Assuming that the image volumes within each R-fMRI sub-series are identically distributed with P-variate Gaussian

Fig. 1 The proposed sparse temporal network-based disease identification framework

distribution with zero mean and positive definite covariance matrix $\sum^{(k)}$. Due to sparse nature of brain network, there should be many conditionally independent connection pairs. These null connections are denoted as zero elements in the precision matrix, i.e., network matrix $\Theta^{(k)} = (\sum^{(k)})^{-1}$. If the covariance matrix of each $\mathbf{X}^{(k)}$ is $\mathbf{S}^{(k)} = (1/N)(\mathbf{X}^{(k)})^T\mathbf{X}^{(k)}$, then the negative log likelihood of $\Theta^{(k)}$ is given as

$$\mathscr{L}(\Theta) = \sum_{k=1}^{K} \Big(-\log det(\Theta^{(k)}) + tr(\mathbf{S}^{(k)}\Theta^{(k)}) \Big), \qquad (2)$$

where $\Theta = \{\Theta^{(1)}, \ldots, \Theta^{(K)}\}$. Minimizing (2) leads to the maximum likelihood estimate (MLE) of $\hat{\Theta}^{(k)} = (\mathbf{S}^{(k)})^{-1}$, which is unlikely to be sparse. Equation 2 can then be solved by minimizing the penalized negative log likelihood as

$$\min_{\Theta^{(k)} \succ 0, k=1,\ldots,K} \mathscr{L}(\Theta) + Q(\Theta), \qquad (3)$$

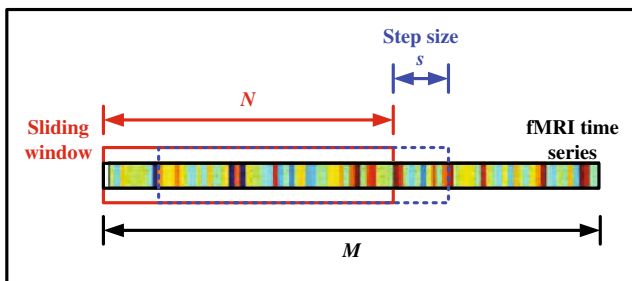with $Q(\Theta)$ to be defined in following subsections.

Fig. 2 Generation of R-fMRI sub-series using sliding window approach

Sparse inverse covariance estimation

To model brain connectivity matrix, which has been proven to be sparse (Kötter and Stephen 2003; Sporns et al. 2004), a "sparsity" constraint can be imposed on the MLE of an precision matrix using sparse inverse covariance matrix (SICE) (Huang et al. 2010), which also known as Gaussian graphical models or graphical LASSO. In SICE, the second term in Eq. 3, $Q(\Theta)$, is defined as (Huang et al. 2010)

$$Q(\Theta) = \lambda \sum_{k=1}^{K} \sum_{i \neq j} \left| \theta_{i,j}^{(k)} \right|, \qquad (4)$$

where λ is a non-negative regularization parameter. ℓ_1-norm in Eq. 4 is imposed to the elements of K networks, producing matrices with different non-zero locations.

Group graphical lasso

Estimating multiple networks individually fails to capture the temporal smoothness among adjoining networks. This problem can be partially solved by using joint estimation methods such as Group Graphical Lasso (GGL) (Danaher et al. 2012), which estimates multiple networks simultaneously. In GGL, the last term of Eq. 3 is defined as (Danaher et al. 2012)

$$Q(\Theta) = \lambda_1 \sum_{k=1}^{K} \sum_{i \neq j} \left| \theta_{i,j}^{(k)} \right| + \lambda_2 \sum_{i \neq j} \sqrt{\sum_{k=1}^{K} (\theta_{i,j}^{(k)})^2}, \qquad (5)$$

where λ_1 and λ_2 are non-negative regularization parameters. The second term of Eq. 5, a $\ell_{2,1}$-norm constraint, is applied to the (i, j) elements across all networks. This penalization

forces an identical pattern of sparsity across all networks, i.e., same non-zero locations in K estimated networks, overlooking temporal smoothness between adjoining temporal networks.

Fused multiple graphical lasso

Drawbacks of SICE and GGL methods can be solved by employing our recently proposed sparse learning algorithm, i.e., Fused Multiple Graphical LASSO (FMGL) algorithm (Yang et al. 2012). FMGL simultaneously estimates a series of temporal networks by replacing the last term in Eq. 3 with following term (Yang et al. 2012)

$$Q(\Theta) = \lambda_1 \sum_{k=1}^{K} \sum_{i \neq j} \left| \theta_{i,j}^{(k)} \right| + \lambda_2 \sum_{k=1}^{K-1} \sum_{i \neq j} \left| \theta_{i,j}^{(k)} - \theta_{i,j}^{(k+1)} \right|.$$

In particular, an ℓ_1-norm penalization is imposed to induce network sparsity (Achard and Bullmore 2007; Chen et al. 2012; Friedman et al. 2008) and a fused regularization is imposed to preserve the temporal smoothness by encouraging $\Theta^{(k)}$ to have similar topology and correlations strengths to its adjoining networks (Tibshirani et al. 2005). FMGL was implemented using an in-house software.

Network analysis

One of the important properties of human brain network is small-worldness. Since multiple sparse networks are generated to extract the temporally dynamic characteristics of brain functional connectivity, it is of great important to ensure that the generated networks are small-world networks instead of random or lattice-like networks. Accordingly, we use two network measures, i.e., small-world coefficient (Humphries et al. 2006; Humphries and Gurney 2008) and small-world measurements (Telesford et al. 2011), to evaluate the small-worldness of the sparse temporal networks. For evaluation based on the small-world coefficient, we compared the clustering coefficient and characteristic path length of each sparse temporal network to the same metrics from equivalent random networks. The small-world coefficient (σ) is defined as (Humphries et al. 2006; Humphries and Gurney 2008)

$$\sigma = \frac{C/C_{rand}}{L/L_{rand}} = \frac{\gamma}{\lambda},$$

where C and L denote the clustering coefficient and characteristic path length of a sparse temporal network, and C_{rand} and L_{rand} denote the clustering coefficient and characteristic path length of equivalent random networks, and γ and λ denote the normalized clustering coefficient and normalized

characteristic path length, respectively. The absolute clustering coefficient for a network G is defined as

$$C = \frac{1}{N} \sum_{i \in G} C_i,$$

where $C_i = \frac{2E_i}{d_i(d_i-1)}$ is the ratio of the number of existing connections to the number of all possible connections, E_i is the number of edges, and d_i is the number of nodes that are directly connected to i-th node with an edge in a subnetwork G_i.

The characteristic path length, which measures the overall routing efficiency, of a network G is defined as the average of the shortest absolute path lengths between the nodes

$$L = \frac{1}{N(N-1)} \sum_{i \in G} \sum_{j \neq i \in G} \min\{L_{i,j}\},$$

where $\min\{L_{i,j}\}$ denotes the shortest minimum path length between the i-th node and the j-th node. The absolute path length is the number of edges included in the path connecting two nodes. A network is considered as small-world if $C \gg C_{rand}$ and $L \approx L_{rand}$, resulting in $\sigma > 1$.

For evaluation based on the small-world measurement (ω), we compared the clustering coefficient of each sparse temporal network to that of equivalent lattice networks, and compared characteristic path length of each sparse temporal network to that of equivalent random networks (Telesford et al. 2011)
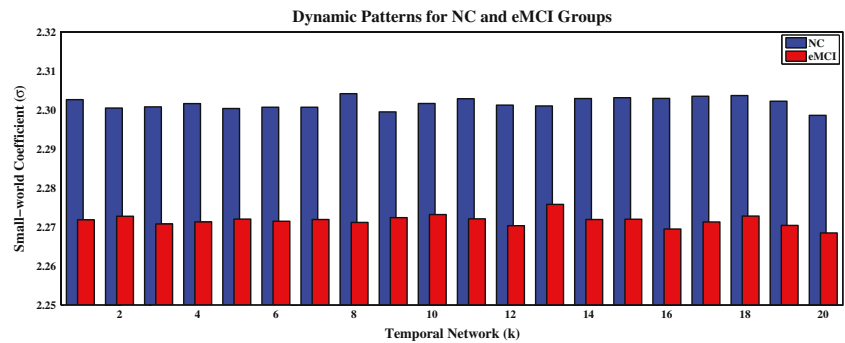
$$\omega = \frac{L_{rand}}{L} - \frac{C}{C_{latt}},$$

where C_{latt} denotes the clustering coefficient of a lattice network. The value of ω is restricted to the interval [-1 1], and a network is considered small-world if $\omega \approx 0$. Positive value indicates a network with more random characteristics (i.e., $L \approx L_{rand}$, and $C \ll C_{latt}$), while negative value indicates a network with more lattice-like characteristics (i.e., $L \gg L_{rand}$, and $C \approx C_{latt}$). It is worth noting that this metric is independent of network size and is more accurate in quantifying small-worldness of a network compared to the small-world coefficient (Telesford et al. 2011).

Evaluation via leave-one-out cross-validation

In this study, we employed SVM with a simple linear kernel based on the LIBSVM library (Chang and Lin 2011) to evaluate the discriminative power of the features derived from sparse temporal networks. The optimal SVM models as well as an unbiased estimation of the generalization classification performance were achieved via a nested leave-one-out cross-validation scheme due to limited number of sample size. Specifically, for N total number of subjects involved in the study, one was first left out for testing, and the remaining

Fig. 3 Small-world coefficient (σ) of dynamic sparse temporal networks for eMCI and NC groups with respect to random networks

$N-1$ were used for constructing the optimal SVM model. From these $N-1$ samples, $N-1$ different training subsets were formed by each time leaving one more sample out, i.e., $N-2$ subjects in each training subset. For each training subset, functional connectivity construction, feature extraction and feature selection were performed. The performance of each combination of SVM parameters along with the selected features was evaluated using the second left out subject. The combination that gives the best performance was used to construct the optimal SVM model for future classification. This procedure was repeated $N-1$ times, once for each training subset. When the completely unseen test sample was to be classified, all $N-1$ classifiers were used, and the final classification decision was determined via majority voting. This process was repeated N times, each time leaving out a different subject, finally leading to an overall cross-validation classification accuracy. In this study, the optimal λ values for all lasso-based methods were determined via grid search.

Experimental results

Temporal network analysis

Figures 3 and 4 show the small-world coefficient and small-world measurement, respectively, of temporal networks with $N = 90$ and $s = 2$ for the eMCI and the NC groups with respect to random and lattice networks of equivalent

distribution. The values $\sigma > 1$ and $\omega \approx 0.3$ (Telesford et al. 2011), indicate the small-worldness of the temporal networks. The small-world coefficients in the eMCI group are smaller than the NC group and show different dynamic patterns across temporal networks (Fig. 3). On the other hand, the eMCI group shows larger small-world measurement across temporal networks than the NC group (0.31 *vs.* 0.30). However, different from small-world coefficient, both the eMCI and NC groups do not demonstrate any dynamic pattern across temporal networks in terms of small-world measurement (Fig. 4).

To further analyze network properties, we computed the clustering coefficients and characteristic path lengths, and their normalized counterparts, for every temporal network in both groups. The results are shown in Figures 5, 6, 7, 8, respectively, for the temporal networks with $N = 90$ and $s = 2$. We observed distinct dynamic patterns for the eMCI and NC groups in terms of clustering coefficient and its normalized counterpart when BOLD signals were studied at a smaller time scale. Specifically, clustering coefficients for the eMCI group are smaller than NC groups for the first eleven temporal networks, and are larger for the subsequent networks, except for the thirteenth network. On the other hand, the normalized clustering coefficient also showed dynamic pattern but with consistently smaller values in the eMCI group than in the NC group. However, we cannot observe any dynamic pattern for characteristic path length and its normalized counterpart. Characteristic path length is consistently larger in the eMCI group compared with

Fig. 4 Small-world measurement (ω) of dynamic sparse temporal networks for eMCI and NC groups with respect to random and lattice networks
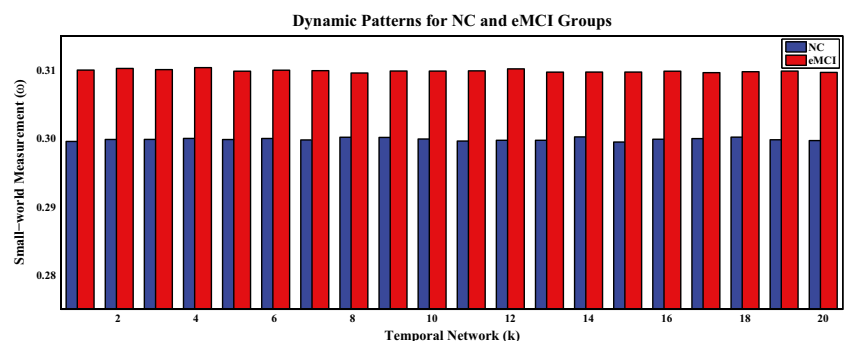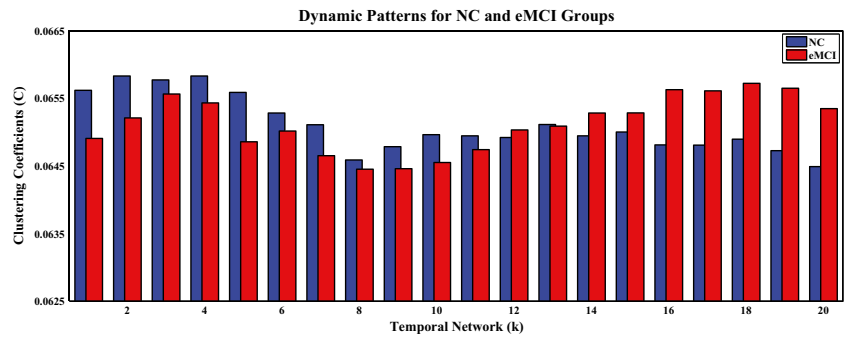
Fig. 5 Clustering coefficients of temporal networks for eMCI and NC groups with $s = 2$ and $N = 90$

the NC group whereas its normalized counterpart showed an opposite trend with smaller values in the eMCI group. According to these observations, we have decided to only extract the local clustering coefficient from the temporal networks as our features for constructing SVM classifiers in this study.

Classification performance

We compared the eMCI identification performance of the proposed framework with the conventional temporally stationary-based (i.e., Partial correlation (PAC) and Pearson correlation-based (PEC)) methods, sparse inverse covariance matrix (SICE), and the graph graphical LASSO (GGL). For PAC and PEC, entire R-fMRI time series is used to compute one connectivity network. For the SICE and GGL approaches, a series of temporal networks are constructed using ℓ_1 and $\ell_1 + \ell_{2,1}$ penalizations, respectively. Performances of the comparison methods are summarized in Table 1.

The proposed framework performed the best by yielding an accuracy of 79.7 %, which is at least 6.5 % improvement compared to the second best GGL-based framework. The conventional PAC- and PEC-based frameworks performed the worst with classification accuracies of 62.7 % and 66.1 %, respectively. The proposed, GGL-, and SICE-based frameworks showed a significantly higher sensitivity value when compared to PAC- and PEC-based frameworks. A cross-validation estimation of the

generalization performance shows an area of 0.792 under the receiver operating characteristic curve (AUC), indicating good diagnostic power of the proposed framework.

Robustness of the proposed framework

In the proposed framework, there are two factors that may influence the network topology and correlation strengths of the estimated sparse temporal networks, i.e., sliding window length (N) and translation step size (s), finally the classification performance. To evaluate the robustness of the proposed framework with respect to these two parameters, we repeated eMCI identification using different window lengths ($N = 50, 70, 90, 110$) and step sizes ($s = 1, 2, 4, 8, 10$), and their results are summarized in Fig. 9. The proposed framework preformed relatively consistent with respect to the window length and step size. When large N and s are used, the proposed framework performed similarly to the conventional PAC- and PEC-based frameworks.

Most discriminative regions

Brain regions that were selected with the highest frequency during the feature selection step based on the training subjects are considered as the most discriminative regions for eMCI identification. These regions (for $N = 90$ and $s = 2$) are mostly the components of the Default Mode Network (DMN), i.e., middle frontal gyrus, bilateral

Fig. 6 Normalized clustering coefficients of temporal networks for eMCI and NC groups with $s = 2$ and $N = 90$
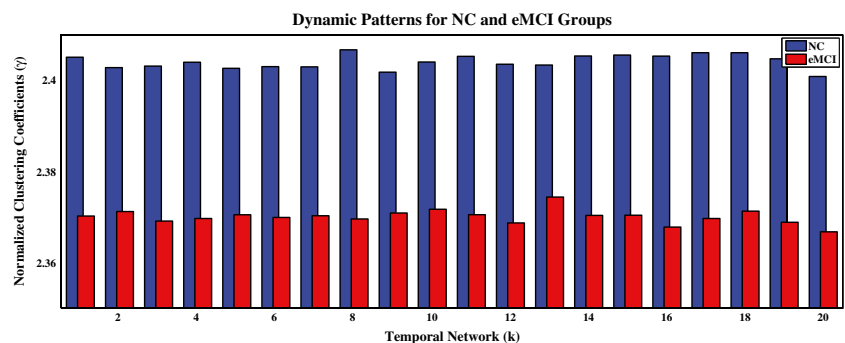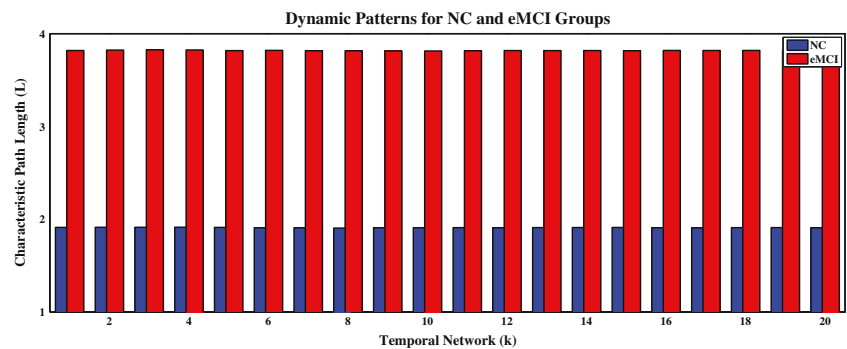
Fig. 7 Characteristic path lengths of temporal networks for eMCI and NC groups with $s = 2$ and $N = 90$

superior orbitofrontal cortex, posterior cingulate gyrus, precuneus, and hippocampus. Other selected regions include the frontal lobe (superior frontal gyrus (dorsal) and inferior orbitofrontal cortex), the temporal lobe (superior temporal pole and inferior temporal), the occipital lobe (inferior occipital gyrus, cuneus and bilateral fusiform gyrus), parietal lobe (postcentral gyrus), subcortical regions (caudate and amygdala), and the cerebellum regions (lobule IV, V of cerebellar hemisphere, Lobule VI of cerebellar hemisphere, and lobule X of cerebellar hemisphere (flocculus)). These selected regions are shown in Fig. 10. Note that many regions across the brain have been selected for good classification accuracy, particularly the cerebellum regions which are seldom detected in previous studies. For the case of the fully-connected, stationary-based Pearson correlation method, the selected regions include parahippocampus, fusiform gyrus, olfactory, superior frontal gyrus (dorsal), inferior temporal, middle frontal gyrus, and temporal pole (middle). All these regions were selected by our method, although some of them, namely parahippocampus, olfactory, and temporal pole (middle), were chosen with relatively low selection frequency.

Discussion

We propose a novel framework that utilizes the temporally dynamic information embraced within an R-fMRI scan to improve disease identification performance. Sparse

temporal networks are constructed from the R-fMRI time series, and graph theoretical analysis has been performed to explore their network properties. Clustering coefficient, which shows different dynamic patterns between two clinical groups across sparse temporal networks, was then used to train a linear SVM classifier. Performance of the proposed framework was evaluated on 59 subjects obtained from the ADNI 2 dataset, demonstrating the great importance of including temporally dynamic information for better characterization of brain activity patterns when compared to the conventional approaches that are based on the temporal stationary assumption.

Temporal networks analysis

The human brain is a complex and highly connected dynamic system with various important topological properties for its optimal functioning (Bullmore and Sporns 2009; He and Evans 2010; Rubinov and Sporns 2010; Stam 2010). In the brain, clustered regions favor modular information processing; however, several regions in different clusters may have long-range connections that support efficient functional routing (Watts and Strogatz 1998). Optimal brain networks, which show small-worldness property, are normally characterized by high clustering coefficients like regular lattices, yet low characteristic path lengths like randomly generated networks (Watts and Strogatz 1998).

Unlike conventional stationary correlation analysis, which generates only one connectivity network per

Fig. 8 Normalized characteristic path lengths of temporal networks for eMCI and NC groups with $s = 2$ and $N = 90$
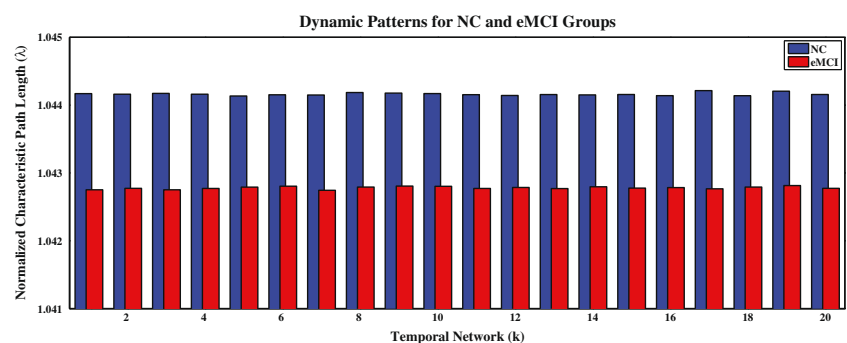
Table 1 eMCI identification performance of the PAC-, PEC-, SICE-, GGL- and the proposed sparse temporal network-based frameworks with $N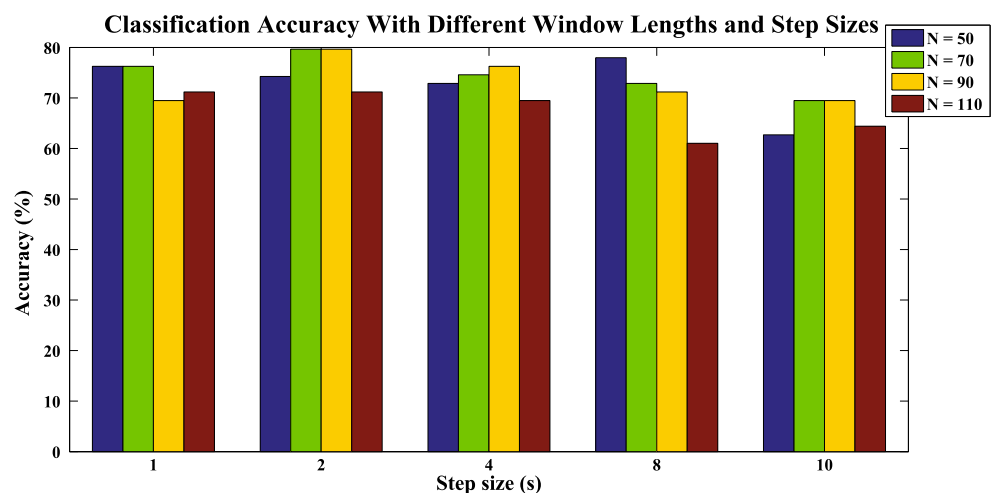 = 90$ and $s = 2$. (ACC = ACCuracy; AUC = Area Under ROC Curve; SEN = SENsitivity; SPE = SPEcificity; Youden = Youden Index; F-score = F score; BAC = Balance ACcuracy)

Method	ACC	AUC	SEN	SPE	Youden	F-score	BAC
PAC	0.6271	0.6598	0.6552	0.6000	0.2552	0.6333	0.6276
PEC	0.6610	0.6138	0.5517	0.7667	0.3184	0.6154	0.6592
SICE	0.6780	0.7189	0.6897	0.6667	0.3563	0.6780	0.6782
GGL	0.7288	0.7315	0.7586	0.7000	0.4586	0.7333	0.7293
Proposed	0.7966	0.7920	0.7586	0.8333	0.5920	0.7857	0.7920

R-fMRI scan, sliding window correlations at different time scales exhibits temporal variations during the R-fMRI scanning session (Allen et al. 2014; Chang and Glover 2010; Handwerker 2012; Hutchison et al. 2013a; Smith et al. 2012), implying constant activations of the brain even at rest. However, it remains unanswered whether 1) the generated temporal networks are biophysically meaningful and exhibit small-world properties as actual human brain networks, 2) there is any between-group difference in terms of temporal network properties, and 3) the differences, if any, can be employed to distinguish eMCI patients from NCs. Therefore, we evaluated the small-world property of the sparse temporal networks by computing the small-world coefficient and the small-world measurement. Consistently large small-world coefficients and close to 0.3 small-world measurements, for both eMCI and NC groups, demonstrate the reliability of the proposed method in generating networks with properties similar to actual brain networks (Telesford et al. 2011). It is worth noting that smaller small-world coefficients and larger small-world measurements indicate that functional connectivity networks in the MCI group are more random compared with the NC group, a common characteristic of most neurodegenerative disorders.

In addition, we observed different dynamic patterns for eMCI and healthy brains in the clustering coefficient and its normalized counterpart. The variations are more obvious in the original clustering coefficient compared to its normalized counterpart. Smooth yet different dynamic patterns in the eMCI and NC groups indicates that 1) changing of brain activity is relatively smooth when the sliding window approach is used due to the overlapping of the R-fMRI time frames, and 2) different dynamic patterns for the eMCI and NC groups can potentially be used as effective biomarkers to identify eMCI patients at individual level. Consistently smaller normalized clustering coefficients are observed in the eMCI group compared with NC group, in line with previous R-fMRI studies (Binnewijzend et al. 2012; Han et al. 2012a; Supekar et al. 2008), although the findings in these studies were obtained based on the conventional stationary assumption.

Temporal networks of the eMCI group showed larger characteristic path length when compared with the NC group, although an opposite trend was observed for its normalized counterpart. Theoretically, a short path length ensures the effective integrity and rapid information propagation between and across distinct regions of the brain that are believed to constitute the basis of cognitive processing

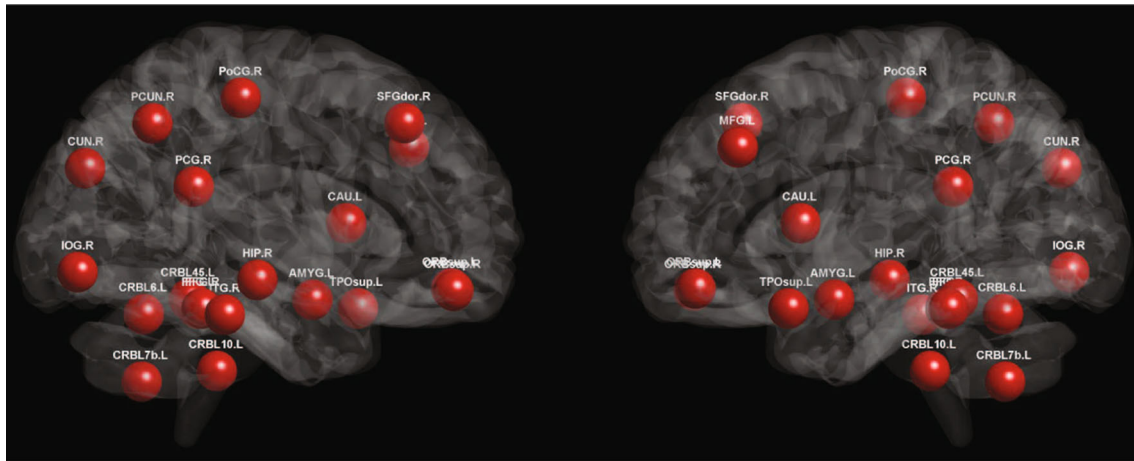**Fig. 9** eMCI identification performance with respect to different window lengths and step sizes

Fig. 10 Most discriminative regions that were selected for eMCI classification. The colors indicate different ROIs

(Sporns and Zwi 2004). Since neurodegenerative diseases such as AD and MCI are considered to be disconnection syndromes (Delbeuck et al. 2003), larger characteristic path length reflects disrupted neuronal integration between distant regions, in line with previous R-fMRI reports, which show relatively spare long-distant and relatively dense short-distant functional connections in MCI and AD groups (Buldú et al. 2011; Liu et al. 2012; Seo et al. 2013; Sorg et al. 2007; Wang et al. 2007). Progression of clinical stages, from mild to severe, may cause more impairment or disconnection of long-distant connections, and thus possibly encourage the establishment of short-distant connections within cluster as alternative paths to preserve information transmission between two distant regions. However, establishment of short-distant connections may introduce abnormal clusters that increases the risk of generating an uncontrolled or random flow of information through the entire network (Kaiser and Hilgetag 2006; Kaiser 2007; Sanz-Arigita et al. 2010). The decrease in normalized clustering coefficient and characteristic path length, is also reflected by the larger small-world measurement in the eMCI group, suggesting a more random-like network topology.

Interpretation of classification results

Compared with the research (Greicius et al. 2004; Rombouts et al. 2005; Wang et al. 2007) that focuses on group-wise differences, to the best of our knowledge, the current work is the first study that uses temporal dynamics of an R-fMRI scan for disease identification. Although group-wise analysis may identify network changes associated with disease, it is of limited utility at an individual level. In contrast, our goal is to develop an individual-based framework that utilizes the temporal dynamics extracted using one well-known network measure, i.e., clustering

coefficients, for disease identification. Promising results using leave-one-out cross-validation indicate the superiority of the proposed temporal network-based framework in eMCI identification when compared with the PAC-, PEC-, SICE, and GGL-based frameworks. Also, the proposed and GGL-based frameworks, both of which include temporally dynamic information, performed better than all other methods, indicating the importance of using temporally dynamic patterns for disease classification, at least in the context of a typical R-fMRI scanning session lasting 5 to 10 minutes.

It has been suggested that correlation values within and between connectivity networks stabilize within 4 ∼ 5 minutes of fMRI data (Van Dijk et al. 2010), implying that most studies are adequately sampling the network activity despite relatively few data points, even under variety of behavioral states (e.g., eyes closed, open, or open and fixating) during R-fMRI acquisition. Several studies (Greicius 2008; Menon 2011) demonstrated subtle yet significant functional connectivity differences between normal and diseased individuals within similar scanning duration. These findings support, to some extent, our findings where the proposed framework always performs better when sliding window length range is around 3.5 minutes ($N = 70$) and 4.5 minutes ($N = 90$), and with relatively smaller step size (s). The proposed framework performs slightly inferior when the window length (N) and step size (s) are relatively large. With large N and s, it is possible that neural activity changes that occur at time scales shorter than the window length and/or step size are overlooked, causing the performance of the proposed framework approaching to the conventional stationary-based approaches. This situation is particularly obvious when the window length is similar to the length of entire R-fMRI scan ($M = 130$ in this study). These findings may imply that individuals who are affected by the disease may experience temporal changes, which

might be overlooked by the conventional stationary-based approaches.

Most discriminative regions

All the regions that were selected by the conventional fully-connected, stationary-based PC method were also selected by our method as discriminative regions, although with varying selection frequency. More precisely, we observed three different conditions for region selection: a) Selection by the PC method and with relatively high frequency by our method, which indicates a region has informative stationary features; b) Selection by the PC method but with relatively low frequency by our method, which indicates a region's stationary and dynamic features are less informative; and c) Selection with high frequency by our method, without selection by the PC method. In the last case, the dynamic features derived from selected regions greatly contribute to achieving high eMCI identification accuracy. These observations suggest that dynamic patterns of brain regions, which are distributed across sparse temporal networks, provide novel features for detecting possible functional connectivity abnormalities in eMCI patients within an R-fMRI scan. The phenomena of identifying new network nodes (brain regions) that were unobserved at full time series has been reported in anesthetized macaques and awake humans study at shorter window lengths (Hutchison et al. 2013b).

DMN components, which are selected as the most discriminative regions in eMCI identification, exhibit high correlation with the regions in the higher-level cognitive functional networks (e.g., temporo-parietal cortex, inferior frontal cortex, anterior cingulate, and orbital fronto-insular cortices) (Corbetta and Shulman 2002; Damoiseaux et al. 2006; Seeley et al. 2007). Posterior cingulate cortex, an important component of DMN, demonstrates reduced functional connectivity with the orbitofrontal cortices and middle frontal gyrus, while increased connectivity with the inferior frontal, fusiform, and precentral regions (Han et al. 2012a) in MCI patients. In addition, MCI patients experience reduced activation compared to NCs in several regions of the occipito-temporal cortex, a small portion of the left inferior frontal gyrus, while increased activation in bilateral anterior cingulate and precueus, with the largest activation differences centered on the fusiform gyri (Gold et al. 2010). Similar fMRI connectivity disruptions have been observed in the posterior hippocampal, and parahippocampal (Hämäläinen et al. 2007). Also, correlation between precuneus and DMN regions has been previously used to distinguish MCI patients from NCs, achieving relatively good accuracy (Rombouts et al. 2005). In our previous R-fMRI study, orbitofrontal cortex, temporal pole, posterior cingulate gyrus, precuneus, and amygdala have been selected as important regions for accurate MCI identification (Wee et al. 2012a). All these findings, either obtained at task-based or task-free fMRI analysis, demonstrate widespread of functional connectivity disruptions across MCI brains even at its prodromal phase (Fennema-Notestine et al. 2009).

Parts of the cerebellum regions of MCI patients underwent structural degenerative (Baloyannis et al. 2000; Fukutani et al. 1997; Li et al. 1994; Sjöbeck and Englund 2001; Wang et al. 2002), including reduction of Purkinje cell density, atrophy of the molecular and granular cell layer, and concentration of amyloid plaques in the cerebellar cortex (Wegiel et al. 1999). In a study that manually partitioned human cerebellum into four substructures, i.e., anterior, superior posterior, inferior posterior lobes, and corpus medullare on each hemisphere, the posterior cerebellar lobe experience significant volume reduction in AD patients (Thomann et al. 2008), leading to poorer cognitive performance. In terms of functional study, the left cerebellum is activated in memory encoding task (Kircher et al. 2007), implying its association with recognition memory. Another event-related study suggested that the cerebellar regions contribute to human recognition (Weis et al. 2004). Compared to the structural and task-based functional studies, the cerebellum regions were always been ignored in the R-fMRI-based AD/MCI analysis. Our findings of detecting cerebellum regions as the important features for accurate eMCI identification may suggest the disruption of cerebellum functions in AD/MCI brains, thus merit for more attention in the future R-fMRI-based AD/MCI study.

Methodological issues/limitations

The current study has three main limitations. First, the proposed framework is based on the temporal dynamics in the R-fMRI scan, however, neurobiological underpinnings and mechanisms of these temporal dynamics are still largely unknown (Smith et al. 2011). It is possible that sliding window correlation variations in the sparse temporal networks are associated with changes in brain state or arousal. However, we cannot rule out the possibility that these changes are driven simply by the noise (e.g. motion, physiological, and scanner noises) (Handwerker et al. 2012). Hence, careful interpretations of the results are required, and more works shall be performed to improve our understanding on these issues.

Second, the choice of window size is important for the sliding window approach. The window length should be long enough to permit robust estimation of functional connectivity and short enough to allow the detection of interesting short-term fluctuations (Sakoğlu et al. 2010). As the window shrinks, the signal-to-noise ratio decreases

since there are fewer time points available for connectivity estimation. The estimation variability may increase due to signals of higher frequencies, physiological noise, and head motion (Hutchison et al. 2013a). On the other hand, as the window expands, brief neural activities tend to be smoothed out. Since brain activation patterns vary with diseases, it is therefore important to vary the window size to capture brain activities of different time scales.

Third, the diagnosis criteria of eMCI are still unclear and standard where they may be similar to the early diagnosis of other neurodegenerative dementia such Parkinson disease and frontotemporal dementia. This may rise reliability issue of the eMCI diagnosis. Additionally, subjects used in this study are at the very beginning stage of AD and there are very few studies focused in this type of cohort. Due to this, the selected most discriminative regions can only be compared with the AD/MCI findings using the conventional stationary approach. However, these selected brain regions spread over the whole brain, thus future analysis is required to verify these findings.

Conclusions

In this paper, we propose to utilize the temporally dynamic information of an R-fMRI scan for disease identification. This work is motivated by the variabilities of sliding window correlation strengths induced by the neural interactions that happen within time scales that are shorter than a typical R-fMRI scanning session. Unlike the conventional approach that assumes functional connectivity is temporally stationary, we estimate temporal networks from a set of R-fMRI sub-series using a sliding window approach to extract the between-region neural interactions across time. We utilize a fused LASSO regularization-based sparse learning algorithm to jointly estimate temporal networks, encouraging similar network topology and correlation strengths between temporally adjoining networks. Clustering coefficients are calculated from each network to extract temporally dynamic patterns. Promising results on eMCI identification suggest that the temporal dynamic information encapsulated in an R-fMRI scan is crucial for accurate diagnosis of neurological disorders. In the future, we will extend the proposed disease identification framework for diagnosis of neurodevelopmental disorders such as autism and attention deficit hyperactivity disorder (ADHD), which can benefit greatly from early detection.

Financial disclosures

There are no conflicts of interest including any financial, personal, or other relationships with people or organizations for any of the coauthors (Chong-Yaw Wee, Sen Yang, Pew-Thian Yap, and Dinggang Shen) related to the work described in the article.

Appendix A: The Alzheimer's disease neuroimaging initiative (ADNI)

The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million as a landmark study that gathered and analyzed thousands of brain scans, genetic profiles and biomarkers in blood and cerebrospinal fluid (CSF). Although the original goal was to define biomarkers for use in clinical trials to determine the best way to measure treatment effects of AD, the goal has been expanded to find more sensitive and accurate methods to detect AD at a pre-dementia stage and mark its progress through biomarkers. ADNI 1 involves scientists at 59 research centers, 54 in USA and 5 in Canada. The Principal Investigator of this initiative is Dr. Michael W. Weiner, MD, VA Medical Center and University of California, San Francisco. Originally 800 participants were enrolled, which comprised of 200 with AD patients, 400 with MCI and 200 with normal cognition. In 2010, the ADNI study moved into the ADNI GO[3] phase to focus on participants who

---

[3] http://www.adcs.org/studies/imagineadni.aspx

exhibit the earliest signs of memory loss in MCI. While the ADNI GO project work continues, the overall ADNI effort is rapidly moving into a third phase, i.e., ADNI 2.[4] ADNI 2 will build upon the successes of earlier ADNI phases to identify the earliest signs of AD. Researchers are eager to determine when damage to the brain begins. The ADNI 2 phase of the study includes a large number of new volunteers in the earliest stages of cognitive impairment.

References

Achard, S., & Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput Biol*, *3*(2), e17.

Achard, S., Salvador, R., Whitcher, B., Suckling, J., & Bullmore, E.T. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J Neurosci*, *26*(1), 63–72.

Allen, E.A., Damaraju, E., Plis, S.M., Erhardt, E.B., Eichele, T., & Calhoun, V.D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb Cortex*, *24*(3), 663–676.

Baloyannis, S., Manolidis, S., & Manolidis, L. (2000). Synaptic alterations in the vestibulocerebellar system in Alzheimer's disease - a Golgi and electron microscope study. *Acta Otolaryngol*, *120*(2), 247–250.

Binnewijzend, M.A., Schoonheim, M.M., Sanz-Arigita, E., Wink, A.M., van der Flier, W.M., Tolboom, N., Adriaanse, S.M., Damoiseaux, J.S., Scheltens, P., van Berckel, B.N., & Barkhof, F. (2012). Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment. *Neurobiol Aging*, *33*(9), 2018–2028.

Buldú, J.M., Bajo, R., Maestú, F., Castellanos, N., Leyva, I., Gil, P., na Nadal, I.S., Almendral, J.A., Nevado, A., del Pozo, F., & Boccaletti, S. (2011). Reorganization of functional networks in mild cognitive impairment. *PLoS ONE*, *6*(5), e19,584.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, *10*, 186–198.

Chang, C., & Glover, G.H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage*, *50*(1), 81–98.

Chang, C., Liu, Z., Chen, M.C., Liu, X., & Duyn, J.H. (2013). EEG correlates of time-varying BOLD functional connectivity. *Neuroimage*, *72*, 227–236.

Chang, C.C., & Lin, C.J. (2011). LIBSVM: A library for support vector machines. *ACM TIST*, *2*(3), 27:1–27. software available at, http://www.csie.ntu.edu.tw/cjlin/libsvm.

Chen, X., He, J., Lawrence, R., & Carbonell, J. (2012). Adaptive multi-task sparse learning with an application to fMRI study. In *SIAM International Conference on Data Mining (SDM), SIAM, California,USA* (pp. 212–223).

Corbetta, M., & Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*, *3*(3), 201–215.

Damoiseaux, J.S., Rombouts, S. A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., & Beckmann, C.F. (2006). Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci U S A*, *103*(37), 13, 848–13, 853.

Danaher, P., Wang, P., & Witten, D.M. (2012). The joint graphical lasso for inverse covariance estimation across multiple classes. arXiv:1111.0324.

Delbeuck, X., Van der Linden, M., & Collette, F. (2003). Alzheimer's disease as a disconnection syndrome *Neuropsychol Rev*, *13*(2), 79–92.

Dosenbach, N.U.F., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., Dosenbach, R.A.T., Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., Schlaggar, B.L., & Petersen, S.E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci U S A*, *104*(26), 11, 073–11,078.

Dubois, B., & Albert, M.L. (2004). Amnestic MCI or prodromal Alzheimer's disease *Lancet Neurol*, *3*(4), 246–248.

Economou, A., Papageorgiou, S.G., Karageorgiou, C., & Vassilopoulos, D. (2007). Nonepisodic memory deficits in amnestic MCI. *Cogn Behav Neurol*, *20*(2), 99–106.

Fennema-Notestine, C., Hagler, D.J. Jr, McEvoy, L.K., Fleisher, A.S., Wu, E.H., Karow, D.S., & Dale, A.M. (2009). Structural MRI biomarkers for preclinical and mild Alzheimer's disease. *Hum Brain Mapp*, *30*(10), 3238–3253.

Fox, M.D., Zhang, D., Snyder, A.Z., & Raichle, M.E. (2009). The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol*, *101*(6), 3270–3283.

Fransson, P., & Marrelec, G. (2008). The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *Neuroimage*, *42*, 1178–1184.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441.

Fukutani, Y., Cairns, N.J., Rossor, M.N., & Lantos, P.L. (1997). Cerebellar pathology in sporadic and familial Alzheimer's disease including APP 717 (Val⟶Ile) mutation cases: a morphometric investigation. *J Neurol Sci*, *149*(2), 177–184.

Gold, B.T., Jiang, Y., Jicha, G.A., & Smith, C.D. (2010). Functional response in ventral temporal cortex differentiates mild cognitive impairment from normal aging. *Hum Brain Mapp*, *31*(8), 1249–1259.

Greicius, M. (2008). Resting-state functional connectivity in neuropsychiatric disorders. *Curr Opin Neurol*, *21*, 424–430.

Greicius, M.D., Srivastava, G., Reiss, A.L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional mri. *Proc Natl Acad Sci U S A*, *101*(13), 4637–4642.

Hämäläinen, A., Pihlajamäki, M., Tanila, H., Hänninen, T., Niskanen, E., Tervo, S., Karjalainen, P.A., Vanninen, R.L., & Soininen, H. (2007). Increased fMRI responses during encoding in mild cognitive impairment. *Neurobiol Aging*, *28*(12), 1889–1903.

Han, S.D., Arfanakis, K., Fleischman, D.A., Leurgans, S.E., Tuminello, E.R., Edmonds, E.C., & Bennett, D.A. (2012a). Functional connectivity variations in mild cognitive impairment: Associations with cognitive function. *J Int Neuropsychol Soc*, *18*(1), 39–48.

Han, Y., Lui, S., Kuang, W., Lang, Q., Zou, L., & Jia, J. (2012b). Anatomical and functional deficits in patients with amnestic mild cognitive impairment. *PLoS ONE*, *7*(2), e28,664.

Handwerker, D.A., Roopchansingh, V., Gonzalez-Castillo, J., & Bandettini, P.A. (2012). Periodic changes in fMRI connectivity. *Neuroimage*, *63*(3), 1712–1719.

He, Y., & Evans, A. (2010). Graph theoretical modeling of brain connectivity. *Curr Opin Neurol*, *23*, 341–350.

Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., Chen, K., & Reiman, E. (2010). Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. *Neuroimage*, *50*(3), 935–949.

Humphries, M.D., & Gurney, K. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLoS ONE*, *3*(4), e0002, 051.

Humphries, M.D., Gurney, K., & Prescott, T.J. (2006). The brainstem reticular formation is a small-world, not scale-free, network. *Proc Biol Sci, 273*(1585), 503–511.

Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Penna, S.D., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., Handwerker, D.A., Keilholz, S., Kiviniemi, V., Leopold, D.A., de Pasquale, F., Sporns, O., Walter, M., & Chang, C. (2013a). Dynamic functional connectivity: Promises, issues, and interpretations. *Neuroimage, 80*, 360–368.

Hutchison, R.M., Womelsdorf, T., Gati, J.S., Everling, S., & Menon, R.S. (2013b). Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Hum Brain Mapp, 34*(9), 2154 –4 2177.

Jia, H., Hu, X., & Deshpande, G. (2013). Finite number of brain network configurations revealed from time-varying connectivity assessment of resting state fMRI. In *Proc Intl Soc Mag Reson Med, p 0036*.

Johnson, S.C., Schmitz, T.W., Moritz, C.H., Meyerand, M.E., Rowley, H.A., Alexander, A.L., Hansen, K.W., Gleason, C.E., Carlsson, C.M., Ries, M.L., Asthana, S., Chen, K., Reiman, E.M., & Alexander, G.E. (2006). Activation of brain regions vulnerable to Alzheimer's disease: The effect of mild cognitive impairment. *Neurobiol Aging, 27*(11), 1604–1612.

Jones, D.T., Vemuri, P., Murphy, M.A.C., Gunter, J.L., Senjem, M.L., Machulda, M.M., Przybelski, S.A., Gregg, B.E., Kantarci, K., Knopman, D.S., Boeve, B.F., Petersen, R.C., & Jack, C.R. (2012). Non-stationarity in the "resting brain's" modular architecture. *PLoS ONE, 7*, e39,731.

Kaiser, M. (2007). Brain architecture: A design for natural computation. *Philos Trans A Math Phys Eng Sci, 365*(1861), 3033–3045.

Kaiser, M., & Hilgetag, C.C. (2006). Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Comput Biol, 2*(7), e95.

Kircher, T.T., Weis, S., Freymann, K., Erb, M., Jessen, F., Grodd, W., Heun, R., & Leube, D.T. (2007). Hippocampal activation in patients with mild cognitive impairment is necessary for successful memory encoding. *J Neurol Neurosurg Psychiatry, 8*, 812–818.

Kötter, R., & Stephen, K.E. (2003). Network participation indices: Characterizing component roles for information processing in neural networks. *Neural Netw, 16*(9), 1261–1275.

Li, C., Wang, J., Gui, L., Zheng, J., Liu, C., & Du, H. (2011). Alterations of whole-brain cortical area and thickness in mild cognitive impairment and Alzheimer's disease. *J Alzheimers Dis, 27*(2), 281–290.

Li, Y., Wang, Y., Wu, G., Shi, F., Zhou, L., Lin, W., Shen, D., & Alzheimer's Disease Neuroimaging Initiative (2012). Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. *Neurobiol Aging, 33*(2), 427.e15–e30.

Li, Y.T., Woodruff-Pak, D.S., & Trojanowski, J.Q. (1994). Amyloid plaques in cerebellar cortex and the integrity of Purkinje cell dendrites. *Neurobiol Aging, 15*(1), 1–9.

Liu, Z., Zhang, Y., Bai, L., Yan, H., Dai, R., Zhong, C., Wang, H., Wei, W., Xue, T., Feng, Y., You, Y., & Tian, J. (2012). Investigation of the effective connectivity of resting state networks in Alzheimer's disease: A functional MRI study combining independent components analysis and multivariate Granger causality analysis. *NMR in Biomedicine, 25*(1), 1311–1320.

Lynall, M.E., Bassett, D.S., Kerwin, R., McKenna, P.J., Kitzbichler, M., Muller, U., & Bullmore, E.T. (2010). Functional connectivity and brain networks in Schizophrenia. *J Neurosci, 30*, 9477–9487.

McEvoy, L.K., Fennema-Notestine, C., Roddey, J.C., Jr, D.J.H., Holland, D., Karow, D.S., Pung, C.J., Brewer, J.B., & Dale, A.M. (2009). Alzheimer disease: Quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology, 251*, 195–205.

Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends Cogn Sci, 15*, 483–506.

Murphy, K., Birn, R.M., Handwerker, D.A., Jones, T.B., & Bandettini, P.A. (2009). The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *Neuroimage, 44*(3), 893–905.

Rombouts, S. A.R.B., Barkhof, F., Goekoop, R., Stam, C.J., & Scheltens, P. (2005). Altered resting state networks in mild cognitive impairment and mild Alzheimer's disease: An fMRI study. *Hum Brain Mapp, 26*(4), 231–239.

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

Saköglu, U., Pearlson, G.D., Kiehl, K.A., Wang, Y., Michael, A., & Calhoun, V.D. (2010). A method for evaluating dynamic functional network connectivity and task-modulation: Application to Schizophrenia. *MAGMA, 23*(5 –6), 351–366.

Sanz-Arigita, E.J., Schoonheim, M.M., Damoiseaux, J.S., Rombouts, S. A.R.B., Maris, E., Barkhof, F., Scheltens, P., & 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), e13,788.

Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., & Greicius, M.D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci, 27*(9), 2349 –2356.

Seo, E.H., Lee, D.Y., Lee, J.M., Park, J.S., Sohn, B.K., Lee, D.S., Choe, Y.M., & Woo, J.I. (2013). Whole-brain functional networks in cognitively normal, mild cognitive impairment, and Alzheimer's disease. *PLoS ONE, 8*(1), e53,922.

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

Sjöbeck, M., & Englund, E. (2001). Alzheimer's disease and the cerebellum: A morphologic study on neuronal and glial changes. *Dement Geriatr Cogn Disord, 12*(3), 211–218.

Smith, C.D., Chebrolu, H., Wekstein, D.R., Schmitt, F.A., Jicha, G.A., Cooper, G., & Markesbery, W.R. (2007). Brain structural alterations before mild cognitive impairment. *Neurology, 68*(16), 1268–1273.

Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., & Woolrich, M.W. (2011). Network modelling methods for fMRI. *Neuroimage, 54*(2), 875–891.

Smith, S.M., Miller, K.L., Moeller, S., Xu, J., Auerbach, E.J., Woolrich, M.W., Beckmann, C.F., Jenkinson, M., Andersson, J., Glasser, M.F., Van Essen, D.C., Feinberg D.A., Yacoub, E.S., & Ugurbil, K. (2012). Temporally-independent functional modes of spontaneous brain activity. *Proc Natl Acad Sci U S A, 109*(8), 3131.

Sorg, C., Riedl, V., Perneczky, R., Kurz, A., & Wohlschlager, A.M. (2007). Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. *Proc Natl Acad Sci USA, 104*(47), 18,760–18,765.

Sporns, O. (2011). The human connectome: A complex network. *Ann N Y Acad Sci, 1224*, 109–125.

Sporns, O., & Zwi, J.D. (2004). The small world of the cerebral cortex. *Neuroinformatics, 2*, 145–161.

Sporns, O., Chialvo, D.R., Kaiser, M., & Hilgetag, C.C. (2004). Organization, development and function of complex brain networks. *Trends Cogn Sci, 8*(9), 418–425.

Stam, C.J. (2010). Characterization of anatomical and functional connectivity in the brain: A complex networks perspective. *Int J Psychophysiol, 77*(3), 186–194.

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

Suk, H.I., Wee, C.Y., & Shen, D. (2013). Discriminative group sparse representation for mild cognitive impairment classification. In *Machine Learning in Medical Imaging* (pp. 131–138).

Suk, H.I., Wee, C.Y., Lee, S.W., & Shen, D. (2014). Supervised discriminative group sparse representation for mild cognitive impairment diagnosis. Neuroinformatics Epub.

Supekar, K., Menon, V., Rubin, D., Musen, M., & Greicius, M.D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Comput Biol*, *4*, e1000,100.

Tabert, M.H., Manly, J.J., Liu, X., Pelton, G.H., Rosenblum, S., Jacobs, M., Zamora, D., Goodkind, M., Bell, K., Stern, Y., & Devanand, D.P. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Arch Gen Psychiatry*, *63*(8), 916–924.

Telesford, Q.K., Joyce, K.E., Hayasaka, S., Burdette, J.H., & Laurienti, P.J. (2011). The ubiquity of small-world networks. *Brain Connect*, *1*(5), 367–375.

Thomann, P.A., Schläfer, C., Seidl, U., Santos, V.D., Essig, M., & Schröder, J. (2008). The cerebellum in mild cognitive impairment and Alzheimer's disease - A structural mri study. *J Psychiatr Res*, *42*(14), 198–202.

Thompson, P.M., & Apostolova, L.G. (2007). Computational anatomical methods as applied to ageing and dementia. *Br J Radiol*, *80*(2), S78–S91.

Tibshirani, R., Sauders, M., Zhu, J., & Knight, K. (2005). Sparsity and smoothness via the fused lasso. *J R Statist Soc B*, *67*(1), 91–108.

Tomasi, D., Wang, R., Wang, G.J., & Volkow. N.D. (2014). Functional connectivity and brain activation: A synergistic approach. *Cereb Cortex*, *24*(10), 2619–2629.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*, 273–289.

Van Dijk, K.R.A., Hedden T., Venkataraman, A., Evans, K.C., Lazar, S.W., & Buckner, R.L. (2010). Intrinsic functional connectivity as a tool for human connectomics: Theory, properties and optimization. *J Neurophysiol*, *103*, 297–321.

Wang, H.Y., D'Andrea, M.R., & Nagele, R. (2002). Cerebellar diffuse amyloid plaques are derived from dendritic Abeta42 accumulations in Purkinje cells. *Neurobiol Aging*, *23*(2), 213–223.

Wang, J., Zuo, X., & He, Y. (2010). Graph-based network analysis of resting-state functional MRI. *Front Syst Neurosci*, *4*, 16.

Wang, K., Liang, M., Wang, L., Tian, L., Zhang, X., Li, K., & Jiang, T. (2007). Altered functional connectivity in early Alzheimer's disease: A resting-state fMRI study. *Hum Brain Mapp*, *28*(10), 967–978.

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

Wee, C.Y., Yap, P.T., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., & Shen, D. (2012a). Resting-state multi-spectrum functional connectivity networks for identification of MCI patients. *PLoS ONE*, *7*(5), e37,828.

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

Wee, C.Y., Yap, P.T., Shen, D., & for the Alzheimer's Disease Neuroimaging Initiative (2013). Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns. *Hum Brain Mapp*, *34*(12), 3411–3425.

Wee, C.Y., Yap, P.T., Zhang, D., & Shen, D. (2014). Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. *Brain Struct Funct*, *219*(2), 641–656.

Wegiel, J., Wisniewski, H.M., Dziewiatkowski, J., Badmajew, E., Tarnawski, M., Reisberg, B., Mlodzik, B., Leon, M.J.D., & Miller, D.C. (1999). Cerebellar atrophy in Alzheimer's disease - Clinicopathological correlations. *Brain Res*, *818*(1), 41–50.

Weis, S., Klaver, P., Reul, J., Elger, C.E., & Fernández, G. (2004). Temporal and cerebellar brain regions that support both declarative memory formation and retrieval. *Cereb Cortex*, *14*(3), 256–267.

Whitwell, J.L., Przybelski, S.A., Weigand, S.D., Knopman, D.S., Boeve, B.F., Petersen, R.C., & Jack, J.r.C.R. (2007). 3d maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. *Brain*, *130*(7), 1777–1786.

Yang, S., Pan, Z., Shen, X., Wonka, P., & Ye, J. (2012). Fused multiple graphical lasso. arXiv:1209.2139.

Zhang, D., Shen, D., & Alzheimer's Disease Neuroimaging Initiative (2012). Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLoS ONE*, *7*(3), e33,182.