Published in final edited form as:
Brain Struct Funct. 2014 March ; 219(2): 641–656. doi:10.1007/s00429-013-0524-8.

Group-Constrained Sparse FMRI Connectivity Modeling for Mild Cognitive Impairment Identification

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

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

Daoqiang Zhang,
Image Display, Enhancement, and Analysis (IDEA) Laboratory, Biomedical Research Imaging Center (BRIC) and Department of Radiology, University of North Carolina at Chapel Hill, NC 27599, USA. Department of Computer Science & Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, PR China

Lihong Wang, and
Brain Imaging and Analysis Center (BIAC), Duke University Medical Center, Durham, NC 27705, USA

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

For the Alzheimer's Disease Neuroimaging Initiative[‡]

Abstract

Emergence of advanced network analysis techniques utilizing resting-state functional Magnetic Resonance Imaging (R-fMRI) has enabled a more comprehensive understanding of neurological disorders at a whole-brain level. However, inferring brain connectivity from R-fMRI is a challenging task, particularly when the ultimate goal is to achieve good control-patient classification performance, owing to perplexing noise effects, curse of dimensionality, and inter-subject variability. Incorporating sparsity into connectivity modeling may be a possible solution to partially remedy this problem since most biological networks are intrinsically sparse. Nevertheless, sparsity constraint, when applied at an individual level, will inevitably cause inter-subject variability and hence degrade classification performance. To this end, we formulate the R-fMRI time series of each region-of-interest (ROI) as a linear representation of time series of other ROIs to infer sparse connectivity networks that are topologically identical across individuals. This formulation allows simultaneous selection of a common set of ROIs across subjects so that their

[*]Corresponding author: Dinggang Shen, dgshen@med.unc.edu, Tel.: +1 919-843-3535, Fax: +1 919-843-2641.
[‡]ADNI
Data used in preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.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.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

linear combination is best in estimating the time series of the considered ROI. Specifically, l_1-norm is imposed on each subject to filter out spurious or insignificant connections to produce sparse networks. A group-constraint is hence imposed via multi-task learning using a l_2-norm to encourage consistent non-zero connections across subjects. This group-constraint is crucial since the network topology is identical for all subjects while still preserving individual information via different connectivity values. We validated the proposed modeling in mild cognitive impairment (MCI) identification and promising results achieved demonstrate its superiority in disease characterization, particularly greater sensitivity to early stage brain pathologies. The inferred group-constrained sparse network is found to be biologically plausible and is highly associated with the disease-associated anatomical anomalies. Furthermore, our proposed approach achieved similar classification performance when finer atlas was used to parcellate the brain space.

Keywords

Mild cognitive impairment (MCI); Group-constrained sparse modeling; Resting-state fMRI; Sparse linear regression; Inter-subject variability; Multi-task learning

1 Introduction

Alzheimer's disease (AD) is commonly characterized by cognitive and intellectual deficits, which is serious enough to interfere daily life, without effective treatment until now. It gets worse over time by gradually destroying brain cells, causing loss in memory and ability to reason, make judgments and communicate, eventually causing death. AD is the most prevalent dementia which accounts for an estimate of 60% to 80% of dementia cases among elderly population [3] and definitive diagnosis can only be made with histopathological confirmation of amyloid plaques and neurofibrillary tangles. It has been reported that the incidence of AD doubles every five years after age of 65 [6] and 1 in every 85 persons will be affected by the disease by year 2050 [12]. Based on the latest statistics, AD is the sixth leading cause of death in United States and is the only among the the top ten causes of death that cannot be prevented, cured or slowed down [3]. Recenly, there are 5.4 million or approximately 1 in 8 senior Americans who are diagnosed with AD [3]. It is estimated that around 44% Americans with age 75 to 84 and 46% with age greater than 85 are AD patients. This becomes worse as life expectancy increases. The average life expectancy of AD patients varies between 3 to 10 years. The median life span is up to 7 to 10 years for AD patients whose conditions are diagnosed when they are in their 60s and early 70s. The median reduces to about 3 years or less for patients whose conditions are diagnosed when they are in their 90s [75]. With the aging world population, this disease has been a serious problem and a huge burden to the healthcare system, particularly in developed countries. Recognizing the urgent need to slow down or completely prevent the occurrence of a healthcare crisis worldwide, effort has been underway to administer and to develop effective pharmacological and behavioral interventions for delaying the onset and progression of the disease, particularly at its early stage.

Mild cognitive impairment (MCI), an intermediate stage of brain cognitive decline between normal aging and dementia, is associated with increased risk of developing AD, especially when memory loss is the predominant symptom. This type of MCI is commonly referred as "amnestic MCI". Some individuals with MCI remain stable or return to normal over time, but more than half progress to dementia within 5 years [25]. Recent studies show that individuals with MCI tend to progress to probable AD at a rate of approximately 10% to 15% per year [29,39], compared with healthy controls who develop dementia at a rate of 1% to 2% per year [11]. According to a latest, long-term study of nearly 4000 participants, cognitive impairment has a significant impact on life expectancy similar to chronic

conditions such as diabetes or chronic heart failure [52]. Early detection is thus important for possible delay of the progression of MCI to moderate and severe stages. However, diagnosis of MCI is difficult due to its mild symptoms of cognitive impairment, causing most computer-aided diagnosis to achieve lower than desired performance.

Constructing functional brain connectivity from neuroimaging data holds great promise for identifying image-based biomarkers that are crucial for distinguishing MCI from normal aging. Many functional connectivity modeling approaches have been proposed in the literature including the correlation-based approaches, Granger causality (or lag-based measures) analysis, and regularized inverse covariance estimation [58]. With a comprehensive comparison on different connectivity estimation approaches using realistic simulated functional magnetic resonance imaging (fMRI) data, it was reported that in general the correlation-based approaches can provide relatively high sensitivity to network connection detection when compared with methods that are based on higher-order statistics and lag-based approaches [58]. A large body of work on functional connectivity were based on the correlation analysis [5, 32, 60, 61, 63, 69, 72] among some. However, correlation analysis only captures pairwise information and is unable to provide an adequate and complete account of the interaction between many brain regions [33]. Also, fully-connected correlation-based network structure is difficult to interpret due to many spurious or insignificant connections. Many spurious connections arise due to the low frequency (< 0.1Hz) spontaneous fluctuation of blood oxygen level dependent (BOLD) signals and physiological noise, e.g., cardiac and respiratory cycles. Inferring brain connectivity from resting-state fMRI (R-fMRI) data for better classification performance is hence a challenging task. Recent work [63] has shown that certain sparsity constraints can be imposed to elucidate robust connections from a set of noisy connections. The sparsity constraint correlates with the fact that, neurologically, a brain region predominantly interacts only with a small number of other regions. Furthermore, it has been demonstrated that the effects of inaccurate ROIs arisen, particularly when using atlas-based ROIs, can almost be ignored if the ROIs of the constructed R-fMRI-based connectivity are spread sparsely across the brain and not spatially neighboring [58].

Partial correlation measures, which are associated with the time series similarity of two regions after factoring out the effects of other regions, correspond to the off-diagonal entries of inverse covariance matrix (or precision matrix) of the data. Estimation of partial correlation is normally achieved via maximum likelihood estimation (MLE) of the inverse covariance matrix. One important requirement for reliable estimation of inverse covariance matrix using MLE is the sample size of the data (time series) must be substantially larger than the number of brain regions modeled [33, 73]. Sparsity in terms of the entries of the inverse covariance matrix [22, 50, 73], so-called the sparse inverse covariance estimation (SICE), has been applied on PET data to learn brain connectivity of AD, MCI and healthy controls [33]. SICE imposes the sparsity constraint through l_1-norm regularization on the MLE to reliably estimate the inverse covariance matrix with the sample size close to or lesser than brain regions. Although this method is effective in identifying the zero and non-zero entries, i.e., adjacency structure, in an inverse covariance matrix, it is not suitable at estimating the magnitude of non-zero entries due to the shrinking effect. To partially solve this problem, Huang *et al.* proposed a “quasi-measure” approach to determine the “strength” of non-zero connections by using a series of different regularization parameters that determine the sparseness of the inverse covariance matrix [33]. By using this approach, the “strength” of a non-zero connection is assigned with the largest regularization parameter value that preserves the existence of connection. However, this approach unable to provide a complete and accurate account of the strength of connections. It also been reported that only a few of the regularization parameter values can provide reasonably good estimation of

network connection [58]. Note that the SICE without regularization is equivalent to the partial correlation, a fully-connected network.

Sparse modeling based on penalization of the l_1-norm linear regression model as in the least absolute shrinkage and selection operator (LASSO) has been proposed for constructing sparse functional connectivity network of autism [34]. However, this method was performed on PET data across different subjects of same group (autism patients) and hence is not suitable to infer the functional connectivity at individual level for all subjects (patients and healthy subjects). In order to infer sparse brain connectivity at individual subject level on R-fMRI data, we suggest to modify the method in [34] by penalizing the l_1-norm linear regression model across R-fMRI time series of each subject. By using this regression model, a sparse representation of brain connectivity can be obtained with only a few of significant connections. The insignificant or spurious connections are forced to have zero contribution and hence makes the constructed sparse connectivity relatively easier to be interpreted. The linear regression model enables the representation of a brain region (in terms of time series) by the linear combination of other brain regions, with the contribution of every region is reflected by the magnitude of the regression coefficient (or connection strength). Zero connection strength between two regions, for instance regions A and B, implies that the region B is redundant in accurate estimation of the time series of region A and vice versa. This provides a view on how a brain region is correlated with the rest of brain regions by filtering out the insignificant or spurious connections. It is noteworthy that this sparse representation-based functional connectivity inferring model is different from SICE since no computation of partial correlation-based connectivity is required before determining the non-zero connection locations via inverse covariance matrix. Hence, errors introduced during partial correlation computation and drawbacks of estimating the magnitude of non-zero entries in SICE can be avoided.

However, the sparse modeling is unable to deal with inter-subject variability problem since l_1-norm penalization at an individual level will result in different network topological structures across subjects [73], i.e., the adjacency structure is different for every subject. This will inevitably make the comparison between subjects difficult and thus possibly degrade generalization performance of trained classifiers. To mitigate the effects of inter-subject variability, we propose in this paper a group-constrained sparse linear regression model, which follow the idea of joint feature selection concept in group-lasso for regression problems [73], to jointly estimate the nonzero connections across subjects via multitask learning. Through multi-task learning, the connection topology is kept identical among subjects, while at the same time allowing individual connection strength to vary between subjects. This modeling enables easier and direct comparison among subjects in MCI identification. Recently, a work based on the similar joint feature selection approach has been proposed to address the problem of inter-subject variability [68]. It is noteworthy that this work is different from our work since it was developed based on SICE with the assumptions that the data are multivariate Gaussian distributed and the covariance has to be positive definite. In [68], partial correlation or covariance matrices were firstly estimated from R-fMRI time series of a group of subjects before a mixed $l_{2,1}$-norm was applied to penalize their inverse covariance matrices to determine the locations of non-zero connections. Although this method can be used to minimize the inter-subject variability problem, it inevitably inherits the drawbacks of SICE such as the requirement of covariance matrix estimation and unreliable at estimating the magnitude of non-zero entries in the sparse inverse covariance matrix. Another group-level modeling, which does not correspond to group-lasso formulation, also been proposed to identify the causal connections of functional connectivity [46]. To the best of our knowledge, the current study is the first attempt to infer sparse functional brain network, based on the linear combination of R-fMRI time series of different brain regions and group-constrained multi-task learning, for the

purpose of MCI identification. We seek to validate if this new connectivity network modeling strategy can be used to better interpret MCI-related pathological anomalies and hence to improve classification performance. We explore the biologically meanings of the inferred sparse network through its significant connections and the crucial brain regions that contribute to classification performance. We also investigate the robustness of the proposed method with respect to different parcellation scales. This paper sheds new light on group-constrained sparse connectivity modeling approach for R-fMRI data and its effectiveness when applied to MCI identification.

The rest of the paper is organized as follows: Section 2 furnishes information on the image dataset, and post-processing pipeline. This is followed by a comprehensive description on the construction of the proposed constrained sparse and Pearson correlation-based functional brain connectivity networks. We also briefly describe the feature extraction, feature selection and classifier training procedures used in the study. Performance of the proposed constrained sparse connectivity modeling is validated extensively in Section 3 by comparing with other functional connectivity inferring approaches, as well as with different parcellation scales. Findings and methodological issues are discussed in Section 4. Section 5 concludes this paper.

2 Material and Methods

2.1 Participants

In this study, subjects from two cohorts were combined to create a balanced dataset for better evaluation of classification performance. The first cohort contains 13 MCI subjects who were randomly selected from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Only the baseline scan of each subject was utilized. The second cohort contains 37 subjects (12 MCI individuals and 25 socio-demographically matched healthy controls) recruited by the Duke-UNC Brain Imaging and Analysis Center (BIAC), Durham, North Carolina, USA. Written consent was obtained from all participants, and the experimental protocols were approved by the institutional ethics board at Duke University Medical Center in compliance with the Health Insurance Portability and Accountability Act. Diagnosis of all the recruited subjects was performed by expert consensus panels at the Joseph and Kathleen Bryan Alzheimer's Disease Research Center (Bryan ADRC) and the Department of Psychiatry at Duke University Medical Center. Diagnosis confirmation was made by consensus with the ultimate decision by a board-certified neurologist in concert with available data from a battery of general neurological examination, neuropsychological assessment evaluation, collateral subject symptom and functional capacity reports. The neuropsychological battery used was a revised Consortium to Establish a Registry for Alzheimer's Disease (CERAD) [40] which included: 1) Mini-Mental State Examination (MMSE) [20]; 2) immediate and delayed verbal memory (Logical Memory subtest of the Wechsler Memory Scale-Revised [71]); 3) visual immediate memory (Benton Visual Retention Test [9]); 4) verbal initiation/lexical fluency (Controlled Oral Word Association Test from the Multilingual Aphasia Examination [10]); 5) attentional/executive functions (Trail Making Test [47], Symbol Digit Modality Test [56], Digit Span sub-test of the Wechsler Adult Intelligence Scale-Revised [70], and a separate ascending Digit Span task modeled after the Digit Ordering Test [15]); 6) premorbid verbal ability (Shipley Vocabulary Test [55]); 7) Finger Oscillation [48] and Grooved Pegboard [37] Tests; and 8) Self Rating of Memory Function [59].

Conformation of diagnosis for MCI if subjects met the following inclusion criteria: 1) age > 55 years and any race; 2) recent worsening of cognition, but still functioning independently; 3) MMSE score between 24 and 30; 4a.) score ≤ −1.5 SD on at least two Bryan ADRC cognitive battery memory tests for single-domain amnestic MCI; or 4b.) score ≤ −1.5 SD on

at least one of the formal memory tests and score ≤ −1.5 SD on at least one other cognitive domain task (e.g., language, visuospatial-processing, or judgment/executive function) for multi-domain MCI; 5) 4 or lower for baseline Hachinski score; 6) does not meet the NINCDS-ADRDA [38] or DSM-IV-TR [4] criteria for dementia; 7) no psychological symptoms or history of depression; and 8) capacity to give informed consent and follow study procedures.

Similarly, all healthy controls met the following criteria: 1) age > 55 years and any race; 2) adequate visual and auditory acuity to properly complete neuropsychological testing; 3) no self-report of neurological or depressive illness; 4) shows no evidence of depression based on the Diagnostic Interview Schedule port of the Duke Depression Evaluation Schedule; 5) normal score on a non-focal neurological examination; 6) a score > −1 SD on any formal memory tests and a score > −1 SD on any formal executive function or other cognitive test; and 7) demonstrates a capacity to give informed consent and follow study procedures. In order for safety purposes and minimizing biases, subjects were excluded from the study if they have: 1) any of the traditional MRI contraindications, such as foreign metallic implants or pacemakers; 2) a past head injury or neurological disorder associated with MRI abnormalities, including dementia, brain tumors, epilepsy, Parkinson's disease, demyelinating diseases, *etc.*; 3) any physical or intellectual disability affecting completion of assessments; 4) documentation of other Axis I psychiatric disorders; and 5) any prescription medication (or nonprescription drugs) with known neurological effects. It is noteworthy that the diagnosis of all cases were made on clinical grounds without reference to MRI. Demographic and clinical information of the participants is summarized in Table 1.

2.2 Data Acquisition

Subjects from the first cohort, i.e., the ADNI dataset, were scanned using 3.0T Philips scanner (Achieva, Philips Healthcare) from different sites with the following parameters: echo time (TE) = 30 ms, repetition time (TR) = 3000 ms and flip angle = 80°, imaging matrix = 64 × 64, 48 slices, 140 volumes, and voxel thickness = 3.3 mm. Readers are referred to www.adni-info.org for more data acquisition information. For the second cohort from Duke University Medical Center, an 3.0T GE scanner (Signa EXCITE, GE Healthcare) was used to acquire R-fMRI volumes from all participants. The R-fMRI scans were acquired axially parallel to the horizontal plane connecting the anterior and posterior commissures (AC-PC line) with TE = 32 ms, TR = 2000 ms, flip angle = 77°, imaging matrix = 64 × 64 with a rectangular FOV of 256 × 256 mm^2, 150 volumes, 34 slices, and voxel thickness = 4.0 mm. The scan was acquired using a SENSE inverse-spiral pulse sequence in the same plane as the low resolution T1-weighted images. All the subjects were told to keep their eyes open and stare at a fixation cross in the middle of the screen during scanning, which lasted for 5 minutes, to prevent neurons excitation problem caused by changing stimuli across time. If stimuli such as the little cross sign in this study is presented steadily without changing across the five minutes period, the neural excitation related to the stimuli can vanish quickly. This procedure prevents the subjects from falling into sleep and avoids saccade-related activation which is unavoidable if eyes are closed. The same scanner was used to acquire the T1-weighted anatomical MRI images using the following parameters: TE = 2.976 ms, TR = 7.460 ms and flip angle = 12°. The acquisition matrix was (256 × 224) with a rectangular FOV of (256 × 256 mm^2), resulting in slice thickness of 1 mm. A total of 216 slices were acquired using the FSPGR ASSET sequence.

2.3 Data Post-Processing

Since the number of time points acquired for Duke and ADNI datasets are different, we used only the first 130 time points for functional connectivity network construction. Postprocessing of the R-fMRI images, including slice timing correction and head-motion

correction, were performed using the Statistical Parametric Mapping (SPM8[1]) software package. 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 [23]. 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.

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 [67]. Given the controversy of removing the global signal in the post-processing of R-fMRI data [21, 41], we did not regress the global signal out [2, 36, 63]. We then parcellated the brain space into 116 ROIs by warping the automatic anatomical labeling (AAL) atlas [66] to the R-fMRI images using deformation fields estimated from T1-weighted images via a deformable registration method called HAMMER [54]. For each subject, the mean time series of each individual ROI was obtained by averaging the regressed R-fMRI time series over all voxels in that particular ROI. Temporal band-pass filtering with frequency interval ($0.025 \leq f \leq 0.100$Hz) was then performed on the mean time series of each individual ROI. It provides a reasonable trade-off between avoiding the physiological noise associated with higher frequency oscillations [16], the measurement error associated with estimating very low frequency correlations from limited time series [1], and the magnetic field drifts of the scanner [64]. This frequency interval was further decomposed into five equal-length spectral, enabling a more frequency specific analysis of the regional mean time series [72]. For each frequency sub-band, we inferred a functional connectivity network by utilizing three different approaches: 1) Pearson correlation between the regional mean time series of all possible pairs of ROIs, 2) Sparse regression without group-constraint via l_1-norm regularization, and 3) Sparse regression with group-constraint via $l_{2,1}$-norm regularization. These three approaches will be discussed in detail in the following subsections.

2.4 Pearson Correlation-Based Functional Brain Connectivity

Assuming we have N training subjects with each of them having M ROIs, and the denoting regional mean time series of the p-th ROI for the n-th subject as $\mathbf{y}_p^n$. By considering the brain regions as a set of nodes and the correlation coefficients as signed weights on the set of edges, functional connectivity that examines interregional correlations in neuronal variability [24] can be measured using Pearson correlation coefficient between a given pair of $\mathbf{y}_p^n$ and $\mathbf{y}_q^n$, given that $p \neq q$. The Pearson correlation matrix is a symmetric matrix in which each off-diagonal element is the correlation coefficient (similarity in terms of time series) between a pair of ROIs. A Fisher's r-to-z transformation was then applied on the Pearson correlation matrix to improve the normality of the correlation coefficients as

$$z = \frac{1}{2}\left[\ln(1+r) - \ln(1-r)\right] \quad (1)$$

[1] http://www.fil.ion.ucl.ac.uk.spm

where r is the Pearson correlation coefficient and z is normal with standard deviation $\sigma_z = 1/\sqrt{N-3}$. The inferring of Pearson correlation-based functional connectivity map is graphically shown in Figure 1.

2.5 Sparse Functional Brain Connectivity Without Group-Constraint

With a total of M ROIs, the regional mean time series of p-th ROI for n-th subject, $\mathbf{y}_p^n$, can be regarded as a response vector that can be estimated as a linear combination of time series of other ROIs as

$$\mathbf{y}_p^n = \mathbf{A}_p^n \alpha_p^n + e_p^n, \quad (2)$$

where e_p^n is the error, $\mathbf{y}_p^n = [y_p^n(1); y_p^n(2); \ldots; y_p^n(T)]$ with T being the number of time points in the time series, $\mathbf{A}_p^n = [\mathbf{y}_1^n, \ldots, \mathbf{y}_{p-1}^n, \mathbf{y}_{p+1}^n, \ldots, \mathbf{y}_M^n]$ is data matrix of the p-th ROI (all time series except for p-th ROI), and $\alpha_p^n = [\alpha_1^n; \ldots; \alpha_{p-1}^n; \alpha_{p+1}^n; \ldots; \alpha_M^n]$ is the weight vector that quantifies the degree of influence of other ROIs to p-th ROI. The formulation of this inferring model is graphically illustrated in Figure 2.

The sparse brain functional connectivity modeling of the n-th subject and p-th ROI can be considered as a standard l_1-norm regularized optimization problem with the following objective function

$$f(\alpha_p^n) = \frac{1}{2}\left|\left|\mathbf{y}_p^n - \mathbf{A}_p^n \alpha_p^n\right|\right|_2^2 + \lambda\left|\left|\alpha_p^n\right|\right|_1, \quad (3)$$

where $\lambda > 0$ is the regularization parameter controlling the "sparsity" of the model, with a higher value corresponding to a sparser model, i.e., more elements in α^n are zero. It is noteworthy that the l_1-norm penalization of α_p^n is imposed individually on different subjects. By using this approach, topological structure of the generated sparse functional connectivity differs for each subject. This causes significant inter-subject variability which may possibly incur various problems in group analysis and classification.

2.6 Sparse Functional Brain Connectivity With Group-Constraint

To minimize the inter-subject variability, we force the inferred connectivity networks to have identical topological structure across all subjects. This is accomplished by imposing a group-constraint into the sparse model in Eq. (3) via an additional l_2-norm penalization across all subjects. Hence, the problem becomes minimizing the following objective function via multi-task learning as

$$f(\alpha_p) = \sum_{n=1}^{N}\left(\frac{1}{2}\left|\left|\mathbf{y}_p^n - \mathbf{A}_p^n \alpha_p^n\right|\right|_2^2\right) + \lambda\left|\left|\alpha_p\right|\right|_{2,1} \quad (4)$$

where $\|\alpha_p\|_{2,1}$ is the summation of l_2-norms of α_p^n, i.e., $\sum_p \left|\left|\alpha_p^n\right|\right|_2$. Specifically, the l_2-norm penalization is imposed on the same elements across different matrices α_p which forces the weights corresponding to certain connections across different subjects to be grouped together. This constraint promotes a common connection topology among subjects, while at the same time allows variation of connectivity values (connection weights) between subjects. This mitigates the inter-subject variability problem and hence allows for easier and more consistent inter-subject comparison, particularly for patient identification. The nonzero

coefficients in α_p matrix are treated as an indicator on how significant other ROIs influence the currently considered ROI. We use the SLEP toolbox [35] to solve the objective function in Eq. (4). The modeling of group-constrained sparse functional connectivity via multi-task learning is graphically shown in Figure 3.

2.7 Feature Extraction and Feature Selection

Local clustering coefficient [51], a measure that quantifies the cliquishness of the nodes, was extracted as feature to reflect network property from all functional connectivity maps. Since the AAL atlas with 116 ROIs was utilized in brain parcellation, a feature vector consisting of 116 clustering coefficients, one for each ROI, was generated from each map. The feature vectors from all frequency sub-bands of each individual subject were concatenated to generate a long feature vector with 580 elements. We then utilized a hybrid method to effectively select the most relevant features from the long feature vector for classification. Two filter-based approaches were firstly used to reduce the number of features, followed by a wrapper-based approach to further select a subset of features that is favorable for MCI classification. Specifically, in the first filter-based approach, only features that differ significant between the MCI and the healthy control groups, measured using *t*-tests, were retained for subsequent feature selection. Despite the reduction in dimensionality, the features retained by this simple approach may still inevitably be inter-correlated. Therefore, we employed another filter-based approach, named as minimum redundancy and maximum relevance (mRMR) [18, 44], to further reduce the feature dimension. The mRMR model provides a balance between two aspects of feature selection, i.e., efficiency and broadness. Efficiency ensures that characteristics can be represented with a minimal number of features without significant reduction in prediction performance. Broadness ensures that the selected feature subset can be the maximally representative of original space covered by the entire dataset. Mutual information is utilized in mRMR to measure the relevance of every feature pair and between features and classes. Specifically, we minimized the total relevance of each feature pair to achieve minimum redundancy, while simultaneously maximized the total relevance of each feature-class pair to achieve maximum relevance.

Finally, the support vector machine (SVM) recursive feature elimination (SVM-RFE) [30, 45], a wrapper-based method, was used to select a subset of features that is most discriminative and favorable for MCI classification. SVM-RFE removes features that make the classification error smallest one by one. In this study, we employed SVM with a simple linear kernel to evaluate the discriminative power of the selected features. It is noteworthy that the raw features were firstly scaled individually to range [−1, +1] before feature selection is performed. Every scaled feature was normalized across all training subjects to obtain its standard score (*z*-value). These steps ensured that all the extracted features were within the same scale, minimizing possible bias that may occur when performing selection on features with different dynamic ranges.

2.8 Classifier Training

In this study, we employed SVM with a simple linear kernel to evaluate the discriminative power of the selected features derived from three different functional connectivity inferring approaches. 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 $N - 1$ 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 value in Eqs. (3) and (4) was determined via grid search.

3 Experimental Results

3.1 Group-Constrained Sparse Functional Brain Connectivity

Functional connectivity maps inferred using the proposed group-constrained sparse linear regression model of one healthy normal control (NC) and one MCI patient are shown in Figure 4. Spatial connection topology of the constrained sparse and the fully-connected Pearson correlation-based networks are projected on a 3D brain as shown in Figure 5. Connectogram of group-constrained sparse functional connectivity for single and multi-spectrum are provided in Figure 6. Links with red and black colors indicate inter- and intrahemispheric connections. Generally, the significant connections for single spectrum case were relatively balance between inter- and intrahemispheres. However, there were more inter-hemispheric connections have been selected via the group-constrained sparse regularization when the BOLD time series was decomposed into several frequency bands.

It can be observed that the generated group-constrained sparse connectivity networks are significantly sparser than the fully-connected correlation-based networks. This sparse connectivity network shows a number of biologically interesting findings. First, the bilateral temporal lobes show a smaller amount of intra-lobe connections than other lobes of the brain for all five frequency sub-bands. Second, there are significantly more inter-lobe connections between parietal and occipital lobes than any other lobe pairs. Third, there are more negative connections, which are widely distributed in sparse network of different frequency bands, in the MCI subjects than the NC subjects. Fourth, MCI subjects shown markedly reduced (or negative) functional connectivity in the frontal lobes.

3.2 Comparison Between Group-Constrained Sparse and Correlation-Based Functional Connectivity

Performance of the proposed group-constrained sparse connectivity network was compared with the conventional Pearson correlation-based connectivity network (after correlation correction with a cut-off FDR < 0.2) using single and multi-spectral BOLD signals. In the single spectrum case, feature extraction was directly performed on the band-pass filtered BOLD signals without further frequency sub-band decomposition. Sub-band decomposition in the multi-spectral case enables more frequency specific characterization of subtle changes in BOLD signals, and hence provides better discriminative power [72]. During evaluation, SVM classifiers with the same linear kernel but different hyperparameters were used in a leave-one-out fashion due to the limited number of available samples. Classification performance of both the group-constrained sparse and correlation-based networks using single and multi-spectral BOLD signals are summarized in Table 2. The proposed group-constrained sparse network with multi-spectral BOLD signals yields the best classification performance with an accuracy of 84.0%, which is an increment of at least 8.0% from that of the Pearson correlation-based approach. A cross-validation estimation of the generalization performance shows an area of 0.8656 under the receiver operating characteristic curve (AUC), indicating good diagnostic power. It is also observed that the multi-spectral case not

only shows better classification accuracy, but also provides higher AUC and sensitivity values for the sparse connectivity networks.

3.3 Comparison Between Sparse Functional Connectivity With and Without Group-Constraint

MCI classification performance of two sparse functional connectivity networks, i.e., with and without group-constraint, were compared using similar setting that was used in the previous experiment. Note that the optimal λ value for the sparse functional connectivity networks with and without group-constraint were determined and the best classification performance of each network are summarized in Table 3. The classification performance of using the multi-spectral approach was significantly better than the single spectrum approach in both the sparse functional connectivity networks. In addition, the proposed sparse network with group-constraint performed better than those without group-constraint in both the single and multi-spectral cases. With group-constraint, the inter-subject variability problem is minimized through the generated identical network topology among subjects and thus enables relatively easier differentiation between MCI subjects and healthy controls, as given by the significantly higher sensitivity value. It is interesting to observe that the sparse functional connectivity without group-constraint performed better than the Pearson correlation-based approach in terms of AUC and sensitivity values in multi-spectral case.

3.4 Classification Performance Using Different Atlas Scales

In order to investigate the effects of brain parcellation scale on our proposed modeling, we repeated the MCI classification procedure using a finer atlas with 238 ROIs. This finer atlas was generated by subdivision of the ROIs of the AAL atlas, by setting the size of every new ROI to be approximately equal. Results of using the original 116 ROIs and the finer 238 ROIs are summarized in Table 4.

It can be observed that all the compared methods did not show significant difference in terms of classification accuracy when finer parcellation was used. However, with the finer parcellation all the compared methods experienced decrease in sensitivity value, although this decrease was relatively smaller for the proposed group-constrained method. We also observed slight improvement of AUC value when the proposed method was used together with the finer scale of atlas. In addition, the proposed method outperformed the other compared methods in various metrics for both the coarse and fine parcellation scales. This implies that the proposed method is relatively robust with respect to parcellation scale, at least in terms of classification accuracy.

3.5 Most Discriminant Regions

Since the proposed approach was evaluated in a nested leave-one-out fashion, the selected subset of features might be different for each leave-one-out fold. We hence defined the most significant features (or ROIs) as the regions that were most frequently selected during the construction of optimal SVM models in the training stage. For easier interpretation, the reported most discriminant regions are based on the original AAL atlas with 116 ROIs. The most discriminant regions that were selected for the best classification performance using the proposed group-constrained sparse functional connectivity include regions located in frontal lobe (e.g. orbitofrontal cortex, frontal gyri and rectus gyrus), temporal lobe (e.g. temporal gyri and temporal pole), and other regions such as cingulate gyri, amygdala, angular gyrus, and occipital gyri. These selected regions, as shown graphically in Figure 7, are found to be associated with MCI pathology.

4 Discussion

This paper investigated a novel brain connectivity inferring method for R-fMRI data based on the fact that a brain region predominantly interacts only with a small number of other regions. Thus, each region can be accurately and effectively represented or modeled by using a small number of regions according to their contributions. Compared to the fully-connected correlation-based functional connectivity, sparse brain connectivity have been successfully applied to elucidate robust and meaningful connections from a set of noisy or spurious connections. It has be recently shown that the sparse functional brain connectivity in terms of sparse inverse covariance matrix or compress sensing can be inferred via l_1-norm penalization across subjects to provide biological meaningful information on AD [33] and autism [34] datasets. The major advantages of using l_1-norm penalization including its sparsity-inducing property, convenient convexity, and strong theoretical guarantees [13]. The l_1-norm penalization automatically finds significant network connections with the weights of insignificant connections automatically driven to zero, i.e., minimizing spurious connections in constructed network. The derived solution is helpful in interpreting the estimated networks, without further statistical thresholding in inferring the significant network connections. Although l_1-norm penalization has been used successfully, it is unable to deal with inter-subject variability problem at an individual level [73] and inevitably causes the comparison between subjects in classification difficult. The variations of network topologies across subjects without any constraint, i.e., non-zero connections are distributed randomly across different subjects, may confuse classifiers during the training process and eventually degrade the generalization performance of trained classifiers.

To address the inter-subject variability problem, a group-based constrained sparse linear regression model, which is implemented via multi-task learning, was proposed in this study to enforce an identical connection topology among subjects, while at the same time allowing variations of individual connectivity values between subjects. Modeling with group-constraint sparsity provides an easier and more direct comparison among subjects in classification. This group-constraint is accomplished by imposing an extra l_2-norm penalization on the sparse modeling (l_1-norm penalization) of each individual subject. This extra constraint determines connection topology that is crucial for MCI classification by using a common set of significant sparse functional connections that are learned from different individuals. The elimination of spurious and redundant connections between ROIs pairs improves the effectiveness and capability of the inferred sparse functional connectivity network in conveying crucial and disease-related information. This contribution is particularly crucial since an identical disease-related sparse topological structure can be applied to all new subjects, no matter they are patients or healthy controls, for providing good disease diagnosis rate.

It is interesting to know that the proposed group-constrained sparse functional connectivity network results in a number of biologically meaningful findings which provide better interpretation when compared with the conventional correlation-based network. Compared with other lobes, the bilateral temporal lobes demonstrate a relatively smaller amount of intra-lobe connections, which has been extensively reported in the literature [63, 69]. In addition, significantly more inter-lobe connections between the parietal and occipital lobes are observed when compared with other lobe pairs [63]. Also, many inter- and intra-hemispheric connections were preserved in the inferred sparse network. The observation regarding affected inter- and intrahemispheres, either functionally or anatomically, has been reported recently in literature [7, 31, 65]. This can be partially explained by the progressive loss of synapses, as well as atrophy of corpus callosum that links both hemispheres. These topology patterns are identified via multi-task learning across R-fMRI time series of different subjects. Most of the retained connections have been found to be significantly

altered in AD and/or MCI patients, thus providing a higher discriminative power than other connections. Taken together, these findings suggest that the proposed group-constrained sparse modeling method reveals anatomically-plausible resting-state brain architecture that is associated with AD or MCI. Hence, the SVM classifiers trained using these small amount of disease-related connections perform much better than the conventional correlation-based networks in MCI classification. It is noteworthy that even without group-constraint, the sparse connectivity network still performs better than the conventional correlation-based network. This can be partially explained by the elimination of spurious or insignificant connections that reduces the curse of dimensionality in high-dimensional classification.

Classification performance evaluated via a nested leave-one-out cross-validation, which ensures performance generalization, using the proposed group-constrained sparse functional connectivity demonstrated a significantly higher accuracy of 84.0% when compared with the conventional Pearson correlation-based and the sparse fuctional connectivity without group-constraint. The proposed method also yields an AUC value that is larger than 0.86, indicating good diagnostic power, especially in view of the relatively limited number of samples available in this study. The conventional correlation-based approach can only provide low to moderate performance as indicated by their relatively smaller AUC and sensitivity values. Significant improvement in sensitivity value indicates the superiority of the proposed approach in facilitating MCI classification using the conveyed disease-associated information. It is tremendously important since misclassifying a patient to be a healthy person may cause severe consequences, particularly necessary treatments that are required to delay or cure the disease may not be provided within critical treatment period. This may accelerate the progression of disease from mild to severe stage, a point where no effective treatment is available, and eventually causing death of patient.

The most discriminant regions that were selected for the best classification performance from the proposed group-constrained sparse functional connectivity networks include several regions that have been reported in previous studies such as regions located in the frontal lobes (e.g. orbitofrontal cortex [28], frontal gyri [8] and rectus gyrus [19]), temporal lobes (e.g. temporal gyri [14, 19, 57] and temporal pole [43]), and other regions such as cingulate gyri [26], amygdala [17], angular gyrus [57], and occipital gyri [42], which is in line with findings that AD, strongly related to episodic memory impairment, causes atrophies in temporal and frontal lobes at the beginning stages of the disease. It has been reported that compared with control group, the AD group has decreased functional connectivity values within the temporal lobe, but increased functional connectivity within the frontal lobe, including the prefrontal areas [63,69]. Based on fMRI studies, increase of prefrontal activation in AD during task performance may rely on increased prefrontal connectivity to compensate for reduced temporal connectivity [27]. This observation can be explained in part by the “cognitive reserve” phenomenon [62] that the ability to make such compensatory changes in the frontal lobe connectivity allows some patients to perform better than others despite equivalent pathological burdens.

Recent evidence suggests that topological organization of anatomical brain networks are critically affected by a priori atlases or spatial scales derived from random nodal parcellation [49, 53, 74]. We investigated the classification performance of the proposed method using finer parcellation scale with 238 ROIs compared with 116 ROIs in the original AAL atlas, and found that it performed with greater stability than other methods that were used for comparison.

Another important issue of the current study is the limited number of samples which may cause statistical power to be a potential concern. Leave-one-out cross-validation, as employed in this study, provides an optimistic estimate of the classification accuracy since

all except one of the subjects are used to train the classifier. Almost all information available in the dataset is used in classifier model construction. Other approaches such as *k*-fold cross-validation might be more precise when compared to leave-one-out cross-validation provided that there are sufficient data to accurately train the classifier. Hence, the results in this study have to be verified in the future with larger datasets to validate the effectiveness of the proposed technique.

5 Conclusions

We proposed a novel approach to infer functional connectivity networks from R-fMRI data for classification purposes by imposing group-based sparsity constraint via multi-task learning on training subjects. This approach minimizes spurious and redundant connections, as well as the inter-subject variability problem that may influence the performance of MCI classification. Group-based sparsity is accomplished by considering a constrained sparse linear regression model via $l_{2,1}$-norm penalization. Specifically, R-fMRI times series of a ROI is modeled as a sparse linear regression of time series of remaining ROIs via a l_1-norm penalization to filter out spurious or insignificant connections. The inter-subject stability of network structure is encouraged via a multi-task learning using l_2-norm penalization across subjects. This group-constrained sparse representation generates topologically consistent functional connectivity networks that allow direct comparison between subjects in classification. The experiment results demonstrate the capability of the proposed approach in constructing functional connectivity network that yields markedly improved classification performance compared with the conventional Pearson correlation-based network.

Acknowledgments

This work was supported in part by National Institute of Health (NIH) grants EB006733, EB008374, AG041721, EB009634, MH088520, K23-AG028982, as well as a National Alliance for Research in Schizophrenia and Depression Young Investigator Award (LW).

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: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

References

1. Achard S, Bassett DS, Meyer-Lindenberg A, Bullmore ET. Fractal connectivity of long-memory networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2008; 77(3 Pt 2):036, 104.
2. Achard S, Salvador R, Whitcher B, Suckling J, Bullmore ET. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J Neurosci. 2006; 26(1):63–72. [PubMed: 16399673]
3. Alzheimer's Association. 2012 Alzheimer's disease facts and figgues. Alzheimers Dement. 2012; 8(2):1–72. [PubMed: 22265587]
4. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4. American Psychiatric Association; 2000. Text Revision (DSMIV-TR)

5. Azari NP, Rapoport SI, Grady CL, Schapiro MB, Salerno JA, Gonzalez-Aviles A, Horwitz B. Patterns of interregional correlations of cerebral glucose metabolic rates in patients with dementia of the Alzheimer type. Neurodegeneration. 1992; 1:101–111.
6. Bain LJ, Jedrziewski K, Morrison-Bogorad M, Albert M, Cotman C, Hendrie H, Trojanowski JQ. Healthy brain aging: A meeting report from the Sylvan M. Cohen annual retreat of the University of Pennsylvania Institute on aging. Alzheimers Dement. 2008; 4(6):443–446. [PubMed: 18945646]
7. Bajo R, Maestú F, Nevado A, Sancho M, Gutiérrez R, Campo P, Castellanos NP, Gil P, Moratti S, Pereda E, Del-Pozo F. Functional connectivity in mild cognitive impairment during a memory task: implications for the disconnection hypothesis. J Alzheimers Dis. 2010; 22(1):183–193. [PubMed: 20847450]
8. Bell-McGinty S, Lopez OL, Meltzer CC, Scanlon JM, Whyte EM, Dekosky ST, Becker JT. Differential cortical atrophy in subgroups of mild cognitive impairment. Arch Neurol. 2005; 62(9): 1393–1397. [PubMed: 16157746]
9. Benton AL. The visual retention test as a constructional praxis task. Confin Neurol. 1962; 22:141–155. [PubMed: 13967555]
10. Benton, AL.; Hamsher, K. Multilingual Aphasia Examination manual. University of Iowa, Iowa City; 1976.
11. Bischkopf J, Busse A, Angermeyer MC. Mild cognitive impairment - a review of prevalence, incidence and outcome according to current approaches. Acta Psychiatr Scand. 2002; 106:403–414. [PubMed: 12392483]
12. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. Alzheimers Dement. 2007; 3(3):186–191. [PubMed: 19595937]
13. Candés EJ, Wakin MB. An introduction to compressive sampling–a sensing/sampling paradigm that goes against the common knowledge in data acquisition. IEEE Signal Process Mag. 2008; 25(2):21–30.
14. Convit A, de Asis J, de Leon MJ, Tarshish CY, De Santi S, Rusinek H. Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimers disease. Neurobiol Aging. 2000; 21(1):19–26. [PubMed: 10794844]
15. Cooper JA, Sagar HJ, Jordan N, Harvey NS, Sullivan EV. Cognitive impairment in early, untreated parkinsons disease and its relationship to motor function. Brain. 1991; 114(5):2095–2122. [PubMed: 1933236]
16. Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, Meyerand ME. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. Am J Neuroradiol. 2001; 22:1326–1333. [PubMed: 11498421]
17. Dai W, Lopez OL, Carmichael OT, Becker JT, Kuller LH, Gach HM. Mild cognitive impairment and Alzheimer disease: Patterns of altered cerebral blood flow at MR imaging. Radiology. 2009; 250:856–866.10.1148/radiol.2503080751 [PubMed: 19164119]
18. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. J Bioinform Comput Biol. 2005; 3(2):185–205. [PubMed: 15852500]
19. Fleisher AS, Sherzai A, Taylor C, Langbaum JB, Chen K, Buxton RB. Resting-state BOLD networks versus task-associated functional mri for distinguishing Alzheimer's disease risk groups. Neuroimage. 2009; 47(4):1678–1690. [PubMed: 19539034]
20. Folstein MF, Folstein SE, McHugh PR. "mini-mental state". A practical method for grading the cognitive state of patient for the clinician. J Psychiatr Res. 1975; 12(3):189–198. [PubMed: 1202204]
21. Fox MD, Zhang D, Snyder AZ, Raichle ME. The global signal and observed anticorrelated resting state brain networks. J Neurophysiol. 2009; 101(6):3270–3283. [PubMed: 19339462]
22. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. Biostat. 2008; 9(3):432–441.
23. Friston KJ, Frith C, Frackowiak RSJ, Turner R. Characterizing dynamic brain responses with fMRI: A multivariate approach. Neuroimage. 1995; 2:166–172. [PubMed: 9343599]
24. Friston KJ, Frith CD, Liddle PF, Frackowiak RS. Functional connectivity: The principal-component analysis of large (PET) data sets. J Cereb Blood Flow Metab. 1993; 13:5–14. [PubMed: 8417010]

25. Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, Belleville S, Brodaty H, Bennett D, Chertkow H, Cummings JL, de Leon M, Feldman H, Ganguli M, Hampel H, Scheltens P, Tierney MC, Whitehouse P, Winblad B. on behalf of the participants of the International Psychogeriatric Association. Expert Conference on mild cognitive impairment: Mild cognitive impairment. Lancet. 2006; 367:1262–1270. [PubMed: 16631882]
26. Gold BT, Jiang Y, Jicha GA, Smith CD. Functional response in ventral temporal cortex differentiates mild cognitive impairment from normal aging. Hum Brain Mapp. 2010; 31(8):1249–1259. [PubMed: 20063353]
27. Gould RL, Arroyo B, Brown RG, Owen AM, Bullmore ET, Howard RJ. Brain mechanisms of successful compensation during learning in Alzheimer disease. Neurology. 2006; 67(7):1011–1017. [PubMed: 17000970]
28. Grady CL, McIntosh AR, Beig S, Keightley ML, Burian H, Black SE. Evidence from functional neuroimaging of a compensatory prefrontal network in Alzheimer's disease. J Neurosci. 2003; 23(3):986–993. [PubMed: 12574428]
29. Grundman M, Petersen RC, Ferris SH, Thomas RG, Aisen PS, Bennett DA, et al. Mild cognitive impairment can be distinguished from Alzheimer's disease and normal aging for clinical trials. Arch Neurol. 2004; 61(1):59–66. [PubMed: 14732621]
30. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Machine Learning. 2004; 46(1–3):389–422.
31. Haller S, Missonnier P, Herrmann FR, Rodriguez C, Deiber MP, Nguyen D, Gold G, Lovblad KO, Giannakopoulos P. Individual classification of mild cognitive impairment subtypes by support vector machine analysis of white matter dti. AJNR Am J Neuroradiol. 2012 Epub ahead of print.
32. Horwitz B, Grady CL, Schlageter NL, Duara R, Rapoport SI. Intercorrelations of regional cerebral glucose metabolic rates in Alzheimer's disease. Brain Res. 1987; 407(2):294–306. [PubMed: 3494486]
33. Huang S, Li J, Sun L, Ye J, Fleisher A, Wu T, Chen K, Reiman E. the Alzheimer's Disease NeuroImaging Initiative. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. Neuroimage. 2010; 50(3):935–949. [PubMed: 20079441]
34. Lee H, Lee DS, Kang H, Kim BN, Chung MK. Sparse brain network recovery under compressed sensing. IEEE Trans Med Imaging. 2011; 30(5):1154–1165. [PubMed: 21478072]
35. Liu, J.; Ji, S.; Ye, J. SLEP: Sparse Learning with Efficient Projections. Arizona State University; 2009. URL http://www.public.asu.edu/jye02/Software/SLEP
36. Lynall ME, Bassett DS, Kerwin R, McKenna PJ, Kitzbichler M, Muller U, Bullmore ET. Functional connectivity and brain networks in schizophrenia. J Neurosci. 2010; 30:9477–9487. [PubMed: 20631176]
37. Matthews, CG.; Klove, H. Instruction Manual for the Adult Neuropsychology Test Battery. University of Wisconsin Medical School; Madison, WI: 1964.
38. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS–ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984; 34(7):939–944. [PubMed: 6610841]
39. Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI. Neuroimage. 2009; 44:1414–1422.
40. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology. 1989; 39(9):1159–1165. [PubMed: 2771064]
41. Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage. 2009; 44(3):893–905. [PubMed: 18976716]
42. Nobili F, Mazzei D, Dessi B, Morbelli S, Brugnolo A, Barbieri P, Girtler N, Sambuceti G, Rodriguez G, Pagani M. Unawareness of memory deficit in amnestic mci: FDG-PET findings. J Alzheimers Dis. 2010; 22(3):993–1003. [PubMed: 20858977]

43. Nobili F, Salmaso D, Morbelli S, Girtler N, Piccardo A, Brugnolo A, Dessi B, Larsson SA, Rodriguez G, Pagani M. Principal component analysis of FDG PET in amnestic MCI. Eur J Nucl Med Mol Imaging. 2008; 35(12):2191–2202. [PubMed: 18648805]
44. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. 2005; 27(8):1226–1285. [PubMed: 16119262]
45. Rakotomamonjy A. Variable selection using SVM based criteria. J Mach Learn Res. 2003; 3:1357–1370.
46. Ramsey JD, Hanson SJ, Glymour C. Multi-subject search correctly identifies causal connections and most causal directions in the dcm models of the smith et al. simulation study. Neuroimage. 2011; 58(3):838–848. [PubMed: 21745580]
47. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept Mot Skills. 1958; 8:271–276.
48. Reitan, RM.; Wolfson, D. Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation. Neuropsychological Press; Tucson, AZ: 1993.
49. Romero-Garcia R, Atienza M, Clemmensen LH, Cantero JL. Effects of network resolution on topological properties of human neocortex. Neuroimage. 2012; 59(4):3522–3532. [PubMed: 22094643]
50. Rothman AJ, Bickel PJ, Levina E. Sparse permutation invariant covariance estimation. Electronic Journal of Statisics. 2008; 2:494–515.
51. Rubinov M, Sporns O. Complex networks measures of brain connectivity: Uses and interpretations. Neuroimage. 2010; 52(3):1059–1069.10.1016/j.neuroimage.2009.10.003 [PubMed: 19819337]
52. Sachs GA, Carter R, Holtz LR, Smith F, Stump TE, Tu W, Callahan CM. Cognitive impairment: An independent predictor of excess mortality: A cohort study. Ann Intern Med. 2011; 155(5):300–308. [PubMed: 21893623]
53. Sanabria-Diaz G, Melie-García L, Iturria-Medina Y, Alemán-Gómez Y, Hernández-González G, Valdés-Urrutia L, Galán L, Valdés-Sosa P. Surface area and cortical thickness descriptors reveal different attributes of the structural human brain networks. Neuroimage. 2010; 50(4):1497–1510. [PubMed: 20083210]
54. Shen D, Davatzikos C. HAMMER: Heirarchical attribute matching mechanism for elastic registration. IEEE Trans Med Imaging. 2002; 21(11):1421–1439. [PubMed: 12575879]
55. Shipley, WC. Institute of Living Scale. Western Psychological Services; Los Angeles, Calif: 1946.
56. Smith A. The symbol-digit modalities test: a neuropsychologic test of learning and other cerebral disorders. Learning Disorders. 1968; 3:83–91.
57. Smith CD, Chebrolu H, Wekstein DR, Schmitt FA, Jicha GA, Cooper G, Markesbery WR. Brain structural alterations before mild cognitive impairment. Neurology. 2007; 68(16):1268–1273. [PubMed: 17438217]
58. Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, Ramsey JD, Woolrich MW. Network modelling methods for fmri. Neuroimage. 2011; 54(2):875–891. [PubMed: 20817103]
59. Squire LR, Zouzounis JA. Self-ratings of memory dysfunction: different findings in depression and amnesia. J Clin Exp Neuropsychol. 1988; 10(6):727–738. [PubMed: 3235647]
60. Stam CJ, de Haan W, Daffertshofer ABFJ, Manshanden I, van Cappelen van Walsum AM, Montez T, Verbunt JPA, de Munck JC, van Dijk BW, Berendse HW, Scheltens P. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. Brain. 2009; 132:213–224. [PubMed: 18952674]
61. Stam CJ, Jones BF, Nolte G, Breakspear M, Scheltens P. Small-world networks and functional connectivity in Alzheimer's disease. Cereb Cortex. 2007; 17:92–99. [PubMed: 16452642]
62. Stern Y. Cognitive reserve and Alzheimer disease. Alzheimer Dis Assoc Disord. 2006; 20(3 Suppl 2):S69–S74. [PubMed: 16917199]
63. Supekar K, Menon V, Rubin D, Musen M, Greicius MD. Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. PLoS Comput Biol. 2008; 4:e1000, 100.

64. Tomasi D, Volkow ND. Functional connectivity density mapping. Proc Natl Acad Sci U S A. 2010; 107(21):9885–9890. [PubMed: 20457896]

65. Tsutsumi R, Hanajima R, Hamada M, Shirota Y, Matsumoto H, Terao Y, Ohminami S, Yamakawa Y, Shimada H, Tsuji S, Ugawa Y. Reduced interhemispheric inhibition in mild cognitive impairment. Exp Brain Res. 2012; 218(1):21–26. [PubMed: 22234435]

66. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 2002; 15(1):273–289. [PubMed: 11771995]

67. Van Dijk KRA, Hedden T, Venkataraman A, Evans KC, Lazar SW, Buckner RL. Intrinsic functional connectivity as a tool for human connectomics: Theory, properties and optimization. J Neurophysiol. 2010; 103:297–321. [PubMed: 19889849]

68. Varoquaux G, Gramfort A, Poline JB, Thirion B. Brain covariance selection: better individual functional connectivity models using population prior. NIPS'10. 2010:2334–2342.

69. Wang K, Liang M, Wang L, Tian L, Zhang X, Li K, Jiang T. Altered functional connectivity in early Alzheimer's disease: A resting-state fMRI study. Hum Brain Mapp. 2007; 28(10):967–978. [PubMed: 17133390]

70. Wechsler, D. Manual for the Wechsler Adult Intelligence Scale - Revised. Psychological Corporation; New York: 1981.

71. Wechsler, D. WMS-R: Wechsler Memory Scale-Revised Manual. The Psychological Corporation; 1987.

72. Wee CY, Yap PT, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D. Resting-state multi-spectrum functional connectivity networks for identification of MCI patients. PLoS ONE. 2012; 7(5):e37, 828.

73. Yuan M, Lin Y. Model selection and estimation in regression with grouped variables. J Roy Stat Soc B. 2006; 68(1):49–67.

74. Zalesky A, Fornito A, Harding IH, Cocchi L, Yücel M, Pantelis C, Bullmore ET. Whole-brain anatomical networks: Does the choice of nodes matter? Neuroimage. 2010; 50(3):970–983. [PubMed: 20035887]

75. Zanetti O, Solerte SB, Cantonni F. Life expectancy in Alzheimer's disease (AD). Arch Gerontol Geriatr. 2009; 49:237–243. [PubMed: 19836639]

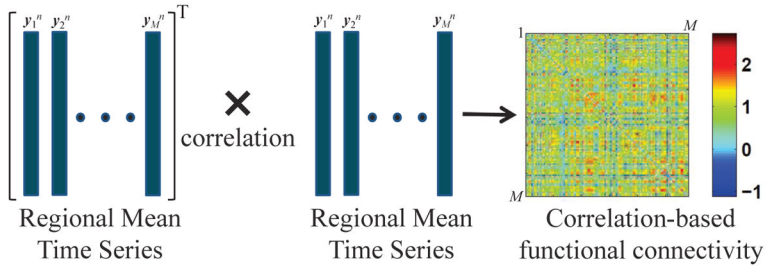

Fig. 1.
Inferring of Pearson correlation-based functional connectivity map, including the Fisher's *r*-to-*z* transformation.

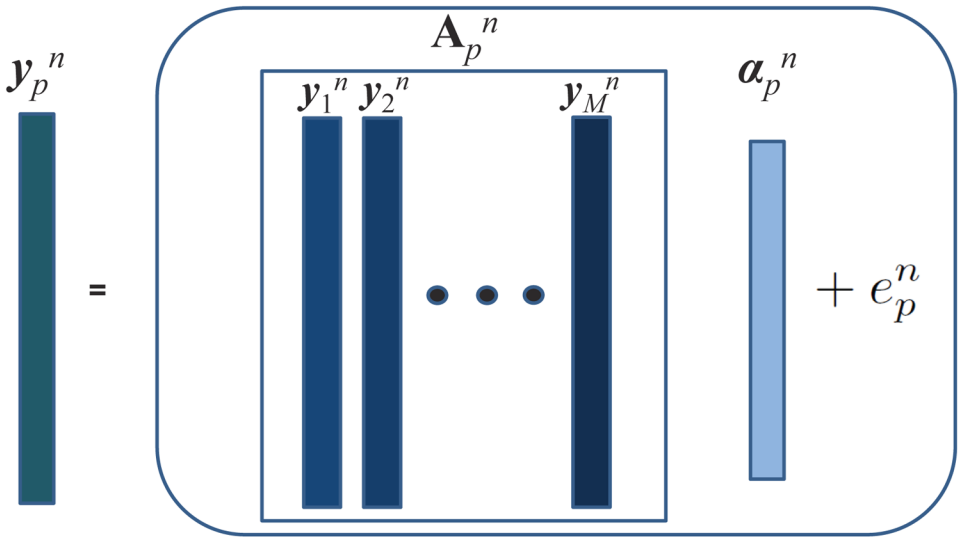

Fig. 2.

Modeling of time series $\mathbf{y}_p^n$ as a linear combination of time series of other ROIs.

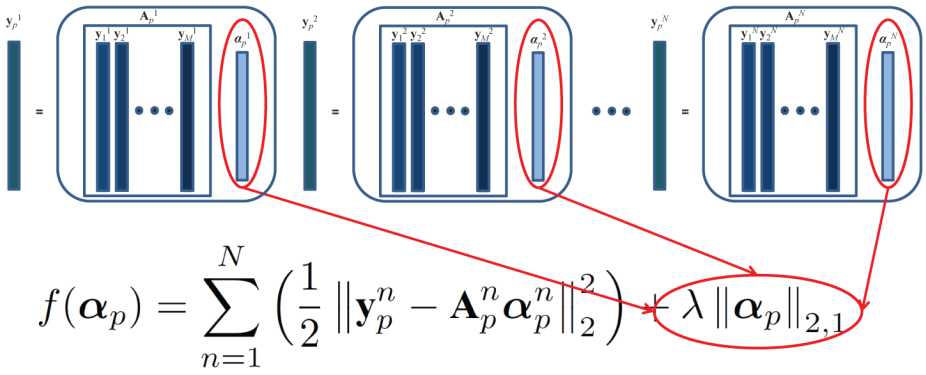

Fig. 3.
Modeling the group-constrained sparse functional connectivity of the *p*-th ROI, $f(\alpha_p)$ via multi-task learning on the same elements across all training subjects.

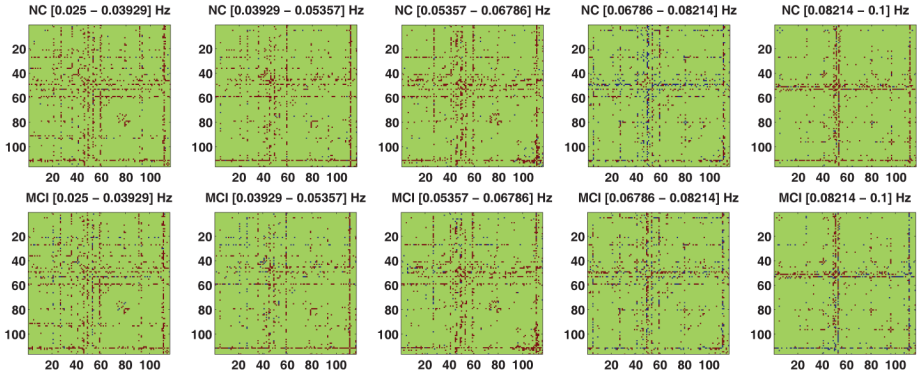

Fig. 4.

Group-constrained sparse connectivity maps with $\lambda = 0.2$ for one healthy control (NC) and one MCI patient. (Red = positive connection, blue = negative connection, green = no connection)

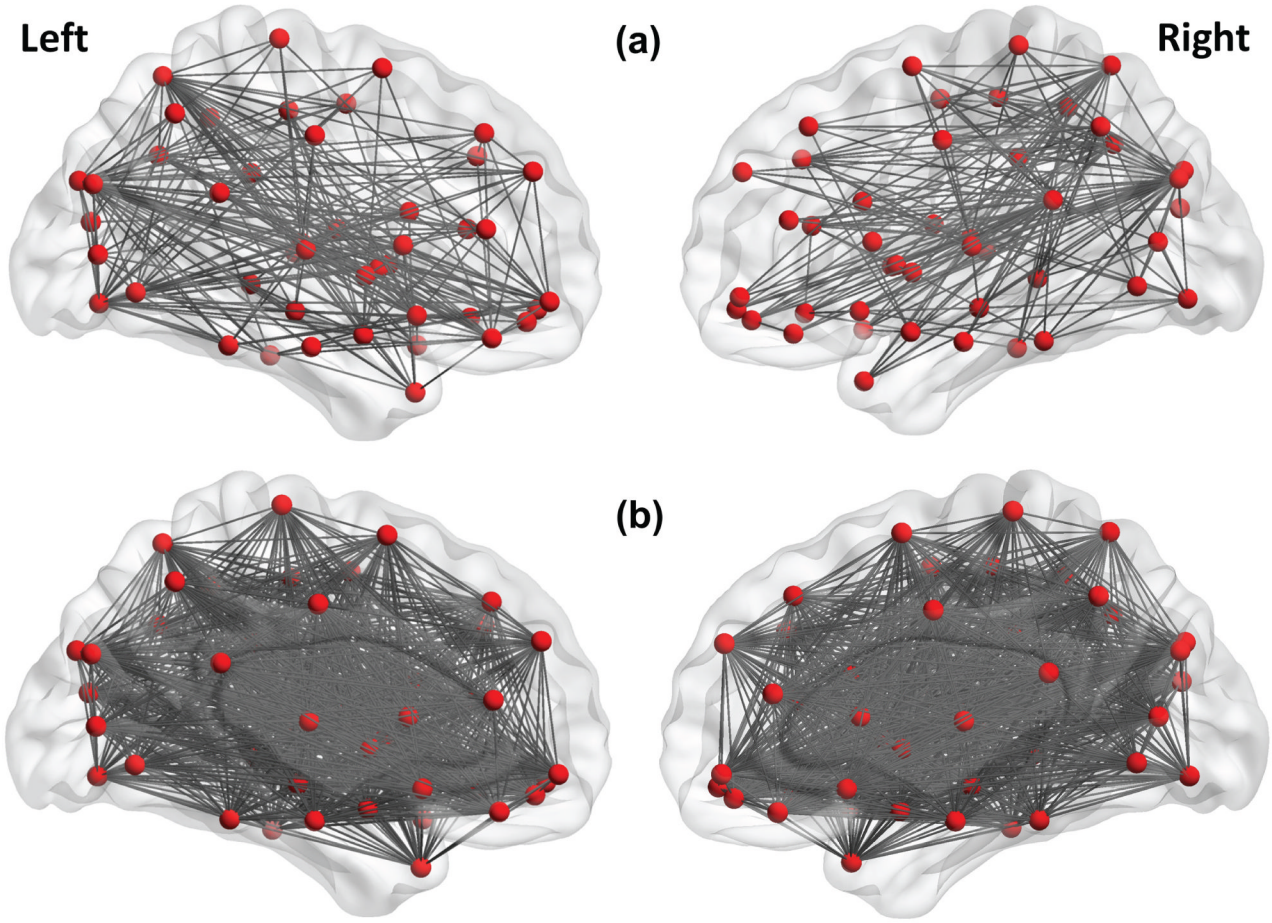

Fig. 5.

Topology structure of constrained sparse functional connectivity networks with $\lambda = 0.2$ (a) and the fully-connected Pearson correlation-based functional connectivity networks (b), after excluding the cerebellum. This visualization was created using the BrainNet Viewer (http://nitrc.org/projects/bnv/).

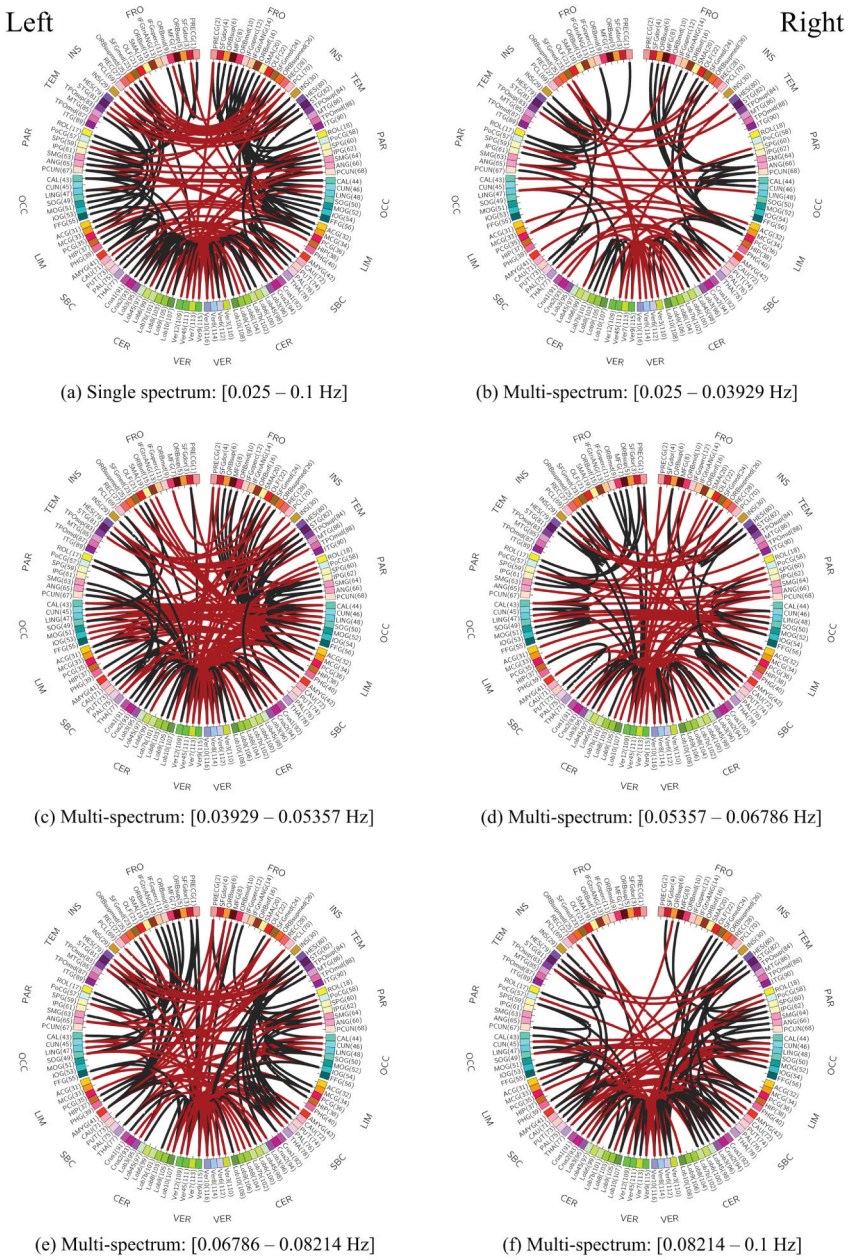

Fig. 6.

Connectogram of the group-constrained sparse functional connectivities for single spectrum and multi-spectrum cases, after removing insignificant connections (small amplitude) for better visualization. Red: inter-hemisphere connections, Black: intra-hemisphere connections. (FRO: Frontal, INS: Insula, TEM: Temporal, PAR: Parietal, OCC: Occipital, LIM: Limbic, SBC: Subcortical, CER: Cerebellum, VER: Vermis). This visualization was created using the Circos (http://circos.ca/).

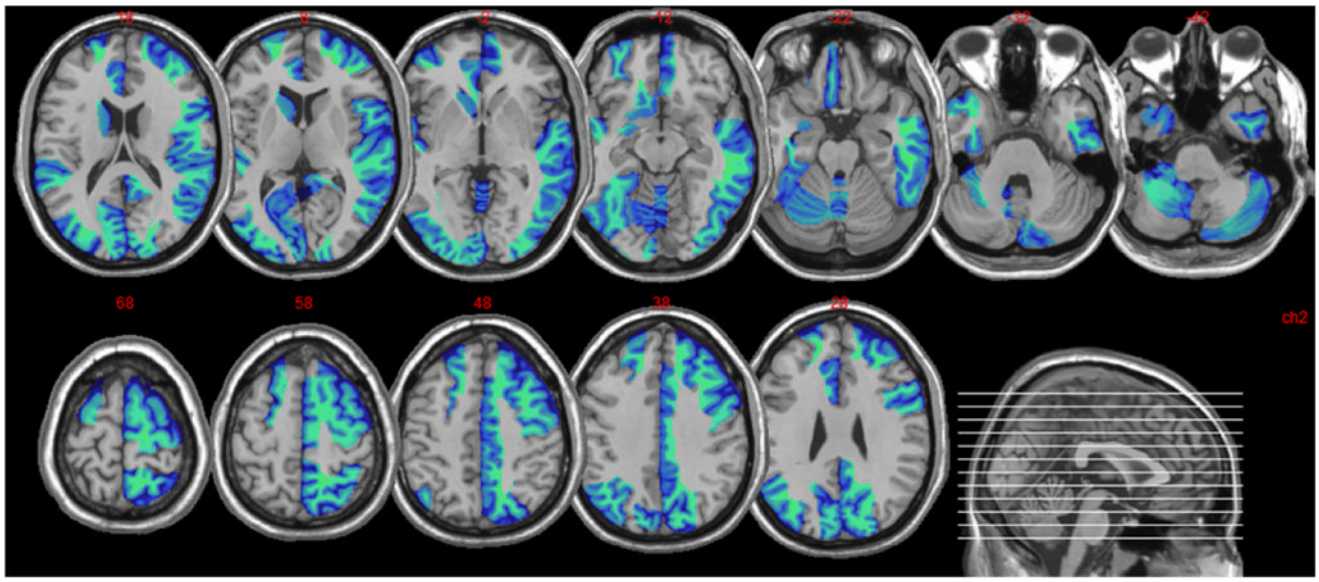

Fig. 7.
Most discriminant regions that were selected during MCI classification from the group-constrained sparse functional connectivity.

Table 1

Demographic and clinical information of the participants.

Group	MCI	Control	*p*-value
No. of subjects (Male/Female)	12/13	9/16	0.3927[a]
Age (mean ± SD)	73.0 ± 7.8	72.3 ± 8.3	0.7655[b]
MMSE (mean ± SD)	28.2 ± 1.6[c]	29.2 ± 1.1	0.0218[b]

[a] The *p* value was obtained by two-sample Chi-square (χ^2) test.

[b] The *p* value was obtained by two-sample two-tailed *t*-test.

[c] Sixteen out of 25 MCI subjects do not have MMSE score.

Table 2

Classification performance of constrained sparse and Pearson correlation-based connectivity networks for single and multi-spectral BOLD signals.

Approach	ACC (%)	AUC	SEN	SPE
Correlation + Single Spectrum	72.00	0.6912	0.6400	0.8000
Constrained + Single Spectrum	78.00	0.7904	0.7600	0.8000
Correlation + Multi-Spectral	76.00	0.7536	0.6800	0.8400
Constrained + Multi-Spectral	84.00	0.8656	0.8400	0.8400

ACC = Accuracy; SEN = SENsitivity; SPE = SPEcificity; Correlation = Pearson correlation-based connectivity; Constrained = Sparse connectivity with group-constraint

Table 3

Classification performance of sparse functional connectivity networks with and without group-constraint for single and multi-spectral BOLD signals.

Approach	ACC (%)	AUC	SEN	SPE
Sparse + Single Spectrum	72.00	0.7376	0.6000	0.8400
Constrained + Single Spectrum	78.00	0.7904	0.7600	0.8000
Sparse + Multi-Spectral	76.00	0.7840	0.8000	0.7200
Constrained + Multi-Spectral	84.00	0.8656	0.8400	0.8400

ACC = Accuracy; SEN = SENsitivity; SPE = SPEcificity; Sparse = Sparse connectivity without group constraint; Constrained = Sparse connectivity with group-constraint

Table 4

Classification performance of various functional connectivity network construction methods using the original version and finer version of AAL atlas.

Approach	No. of ROIs = 116				No. of ROIs = 238			
	ACC	AUC	SEN	SPE	ACC	AUC	SEN	SPE
Correlation + Multi-Spectral	76.00	0.7536	0.6800	0.8400	72.00	0.7556	0.5600	0.8800
Sparse + Multi-Spectral	76.00	0.7840	0.8000	0.7200	78.00	0.7024	0.6400	0.9200
Constrained + Multi-Spectral	84.00	0.8656	0.8400	0.8400	84.00	0.8960	0.8000	0.8800

ACC = Accuracy; SEN = SENsitivity; SPE = SPEcificity; Correlation = Pearson correlation-based connectivity; Sparse = Sparse connectivity without group constraint; Constrained = Sparse connectivity with group-constraint