Published in final edited form as:

Stat Med. 2016 November 30; 35(27): 4994–5008. doi:10.1002/sim.7039.

Edgewise and subgraph level tests for brain networks

Huaihou Chen^{1,2}, Bingxin Zhao¹, Eric C. Porges², Ronald A. Cohen², and Natalie C. Ebner^{2,3}

¹Department of Biostatistics, University of Florida, Gainesville, FL, USA

²Center for Cognitive Aging and Memory (CAM), Institute on Aging, McKnight Brain Institute, Department of Aging and Geriatric Research, University of Florida, Gainesville, FL, USA

³Department of Psychology, University of Florida, Gainesville, FL, USA

Abstract

Resting-state functional magnetic resonance image (rsfMRI) is a useful technique for investigating brain functional connectivity at rest. In this work, we develop flexible regression models and methods for determining differences in resting-state functional connectivity as a function of age, gender, drug intervention, or neuropsychiatric disorders. We propose two complementary methods for identifying changes of edges and subgraphs. 1) For detecting changes of edges, we select the optimal model at each edge, and then conduct contrast tests to identify the effects of the important variables while controlling the familywise error rate (FWER). 2) We adopt the network based statistics (NBS) method to improve power by incorporating the graph topological structure. Both methods have wide applications for low signal-to-noise ratio data. We propose stability criteria for the choice of threshold in the NBS procedure and utilize efficient massive parallel procedure to speed up the estimation and inference procedure. Results from our simulation studies show that the thresholds chosen by the proposed stability criteria outperform the Bonferroni threshold. To demonstrate applicability, we use both methods in the context of the Oxytocin and Aging Study to determine effects of age, gender, and drug treatment on resting-state functional connectivity, as well as in the context of the Autism Brain Imaging Data Exchange Study to determine effects of autism spectrum disorder on functional connectivity at rest.

Keywords

network based statistics; graphical structure; resting-state brain functional connectivity

1. Introduction

Graphical model and network analysis have attracted considerable attention in the statistical and mathematical fields as well as in computer science over the past few decades. One application of graphical model and network analysis is in brain functional connectivity analysis. For region-to-region functional connectivity analysis, there are several levels of approaches. In whole brain level analysis, the multivariate distance matrix regression (MDMR; [1]) has been used to study differences in functional connectivity between diagnostic groups [2]. Using pseudo- F tests based on a distance matrix, MDMR provides analysis of variance (ANOVA) test results without multivariate normality assumption for the

outcomes. For edgewise analysis, the sum of powered score test has been developed [3] with edges defined as the region-to-region connections. For subgraph level analysis, the network based statistics (NBS) testing procedure has been proposed for thresholded subgraphs [4], with subgraphs defined as topologically connected edges. However, those aforementioned methods for edgewise and subgraph analysis focus on two-group comparison, and their extension to general regression analysis is nontrivial.

In this work, we develop flexible regression models and methods for detecting functional connectivity differences associated with variables of interest. Our goals are twofold: 1) to test whether there is an overall functional connectivity difference among different groups, such as male versus female participants, older versus younger participants, participants who have received oxytocin treatment versus participants who have received placebo, or patients with autism versus healthy participants, using the MDMR method; 2) to identify the specific edges and subgraphs that drive the overall difference under a general regression model while correcting for multiple comparison and incorporating graphical structure.

For detecting the effects of variables of interest on the edges, we first select the optimal model from a set of candidate models at each edge. We then use linear contrast tests or ANOVA F tests to detect the effects of the variables of interest based on the selected model, while controlling the familywise error rate (FWER) for all the edges. For the subgraph level test, we adopt the NBS method to improve power by incorporating the graph's topological structure. An advantage of the edgewise testing method is that it generates very specific findings. A disadvantage of this method is that it is conservative and does not account for the graphical structure. In contrast, the NBS method has the advantage of high power. However, its results are not specific and depend on the choice of threshold. Thus, use of both these two methods in a complementary fashion offers potential for a wide range of applications to low signal-to-noise ratio data. To fully take advantage of the NBS method, we propose stability criteria for the choice of threshold. Furthermore, we implement massive parallel estimation and inference procedure to reduce the computational time, as both the edgewise and the subgraph methods are based on permutations.

The remainder of the paper is structured as follows. Section 2 introduces the two motivating examples. Section 3 presents the proposed methods. Sections 4 and 5 show the results of the simulation studies and applications in two neuroimaging studies. Section 6 concludes with a discussion of our work.

2. Two motivating examples

Resting-state functional magnetic resonance image (rsfMRI) is a functional neuroimaging modality for evaluating regional interactions occurring when a participant is not conducting any specific task [5]. The resting brain activity is quantified by the blood-oxygen-level dependent (BOLD) signal, which reflects changes in blood flow in the brain. rsfMRI is useful for investigating functional connectivity of the brain at rest and for examining alterations in the brain connectivity as a function of, for instance, age, gender, drug treatment, or neuropsychiatric diseases. Based on well-established anatomical or functional atlases, the whole brain is parcellated into approximately one hundred regions of interest

(ROIs). The BOLD time series are first averaged across the voxels within each ROI. Next, the intrinsic functional connectivity (iFC) map is obtained by computing the pairwise Pearson correlation of those averaged time series. Let R be the number of ROIs. The iFC map for a participant is an $R \times R$ correlation matrix. Usually Fisher transformation is applied to achieve normal distribution of the correlations.

2.1. The Oxytocin and Aging Study

The Oxytocin and Aging Study is a double-blind, placebo-controlled, randomized trial of healthy young and older men and women [6]. The study aims at investigation of the effect of intranasal oxytocin administration on modulating processing of social-cognitive and affective information as well as to determine oxytocin's mechanisms in the brain that underlie this modulation by use of resting-state functional connectivity. The study comprises a full factorial design with three factors: age (young; older), gender (female; male), and treatment (oxytocin; placebo). The mean ages for the young and older groups were 22.7 (SD=3.02) and 71.2 (SD=5.19) years, respectively. Participants were randomly assigned to intranasally self-administer either oxytocin or placebo. rsfMRI data were collected for 83 participants. The rsfMRI scan lasted about 8 minutes, and 240 time points were acquired. After exclusion of four participants with excessive head motion and two left-handed participants, the final sample comprised 77 participants with 38 young (50% female) and 39 older (56% female) participants.

We applied the standard preprocessing procedures including slice time correction, motion correction, spatial normalization and smoothing as implemented in the Functional Connectivity Toolbox [7]. Using the atlas developed by [8], we parcellated the brain into 132 ROIs. We computed the iFC map for each participant and applied Fisher transformation. This resulted in a 132×132 Fisher transformed iFC map (correlation matrix) for each participant.

In the current study, we are particularly interested in testing the effects of age, gender, and treatment on functional connectivity at rest. However, existing methods developed for two-group comparison cannot be directly applied, when there is potential for interaction effects. For example, it has shown that there is a strong age by gender interaction in the lifespan developmental trajectories of functional homotopy [9]. Therefore, we extend the methods to test the effects of age, gender, and treatment on resting-state functional connectivity under a general regression model.

2.2. The Autism Brain Imaging Data Exchange (ABIDE) Study

The ABIDE Study [10] is a multi-site observational study on autism spectrum disorder (ASD). For developmental disorders such as ASD, researchers are particularly interested in examining the abnormality of neurodevelopmental trajectories [11]. Thus one aim of the ABIDE study is to determine the extent to which ASD impacted brain functional connectivity at rest and to investigate how age interacted with an effect of ASD on resting-state functional connectivity. There were data from 674 male participants (across sites) after exclusion of participants with excessive head motion. The final sample comprised 304 ASD patients and 370 typical controls with an age range from 6 to 40 years.

The following standard preprocessing steps were applied to the rsfMRI data: slice timing correction, motion realignment, intensity normalization, nuisance signal correction, temporal filtering, registration to the Montreal Neurological Institute template, and spatial filtering. Participants' iFC maps were computed based on the Harvard-Oxford Atlas [12], which includes 112 ROIs with 48 cortical and 8 subcortical regions in each hemisphere of the brain. Applying the Fisher transformation resulted in a 112×112 matrix of Fisher transformed correlations for each individual as a measure of resting-state functional connectivity between pairs of ROIs.

Given the large age range (6 to 40 years) in the ABIDE study, the potential nonlinearity for the age trajectories has to be considered [13]. To avoid bias due to misspecified linear models, we will consider varying-coefficient models [14].

3. Methods

In this section, we propose methods for detecting edgewise and subgraph level resting-state functional connectivity changes.

3.1. Edgewise model selection and testing

We consider the following univariate generic linear model at the (r, s) th edge, for $1 \leq r < s \leq R$,

$$Y_{i,rs} = \beta_{x,rs} X_i + \theta_{z,rs} Z_i + \epsilon_{i,rs},$$

$$\epsilon_{i,rs} \sim N(0, \sigma_{rs}^2), \quad i = 1, \dots, n, \quad (1)$$

where X_i is a $q_x \times 1$ vector of variables of interest, Z_i is a $q_z \times 1$ vector of covariates, $\beta_{x,rs}$ and $\theta_{z,rs}$ are the corresponding coefficients, and n is the number of participants. The measurement errors $\epsilon_{i,rs}$, $i = 1, \dots, n$, at the (r, s) th edge are assumed to be independent and identically distributed and follow normal distribution with mean zero and variance σ_{rs}^2 . As we are considering the univariate model at each edge, the correlations between different edges are not modeled here (see also [15]). In model (1), X_i can include age, gender, and treatment variables as well as their interaction terms in the Oxytocin and Aging sample. Under the generic model, model selection is conducted to improve efficiency for estimation and power for testing. We propose the following simultaneous selection and testing procedure based on p -values for the overall effects of X_i .

Let S_l , $l = 1, \dots, L$, be appropriate subsets of $\{1, \dots, q_x\}$ for X_i , which lead to reasonable sub-models of (1). For example, if a higher order interaction term is included in a model, then the lower order terms should also be included. For each sub-model, we consider the following hypothesis at each edge.

$$H_0: \beta_{x,rs}^{S_l} = 0 \quad \text{versus} \quad H_a: \beta_{x,rs}^{S_l} \neq 0,$$

$$l = 1, \dots, L, \quad 1 \leq r < s \leq R. \quad (2)$$

Denote $RSS_0 = \sum_{i=1}^n (Y_{i,rs} - \theta_{z,rs} \mathbf{Z}_i)^2$ and $RSS_l = \sum_{i=1}^n (Y_{i,rs} - \beta_{x,rs}^{S_l} \mathbf{X}_i^{S_l} - \theta_{z,rs} \mathbf{Z}_i)^2$ as the residual sum of squares under the null and alternative hypothesis, respectively. The

$F = \frac{(RSS_0 - RSS_l)/q_l}{RSS_l/(n - q_z - q_l)}$ statistics can be used for the hypothesis (2), where $q_l = |\mathbf{S}_l|$ is the number of $\mathbf{X}_i$ variable in the l th sub-model. Under the null hypothesis, the F statistics follows the F distribution with degrees of freedom q_l and $n - q_z - q_l$. Let $p_{rs,l}$ be the p -value of the test under the l th sub-model. The optimal model is selected from the candidate pool $\mathbf{S}_l$, $l = 1, \dots, L$, as the one with the smallest p -value $p_{rs}^* = \min_{1 \leq l \leq L} p_{rs,l}$. Next, we use p_{rs}^* as the test statistics for combining model selection and testing. Denote $N_e = R(R-1)/2$ as the total number of unique edges. To control the FWER for the multiple tests at the N_e edges, we derive the null distribution for

$$\min_{1 \leq r < s \leq R} p_{rs}^* = \min_{1 \leq l \leq L, 1 \leq r < s \leq R} p_{rs,l}$$

through a permutation procedure (see also chapter 2 in [16]), which is denoted as $\mathcal{P}_0$. Let c_α be the lower α th quantile of $\mathcal{P}_0$. The null hypothesis is rejected at the (r, s) th edge, if $p_{rs}^* < c_\alpha$. The FWER is maintained at the nominal level α across all edges, that is

$$Pr \left(\bigcup_{1 \leq r < s \leq R} \{p_{rs}^* \geq c_\alpha\} \middle| H_0 \right) = Pr \left(\min_{1 \leq r < s \leq R} p_{rs}^* \geq c_\alpha \middle| H_0 \right) = \alpha.$$

Here, ' $|H_0$ ' means 'assuming all the null hypotheses are true.'

We obtain the null distribution $\mathcal{P}_0$ by permuting the participant ID of the predictors $\mathbf{X}_i$. Specifically, we obtain the permuted predictors $\{\mathbf{X}_{\sigma(1)}, \dots, \mathbf{X}_{\sigma(n)}\}$ by permuting the participant ID $\{1, \dots, n\}$, where $\sigma(\cdot)$ is the permutation operator and the predictors $\mathbf{X}_i$ is permuted as a block. During the permutation we keep the structure between the outcomes and the covariates $\mathbf{Z}_i$ fixed. This nonparametric estimation of $\mathcal{P}_0$ does not assume independence among the N_e tests. The p -value at each edge can be calculated based on the estimate of $\mathcal{P}_0$. We will show in both data examples that the selection improves power for testing the overall effect compared to a full model based approach. For those edges with significant overall effects, we conduct further contrast tests or analysis of variance (ANOVA) tests for each variable of interest. The Bonferroni correction is applied for those follow-up tests.

Efficient massive parallel algorithm—As the number of regions R increases, the number of edges increases quadratically. For instance, when R increases from 100 to 200, the total number of pairs increases from 4,950 to 19,900. In practice, to obtain accurate p -values, thousands of permutations have to be conducted for the derivation of the empirical null distribution. Consequently, the edgewise estimation and inference procedure is computationally intensive when the number of edges is large. By taking advantage of the identical design matrix for all the edges, we use efficient massive parallel algorithm to speed up the estimation and inference procedure. Let $\mathbf{X}^* = [\mathbf{X} : \mathbf{Z}]$ be the stacked design matrix.

Denote $\gamma_{rs} = (\beta_{x,rs}^T, \theta_{z,rs}^T)^T$ and denote $\mathbf{Y}_{rs}$ as the vector of outcomes at the (r, s) th edge. We have

$$\begin{aligned}\hat{\gamma}_{rs} &= (\mathbf{X}^{*T} \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{Y}_{rs}, \quad 1 \leq r < s \leq R, \quad \text{and} \\ \hat{\epsilon}_{rs} &= (\mathbf{Y}_{rs} - \hat{\mathbf{Y}}_{rs}) = (\mathbf{I} - \mathbf{H}) \mathbf{Y}_{rs}, \quad 1 \leq r < s \leq R,\end{aligned}$$

where $\mathbf{H} = \mathbf{X}^*(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{X}^{*T}$ is the projection matrix, and $\mathbf{I}$ is an identity matrix. The residual sum of squares is given by $RSS_{rs} = \hat{\epsilon}_{rs}^T \hat{\epsilon}_{rs} = \mathbf{Y}_{rs}^T (\mathbf{I} - \mathbf{H}) \mathbf{Y}_{rs}$. We rewrite the formula in matrix format for all the edges as follows.

$$\begin{aligned}\hat{\boldsymbol{\gamma}} &= (\mathbf{X}^{*T} \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{Y}, \\ \hat{\mathbf{E}} &= (\mathbf{Y} - \hat{\mathbf{Y}}) = (\mathbf{I} - \mathbf{H}) \mathbf{Y}, \\ \mathbf{RSS} &= \text{diag} \{ \hat{\mathbf{E}}^T \hat{\mathbf{E}} \},\end{aligned}$$

where $\mathbf{Y}$ is the $n \times N_e$ outcome matrix and $\mathbf{RSS}$ is the $N_e \times 1$ vector of residual sum of squares at all the edges. Analogously, under the null hypothesis, we compute the estimate of $\boldsymbol{\theta}$ and the residual sum of squares denoted as $\mathbf{RSS}_0$. Based on the two vectors of residual sum of squares $\mathbf{RSS}_0$ and $\mathbf{RSS}$, we calculate the F test statistics and the p -values at each of the N_e edges. We utilize the R package *vows* [15] to implement the massive parallel method.

3.2. Network based statistics (NBS) for subgraph-level testing

Let c be a predetermined small threshold, and let p_{rs} be the p -value of a test for hypothesis (2) at the (r, s) th edge. The NBS testing procedure [4] assumes that the edges associated with the variable of interest are not isolated from each other and form a cluster. More specifically, based on massive univariate models (1), the NBS method selects edges with p -values of tests for hypothesis (2) that are smaller than the given threshold c . Subgraphs are determined on those selected edges, which comprise topologically connected edges. The number of edges in each subgraph (i.e., the size of the subgraph) is used as the test statistics. Permutation is utilized to obtain the null distribution of the test statistics [17]. The procedure comprises the following steps:

In Step 1, we screen edges with p -value below the threshold c , that is

$$\mathcal{E} = \{(r, s) : p_{rs} < c, \quad 1 \leq r < s \leq R\}.$$

In Step 2, we determine connected subgraphs $G_1, \dots, G_m$ among those selected edges and compute the number of edges for each of the subgraphs.

Isolated edges are ignored, because they are likely due to randomness. In contrast to the edgewise hypotheses (2), the subgraph level hypotheses of interest are

H_0 : there is no effect of the predictors X_i on G_j versus H_a :
there is an effect of the predictors X_i on G_j ,
 $j=1, \dots, m$.

(3)

Denote $G_j = (V_j, E_j)$, where V_j are the nodes and E_j are the edges. Denote A_j as the adjacency matrix for G_j such that its element $a_{j,rs}$ equals one when there is an edge from node r to node s , and equals zero when there is no edge [18]. Here, $\cup_{j=1}^m V_j \subseteq \{1, \dots, R\}$ and $\cup_{j=1}^m E_j \subseteq \mathcal{E}$. For each subgraph G_j , all the nodes are connected. That is, there is at least one path in the subgraph between two nodes, while the nodes in different subgraphs are disconnected. We use the *clusters* function in the R package *igraph* for searching the subgraphs in $\mathcal{E}$ [19]. We compute the number of edges in each subgraph and denote it as $T_j = |E_j| = \sum_{r < s} a_{j,rs}$, where $a_{j,rs}$ is the (r, s) th element of the adjacency matrix A_j . $T_G = \max_{1 \leq j \leq m} T_j$ is defined as the maximum size of the subgraphs. Analogous to the edgewise test, we use T_j as the test statistics for hypothesis (3) and use the null distribution of T_G to control FWER for the multiple subgraphs.

In Step 3, we permute the participant ID for the corresponding variables and repeat Steps 1 and 2, to generate the null distribution of T_G . The corresponding variables are permuted as a block.

Let $\{T_1^b, \dots, T_{m_b}^b\}$ be the size of the m_b subgraphs derived from the b th permuted data, for $b = 1, \dots, B$, where B is the number of conducted permutations. The estimated null distribution of T_G is

given by $P(T_G \leq k) = \frac{1}{B} \sum_{b=1}^B I\left(\max_{1 \leq j \leq m_b} T_j^b \leq k\right)$, for $2 \leq k \leq N_{\mathcal{E}}$, where $I(\cdot)$ is the indicator function. Denote c_α as the $(1 - \alpha)$ th sample quantile for the null distribution of T_G .

In Step 4, we reject the null hypothesis for subgraph G_j if $T_j > c_\alpha$. The adjusted p -values can be obtained by $Pr(T_j > T_G | H_0)$.

The FWER is controlled for the subgraph level tests, because

$$Pr\left(\cup_{1 \leq j \leq m} \{T_j > c_\alpha\} \mid H_0\right) = Pr\left(\max_{1 \leq j \leq m} T_j > c_\alpha \mid H_0\right) = \alpha.$$

In general, as a subgraph level test, the NBS test is more powerful for functional connectivity analysis than the edgewise testing described in Section 3.1, because the number of edges $N_{\mathcal{E}}$ is usually much larger than the number of subgraphs under consideration m . By

reducing the number of tests, the power of the NBS test is improved. In addition, NBS also has visualization advantages for the identified subgraphs, because topologically all the nodes in each subgraph are connected.

Stability selection for the threshold—Stability selection was first used in clustering analysis for the choice of number of clusters [20, 21]. It was then further developed for variable selection in regression models [22] and graphical models [23, 24] for selecting the tuning parameter with the penalty term. Random sampling of the subjects such as bootstrap or subsample without replacement is often used to generate the stability criteria. In the recent work [25], bootstrap of the time series was used to obtain the variation of the graphical summary indices. Stability criteria have been shown to outperform other methods such as cross-validation, AIC, and BIC [24].

Despite the aforementioned advantages, the power of the NBS test depends on the arbitrarily specified threshold c . To our knowledge, there is no rigorous selection criterion for the choice of the threshold. To take full advantage of the NBS method, we developed stability criteria for the choice of the threshold. In particular, bootstrapped or Jackknife samples are drawn first. Next, the samples are pursued with edgewise testing and thresholding for obtaining the graphs. The graph size, or the similarity among the graphs, are computed. The coefficient of variation of the graph size or the graph similarity is used as the stability criterion for the choice of threshold. The basic idea is to choose the graph with a high chance of reproducibility. The procedure comprises the following steps:

In Step 1, we bootstrap the original sample. Denote $(Y^{(1)}, X^{(1)}, Z^{(1)}), \dots, (Y^{(B)}, X^{(B)}, Z^{(B)})$ as the bootstrap samples, where B is the number of bootstrap samples. The Jackknife method can also be used, especially when the sample size is moderate.

In Step 2, for a sequence of pre-specified thresholds $\{c_1, \dots, c_K\}$, we apply the edgewise testing procedure and threshold the p -values for each bootstrapped sample to obtain the graphs $G(b, c_j)$, $b = 1, \dots, B, j = 1, \dots, K$, where $G(b, c_j)$ is the graph corresponding to threshold c_j in a bootstrap sample b .

In Step 3, we compute the graph size or similarity among the graphs $G(b, c_j)$, $b = 1, \dots, B$, for each threshold c_j , $j = 1, \dots, K$. The coefficient of variation (CV) of the subgraph size or similarity is computed and used as the criterion for the choice of threshold.

Denote $A(b, c_j)$ and $T(b, c_j)$ as the adjacency matrix and the number of edges for $G(b, c_j)$, respectively. In Step 3, we consider two criteria. We choose c^* by minimizing the coefficient of variation of the subgraph size

$$CV[T(c_j)] = \frac{\sqrt{\frac{1}{B-1} \sum_{b=1}^B [T(b, c_j) - \bar{T}_j]^2}}{\bar{T}_j},$$

where $\bar{T}_j = \frac{1}{B} \sum_{b=1}^B T(b, c_j)$ is the mean subgraph size averaged across all the B samples. We denote the method as SM1. Alternatively, we consider the following similarity measure between two graphs with adjacency matrices $A(b, c_j)$ and $A(b', c_j)$.

$$d[A(b, c_j), A(b', c_j)] = \sum_{k=1}^{\infty} \sum_r \sum_s [\lambda A(b, c_j) \circ A(b', c_j)]_{rs}^k = [I - \lambda A(b, c_j) \circ A(b', c_j)]^{-1},$$

where λ is a small scalar for controlling the convergence of the series, and $\circ$ is the Hadamard product operator for matrices. $d[A(b, c_j), A(b', c_j)]$ counts the steps in the random walk of the common subgraphs between $G(b, c_j)$ and $G(b', c_j)$, which is an indicator of similarity between the two graphs [26]. $D(c_j)$ is the similarity matrix among the graphs obtained under the threshold c_j . We notice that normalizing by the graph size improves the similarity measure. $\bar{d}_j = \frac{1}{B(B-1)} \sum_{b \neq b'} d[A(b, c_j), A(b', c_j)]$ and

$$CV[D(c_j)] = \frac{1}{\bar{d}_j} \sqrt{\frac{2}{B(B-1)} \sum_{b < b'} \left\{ d[A(b, c_j), A(b', c_j)] - \bar{d}_j \right\}^2}$$

are the mean and the coefficient of variation of the upper triangular elements of $D(c_j)$, respectively. We choose the optimal c^* by minimizing $CV[D(c_j)]$ with respect to c_j . We denote the method as SM2.

4. Simulation studies

In this section, we conduct simulations to evaluate the proposed stability selection for the threshold in the NBS method. We compare the proposed selection methods SM1 and SM2 with fixed thresholding including the Bonferroni threshold and a most liberal threshold of 0.05. The Bonferroni threshold is 0.05 divided by the number of tests, which controls the FWER. To evaluate the performance of different methods, we compute the number of false positive and false negative edges.

We first generate a binary predictor X_i , $i = 1, \dots, n$, from a Bernoulli distribution with success probability $p = 0.5$. We then generate the outcomes as the mimic Fisher-transformed correlations from multivariate normal distribution based on the following model.

$$Y_i = X_i \beta + \epsilon_i, \\ \epsilon_i \sim N(0, \Sigma), \quad i = 1, \dots, n.$$

Here Σ is a $N_e \times N_e$ covariance matrix with compound symmetry correlation structure, and N_e is the total number of edges. The diagonal and off-diagonal components of Σ are 0.3 and 0.1, respectively. We consider $R = 40$, $n = 100$ or 150, and three sparsity cases. For cases 1, 2, and 3, we choose the last 5% (39), 10% (78) and 15% (117) edges to have nonzero coefficients to mimic the sparse effects. The nonzero coefficients β are specified between 0.45 and 0.6 for case 1, between 0.4 and 0.5 for case 2, and between 0.35 and 0.45 for case

3, which vary across different edges. Figure 1 displays the β coefficients for the edges and graph plots for the connected ROIs.

Table 1 summarizes the simulation results. The Bonferroni threshold is conservative and leads to close to zero false positive edges for all the scenarios. However, the number of false negative edges is large especially for cases 2 and 3. In contrast, the hard threshold of 0.05 is the most liberal one. It results in the smallest number of false negative findings, while the number of false positive findings is large, as expected, because the threshold of 0.05 does not account for the multiple testing. As shown in Table 1, the proposed methods are better than both fixed thresholds in the overall control of false positive and false negative findings for all the settings. The SM2 is better than the SM1, as the SM2 accounts for the graph structure while SM1 only considers the graph size. Also, when the sample size is increased from 100 to 150 the performances of the proposed methods improve.

5. Two applications

5.1. The Oxytocin and Aging Study

We apply the MDMR method [1, 2] and our newly developed methods to investigate the effects of age, gender, and treatment on brain functional connectivity at rest. We first compute the distance between two participants' iFC maps by calculating the Frobenius norm of the difference between the two matrices. We then apply the pseudo-F test based on the $n \times n$ distance matrix to test the main effects and interaction effects of age, gender, and treatment on iFC. There is a significant age $\times$ gender interaction effect ($p=0.015$), as well as significant main effects for age ($p=0.002$), gender ($p=0.003$) and treatment ($p=0.024$).

Next, we conduct edgewise regression analysis for the Fisher-transformed correlation outcomes using age, gender, treatment, and their interactions as predictors. We consider eight candidate models at each edge. In particular, one model includes only the main effects of age, gender, and treatment as predictors. Additionally, six models include the two-way interactions treatment $\times$ age, treatment $\times$ gender, and/or age $\times$ gender interactions together with the main effects. The saturated model (4) includes all the main and two-way and three-way interaction effects, which is a special case of the generic model (1).

$$\begin{aligned}
 Y_{i,rs} = & \beta_{A,rs} \text{age}_i + \beta_{S,rs} \text{gender}_i + \beta_{T,rs} \text{treatment}_i \\
 & + \beta_{AS,rs} \text{age}_i \times \text{gender}_i + \beta_{AT,rs} \text{age}_i \times \text{treatment}_i + \beta_{ST,rs} \text{gender}_i \times \text{treatment}_i \\
 & + \beta_{AST,rs} \text{age}_i \times \text{gender}_i \times \text{treatment}_i + \theta_{z,rs} Z_i + \epsilon_{i,rs}, \\
 i=1, \dots, n, \quad \epsilon_{i,rs} \sim N(0, \sigma_{rs}^2), \quad 1 \leq r < s \leq R,
 \end{aligned} \tag{4}$$

where the outcome $Y_{i,rs}$ is the Fisher-transformed Pearson correlation between the r th and the s th ROIs for the i th participant, and Z_i is the mean framewise displacement (mFD; [27]) variable measuring head motion. We apply the method described in Section 3.1 to select the optimal model from the eight candidate models and to test the overall effect of age, gender, or treatment on resting-state connectivity at each edge. Our edgewise hypotheses state H_0 : there is no age, gender, or treatment effect on resting-state connectivity, versus, H_I : there is

an age, gender, or treatment effect on resting-state connectivity under the model I , $I = 1, \dots, 8$.

We permute the participant IDs for the age, gender, and treatment variables. For each permutation test, we carry out 1000 permutations to derive the null distribution of the test statistics with the efficient massive parallel algorithm. Applying the null distribution of the minimum p -value, there are 137 edges with a significant overall effect at the level of 0.05, while controlling the FWER. For comparison, we conduct permutation tests under the saturated model (4) without model selection and find 73 significant edges. This is remarkably smaller than the number of significant edges of the selection procedure. We therefore conclude that the proposed procedure is more powerful in detecting the overall effect by selecting the optimal candidate model. Among those 137 selected models, 57% are the main effects model and 32% are the model with age $\times$ gender interaction. The saturated model (4) is not selected in any of those 137 models. Following up those edges with significant overall effects, we conduct ANOVA tests to identify the effects for age and treatment, respectively. For the sake of space, we subsequently skip the individual test for the gender variable, as those results can be obtained similarly to age and treatment. Overall, age has a dominant negative effect on most of the ROI-to-ROI connectivity in that older age is associated with lower functional connectivity at rest. There are 113 edges with significant age effect after Bonferroni correction. Due to the small effect size, only two edges survive the Bonferroni correction for the treatment effect.

To improve power, we apply the NBS methods described in Section 3.2 to detect subgraph level differences. We use the Jackknife procedure to select the thresholds, as the sample size is moderate and bootstrap samples tend to choose less unique participants. The lower limit of the threshold is fixed at the Bonferroni threshold $0.05/8646 = 5.783021 \times 10^{-6}$, while the upper limit is fixed at 0.01. For the overall effect, the SM2 method select a threshold of 4.7×10^{-5} . The instability measure displays a monotone decreasing trend in the thresholds. Instead of using the minimum value (i.e., the Bonferroni threshold), we apply the rule of one standard deviation from the minimum, which balances instability and conservativeness. A similar rule is often used in tuning parameter selection in penalized regression methods (see e.g., [28]), where the optimal value balances the predicted mean squared error and model complexity. Thresholding the edges with the selected threshold leads to a graph with 109 nodes and 479 edges. The largest connected subgraph for the overall effect contains 106 nodes and 477 edges. We conduct 1000 permutations to obtain the null distribution of the test statistics T_G , with the participant IDs of the age, gender, and treatment variables permuted. The permutation test results in a p -value less than 0.001 for the subgraph. We apply a similar procedure to test the age and treatment effects, with the participant IDs of the age/treatment variable permuted. The SM2 method selects the thresholds 3.0×10^{-5} and 8.3×10^{-4} for the age and treatment effects, respectively, achieving the minimum values for the instability criteria. The largest subgraphs defined by thresholding contains 70 nodes and 145 edges for the age effect (excluding 3 isolated edges) and 15 nodes and 16 edges for the treatment effect. The corresponding p -values are less than 0.001 and 0.234, respectively.

Figure 2 displays the brain map of the subgraph with the age effects using the BrainNet Viewer toolbox [29]. The first row from left to right are lateral view of left hemisphere, top

side view of brain, and lateral view of right hemisphere. The second row from left to right are medial view of left hemisphere, bottom side view of brain, and medial view of right hemisphere. The third row are frontal side and back side view of brain. As shown, age has a strong effect on resting-state functional connectivity in subcortical ROIs including amygdala, caudate, palladium, and cerebellum, as well as in cortical ROIs including cortex, temporal lobe, and planum polare. Furthermore, age has an effect on both intra- and inter-hemisphere connectivity. To summarize the pattern, we utilize the six functional divisions proposed by [30]. These include 10 primary sensory-motor (SM) ROIs, 44 unimodal association ROIs, 23 heteromodal association ROIs, 14 paralimbic ROIs, 4 limbic ROIs, and 37 subcortical ROIs. Table 2 summarizes the age-related resting-state functional connectivity changes within and between these six functional divisions. As shown, connectivity both within and between functional areas is strongly influenced by age, in that low functional connectivity at rest was associated with old age. The connectivity between primary sensory-motor ROIs and the unimodal ROIs, primary sensory-motor ROIs and the paralimbic, unimodal, and paralimbic ROIs is most affected.

Figure 3 displays the graphical plots for the largest subgraph with treatment effect using *igraph*. Although the effect of treatment on the subgraph is not statistically significant, there is a trend that oxytocin relative to placebo intranasal administration influences resting-state functional connectivity in the caudate. The caudate is part of the basal ganglia and is known to play an important role in voluntary movement [31]. Low signal-to-noise ratio for the treatment effect may have made it difficult to detect the potential effect by exhaustive search at the edgewise or subgraph level, given the moderate sample size.

5.2. The ABIDE Study

In the subsection, we investigate the effects of ASD on resting-state functional connectivity, which may explain the neurobehavioral abnormalities observed among ASD patients. The age range (6 to 40 years) in the ABIDE study is quite large. We therefore consider the following varying-coefficient models to avoid bias due to misspecified linear models.

$$Y_{i,rs} = f_{rs}(\text{Age}_i) + \beta_{rs}(\text{Age}_i) \times \text{ASD}_i + \theta_{z,rs} \mathbf{Z}_i + \epsilon_{i,rs}, \quad i=1, \dots, n, \quad \epsilon_{i,rs} \sim N(0, \sigma_{rs}^2), \quad 1 \leq r < s \leq R, \quad (5)$$

where $f_{rs}(\cdot)$ is the mean age trajectory of the connectivity for the typical controls at the (r, s) th edge, $\beta_{rs}(\cdot)$ is the vary-coefficient function indexing the age trajectory difference between the ASD patients and the typical controls, and $\mathbf{Z}_i$ represents the intelligence quotient (IQ) and mFD variables. We assume that both the baseline and varying coefficient functions can be approximated by B-spline basis functions [32] in their support T . Specifically, $f_{rs}(t) = \mathbf{B}(t)\boldsymbol{\beta}_{f,rs}$ and $\beta_{rs}(t) = \mathbf{B}(t)\boldsymbol{\beta}_{v,rs}$, where $\mathbf{B}(t)$ are a set of B-spline basis functions, $\boldsymbol{\beta}_{f,rs}$ and $\boldsymbol{\beta}_{v,rs}$ are the spline coefficients. Based on the spline representation, we can rewrite model (5) in a linear model format. Thus, it can be viewed as a special case of the generic model (1). Under the general model (5), we are particularly interested in testing the following hypotheses (6) and (7) in sequence.

$$H_0^{(1)}: \beta_{rs}(t) = 0 \quad \text{for all } t \in T, \quad \text{versus} \quad H_a^{(1)}: \beta_{rs}(t) \neq 0 \quad \text{for some } t \in T; \quad (6)$$

$$H_0^{(2)}: \beta_{rs}(t) = c_{rs} \quad \text{for all } t \in T, \quad \text{versus} \quad H_a^{(2)}: \beta_{rs}(t) \neq c_{rs} \quad \text{for some } t \in T, \quad (7)$$

where c_{rs} is a constant. The null hypothesis in (6) states that there is no group difference between ASD patients and typical controls at the edge. If the null hypothesis in (6) is rejected, we then test (7) for an age by ASD interaction. The null hypothesis in (7) indicates only a main effect of ASD at the edge, while the corresponding alternative hypothesis states there is an interaction effect between age and ASD, as $\beta_{rs}(t)$ depends on age.

To efficiently study the effects of ASD on resting-state functional connectivity, we consider the following three candidate models: i) the ASD main effect model with $\beta_{rs}(t) = \beta_{v0,rs}$, ii) ASD main effect and linear interaction with age model with $\beta_{rs}(t) = \beta_{v0,rs} + \beta_{v1,rs} \times t$, and iii) the B-spline based general varying coefficient model with $\beta_{rs}(t) = \mathbf{B}(t)^T \boldsymbol{\beta}_{v,rs}$. For all the three models, we fit the baseline function with B-spline basis functions.

We start by testing the hypothesis (6) under each of the three candidate models. We select the optimal model from the three candidate models at each of the 6,216 edges using the p -values for hypothesis (6). Approximately, 37% of the selected models are the nonlinear varying coefficient function model. We apply the selection and testing procedure described in Section 3.1 by permuting the ASD variable. There are 53 edges with significant ASD effects. Among those 53 edges, 42 have ASD main effects only, 10 have nonlinear interaction effects of age, and ASD and one has a linear interaction effect. Moreover, 48 edges of the 53 are connected and form a subgraph with 29 nodes, which are all cortical ROIs. We further test the ASD by age interaction effect, i.e., hypothesis (7), for those 11 edges with (linear or nonlinear) interaction between ASD and age. Bonferroni correction results in 4 significant edges. We plot the mean trajectories and the varying coefficient function for two of those four edges in Figure 4. As shown in the plots, the nonlinear pattern of the varying coefficient functions is convincing and unignorable. The varying coefficient functions show that the difference between ASD patients and typical controls are mainly in the age interval of 12-28 years at the two edges. This result suggests that it is necessary to consider more flexible models than simple linear models. For comparison, we conduct permutation test based on the varying coefficient model without model selection and find 37 significant edges. There is a considerable reduction in the number of significant edges compared to the procedure with model selection, which further supports the advantage of model selection.

Next, we apply the NBS method described in Section 3.2 to improve the power for the tests and for better visualization of the subgraphs. The threshold in the NBS procedure is selected using the proposed stability method SM2. Test of the overall ASD effect, that is test of hypothesis (6), results in the largest subgraph with 81 nodes and 318 edges after thresholding using the selected threshold. Figure 5 shows the brain map of the subgraph. The

subgraph consists of both subcortical and cortical ROIs, in contrast to the one obtained by the method described in Section 3.1. The left and right thalamus, the left caudate, and the right hippocampus are connected with the cortical ROIs. Applying the permutation testing procedure by permuting the ASD variable, we get a p -value of less than 0.001 for the subgraph. Furthermore, we test the interaction effect between ASD and age, which is the hypothesis (7). This results in the largest subgraph with 37 edges and 31 nodes with the left and right thalamus also included as shown in Figure 6. The p -value for the subgraph is less than 0.001. Also, most of those 37 edges show nonlinear deviation in the age trajectory of the ASD patients from the typical controls.

For optimized summary of the NBS results, we check the intra and inter-hemisphere connectivity changes among the 112 left-right corresponding ROIs. Specifically, we summarize within the left/right hemisphere connectivity changes as well as between the left-right hemisphere changes. Among those 318 edges with significant ASD effects, 52 are within the right hemisphere, 91 are within the left hemisphere, and 175 are inter-hemispheric. Among those 37 edges with an ASD by age interaction, 15 are within the right hemisphere, 7 are within the left hemisphere, and 15 are inter-hemispheric. By using the six functional divisions introduced above [30], we have 10 primary sensory-motor (SM) ROIs, 44 unimodal association ROIs, 24 heteromodal association ROIs, 18 paralimbic ROIs, 4 limbic ROIs, and 12 subcortical ROIs. Due to the slightly different parcellation method, the number of ROIs in each functional division in the ABIDE study is different from the number in the Oxytocin and Aging study. Table 3 summarizes the ASD-related connectivity changes in functional connectivity at rest within and between these six functional divisions. As shown, the ASD patients show decreased cortico-cortical connections across all functional divisions, and especially for the unimodal and paralimbic divisions. This is consistent with [10]. The paralimbic area plays an important role in linking emotion with sensory inputs and in learning and memory. The abnormalities of the paralimbic area among the ASD patients in the ABIDE study are in line with other studies highlighting ASD-related social, cognitive and affective dysfunctions [33, 10].

6. Discussions

In this work, we extend existing methods for brain functional connectivity analysis (usually focusing on two-group comparison) to more flexible regression model scenarios. As demonstrated in the context of our two example applications, flexible models are needed to reduce the potential biases due to misspecified models. The flexible models we considered are able to account for potential interaction effects among the predictors and time-varying nonlinear effects of the predictors. We develop methods that can select the optimal model and test the effects of interest while controlling the FWER. We show in two data applications that selection of the optimal model improved the power of the tests. To utilize the more powerful NBS testing procedure for detecting subgraph level difference, we develop stability selection criteria for the choice of threshold, which is usually chosen arbitrarily. Our simulation results show that the thresholds chosen by the stability criteria outperform a fixed threshold such as the Bonferroni threshold. However, as the proposed method is the first attempt to determine the threshold in a data driven instead of an empirical way, further research is needed to optimally choose the threshold and to investigate the

theoretical properties of the resulting subgraphs. Using the parallel algorithm, it took about 33 minutes to carry out 1,000 permutations for the 8,646 edges in the Oxytocin and Aging study using an iMac computer with a processor of 3.5 GHz and a memory of 32GB.

Although motivated and generated in the context of resting-state functional connectivity analysis, the proposed methods can be easily generalized to other data such as gene network tests for different groups [34] or protein network tests under different conditions [35]. The proposed methods can also be applied to detect changes in structural connectivity derived by fibre tractography.

Acknowledgements

The authors thank Dr. Kai Yu at National Cancer Institute for discussion and suggestions on the edgewise regression and Harold Rocha for editorial assistance. This work was supported by the Center for Cognitive Aging and Memory at the University of Florida, the McKnight Brain Research Foundation, and the Claude D. Pepper Center at the University of Florida (P30 AG028740).

References

1. McArdle BH, Anderson MJ. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology*. 2001; 82(1):290–297.
2. Reiss PT, Stevens MHH, Shehzad Z, Petkova E, Milham MP. On distance-based permutation tests for between-group comparisons. *Biometrics*. 2010; 66(2):636–643. [PubMed: 19673867]
3. Kim J, Wozniak JR, Mueller BA, Shen X, Pan W. Comparison of statistical tests for group differences in brain functional networks. *NeuroImage*. 2014; 101:681–694. [PubMed: 25086298]
4. Zalesky A, Fornito A, Bullmore ET. Network-based statistic: identifying differences in brain networks. *Neuroimage*. 2010; 53(4):1197–1207. [PubMed: 20600983]
5. Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*. 2007; 8(9):700–711. [PubMed: 17704812]
6. Ebner NC, Chen H, Porges E, Lin T, Fischer H, Feifel D, Cohen RA. Oxytocin's effect on resting-state functional connectivity varies by age and sex. *Psychoneuroendocrinology*. 2016; 69:50–59. [PubMed: 27032063]
7. Whitfield-Gabrieli S, Nieto-Castanon A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity*. 2012; 2(3):125–141. [PubMed: 22642651]
8. Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT. Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*. 2000; 10(3):120–131. [PubMed: 10912591]
9. Zuo XN, Kelly C, Di Martino A, Mennes M, Margulies DS, Bangaru S, Grzadzinski R, Evans AC, Zang YF, Castellanos FX, et al. Growing together and growing apart: regional and sex differences in the lifespan developmental trajectories of functional homotopy. *The Journal of Neuroscience*. 2010; 30(45):15034–15043. [PubMed: 21068309]
10. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, Anderson JS, Assaf M, Bookheimer SY, Dapretto M, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*. 2014; 19(6):659–667. [PubMed: 23774715]
11. Insel TR. Mental disorders in childhood: shifting the focus from behavioral symptoms to neurodevelopmental trajectories. *Journal of the American Medical Association*. 2014; 311(17):1727–1728. [PubMed: 24794359]
12. Kennedy DN, Lange N, Makris N, Bates J, Meyer J, Caviness VS. Gyri of the human neocortex: an MRI-based analysis of volume and variance. *Cerebral Cortex*. 1998; 8(4):372–384. [PubMed: 9651132]

13. Chen H, Kelly C, Castellanos FX, He Y, Zuo XN, Reiss PT. Quantile rank maps: a new tool for understanding individual brain development. *NeuroImage*. 2015; 111:454–463. [PubMed: 25585020]
14. Hastie T, Tibshirani R. Varying-coefficient models. *Journal of the Royal Statistical Society. Series B*. 1993;757–796.
15. Reiss PT, Huang L, Chen YH, Huo L, Tarpey T, Mennes M. Massively parallel nonparametric regression, with an application to developmental brain mapping. *Journal of Computational and Graphical Statistics*. 2014; 23(1):232–248. [PubMed: 24683303]
16. Westfall, PH.; Young, SS. Resampling-based multiple testing: Examples and methods for p-value adjustment. Vol. 279. John Wiley & Sons; 1993.
17. Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*. 2002; 15(1):1–25. [PubMed: 11747097]
18. Lauritzen, SL. Graphical models. Oxford University Press; 1996.
19. Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal, Complex Systems*. 2006; 1695(5):1–9.
20. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning*. 2003; 52(1-2):91–118.
21. Fang Y, Wang J. Selection of the number of clusters via the bootstrap method. *Computational Statistics & Data Analysis*. 2012; 56(3):468–477.
22. Meinshausen N, Bühlmann P. Stability selection. *Journal of the Royal Statistical Society: Series B*. 2010; 72(4):417–473.
23. Meinshausen N, Bühlmann P. High-dimensional graphs and variable selection with the lasso. *The Annals of Statistics*. 2006:1436–1462.
24. Liu H, Roeder K, Wasserman L. Stability approach to regularization selection (stars) for high dimensional graphical models. *Advances in Neural Information Processing Systems*. 2010:1432–1440. [PubMed: 25152607]
25. Narayan M, Allen GI. Mixed effects models for resampled network statistics improves statistical power to find differences in multi-subject functional connectivity. *Frontiers in Neuroscience*. 2016 doi: 10.3389/fnins.2016.00108.
26. Gärtner T. A survey of kernels for structured data. *ACM SIGKDD Explorations Newsletter*. 2003; 5(1):49–58.
27. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity mri networks arise from subject motion. *Neuroimage*. 2012; 59(3):2142–2154. [PubMed: 22019881]
28. Friedman J, Hastie T, Tibshirani R. glmnet: Lasso and elastic-net regularized generalized linear models. R package version. 2009; 1
29. Xia M, Wang J, He Y, et al. BrainNet Viewer: a network visualization tool for human brain connectomics. *PIOS one*. 2013; 8(7):e68910. [PubMed: 23861951]
30. Mesulam MM. From sensation to cognition. *Brain*. 1998; 121(6):1013–1052. [PubMed: 9648540]
31. Middleton FA, Strick PL. Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition*. 2000; 42(2):183–200. [PubMed: 10744919]
32. Eilers PH, Marx BD. Flexible smoothing with B-splines and penalties. *Statistical Science*. 1996; 11(2):89–102.
33. Rudie JD, Shehzad Z, Hernandez LM, Colich NL, Bookheimer SY, Iacoboni M, Dapretto M. Reduced functional integration and segregation of distributed neural systems underlying social and emotional information processing in autism spectrum disorders. *Cerebral Cortex*. 2011:bhr171.
34. Zhang L, Baladandayuthapani V, Zhu H, Baggerly KA, Majewski T, Czerniak BA, Morris JS. Functional CAR models for spatially correlated functional datasets. *Journal of the American Statistical Association*. 2016 DOI:10.1080/01621459.2015.1042581.
35. Zhu Y, Shen X, Pan W. Structural pursuit over multiple undirected graphs. *Journal of the American Statistical Association*. 2014; 109(508):1683–1696. [PubMed: 25642006]

Figure 1. Simulation settings for the three sparse cases. From left to right, there are 5%, 10%, and 15% edges with nonzero coefficient, respectively.

Figure 2.
Oxytocin and Aging Study: cortical view maps for the largest subgraph with significant age effects using the NBS method.

aSTG=Superior Temporal Gyrus, anterior division; toITG=Inferior Temporal Gyrus, temporooccipital part; Forb= Frontal Orbital Cortex; TOFusC=Temporal Occipital Fusiform Cortex. 'l' or 'r' indicates the left or right hemisphere ROI.

Figure 3.

Oxytocin and Aging Study: the graphical plot of the largest subgraph for the treatment effect using the NBS method.

Figure 4.

ABIDE Study. Left panels: Black and red solid lines represent the mean trajectories for the typical controls and the ASD patients, respectively; black dotted lines represent the pointwise 95% confidence intervals for the mean trajectories of the typical controls. Right panels: Varying coefficient functions together with pointwise 95% confidence intervals.

Figure 5.
ABIDE Study: cortical view maps for the largest subgraph with significant ASD effects using the NBS method.

Figure 6.
 ABIDE Study: The largest subgraph with ASD by age interaction effect using the NBS method.

Table 1

Average number of false positive and false negative edges, based on 1000 simulation runs.

Sparsity	Method	# false positive	# false negative	Sum	# false positive	# false negative	Sum
		<i>n</i> =100			<i>n</i> =150		
5%	SMI	1.08	7.20	8.28	0.02	2.10	2.12
	SM2	0.16	5.33	5.49	0.03	1.84	1.87
	Bonferroni	0.02	10.92	10.94	0.02	2.48	2.50
	<i>c</i> =0.05	32.22	0.13	32.34	34.36	0.00	34.36
10%	SMI	3.18	14.85	18.03	0.13	7.05	7.18
	SM2	1.67	10.69	12.36	0.22	5.37	5.58
	Bonferroni	0.02	39.62	39.65	0.04	15.17	15.21
	<i>c</i> =0.05	29.39	1.17	30.56	34.81	0.12	34.94
15%	SMI	10.21	16.02	26.24	0.45	12.82	13.27
	SM2	9.58	14.69	24.27	0.90	9.64	10.54
	Bonferroni	0.02	78.84	78.86	0.03	41.72	41.75
	<i>c</i> =0.05	27.16	5.06	32.22	31.29	0.85	32.14

SMI: the stability method based on the coefficient of variation of graph size; SM2: the stability method based on the coefficient of variation of graph similarity; Bonferroni: the threshold is determined by Bonferroni method that is $c=0.05/780=6.41 \times 10^{-5}$.

Oxytocin and Aging Study: age-related resting-state functional connectivity changes within and between the six functional divisions.

Table 2

	Primary SM [*] (10)	Unimodal (44)	Heteromodal (23)	Paralimbic (14)	Limbic (4)	Subcortical (37)
Primary SM	3	17	2	9	1	4
Unimodal	17	17	5	19	4	9
Heteromodal	2	5	4	6	0	3
Paralimbic	9	19	6	3	7	13
Limbic	1	4	0	7	2	8
Subcortical	4	9	3	13	8	9

* Primary SM: Primary sensory-motor. Information in parentheses represents the number of ROIs in each of the six functional divisions. The information in the table represents the number of edges within a functional division (diagonal) or between two functional divisions (off diagonal).

Table 3

ABIDE Study: ASD-related resting-state functional connectivity changes within and between the six functional divisions.

	Primary SM * (10)	Unimodal (44)	Heteromodal (24)	Paralimbic (18)	Limbic (4)	Subcortical (12)
Primary SM	1	29	5	11	0	4
Unimodal	29	95	43	44	0	4
Heteromodal	5	43	24	42	0	0
Paralimbic	11	44	42	14	0	1
Limbic	0	0	0	0	0	1
Subcortical	4	4	0	1	1	0

* Primary SM: Primary sensory-motor. Information in parentheses represents the number of ROIs in each of the six functional divisions. The information in the table represents the number of edges within a functional division (diagonal) or between two functional divisions (off diagonal).