NeuroImage: Clinical 9 (2015) 625–639

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Highly adaptive tests for group differences in brain functional connectivity

Junghi Kim, Wei Pan*, for the Alzheimer's Disease Neuroimaging Initiative[1]

Division of Biostatistics, University of Minnesota, Minneapolis, MN 55455, USA

ARTICLE INFO

Article history:
Received 26 June 2015
Received in revised form 14 September 2015
Accepted 5 October 2015
Available online 22 October 2015

Keywords:
Covariance matrix
Graphical lasso
NBS
Precision matrix
rs-fMRI
Sparse estimation
SPU tests
Statistical power

ABSTRACT

Resting-state functional magnetic resonance imaging (rs-fMRI) and other technologies have been offering evidence and insights showing that altered brain functional networks are associated with neurological illnesses such as Alzheimer's disease. Exploring brain networks of clinical populations compared to those of controls would be a key inquiry to reveal underlying neurological processes related to such illnesses. For such a purpose, group-level inference is a necessary first step in order to establish whether there are any genuinely disrupted brain subnetworks. Such an analysis is also challenging due to the high dimensionality of the parameters in a network model and high noise levels in neuroimaging data. We are still in the early stage of method development as highlighted by Varoquaux and Craddock (2013) that "there is currently no unique solution, but a spectrum of related methods and analytical strategies" to learn and compare brain connectivity. In practice the important issue of how to choose several critical parameters in estimating a network, such as what association measure to use and what is the sparsity of the estimated network, has not been carefully addressed, largely because the answers are unknown yet. For example, even though the choice of tuning parameters in model estimation has been extensively discussed in the literature, as to be shown here, an optimal choice of a parameter for network estimation may not be optimal in the current context of hypothesis testing. Arbitrarily choosing or mis-specifying such parameters may lead to extremely low-powered tests. Here we develop highly adaptive tests to detect group differences in brain connectivity while accounting for unknown optimal choices of some tuning parameters.
The proposed tests combine statistical evidence against a null hypothesis from multiple sources across a range of plausible tuning parameter values reflecting uncertainty with the unknown truth. These highly adaptive tests are not only easy to use, but also high-powered robustly across various scenarios. The usage and advantages of these novel tests are demonstrated on an Alzheimer's disease dataset and simulated data.

1. Introduction

Previous studies have shown that neurological illnesses such as Alzheimer's disease and autism are related to altered brain functional networks, or functional connectivity, among distinct and distant brain regions (Greicius et al., 2004; Supekar et al., 2008). However, group-level statistical inference on functional connectivity is both challenging and necessary due to the high dimensionality of the parameters in a network model and the high noise level in neuroimaging data. Even though recent advances in neuroimaging technologies, such as rs-fMRI, offer great potentials for studying brain functional networks (Biswal, 2012), most statistical methods are for point estimation while only few exist in drawing inference for group comparisons in brain networks (Varoquaux and Craddock, 2013). In particular, many studies examined possible differences between two covariance or precision matrix estimates; however, whether there were any genuine differences between the corresponding population matrices was never or inadequately addressed. Without rigorous statistical testing, it is unknown whether any estimated network differences are simply due to estimation errors. In practice, we recommend to first test for differences of the networks between two groups; only if it is confirmed that the differences exist, then proceed to the next step to examine how they are different, possibly in an exploratory analysis.

Functional brain connectivity can be described as a network or graph (Bullmore and Sporns, 2009; Habeck and Moeller, 2011; He and Evans,

* Correspondence author at: Division of Biostatistics, MMC 303, School of Public Health, University of Minnesota, Minneapolis, MN 55455-0392, USA.
E-mail address: weip@biostat.umn.edu (W. Pan).

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

http://dx.doi.org/10.1016/j.nicl.2015.10.004

2010), in which a set of nodes are linked by edges. Nodes stand for brain regions, and brain connectivity (or edges) refer to pairwise associations between every two nodes (Varoquaux and Craddock, 2013). Yet, what measure of association should be used for the weights of network edges remains to be an open question. Currently the two most popular choices are Pearson's correlations and partial correlations. A correlation represents the marginal relationship between two brain regions, while a partial correlation quantifies the conditional association between the two, conditioning on other brain regions' activities. Pearson's (full or marginal) correlations have been adopted in many functional connectivity studies (Azari et al., 1992; Horwitz et al., 1987; Kim et al., 2014; Stam et al., 2007; Supekar et al., 2008). Partial correlations also have been advocated (Marrelec et al., 2006; Salvador et al., 2005; Smith et al., 2011), especially when one is interested in seeking conditional independence between any two regions. However, as shown by Kim et al. (2015), the choice of the association measure would influence statistical power in subsequently detecting group-level network differences. Moreover, since brain networks can be reasonably assumed to be sparse, they can be thus estimated via regularized sparse covariance or precision matrices, which however pose some challenges in the choice of a suitable regularization or tuning parameter. In addition, there may be other tuning parameters to be determined in an analysis. In general, the choice of such parameters may be difficult, and at the same time critical. We propose highly adaptive testing procedures that automatically search over and thus take account of these parameter values to maintain high power across various situations.

Kim et al. (2014) reviewed and compared many statistical methods, and concluded that the Network Based Statistic (NBS) (Zalesky et al., 2012) and an adaptive sum of powered score (aSPU) test (Pan et al., 2014) showed great performance, often complementary to each other, for testing group differences in brain functional connectivity. In particular, mass-univariate testing on each edge and testing on some network summary statistics is in general low-powered: for the former, in addition to possibly small differences on the edges, the stringent significance level after multiple testing adjustment is often too high to achieve; for the latter, it largely depends on the choice of the network summary statistics, which may be too mildly altered to be detected. NBS was originally developed to identify changed subnetworks in the neuroimaging research, but could be employed for global testing as considered here. By assuming that changed network edges form subnetworks, NBS can be potentially powerful when the assumption holds. However, since it is based on mass univariate testing to detect altered edges, it may miss some edges and thus connected subnetworks, leading to loss of power. As an alternative, the aSPU test does not impose any assumption on the structure of altered edges or networks while data-adaptively accumulating evidence against a null hypothesis, and thus yields high power under various situations. These are two representative and state-of-the-art tests we consider here. In particular, we propose two versions of highly adaptive aSPU and NBS tests respectively. The proposed tests utilize and data-adaptively choose multiple parameter values reflecting uncertainty with unknown true data structures, while adjusting for multiple testing automatically. At the end, we obtain two adaptive tests that can maintain high power more robustly across many situations; being data-adaptive to unknown true association patterns, the proposed tests are robust in the sense that they can achieve high power across many situations, in contrast to a possibly high powered test in only one or few situations that may lose much power in many other situations. We will apply the proposed adaptive tests to simulated data and an Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset.

The remainder of the paper is organized as follows. We first review sparse estimation of brain connectivity, then present highly adaptive versions of aSPU and NBS in Section 2. Section 3 is devoted to the analysis of ADNI data. In Section 4, we conduct simulation studies with realistic setups mimicking the ADNI data. We end with some discussions on a few technical aspects and comparisons with the literature in Section 5.

2. Methods

2.1. Data and notation

We focus on a case–control study design with covariates. Suppose there are n unrelated subjects, either healthy or having a disease.

We denote a group indicator $Y_i = 0$ for controls, $Y_i = 1$ for cases, and $Z_i = (Z_{i1}, \ldots, Z_{il})'$ for the covariates for subject i. We consider N brain regions of interests (ROIs), which define the nodes in a network or graph. For each node, fMRI BOLD signals are measured at $t = 1,2,\ldots,M$ time points. The BOLD signals from N nodes at time point t, $D_t = (D_{t1}, \cdots, D_{tN})'$, are assumed to be distributed as multivariate Gaussian $N(\mu, \Sigma)$ with the mean vector $\mu \in \mathbb{R}^N$ and the positive definite covariance matrix Σ.

The next step is to estimate the pairwise association between any two nodes. This measure is stored in a symmetric $N \times N$ adjacency matrix with its (i,j)th element as the association between the ith and jth nodes (Bullmore and Sporns, 2009). Hence a total of $k = N \times (N-1)/2$ unique pairwise associations are estimated for each subject, and these k continuous measures are called brain connectivity or network edges to be used for testing group differences. In our study, both Pearson's correlations and partial correlations were considered. As usual, the correlations (or partial correlations) are normalized via Fisher's z-transformation to generate subject i's brain connectivity, $X_i = (X_{i1}, \cdots, X_{ik})'$. Fisher's transformation is used to alleviate the effects of skewed distributions and outliers, which may negatively influence the power of a subsequent test.

In the following, we use matrix notation: $Y_{n \times 1}$ as a vector for disease indicators, $\mathbf{X}_{n \times k}$ as a matrix of pairwise associations between nodes (with each element as a z-transformed correlation or partial correlation), and $\mathbf{Z}_{n \times l}$ as a covariate matrix.

2.2. Estimating covariance and precision matrices via graphical lasso

Building on the work of Banerjee et al. (2008), Friedman et al. (2008) proposed a graphical lasso (glasso) algorithm to estimate sparse (or non-sparse) covariance and precision matrices, given M observations of dimension N, which are distributed as multivariate Gaussian $N(\mu, \Sigma)$. Let subject i have a possibly subject-specific true precision matrix $\Theta_i = \Sigma_i^{-1}$, and S_i be its empirical covariance matrix of the BOLD signals extracted from N nodes; the problem is to maximize the penalized log-likelihood,

$$L_{P,i}(\Theta_i; \lambda) = L_i(\Theta_i) - P(\Theta_i; \lambda) = \log \det(\Theta_i) - \mathrm{tr}(S_i\Theta_i) - \lambda_i \|\Theta_i\|_1$$

over the semi-positive definite Θ_i, where tr denotes the trace. $\|\Theta_i\|_1 = \sum_{p \neq q} |\theta_{i,pq}|$ is the L_1 norm for non-diagonal elements; $\lambda_i \geq 0$ is a regularization parameter to be determined. The graphical lasso finds the estimate, satisfying $\hat{\Theta}_i = \hat{\Theta}_{ii}(\lambda_i) = \arg\max_{\Theta_i} L_{P,i}(\Theta_i; \hat{\lambda}_i)$. Here $\hat{\Theta}_i$ is a function of λ_i, from which we also obtain a regularized estimate for the covariance matrix $\hat{\Sigma}_i = \hat{\Theta}_i^{-1}$. Full correlations and partial correlations can be estimated with $\hat{\Sigma}_i$ and $\hat{\Theta}_i$ respectively.

In this paper, the graphical lasso was employed to estimate brain connectivity at various sparsity levels. Denote c as the connection density (or proportion of nonzero elements) in a precision matrix estimate $\hat{\Theta}_i$. Using a grid search, λ_i can be chosen to generate the precision matrix estimate at a predefined connection density c, namely $\hat{\Theta}_i(c) = \hat{\Theta}_i(\lambda_i(c))$. Again $\hat{\Sigma}_i(c)$ is obtained by inverting $\hat{\Theta}_i(c)$; we note that, c is the connection density of the precision matrix estimate, not of the corresponding covariance matrix estimate.

For testing between-group differences in brain connectivity, recent studies have chosen λ_i at the group level so that each group should have the same or similar connection density in their estimated precision matrices (Huang et al., 2010; Stam et al., 2007; Supekar et al., 2008). Hence, at a low estimated connection density, only strong connectivity would show up as non-zeros in the resulting estimates, based on