OPEN ACCESS

GAVIN PUBLISHERS

Journal of Biostatistics & Biometrics

Mini Review

Bhaumik DK, et al. J Biostat Biom: JBSB-108.
DOI: 10.29011/JBSB-108. 100008.

Controlling False Disrupted Connectivities in Neuroimaging Studies

Dulal K Bhaumik[1], Yuanbo Song[2], Dan Zhao[2*], Olusola Ajilore[1]

[1]Department of Psychiatry, University of Illinois at Chicago, USA

[2]Department of Epidemiology and Biostatistics, University of Illinois at Chicago, USA

[*]**Corresponding author:** Dulal K. Bhaumik, Department of Epidemiology and Biostatistics, University of Illinois at Chicago, USA. Email: dbhaumik@uic.edu; Tel: +1 (312) 413-4455

Citation: Bhaumik DK, Song Y, Zhao D, Ajilore O (2018) Controlling False Disrupted Connectivities in Neuroimaging Studies. J Biostat Biom: JBSB-108. DOI: 10.29011/JBSB-108. 100008.

Received Date: 15 October, 2018; **Accepted Date:** 30 October, 2018; **Published Date:** 07 November, 2018

Abstract

Detection of disrupted connectivities is an important problem in neuroimaging research for targeting treatment interventions in order to get optimal therapeutic benefits. Disrupted connectivities are generally detected by comparing the disease group with a healthy control group utilizing the whole brain, resulting in thousands of comparisons. However, standard neuroimaging studies devoted to making such large scale multiple comparisons towards discovering statistically significant disrupted connectivities are seen to produce inflated false discoveries. Thus, development of statistical methods for multiple comparisons built upon a solid theoretical foundation in neuroimaging studies for controlling false discoveries is urgently needed. This article explores several approaches to control the false discovery rate, and recommends one best suited for detecting disrupted connectivities. Results are illustrated with a live data set generated from a late life depression study.

Keywords: Functional magnetic resonance imaging; False discovery rate; Local false discovery rate; Mixed-effects linear models; Neuroconnectivity

Introduction

Controlling the False Discovery Rate (FDR) is critically important for a detailed understanding of how a healthy brain differs with neurological diseases (e.g., depression), which is the fundamental requirement to the development and application of treatments for these conditions. Most of the neuroimaging research studies either do not address the issue of Multiple Testing (MT), or inadequately address MT by failing to incorporate the underlying dependency structure. Software packages often used for fMRI analysis (SPM, FSL, AFNI) can result in making false-positive rates of up to 70%, when it is fixed at 5% [1]. Inadequate use of statistical methods for MT is hindering the search for truly disrupted connectivities. Excessive false discoveries lead to wrong conclusion and may mar the importance of the study. Thus, a proper procedure to control the false discovery rate in neuro connectivity research is urgently needed.

Currently functional connectivity data are analyzed using model based methods such as cross correlation analysis or statistical parametric mapping [2-8]. Functional connectivity is also analyzed by data-driven methods such as independent component analysis, or principal component analysis [9-11]. Generally, structural connectivity data are obtained using diffusion tensor imaging, with measurements such as fractional anisotropy, relative anisotropy, mean diffusivity, or axial diffusivity [12-17]. For multimodal analysis, the existing Bayesian approach utilizes structural connectivity as a prior for functional connectivity [18-21], and frequentist statistical methods are also used with or without prior information [11,22-38]. In all neuroconnectivity studies based on either the whole brain, or any network (e.g. Salience Network, Default Mode Network), multiple testing becomes an integral part when two groups (e.g. Depression and Healthy Control) are compared. So, it is important to know the performance of different methods that are commonly used to control the false discovery rate in neuroconnectivity analysis.

Traditional approaches of large sample tests based on likelihood ratios and Wald-type statistics cannot address the issue of multiple comparison in neuroconnectivity analysis. Bonferroni correction for multiple testing becomes over conservative for a large number of comparisons. In rare cases, to control the false discovery rate, the approach by Benjamini and Hochberg (BH) [39] is used. However, the BH approach produces inflated false discovery rates in certain cases, and hence poses a problem in correctly identifying true discoveries with small samples.

The goal of this article is to develop statistical approaches to precisely detect disruptive neuroconnectivity in diseased populations and apply the findings to optimize benefits of clinical interventions. The first novelty of this article is in the development of a statistical model for neuroconnectivity data collected from neurological conditions (e.g., psychiatric, cognitive impairment etc.) and healthy control populations to identify significant changes in neuroconnectivity. For that purpose, we develop a mixed-effects model with some important characteristics in order to address within and between-subject heterogeneities, and also variabilities across links. The second novelty is in providing a simple interpretation to the q-value cut off which represents the minimum FDR at which the test can be called significant. Disrupted findings are next used to guide decisions on where to apply a neuro-modulatory intervention intended to improve behavioural disorders. The optimal approach of this article will indirectly reduce the cost of longitudinal and/or cluster neuroimaging studies, and provide the ability to study populations that are more difficult to recruit, such as minorities and traumatic brain injury patients. In summary, this article addresses modeling of neuroconnectivity data, and controlling of FDR in multiple testing for selecting the target of treatment applications.

We organize the article as follows. In Section 2, we motivate our problem with a study related to Late Life Depression (LLD) and argue that instead of controlling the type I error rate, we should control the false discovery rate for within or between group neuroconnectivity comparisons. In this section, we also discuss mixed-effects models and multiple testing procedures. In Section 3, we perform a simulation study for identifying an FDR approach that suits the best for neuroconnectivity comparisons and apply it to the LLD study for controlling the false discovery rate in section 4. We conclude the paper in Section 5.

Material and Methods

In this section, we discuss and develop some statistical methodologies intended to implement for detecting disrupted connectivities in group comparison studies. Towards that we start with a motivational example.

Motivational Example: Late Life Depression (LLD)

We motivate our research problem with a live data set from a study known as LLD recently conducted by Dr. Ajilore at University of Illinois at Chicago. LLD refers to major depressive episodes in elderly patients (usually over 50 or 60 years of age). This study recruited 23 subjects (13 healthy control (HC), and 10 LLD subjects), and collected data from 87 brain regions (i.e. a whole brain study, see Table 1) of each subject. The inclusion criteria for all subjects were over 55 years of age, medication-naive or anti-depressant free for at least two weeks (in the case of our LLD subjects) and no history of unstable cardiac or neurological diseases. The objective of this study is to detect brain regions and/or connectivity that are impaired during depression, and find out whether the identified regions can inform us about the pathophysiology of LLD Table 1.

Region	Description	Region	Description
1	cerebellum cortex	23	Middle temporal
2	thalamus proper	24	parahippocampal
3	caudate magnetic	25	paracentral
4	putamen	26	parsopercularis
5	pallidum	27	parsorbitalis
6	hippocampus	28	parstriangularis
7	amygdala	29	pericalcarine
8	accumbens area	30	postcentral
9	Ventral DC	31	osteriorcingulate
10	Bankssts	32	precentral
11	caudalanteriorcingulate	33	precuneus
12	caudalmiddlefrontal	34	rostralanterior cingulate
13	cuneus	35	rostralmiddlefrontal
14	entorhinal	36	superiorfrontal
15	fusiform	37	superiorparietal
16	inferiorparietal	38	superiortemporal
17	inferiortemporal	39	supramarginal
18	isthmuscingulate	40	frontalpole
19	lateraloccipital	41	temporalpole
20	lateralorbitofrontal	42	transversetemporal
21	lingual	43	insula
22	medialorbitofrontal	44	brain stem (central)

Table 1: Brain regions (left and right) in whole brain analysis.

Imaging data were collected using a Philips Achieva 3.0T scanner (Philips Medical Systems, Best, The Netherlands) with an 8-channel sensitivity encoded (SENSE) head coil. To reduce noise, headphones and foam pads were used and head movement was minimized. Subjects were instructed to stay still with their eyes closed yet awake without "thinking anything in particular" throughout the scanning process. Other sources of confounding effects such as motion art fact, white matter, and CSF were also regressed out before analysis. High resolution 3D T1-weighted images were attained by a MPRAGE (Magnetization Prepared Rapid Acquisition Gradient Echo) sequence with the following parameters: FOV = 240mm, TR/TE = 8.4/3.9 ms, flip angle = 8 degree, voxel size = 1.1 × 1.1 × 1.1 mm, and 134 contiguous axial slices. Resting-state imaging data were collected using a single-shot gradient-echo EPI sequence (EPI factor = 47, FOV = 23×23×15 cm^3, TR/TE = 2000/30ms, Flip angle = 80, in-plane resolution = 3×3 mm^2, slice thickness/gap = 5/0 mm, slice number = 30, SENSE reduction factor = 1.8, NEX = 200, total scan time = 6: 52. The resting-state functional neuronetworks were configured

using CONN [40], a Matlab toolbox. In brief, we preprocessed our raw EPI images involving four steps namely: realignment, co-registration, normalization, and smoothing.

The adjacency matrix of functional network as shown in (1), it has 87 × 86/2 = 3471 unique measures of functional connectivity. To satisfy the key assumption of normal distribution of our model, we applied the Fisher's Z transformation to these Pearson Correlations (PC).

$$\begin{bmatrix} - & & & & \\ PC_{1-2} & - & & & \\ PC_{1-3} & PC_{2-3} & - & & \\ \cdots & \cdots & \cdots & - & \\ PC_{1-87} & PC_{2-87} & \cdots & \cdots & - \end{bmatrix}_{87\times 87} \quad (1)$$

Modeling of Data

Mixed-effects regression models are frequently used to analyze neuroimaging data, as it can borrow strengths from multiple sources and provides reliable estimates of model parameters [4,18,41-48]. Fixed parameters of mixed-effects regression models provide population related information while random parameters provide subject specific information beyond the population mean [49-53]. Let Y_{ijg}/ be the fisher Z transformed pearson correlation measure of the ith link (i.e. connectivity between two regions) of the j^{th} subject nested within the g^{th} group (e.g. intervention, disease, control etc.).

$$Y_{ijg} = \gamma_{0i} + \gamma_{1i}x_{ijg} + \beta'^{*}w_{jg} + \nu_{jg} + \varepsilon_{ijg} \quad (2)$$

where i (= 1, 2..., m) denotes the ith link (connection between two brain regions), j (= 1, 2...,N_g) denotes the j subject nested within the g^{th} group, x_{ijg} = 0, if the j^{th} subject is from the control group (i.e. g = control), and x_{ijg}= 1, if the j^{th} subject is from the other group (i.e. g = intervention or disease). We assume the random noise ε_{ijg} follows a normal distribution N (0, σ^2_{gi}), in addition, random subject effect ν_{jg} follows a normal distribution N 0, σ^2_{g}, and these two normal distributions are independent. The vector w_{jg} represents measures of additional fixed covariates such as age, race, sex etc. Thus, our decision related to significance of link connectivity will be adjusted by such fixed covariates. In this model, γ_{1i} differentiates the intercept of the intervention group from that of the control group at the i^{th} link. Note that this model allows each link to have its own mean and variance parameters, and those vary from one group to the other group. To compare two groups at the i^{th} link, we will test null hypotheses H_0: γ_{1i} = 0, i = 1, 2,…m. Thus, we have a total number of m hypotheses, which brings the notion of multiple testing.

Parameters in the model are estimated using an EM algorithm. In the E step, with the current values of the other parameters, we compute the "expected posteriori" or Empirical Bayes (EB) estimates of random effects as well as the conditional variances of random effects given data. In the M step, given the current estimated values of random effects, we obtain the Maximum Marginal Likelihood (MML) estimates of the regression coefficients, error variances, and variances of random effects. The algorithm iterates between the EB and MML estimates until convergence.

q-value based FDR Procedure

To detect disrupted connectivities in a neuroimaging study, we generally perform between group comparisons. In whole brain study, the number of Regions of Interest (ROI) in our example is 87 with a total number of 3471 connectivities. In a single hypothesis testing problem, we generally control the type I error rate at a pre-specified level (e. g. α = .05), and determine a rejection region to maximize the probability of detecting the true alternative hypothesis. When multiple hypotheses are tested simultaneously, it is important to control the expected proportion of false rejections among all rejections, and the concept of False Discovery Rate (FDR) [39] prevails there. The ingredients that we need to formally define the FDR are presented in Table 2.

	H_0 Accepted	H_0 Rejected	Total
H_0 True	U	V	m_0
H_0 False	T	S	m-m_0
Total	m-R	R	m

Table 2: Different types of outcomes in multiple testing.

By definition, FDR is the expected proportion of Type I errors among the rejected hypotheses.

Hence, FDR = E(V/(V + S) = E(V/R), where V is the number of false rejections, R is the total number of rejections, and we define V/R=0, if R=0.

Algorithm to Control FDR

We start with the Benjamini and Hochberg [39] procedure to control the FDR. Consider a multiple testing problem consists of m hypotheses H_1, …, H_m with the corresponding p values p_1…, p_m. In the literature of FDR, [54, 55] defined a q-value as an analog of the p value that incorporates FDR-based multiple testing correction. q-value is the minimum positive FDR that can occur when rejecting a statistic with a threshold for the set of nested significance regions. The general procedure of the algorithm that controls the FDR at a threshold value of q is as follows:

1) Obtain p-values of all test statistics and sort those in an increasing order, i.e. $p(1) \leq p(2) \leq \ldots \leq p_{(m)}$.
2) For a given q, find the largest k for which $p_{(k)} \leq kq/m$ holds.
3) Declare discoveries for all $H_{(i)}$, where i = 1, 2…, k.

First the BH procedure needs a cut-off q-value, then starts from the largest p value (i.e. $p_{(m)}$) and moves towards the smallest one, consequently we reject all null hypotheses with p-values less than the aforementioned threshold.

Determination of FDR Level q

In a single hypothesis testing, generally we fix the individual type I error rate α at 5%, and try to achieve 80% power. Similarly, in multiple hypothesis testing, we need to determine a q-value. However, no such standard value of the FDR level q exists in the literature, mainly because of some other factors that play crucial roles in this process. Suppose in a multiple testing scenario with 100 (out of a total of 200) hypotheses are truly null (i.e., the proportion of nulls is $p_0 = 0.50$), we would expect 5 null rejections (out of 100) when α is individually fixed at 5%. With 80 rejections of the hypotheses that are truly non-null (i.e., 80% power), the observed the FDR level q = 5/(5+80) ≈ 0.059. A second scenario also conforms to multiple testing problem, with 90 of the 100 hypotheses being truly null (i.e., p_0 0.90), and 10 of these being truly false. In this scenario, we should expect observed FDR level q close to 4/(4+8) ≈ 0.333.

If this procedure continues for a large value of p_0 (e.g. p_0=.99), q will be close to 1 which will lead to less power and is undesirable. Thus, in a multiple testing problem, the FDR level q should be adjusted with the value of p_0, together with the individual type I error rate and power, such that q does not exceed to a desired threshold (e.g., 0.20 or 0.30). In order to better understand the interrelationship between individual type I error rate α, power, q value, and p_0, consider the concept of marginal FDR (mFDR) [56,57], where mFDR is the expected FDR level q. The mFDR can be approximated by the following expression in terms of p_0, α and power:

$$\text{mFDR} \approx \frac{\alpha p_0}{\alpha p_0 + power\ p_1} \qquad (3)$$

Figure 1 shows how mFDR changes with respect to p_0 for three distinct values of α = 0.05, 0.01, *and* 0 .001. mFDR and p_0 has a monotonicity property, i.e. for a fixed value of α, larger values of p_0 are expected to produce larger values of mFDR.

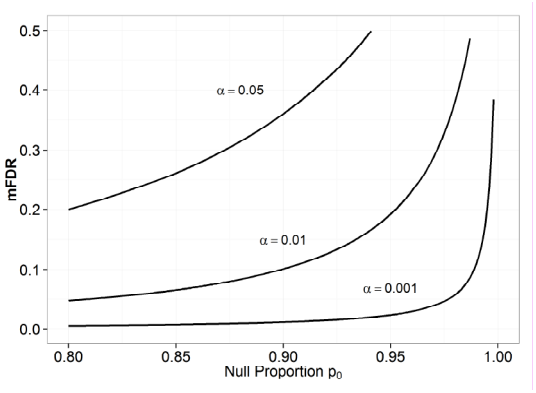

Figure 1: mFDR level vs. p_0 for various values of α.

The top most curve in Figure 2 reveals that in order to control the FDR level q at 0.2, for a very high proportion of null hypotheses (e.g p_0 = 0.98), α should be set at 0.01. On the other hand, for a relatively smaller proportion of null hypotheses (e.g p_0 = 0.8), α should be set at 0.05. Thus, in order to control the FDR level q at a desired value, when the proportion of null is large, α should be adjusted to a smaller value. Figure 2 also shows how the adjusted type I error rate changes at different levels of q-values when the null proportion increases (from 0.80 to 1.00). When the FDR level q is fixed at 0.20, decreases dramatically if 2 increases. When q-value equals to .05 (the bottom curve), the impact of 2 on the adjusted type I error rate is relatively less severe compared to q = 0.20.

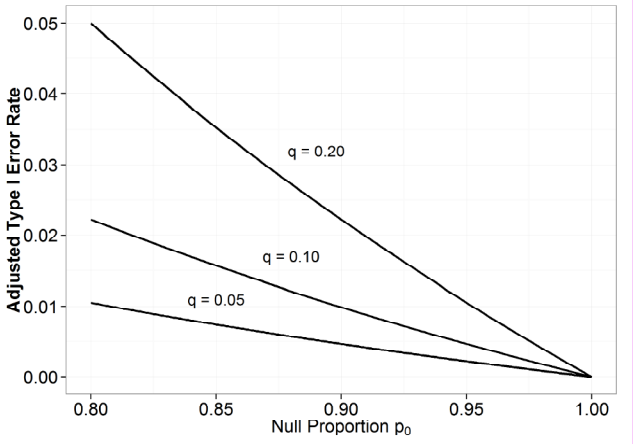

Figure 2: Adjusted type I error rate (α) v.s. p_0 for various values of q.

Finally, we conclude that efficient estimation of the null proportion p_0, and proper adjustment of the type I error rate are two important factors that should be considered carefully before we fix the FDR level q in multiple testing problems.

Adaptive BH (A-BH)

Benjamini and Hochberg [58] proposed an adaptive procedure incorporating the information of estimate of unknown null proportion. The adaptive BH procedure is as follows. In the same setting as Benjamini and Hochberg [39], determine k where,

$$k = \max\left\{i: P_{(i)} \leq \frac{iq}{\hat{p}_0 m}\right\} \qquad (4)$$

Reject all $H_{(i)}$, $i \leq k$, where $\hat{p}_0$ is an estimate of p2. The adaptive procedure is expected to gain more power compared to the original BH procedure. Note that this procedure also requires the q-value.

Non q-value based FDR Procedure

Both BH and adaptive BH are q-value based approaches, in what follows we discuss two more FDR procedures that are z-value based approaches, and are relevant to multiple comparisons in neuroimaging analysis.

In neuroimaging studies, it is common and reasonable to assume that a large majority of connectivities follows the null hypothesis. Generally, the central region of z-values roughly follows a standard normal density. In order to implement z-value based approaches, we not only need to compute the proportion of null hypotheses (p_0), we also need to define the null density (f_0), and the non-null density (f_1). Efron [59] proposed an approach based on z-value to control the false discovery rate. Null hypotheses are assumed to follow a normal distribution. Thus, the t-score (i.e. the test statistics value) of each hypothesis follows a normal distribution for large sample size. For small sample size, the t-score can be conveniently transformed to a z-score, theoretically, each z_i follows a standard normal distribution N (0, 1) under true null hypothesis. On the other hand, z-values with higher densities (compared to a standard normal density) in both tails, implies the presence of non-null cases.

Cai and Sun [57] developed an oracle and adaptive compound decision procedure for large-scale hypothesis testing. Their approach is to use data-driven adaptive procedure based on the z value to learn the distribution of the test statistics and use the information to construct a more efficient test. Jin and Cai [60] proposed an approach which used the empirical characteristic function and Fourier analysis for estimation of null distribution parameters, they demonstrated that the information about the null distribution is well preserved in the high frequency Fourier coefficients, thus provided good estimates of the null.

Connection between q-value based FDR and Non q-value based FDR

BH approach defines the expected probability of null cases conditioned on any observed p values greater than a cut off value. Local False Discovery Rate (LFDR) proposed by Efron [58] is a distinct alternative measure of false positive error rate based on z values. LFDR is defined as

$$\mathrm{LFDR}(z) = \frac{p_0 f_0}{p_0 f_0 + p_1 f_1} = \frac{f_0^+(z)}{f(z)} \qquad (5)$$

However, these two concepts (q-value vs z-value) are closely connected. Revisiting to the local false discovery rate model, and replacing the density functions f_0 and f_1 with the cumulative distribution functions F_0 and F_1, we now define

$$F(z) = p_0 F_0(z) + p_1 F_1(z) \qquad (6)$$

Then a simple application of Bayes theorem should get the posterior probability of being a null case for an observed z score less than or equal to some small z value.

$$\mathrm{Fdr}(z) = \frac{p_0 F_0}{p_0 F_0 + p_1 F_1} = \frac{F_0^+(z)}{F(z)} \qquad (7)$$

where $F_0^+ z = p_0 F_0(z)$ denotes the null sub distribution (this only considers a left-sided case for notational convenience, but it can be readily extended to a right-sided or two-sided case). Fdr and fdr, according to the definition, are mathematically equated by

$$\mathrm{Fdr}(z) = \frac{\int_{-\infty}^{z} \mathrm{fdr}(Z) f(Z) dZ}{\int_{-\infty}^{z} f(Z) dZ} = E_f\{\mathrm{fdr}(Z) | Z \leq z\} \qquad (8)$$

That is, Fdr(z) is conditional expectation of fdr(z) and can be computed by averaging fdr(z) for all $Z \leq z$. Considering that, in regular cases, fdr(z) decreases as |z| increases, Fdr(z) is always expected to be less than fdr(z). Fdr(z) is the Bayesian form of the false discovery rate and it corresponds to Storey's [53,54] "q-value", the value of the tail area false discovery rate attained at a given observed value Z = z.

Results

We analyze the fMRI data of the LLD study using a mixed-effect model discussed in Section 2, and obtain estimates of mean fMRI measures of all 3741 links for each group. A total of 7,506 model parameters (3741 fixed effects, 23 random subject effects, 3741 error covariance components, and 1 random effect covariance component) are iteratively estimated until the convergence criteria are met. The convergence criteria are set at the tolerance level of 10%w. The model converged after 5 iterations.

Table 3 shows summary statistics of parameter estimates. Gamma distributions are fitted to error variances for each study group. We performed a family of t-tests for testing link-specific regression coefficients (i.e. H_0: $\gamma_{1i} = 0$, i = 1, 2, … m) utilizing estimates of γ_{1i}- and corresponding standard errors obtained from the mixed-effect model.

	μ	σ	p_0
Theorical Null	0	1	0.901
Efron	0.038	1.121	0.988
Jin and Cai	-0.025	1.116	0.985

Table 3: Estimates of parameters by different procedures.

The aforementioned mixed-effects model produced 281 (out of 3741 tests) significant results for testing H2: γ$- = 0, whereas independent two sample t-tests produced 173 significants, indicating that the mixed-effects model has more detection power compared to two sample independent t-tests. This is because the proposed mixed-effect model with random subject effects divides the total variability into between-subject and within-subject. In both cases, significance was determined individually based on p-value < 0.05 leading to substantial inflation of type I error rate.

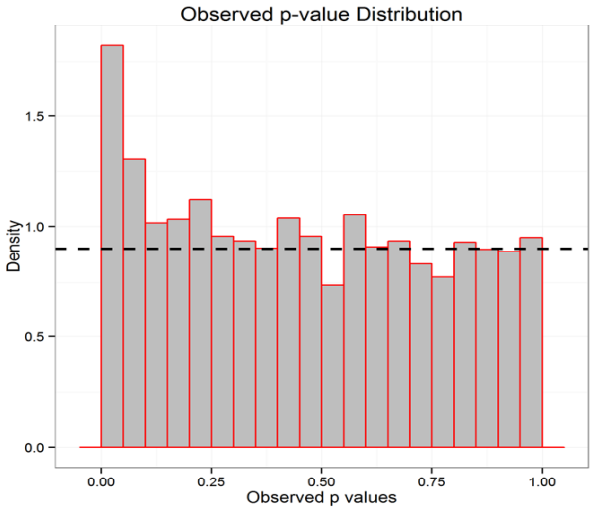

Figure 3: Distribution of p-values.

Next, in order to adjust the type, I error rate, we identify significant disrupted connectivities using the q-value based approach. Figure 3 shows the distribution of observed p-values. Under the null hypothesis, p-values should be uniformly distributed over [0, 1]. As we can see in the histogram, p-values in the range of [0.1, 1] are quite evenly distributed, and those in the range of [0, 0.1] apparently are taller than the rest. This indicates the existence of non-null hypotheses in the experiment. Using the q-value package [55], we estimated the null proportion *p*2 as 0.923, shown by a dashed line. Next we use a z-value based procedure with three different null distributions: theoretical null N (0, 1), empirical null estimated using [58], and empirical null estimated using [59]. Table 4 summarizes estimated parameters by these threes methods.

		Under Connectivity	**Over Connectivity Connectivity**	**Total**
q-value	Adaptive BH	22	37	59
	Theorical Null	18	51	69
z-value	Efron	12	11	12
	Jin and Cai	1	13	14

Table 4: Number of significant connectivities by various fdr procedures.

Table 4. reveals that estimated mean parameter using the theoretical null, Efron, and Jin and Cai are 0.00, 0.038 and -0.025 respectively, which indicates that the assumption of zero mean is plausible. Estimates of standard deviations and null proportions by Efron, and Jin and Cai are comparable, however, those are considerably larger compared to the theoretical null. The estimated null proportion under the theoretical null is indeed close to the p-value based estimate (0.901 vs. 0.923). A Histogram of z values from LLD study is presented in Figure 4.

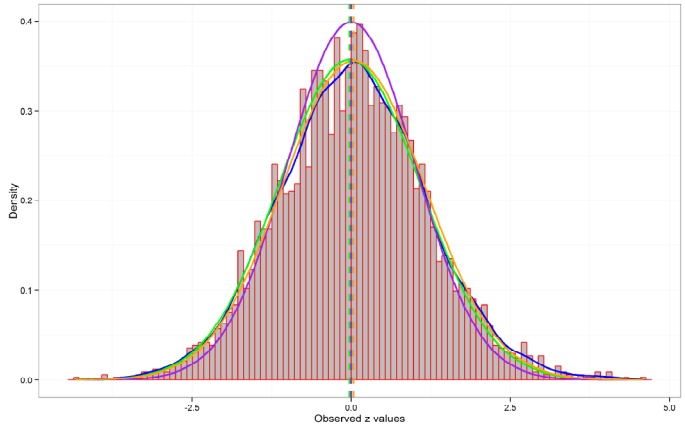

Figure 4: Fitted Empirical Null Densities.

The blue curve is the mixture density, the purple curve is the theoretical null density, the orange curve is the empirical null density by [58]. and the green curve is the empirical null density by [59]. As we see in Figure 4, the theoretical distribution is narrower than the mixture distribution. This may be largely due to the correlation of z values. The mixture density is taller than the

theoretical density at both extreme ends, indicating that there may be both under connectivity (at the left end) and over-connectivity (at the right end) under the theoretical null. We further notice in Figure 4 that the mixture density is taller than the empirical null density only at the right extreme and matches with the central peak and left extreme end very well. This is a sign of over connectivity only. Controlling the tail-area FDR at 0.2, the three methods identified 69 (theoretical), 12 (Efron) and 14 (Sun and Cai) significant connectivities respectively.

Table 5 presents significant connectivities obtained from four different procedures. As we can see in Table 5. underlying assumptions of the null distribution naturally classify the results into two categories: adaptive BH vs theoretical null (i.e. standard normal distribution), empirical null of Efron (a normal distribution with nonzero mean and variance may be different than one) and that of Sun and Cai. For Adaptive BH and Theoretical Null, the significant findings are quite similar yet they vary slightly. However, significant findings detected by theoretical null and empirical null do not differ that much. This reveals that the choice of null distribution can substantially affect the multiple testing results, especially when test statistics are correlated and sample size is small. Regions of interest involving significant overlapping of disrupted connectivities obtained from all three approaches mentioned in Table 5.

	Region 1		Region 2
28	Left isthmuscingulate	56	Right caudalmiddlefrontal
12	Right thalamus_Proper	56	Right caudalmiddlefrontal
19	Right ventralDC	59	Right fusiform
45	Left rostralmiddlefrontal	56	Right caudalmiddlefrontal
15	Right pallidum	56	Right caudalmiddlefrontal
15	Right pallidum	26	Left inferiorparietal
18	Right accumbens_area	28	Left isthmuscingulate
41	Left posteriorcingulate	79	Right rostralmiddlefrontal
55	Right caudalanteriorcingulate	56	Right caudalmiddlefrontal
46	Left superiorfrontal	62	Right isthmuscingulate
56	Right caudalmiddlefrontal	62	Right isthmuscingulate
25	Left fusiform	49	Left supramarginal

Table 5: Significant links.

Right caudal middle frontal denoted by 56 in the network in Figure 5 has been detected as the main hub of disruptions. This hub has a total of six disrupted connectivities. Three other secondary hubs are Right pallidum (#15), Left isthmus cingulate (#28), and Right isthmus cingulate (#62). The identification of the right caudal middle frontal cortex (also referred to as the Dorsolateral Prefrontal Cortex or DLPFC) is consistent with extensive literature on the neuroanatomy of major depression [61]. The DLPFC has been implicated in a number of neuroimaging studies in the depression literature. This region is used as a target for Transcranial Magnetic Stimulation (rTMS) in the treatment of major depression and plays an important role in emotion regulation [62]. It also appears to be the key target of antidepressant action [63].

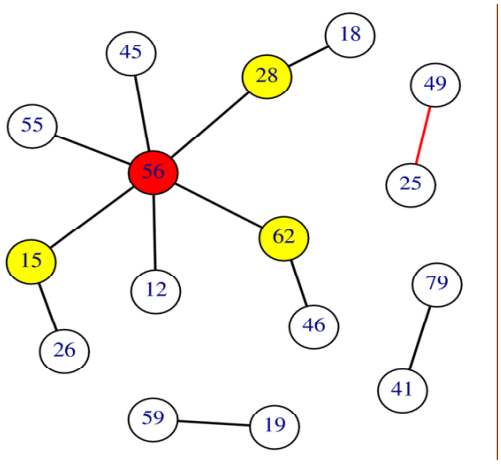

Figure 5: Network Analysis of Disrupted Connectivities.

Discussion

This paper argues that controlling the false discovery rate is essential in neuroconnectivity studies. The impact of small samples is quite evident in terms of controlling the FDR at the desired level. We identify Efron FDR procedures are more suitable for neuroconnectivity studies and provide guidelines of how to implement those in our context. Moreover, Efron's local FDR has a Bayesian interpretation that is useful and easy to understand. Also, implementation of LFDR by Efron is less computationally intensive compared to Sun and Cai as the former is based on normal approximations and the latter uses Fourier transformations. Considering all aspects, we recommend LFDR of [58]. Nonetheless, both BH and adaptive BH are extensively used in practices because of their historical importance.

A hub of disrupted connectivities detected by our approach is a useful target for applications of interventions or stimulus. Another application of detection of disrupted connectivities is in classification of subjects at high risk for early prevention. Investigators may feel comfortable now to use the FDR controlled

hub of disrupted connectivities for targeting treatment interventions in order to get optimal therapeutic benefits.

Neuroimaging studies need more research for correlated hypotheses. High positive dependency among tests, which is predominant in most of the applications of multiple testing in modern scientific investigations, like the one discussed before, has been a major impediment to taking full advantage of the BH and related FDR controlling procedures since those, even though can control the FDR at a desired level [64-66], can be very conservative under such dependency [67]. Authors of this paper are currently investigating an FDR procedure utilizing a general correlation structure for neuroconnectivity studies.

References

1. Eklund A, Nichols TE, Knutson H (2016) Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proceedings of the National Academy of Sciences 113: 7900.
2. Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, et al. (1992) Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging Proceedings of the National Academy of Sciences 89: 5951-5955.
3. Cao J, Worsley K (1999) The geometry of correlation fields with an application to functional connectivity of the brain. The Annals of Applied Probability 9: 1021-1057.
4. Blamire AM, Ogawa S, Ugurbil K, Rothman D, McCarthy G, et al. (1992) Dynamic mapping of the human visual cortex by high-speed magnetic resonance imaging. Proceedings of the National Academy of Sciences 89: 11069-11073.
5. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, et al. (1994) Statistical parametric maps in functional imaging: a general linear approach. Human brain mapping 2: 189-210.
6. Gibbons RD, Lazar NA, Bhaumik DK, Sclove SL, Chen HY, et al. Estimation and classification of fMRI hemodynamic response patterns. NeuroImage 22: 804-814.
7. Kapur K, Roy A, Bhaumik DK, Gibbons RD, Lazar NA, et al. (2009) Estimation and classification of BOLD responses over multiple trials. Communications in Statistics-Theory and Methods 38: 3099-3113.
8. Pape TLB, Rosenow JM, Steiner M, Parrish T, Guernon A, et al. (2015) Placebo-controlled trial of familiar auditory sensory training for acute severe traumatic brain injury: a preliminary report. Neurorehabilitation and neural repair 29: 537-547.
9. Chuang KH, Chiu MJ, Lin CC, Chen JH (1999) Model-free functional MRI analysis using Kohonen clustering neural network and fuzzy C-means. IEEE transactions on medical imaging 18: 1117-1128.
10. Ngan SC, Hu X (1999) Analysis of functional magnetic resonance imaging data using self-organizing mapping with spatial connectivity. Magnetic Resonance in Medicine 41: 939-946.
11. Li K, Guo L, Nie J, Li G, Liu T (2009) Review of methods for functional brain connectivity detection using fMRI. Computerized Medical Imaging and Graphics 33: 131-139.
12. Friston KJ, Frith CD, Frackowiak RSJ (1993) Time-dependent changes in effective connectivity measured with PET. Human Brain Mapping 1: 69-79.
13. Friston KJ, Frith CD, Liddle P, Frackowiak R (1993) Functional connectivity: the principal-component analysis of large (PET) data sets. Journal of Cerebral Blood Flow & Metabolism 13: 5-4.
14. Conturo TE, McKinstry RC, Akbudak E, Robinson BH (1996) Encoding of anisotropic diffusion with tetrahedral gradients: a general mathematical diffusion formalism and experimental results. Magnetic Resonance in Medicine 35: 399-412.
15. Mori S, Zijl PC (2002) Fiber tracking: principles and strategies-a technical review. NMR in Biomedicine 15: 468-480.
16. Lauzon CB, Crainiceanu C, Caffo BC, Landman BA (2013) Assessment of bias in experimentally measured diffusion tensor imaging parameters using SIMEX. Magnetic resonance in medicine 69: 891-902.
17. Rykhlevskaia E, Gratton G, Fabiani M (2008) Combining structural and functional neuroimaging data for studying brain connectivity: a review. Psychophysiology 45: 173-187.
18. Bhaumik DK, Roy A, Lazar NA, Kapur K, Aryal S, et al. (2009) Hypothesis testing, power and sample size determination for between group comparisons in fMRI experiments. Statistical methodology 6: 133-146.
19. Patel RS, Bowman FD, Rilling JK (2006) A Bayesian approach to determining connectivity of the human brain. Human brain mapping 27: 267-276.
20. Xue W, Bowman FD, Pileggi AV, Mayer AR (2015) A multimodal approach for determining brain networks by jointly modeling functional and structural connectivity. Frontiers in computational neuroscience 9: 22.
21. Zhang L, Guindani M, Vannucci M (2015) Bayesian models for functional magnetic resonance imaging data analysis. Wiley Interdisciplinary Reviews: Computational Statistics 7: 21-41.
22. Calhoun VD, Adali T (2009) Feature-based fusion of medical imaging data. IEEE Transactions on Information Technology in Biomedicine 13: 711-720.
23. Calhoun V, Adali T, Liu J (2006) A feature-based approach to combine functional MRI, structural MRI and EEG brain imaging data. Conf Proc IEEE Eng Med Biol Soc 1: 3672-3675.
24. Calhoun VD, Wu L, Kiehl KA, Eichele T, Pearlson GD (2010) Aberrant processing of deviant stimuli in schizophrenia revealed by fusion of fMRI and EEG data. Acta neuropsychiatrica 22: 127-138.
25. Eichele T, Calhoun VD, Debener S (2009) Mining EEG-fMRI using independent component analysis. International Journal of Psychophysiology 73: 53-61.
26. Franco R, Ling J, Caprihan A, Calhoun VD, Jung RE, et al. (2008) Multimodal and multi-tissue measures of connectivity revealed by joint independent component analysis. IEEE journal of selected topics in signal processing 2: 986-997.
27. Shenton ME, Dickey CC, Frumin M, McCarley RW (2001) A review of MRI findings in schizophrenia. Schizophrenia research 49: 1-52.
28. Teipel SJ, Bokde AL, Meindl T, Amaro E, Soldner J, et al. (2010) White matter microstructure underlying default mode network connectivity in

the human brain. Neuroimage 49: 2021-2032.

29. Lin FH, Mc-Intosh AR, Agnew JA, Eden GF, Zeffiro TA, et al. (2003) Multivariate analysis of neuronal interactions in the generalized partial least squares framework: simulations and empirical studies. NeuroImage 20: 625-642.

30. Montes ME, Sosa PA, Miwakeichi F, Goldman RI, Cohen MS (2004) Concurrent EEG/fMRI analysis by multiway partial least squares. NeuroImage 22: 1023-1034.

31. Mc-Intosh R, Lobaugh NJ (2004) Partial least squares analysis of neuroimaging data: applications and advances. Neuroimage 23: 250-263.

32. Chen K, Reiman EM, Huan Z, Caselli RJ, Bandy D (2009) Linking functional and structural brain images with multivariate network analyses: a novel application of the partial least square method. Neuroimage 47: 602-610.

33. Correa NM, Li YO, Adali T, Calhoun VD (2008) Canonical correlation analysis for feature-based fusion of biomedical imaging modalities and its application to detection of associative networks in schizophrenia. IEEE journal of selected topics in signal processing 2: 998-1007.

34. Correa NM, Adali T, Li YO, Calhoun VD (2010) Canonical correlation analysis for data fusion and group inferences. IEEE signal processing magazine 27: 39-50.

35. Sui J, Adali T, Li YO, Yang H, Calhoun VD (2012) A review of multivariate methods in brain imaging data fusion 15: 68-81.

36. Sui J, Adali T, Yu Q, Chen J, Calhoun VD (2009) A review of multivariate methods for multimodal fusion of brain imaging data. Journal of neuroscience methods 204: 68-81.

37. Xu L, Pearlson G, Calhoun VD (2009) Joint source based morphometry identifies linked gray and white matter group differences. Neuroimage 44: 777-789.

38. Correa NM, Eichele T, Adalı T, Li YO, Calhoun VD (2010) Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI. Neuroimage 50: 1438-1445.

39. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society Series B (Methodological) 289-300.

40. Gabrieli SW, Castanon AN (2012) Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain connectivity 2: 125-141.

41. Li Y, Wang Y, Wu J, Gao S, Bukhari L, et al. (2005) Antidepressant effect on connectivity of the mood-regulating circuit: an FMRI study. Neuro psychopharmacology 30: 1334, 2005.

42. Fletcher PT, Whitaker RT, Tao R, DuBray MB, Froehlich A, et al. (2010) Microstructural connectivity of the arcuate fasciculus in adolescents with high-functioning autism. Neuroimage 51: 1117-1125.

43. Mahani NK, Zoethout RM, Beckmann CF, Baerends E, Kam ML, et al. (2012) Effects of morphine and alcohol on functional brain connectivity during resting state: A placebo-controlled crossover study in healthy young men. Human brain mapping 33: 1003-1018.

44. Satterthwaite TD, Wolf DH, Loughead J, Ruparel K, Elliott MA, et al. (2012) Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. Neuroimage 60: 623-632.

45. Chen G, Saad ZS, Britton JC, Pine DS, Cox RW (2013) Linear mixed-effects modeling approach to FMRI group analysis. Neuroimage 73: 176-190.

46. Jones BC, Nair G, Shea CD, Crainiceanu CM, Cortese IC, et al. (2013) Quantification of multiple-sclerosis-related brain atrophy in two heterogeneous MRI datasets using mixed-effects modeling. NeuroImage: Clinical 3: 171-179.

47. Simpson SL, Laurienti PJ (2013) A two-part mixed-effects modeling framework for analyzing whole-brain network data. NeuroImage 113: 310-319.

48. Bhaumik D, Jie F, Nordgren R, Bhaumik R, Sinha BK (2018) A mixed-effects model for detecting disrupted connectivities in heterogeneous data. IEEE Transactions on Medical Imaging.

49. Laird NM, Slasor P (1997) Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. Statistics in medicine 16: 2349-2380.

50. Hedeker D, Gibbons RD (2006) Longitudinal data analysis. John Wiley & Sons.

51. Fitzmaurice GM, Laird NM, Ware JH (2012) Applied longitudinal analysis. John Wiley & Sons.

52. Beunckens C, Molenberghs G, Verbeke G, Mallinckrodt C (2006) Analyzing incomplete discrete longitudinal clinical trial data. Statistical Science 1: 52-69.

53. Seltman HJ (2012) Experimental design and analysis.

54. Storey JD (2002) A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64: 479-498.

55. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences 100: 9440-9445.

56. Genovese CR, Wasserman L (2002) Operating characteristics and extensions of the false discovery rate procedure. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64: 499-517.

57. Sun W, Cai TT (2007) Oracle and adaptive compound decision rules for false discovery rate control. Journal of the American Statistical Association 102: 901-912.

58. Benjamini Y, Hochberg Y (2000) On the adaptive control of the false discovery rate in multiple testing with independent statistics. Journal of educational and Behavioral Statistics 25: 60-83.

59. Tibshirani R, Storey JD, Tusher V (2001) Empirical Bayes analysis of a microarray experiment. Journal of the American statistical association 96: 1151-1160.

60. Jin J, Cai TT (2007) Estimating the null and the proportion of nonnull effects in large-scale multiple comparisons. Journal of the American Statistical Association 102: 495-506.

61. Hamilton JP, Etkin A, Furman DJ, Lemus MG, Johnson RF (2012) Functional neuroimaging of major depressive disorder: a meta-analysis and new integration of baseline activation and neural response data. American Journal of Psychiatry 169: 693-703.

62. Rive M, Rooijen GV, Veltman DJ, Phillips ML, Schene AH, et al. (2013) Neural correlates of dysfunctional emotion regulation in major depressive disorder. A systematic review of neuroimaging studies. Neuroscience & Biobehavioral Reviews 37: 2529-2553.

63. Ma Y (2015) Neuropsychological mechanism underlying antidepressant effect: a systematic meta-analysis. Molecular psychiatry 20: 311.

64. Benjamini Y, Yekutieli D (2011) The control of the false discovery rate in multiple testing under dependency. Annals of statistics 1165-1188.

65. Sarkar SK (2008) Some results on false discovery rate in stepwise multiple testing procedures. Annals of statistics 1: 239-257.

66. Sarkar SK (2008) On methods controlling the false discovery rate. Sankhya: The Indian Journal of Statistics, Series A 70: 135-168.

67. Kim KI, Wiel MA (2008) Effects of dependence in high-dimensional multiple testing problems. BMC bioinformatics 9: 114.