



OPEN ACCESS

EDITED BY

Borbála Hunyadi,
Delft University of Technology, Netherlands

REVIEWED BY

Christos Chatzichristos,
KU Leuven, Belgium
Rajesh Nandy,
University of North Texas Health Science
Center, United States

*CORRESPONDENCE

Pei-Shan Yen
✉ pyen2@uic.edu

RECEIVED 18 January 2023

ACCEPTED 20 April 2023

PUBLISHED 24 May 2023

CITATION

Bhaumik DK, Wang Y, Yen P-S and Ajilore OA (2023) Development of a Bayesian multimodal model to detect biomarkers in neuroimaging studies. *Front. Neuroimaging* 2:1147508. doi: 10.3389/fnimg.2023.1147508

COPYRIGHT

© 2023 Bhaumik, Wang, Yen and Ajilore. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Development of a Bayesian multimodal model to detect biomarkers in neuroimaging studies

Dulal K. Bhaumik^{1,2}, Yue Wang¹, Pei-Shan Yen^{1*} and Olusola A. Ajilore²

¹Division of Epidemiology and Biostatistics, University of Illinois at Chicago, Chicago, IL, United States,

²Department of Psychiatry, University of Illinois at Chicago, Chicago, IL, United States

In this article, we developed a Bayesian multimodal model to detect biomarkers (or neuromarkers) using resting-state functional and structural data while comparing a late-life depression group with a healthy control group. Biomarker detection helps determine a target for treatment intervention to get the optimal therapeutic benefit for treatment-resistant patients. The borrowing strength of the structural connectivity has been quantified for functional activity while detecting the biomarker. In the biomarker searching process, thousands of hypotheses are generated and tested simultaneously using our novel method to control the false discovery rate for small samples. Several existing statistical approaches, frequently used in analyzing neuroimaging data have been investigated and compared via simulation with the proposed approach to show its excellent performance. Results are illustrated with a live data set generated in a late-life depression study. The role of detected biomarkers in terms of cognitive function has been explored.

KEYWORDS

multiple testing, local false discovery rate, multimodal, functional connectivity, mixed-effects model

1. Introduction

Late-life depression (LLD) is a common disorder associated with emotional distress, cognitive impairment, and somatic complaints in older people. Patients with LLD are usually over 55 years old and have major depressive symptoms. As the world population of adults aged 60 and older is expanding rapidly from 900 million in 2015 to 2 billion by 2050 (WHO, 2021), understanding age-related disorders is becoming prominent and necessary to accommodate this demographic shift. The general strategy of treating patients having major depressive disorder (MDD) or anxiety disorders with serotonin reuptake inhibitors has a success rate of 50–70% only, and there is no data to support therapeutic selection for these kinds of diseases (Dunlop et al., 2012). Therefore, the field of psychiatric research needs the discovery of biomarkers to guide treatment selection.

The objective of the article is to provide a guidance for treating LLD patients who require alternative treatments. This investigation aims to identify reliable disrupted brain regions (or biomarkers) in LLD patients that can be used to develop neuropsychotherapy, a safe and enriched environment utilizing a neurological approach to facilitate long-term change in neural function. Our approach detects reliable biomarkers, minimizing the risk of false detection, which is essential to get the optimal result from interventions. An immediate benefit of our approach is the development of interventions, such as cognitive

behavioral therapy, that explore the association of the cognitive role of detected biomarkers with neurobehavioral measures. In addition, detected biomarkers can accurately predict prospective subjects at high risk for early prevention.

Brain connectivity usually refers to a pattern of anatomical links (structural connectivity) and statistical dependencies (functional connectivity). Structural connectivity (SC) is the structural link or neural pathways between two areas. This typically corresponds to white matter tracts that run between pairs of brain regions. Functional connectivity (FC) is defined as the “temporal correlation between spatially remote neurophysiological events (Friston et al., 1993).” In other words, FC indicates the synchronized and correlated patterns of activity. An essential problem in neural connectivity research is detecting disrupted connectivity for targeting treatment interventions to achieve optimal therapeutic benefit. The literature suggests that functional and structural abnormality are neurobiologically correlated in LLD (Tadayonnejad et al., 2014). It is a long-standing interest of researchers to know how FC and SC are related, whether FC is mediated by SC or another way. It is also equally important to investigate whether disruptions in neuro-connectivity can be detected with better accuracy by incorporating the information from functional and structural neuroimaging data. This article develops a Bayesian mixture model utilizing resting-state functional magnetic resonance imaging (rs-fMRI) and diffusion tensor imaging (DTI) data to detect disrupted brain region of interest (ROI) and use it as a biomarker while comparing two distinct groups (e.g., LLD and healthy control) of subjects.

Two significant contributions, namely, building a Bayesian mixture model for FC utilizing SC as auxiliary information and developing a strategy for multiple comparisons to control the false discovery rate (FDR), make this article novel for detecting biomarkers. Our Bayesian multimodal approach based on a Bayesian mixture model takes the advantage of the complementary SC data to enhance the modeling of the density of FC statistics while controlling the local false discovery rate (Lfd; Efron, 2007). Our approach is general and helpful in a broad spectrum of applications with multiple modalities. The utility of our approach is illustrated with extensive simulations to show how it controls the FDR even with small samples under different correlation structures compared to some standard methods that solely consider FC. In a way, our approach attempts to answer whether SC is at all needed to detect biomarkers for functional activity, especially when the study sample size is small.

Biswal et al. (1995) first observed a high coherent synchronous correlation between the blood-oxygen-level dependence (BOLD) signals (i.e., FC) from the regions of the somatic motor system in the left and right hemispheres of the brain in healthy volunteers during resting conditions and demonstrated that BOLD signals reflect neural activities for both resting-state and task-state. Since then, rs-fMRI has been applied extensively in cognitive neuroscience research to study patients with neurological, mental, or psychiatric disorders. Studies show that using rs-fMRI can help identify disruptions in FC in various forms of neurological disorders such as Alzheimer’s disease (Dai et al., 2015), psychiatric disorders including depression (Alexopoulos, 2019), and attention deficit hyperactivity disorder (Uddin et al., 2009). This is the

rationale behind considering rs-fMRI as a valid measure of functional activity.

Despite all available significant technical advantages for the analysis of fMRI data, a few challenges remain open. First, we ask how to develop a statistical model for integrating multimodal neuroimaging data optimally. Previous neuroimaging studies involving FC and SC for quantitative/computational modeling (Honey et al., 2009) revealed that brain regions with substantial SC exhibit strong FC, but strong FC could also occur in regions with weak SC. These results suggest that strong SC can be a predictor for strong FC, but the opposite does not necessarily stand (i.e., weak SC cannot be a predictor for weak FC). Information provided by multimodal imaging techniques can complement each other, and thus integrated multimodal analysis enables us to borrow strength from different modalities. In a review paper of studies combining SC and FC data, Rykhlevskaia et al. (2008) discussed various approaches to integrate FC and SC, including analysis of FC informed by SC. Zhao (2014) employed a bivariate model that incorporates between-modality and within-subject correlations while jointly analyzing FC and SC. Xue et al. (2015) introduced a Bayesian multimodal approach for analyzing FC time series data incorporating SC into modeling. Chiang (2016) developed a Bayesian autoregressive model that combined multimodal neuroimaging data by integrating SC into the prior information for improving the inference on effective connectivity. Bhaumik et al. (2018b) proposed a bivariate linear mixed-effects model with random subject intercepts and heteroscedastic errors to analyze FC and SC data jointly. Zhang et al. (2019) introduced a covariate-adaptive method employing a mixture of the generalized linear model and Gaussian model to optimize the p -value threshold for multiple hypothesis testing and applied it to the fMRI data with Brodmann area (cerebral cortex regions) as the covariate. It remains challenging how to *optimally* utilize the information of SC when the interest is in FC for a two-group (e.g., disease and healthy control) comparison study.

Second, high-throughput (a faster method that can produce a greater number of samples to be processed in the same or less time) neuroimaging technologies generate an incredibly large amount of data, “big data,” which can be very challenging to analyze and interpret. The primary objective of neuroimaging studies for comparing two groups is to detect any differences in connectivity associated with the disease. As such, thousands of hypotheses are tested simultaneously, known as the large-scale simultaneous hypothesis testing problem (Efron, 2004). Most of these hypotheses are null, containing nothing but noise, while only a small number of hypotheses have true signals. This sparsity issue makes the problem of detecting true signals very challenging (Sun and Cai, 2007). Conventional statistical methods used for multiple testing to control the family-wise error rate tend to be overly conservative in large-scale hypothesis testing, leading to a high chance of missing many important significant hypotheses that may be meaningful to researchers. An effective way of controlling FDR is needed for bimodal neuroimaging studies.

Third, the current economic cost of imaging techniques is high; consequently, most neuroimaging studies have a relatively limited number of subjects. Szucs and Ioannidis (2019) conducted a systematic review and evaluated sample sizes across highly cited

published MRI study papers between 1990 and 2012 and reported a median sample size of 12.5 per group based on 107 clinical fMRI studies with more than one group. The high sampling variability associated with a small sample leads to lower power with an inflated probability of falsely detected significant findings. Existing statistical methodologies hardly provide satisfactory results in controlling the FDR for neuroimaging studies with small samples (Bhaumik et al., 2018a). Therefore, developing a novel statistical method with a small sample to control the FDR better is critically important and challenging.

We address all three aforementioned concerns while detecting biomarkers. This article is structured as follows. In Section 2, we introduce a neuroimaging study in LLD, the motivating example for this research, where FC and SC data were collected from each participant. In Section 3, we develop a Bayesian multimodal Lfdr method (BLfdr) utilizing a Bayesian mixture model that integrates FC and SC data. This is followed by extensive simulations designed to evaluate the performance of our proposed method in Section 4. Further, we compare our simulated results with Efron's Lfdr method, which solely considers FC. In Section 5, we compare our results using BLfdr with Lfdr and Dirichlet process mixture model (DPM) in the context of the motivating example. Detailed results involving major brain networks associated with LLD are discussed. While classifying subjects to the LLD group, the excellent performance of BLfdr compared to other methods is shown in a table. In Section 6, we compare our approach with another multimodal method to detect biomarkers. In Section 7, we discuss how our discovery can be used to treat LLD subjects and some directions for future work.

2. Motivational example

In order to understand the relationship between structural abnormality and altered functional brain networks in LLD, the principal investigator Ajilore and his team conducted a study at the University of Illinois at Chicago (UIC). The study was approved by UIC Institutional Review Board and performed in compliance with the Declaration of Helsinki. Ajilore's study selected ten unmedicated and symptomatic subjects with geriatric depression (referred to as the LLD group) and 13 healthy elderly comparison subjects (referred to as the HC group) from the community. A total of 87 ROIs were parcellated by the Freesurfer Desikan atlas (Desikan et al., 2006). The study aimed to examine the relationship between abnormal structural changes in patients with LLD and whole-brain FC alternations. Study results are reported in a research paper (Rykhlevskaia et al., 2008). This study motivates us to develop an appropriate statistical methodology needed to achieve the goal.

3. Methods

Our Bayesian multimodal model utilizes the complementary SC data to enhance the modeling of the density of FC statistics while controlling the Lfdr. We assessed a late-fusion approach (Ramachandram and Taylor, 2017) by collecting individual resting-state average correlations of pairwise BOLD signal for FC and fiber counts for SC. The FC was measured using the FC

toolbox CONN (Whitfield-Gabrieli and Nieto-Castanon, 2012), which performs seed-based correlation analysis by computing the Pearson correlation coefficients between the BOLD time series from a given ROI to all other ROIs in the brain. Prior to data analysis, potential confounding factors including motion artifact, white matter, cerebrospinal fluid, and physiological noise source reduction were regressed out from the signal. Pearson correlation was then transformed into an approximately normal distribution using Fisher's Z transformation (Bartlett, 1993). A link-specific test statistic was developed to compare the corresponding links of the two groups. Three specific examples of such test statistics are (i) t-statistic to compare two correlations (from two groups) for a link (Rykhlevskaia et al., 2008), (ii) z-statistic to compare two Fisher-Z transformed correlations, and (iii) t-statistic to compare two intercepts of the linear mixed-effects model (Bhaumik et al., 2018a). In what follows, we first briefly describe Efron's Lfdr approach and then develop our Bayesian Mixture Model.

3.1. Bayesian multimodal local false discovery rate

FDR has been used to address the multiple testing issue in neuroimaging studies where it is defined as the expected proportion of false positives among all rejected hypotheses (Benjamini and Hochberg, 2000). The concept of FDR was further extended from an empirical Bayesian perspective (Efron, 2007). The Lfdr is defined as the posterior probability of the true null hypothesis given the observed test statistics. Literature suggests that Efron's Lfdr method performs better than other popular approaches for multiple comparisons in neuroimaging studies (Bhaumik et al., 2018a,b). What makes this problem exciting and challenging is whether extra strength can be borrowed from SC to better control the FDR by extending the Lfdr approach while analyzing FC data. In this article, we plan to develop a Bayesian multimodal approach to control the FDR extending the basic concept of Lfdr.

In Efron's Lfdr procedure, the status of the i th hypothesis is denoted by $H_{0,i}$. Assume a total of m hypotheses are tested simultaneously (i.e., $i = 1, 2, \dots, m$). Let $H_{0,i} = 0$ for a true $H_{0,i}$, and $H_{0,i} = 1$ otherwise, and t_i be the corresponding test statistic. $H_{0,i}$ follows i.i.d. Bernoulli distribution with probability p_1 . The Bayesian two-group mixture model (Efron, 2008) is given by,

$$f(t_i) = p_0 f_0(t_i) + p_1 f_1(t_i), \quad (1)$$

where $p_0 = P(H_{0,i} = 0)$ and $p_1 = P(H_{0,i} = 1) = 1 - p_0$ are the prior probabilities that are assumed to be the same across all hypotheses, f_0 and f_1 are the probability density functions of the test statistic t_i under null and alternative hypotheses, respectively. The Lfdr with t_i is defined as

$$Lfdr_i(t_i) = P(H_{0,i} = 0 | T = t_i) = \frac{p_0 f_0(t_i)}{f(t_i)}, \quad (2)$$

where p_0 , $f_0(t)$, and $f(t)$ can be estimated using the Poisson generalized linear model and "central matching" method based on the "zero assumption" that the central peak of the distribution of test statistics is comprised mostly of the null hypotheses (Efron,

2004, 2007). We extend the covariate-modulated Lfdr method integrating SC and FC statistics to improve the FDR for a cross-sectional, multimodal neuroimaging study with a small sample size (Zablocki et al., 2014). Our Bayesian multimodal Lfdr approach has two distinct features. First, for each FC link, the prior probability is modeled using a logistic regression model with corresponding link-specific SC statistics as a covariate instead of assuming a constant p_0 for all FC links in the Lfdr method that solely considers FC. Second, the densities of FC test statistics under null and alternative hypotheses are estimated using parametric models, where the posterior sampling for the model parameters is obtained using Gibbs sampling. The SC information serves as auxiliary information in the mixture model, aiding the identification of difference of FC between the disease and control group.

3.2. Bayesian mixture model

Under “zero assumption,” most of the test statistics around zero are from the null hypothesis. There may not be a sufficient number of test statistics under alternative hypotheses in either lower or upper tail areas that can be used for parametric density estimation. Thus, combining both tail areas provides better information to estimate the alternative density. Therefore, we use absolute values of test statistics compromising directions for modeling the density under alternative hypotheses.

From both theoretical and empirical perspectives, it is assumed that the underlying distribution of test statistics under the null hypothesis is normal, as described in the previous section (Efron, 2004, 2007). Under the null hypothesis, absolute values of test statistics follow a folded normal distribution, in which the probability mass values on the left half of the normal distribution are folded over the right half, as “folded” literally means.

Let $t_i^{(F)}$ and $t_i^{(S)}$ denote the absolute values of test statistics for FC and SC, respectively, to compare mean connectivity measures between the disease group and control group for the i^{th} connectivity link ($i = 1, \dots, m$). The $t_i^{(F)}$ and $t_i^{(S)}$ statistics are obtained by fitting the FC and SC data using a univariate linear mixed-effects regression model with heteroscedastic errors (Bhaumik et al., 2018a,b), separately. And $\mathbf{t}^{(F)} = [t_1^{(F)} \ t_2^{(F)} \ \dots \ t_m^{(F)}]^T$ and $\mathbf{t}^{(S)} = [t_1^{(S)} \ t_2^{(S)} \ \dots \ t_m^{(S)}]^T$ are the vectors of the absolute values of test statistics for FC and SC, respectively.

Assume that for links under the null hypothesis of no FC difference between two groups (referred to as null links), the corresponding test statistic $t_i^{(F)}$ follows a folded normal distribution (f_0) with zero location parameter ($\mu_0 = 0$) and unknown scale parameter ($\sigma_0^2 > 0$). The density function f_0 of the null FC has the following expression.

$$f_0(t_i^{(F)} | \mu_0 = 0, \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(t_i^{(F)} - \mu_0)^2}{2\sigma_0^2}\right] + \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(t_i^{(F)} + \mu_0)^2}{2\sigma_0^2}\right], \quad (3)$$

For a comparison study with two groups, the zero mean assumption under the null hypothesis is justified, and the σ_0^2

assumption provides better flexibility to explain variability. Note that the SC test statistic $t_i^{(S)}$ is not involved in (3), under the assumption that SC statistic has an insignificant influence on the null density of FC statistics (Zhang et al., 2019).

Next, we assume that for links with significant FC differences between two groups (referred to as alternative links), the corresponding test statistic $t_i^{(F)}$ follows a gamma distribution (f_1) with a fixed location parameter μ_1 , unknown shape parameter $\alpha(t_i^{(S)} | \boldsymbol{\alpha}) > 0$ and rate parameter $\beta > 0$. We consider a gamma distribution because of its flexibility in accommodating the skewness and flatness in the density function. Denote the shape parameter of the gamma distribution by α , given the t -value $t_i^{(S)}$ of SC and parameter vector $\boldsymbol{\alpha}$, where $\alpha = \alpha(t_i^{(S)} | \boldsymbol{\alpha}) = \exp(\alpha_0 + \alpha_1 t_i^{(S)})$. The density function of f_1 is as follows.

$$f_1(t_i^{(F)} | \mu_1, \alpha, \beta) = \frac{\beta^\alpha \alpha(t_i^{(S)} | \boldsymbol{\alpha})}{\Gamma(\alpha(t_i^{(S)} | \boldsymbol{\alpha}))} (t_i^{(F)} - \mu_1)^{\alpha(t_i^{(S)} | \boldsymbol{\alpha}) - 1} \exp[-\beta(t_i^{(F)} - \mu_1)]. \quad (4)$$

Further, we assume that $\alpha(t_i^{(S)} | \boldsymbol{\alpha})$ is a log-linear function of SC statistic $t_i^{(S)}$ with unknown parameters $\boldsymbol{\alpha} = [\alpha_0 \ \alpha_1]^T$ and $\alpha(t_i^{(S)} | \boldsymbol{\alpha}) = \exp(\alpha_0 + \alpha_1 t_i^{(S)})$. To avoid the non-convergence issue and gain some efficiency, we assume β does not depend on $t_i^{(S)}$. The location parameter μ_1 in (4) is purposely fixed at 0.674, which is the median of the standard folded normal distribution. Based on the “zero assumption,” it is reasonable to assume that $t_i^{(F)} \leq 0.674$ is null *a priori*, so the alternative density is bounded away from zero. Also, this restriction on μ_1 in f_1 is necessary to address the identifiability issue when strong assumptions on parameters are unavailable.

To complete the mixture model, we define a latent variable w_i to indicate which component of the mixture model (f_0 for null and f_1 for alternative) is used as a density function for each $t_i^{(F)}$. The prior probability of $w_i = 1$ denoted by $\pi_1(t_i^{(S)})$ and modeled using a logistic regression with $t_i^{(S)}$ as a covariate,

$$\pi_1(t_i^{(S)} | \boldsymbol{\gamma}) = P(w_i = 1 | t_i^{(S)}) = \frac{\exp(\gamma_0 + \gamma_1 t_i^{(S)})}{1 + \exp(\gamma_0 + \gamma_1 t_i^{(S)})}, \quad (5)$$

where $\boldsymbol{\gamma} = [\gamma_0 \ \gamma_1]^T$ is a vector of unknown parameters. The prior probability of $w_i = 0$, denoted by $\pi_0(t_i^{(S)})$, is

$$\pi_0(t_i^{(S)} | \boldsymbol{\gamma}) = P(w_i = 0 | t_i^{(S)}) = 1 - \pi_1(t_i^{(S)} | \boldsymbol{\gamma}) = \frac{1}{1 + \exp(\gamma_0 + \gamma_1 t_i^{(S)})}. \quad (6)$$

Then, the density of the mixture model (f) and the likelihood function $L = \mathbf{L}(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta, \sigma_0^2 | \mathbf{w}, \mathbf{t}^{(F)}, \mathbf{t}^{(S)})$ are defined as follows, respectively,

$$f(t_i^{(F)} | t_i^{(S)}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta, \sigma_0^2) = \pi_0(t_i^{(S)} | \boldsymbol{\gamma}) f_0(t_i^{(F)} | \sigma_0^2) + \pi_1(t_i^{(S)} | \boldsymbol{\gamma}) f_1(t_i^{(F)} | \boldsymbol{\alpha}, \beta, t_i^{(S)}), \quad (7)$$

$$L = \prod_{i=1}^m \left\{ \left[\pi_0(t_i^{(S)} | \boldsymbol{\gamma}) f_0(t_i^{(F)} | \sigma_0^2) \right]^{(1-w_i)} \left[\pi_1(t_i^{(S)} | \boldsymbol{\gamma}) f_1(t_i^{(F)} | \boldsymbol{\alpha}, \beta, t_i^{(S)}) \right]^{w_i} \right\}, \tag{8}$$

where $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_m]^T$ is a vector of w_i 's. The posterior probability (denoted by $\widehat{BLfdr}_i(t_i^{(F)}, t_i^{(S)})$) of the i th null connectivity link given $t_i^{(F)}, t_i^{(S)}$ statistics is computed using Bayes' theorem with posterior estimates of the parameters $\{\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}, \hat{\beta}, \hat{\sigma}_0^2\}$.

$$\widehat{BLfdr}_i = \left\{ \pi_0(t_i^{(S)} | \hat{\boldsymbol{\gamma}}) f_0(t_i^{(F)} | \hat{\sigma}_0^2) \right\} \left\{ \pi_0(t_i^{(S)} | \hat{\boldsymbol{\gamma}}) f_0(t_i^{(F)} | \hat{\sigma}_0^2) + \pi_1(t_i^{(S)} | \hat{\boldsymbol{\gamma}}) f_1(t_i^{(F)} | \hat{\boldsymbol{\alpha}}, \hat{\beta}, t_i^{(S)}) \right\}^{-1}. \tag{9}$$

Comparing the right hand expression of (9) with that of (2), we see that the numerator of (9) has a weight function $\pi_0(t_i^{(S)} | \hat{\boldsymbol{\gamma}})$ that depends on SC. The denominator is a convex function of $f_0(t_i^{(F)} | \hat{\sigma}_0^2)$ and $f_1(t_i^{(F)} | \hat{\boldsymbol{\alpha}}, \hat{\beta}, t_i^{(S)})$, where f_1 is a function of SC. Thus, $\widehat{BLfdr}_i(t_i^{(F)}, t_i^{(S)})$ in (9) incorporates the information from SC at each link level, whereas $Lfdr_i(t_i)$ in (2) does not depend on SC and related weights p_0 and p_1 are fixed for all links.

3.2.1. Prior distributions

Gamma (0.001, 0.001) and Inverse-Gamma (0.001, 0.001) are commonly used as non-informative priors in Bayesian inference using Gibbs Sampling (Lunn et al., 2009). However, some concerns are being raised and discussed regarding these priors. First, these priors are improper as they do not have corresponding distribution functions. Second, those priors are unbounded, and the integral over the entire parameter space is infinite.

$$\int_0^\infty f(\theta) d\theta = \int_0^\infty \frac{1}{\theta} d\theta = \ln \theta \Big|_0^\infty = \infty - (-\infty) = \infty. \tag{10}$$

Improper priors also yield improper posterior densities (Gelman, 2006; Gelman et al., 2013). The gamma and inverse gamma prior distributions mentioned earlier have means of one and large variances with a high probability density close to zero and a very long tail (as shown in the dashed red lines in Figure 1). The high prior probability density near zero strongly influences the posterior density, leading to a posterior distribution similar to the prior, with a majority mass near zero and a long and heavy tail. As a result, the posterior distribution is unregularized and diffuses to extremely large values. In that sense, the priors are not non-informative but quite influential. Furthermore, in complex cases, the improper prior often results in non-convergence issues in the Gibbs sampler that draws samples directly from the conditional distributions (Robert et al., 1999).

Due to the undesirable properties of non-informative priors as discussed above, it is recommended to use weakly informative prior that contains a little but sufficient information to provide regularization and ensures that the posterior remains bounded within a reasonable range (Gelman, 2006; Gelman et al., 2008, 2013). In this study, we followed the concept of the method of moments estimation (Kim et al., 2009) and chose hyperparameters,

Gamma ($a_\beta = 1, b_\beta = 1$) and Inverse-Gamma ($a_{\sigma_0} = 3, b_{\sigma_0} = 2$) with means = 1 and variances = 1 as the weakly informative priors for β and σ_0^2 , respectively (as shown in the dash-dotted blue lines in Figure 1). These weakly informative priors offer more stable inferences than non-informative priors by providing regularization in parameter estimation and enough vagueness to ensure that the data dominate posteriors.

For the other unknown parameters $\{\boldsymbol{\alpha}, \boldsymbol{\gamma}\}$ in the mixture model (7), we assume generic weakly informative prior distributions providing more stable results compared to non-informative prior, but still have enough vagueness to ensure that the data dominate the posteriors. For $\boldsymbol{\alpha} = [\alpha_0 \ \alpha_1]^T$ and $\boldsymbol{\gamma} = [\gamma_0 \ \gamma_1]^T$, we assume that α_0 and α_1 are *a priori* independent, and also γ_0 and γ_1 are *a priori* independent, and their values are concentrated between -1 and 1 . Further we assume $\boldsymbol{\alpha} \sim N_2(\mathbf{0}, \Sigma_\alpha)$ and $\boldsymbol{\gamma} \sim N_2(\mathbf{0}, \Sigma_\gamma)$, where $\Sigma_\alpha = \Sigma_\gamma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

3.2.2. Posterior distributions

We obtain the posterior sampling distributions for unknown parameters from Markov Chain Monte Carlo (MCMC) using Gibbs sampler and Metropolis-Hasting algorithm (Robert et al., 1999). Expressions of posterior distributions are:

$$f(\boldsymbol{\alpha} | \beta, \mathbf{w}, \mathbf{t}^{(F)}, \mathbf{t}^{(S)}) \propto \exp \left\{ -\frac{1}{2} \boldsymbol{\alpha}^T \Sigma_\alpha^{-1} \boldsymbol{\alpha} \right\} \prod_{\{i|w_i=1\}} \frac{\beta^{\alpha(t_i^{(S)}|\boldsymbol{\alpha})}}{\Gamma(\alpha(t_i^{(S)}|\boldsymbol{\alpha}))} (t_i^{(F)} - \mu_1)^{\alpha(t_i^{(S)}|\boldsymbol{\alpha})}, \tag{11}$$

$$f(\boldsymbol{\gamma} | \mathbf{w}, \mathbf{t}^{(F)}) \propto \exp \left\{ -\frac{1}{2} \boldsymbol{\gamma}^T \Sigma_\gamma^{-1} \boldsymbol{\gamma} \right\} \prod_{i=1}^m \frac{\exp(\gamma_0 + \gamma_1 t_i^{(S)})^{w_i}}{1 + \exp(\gamma_0 + \gamma_1 t_i^{(S)})}, \tag{12}$$

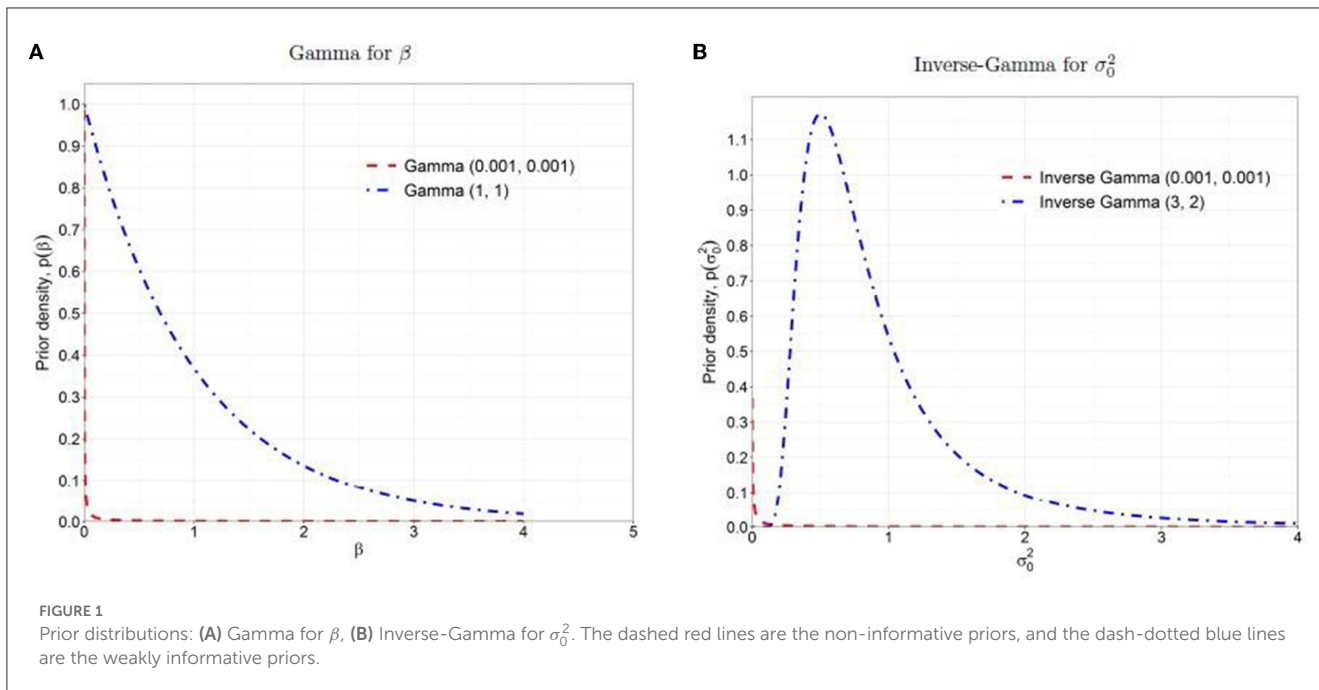
$$f(\beta | \boldsymbol{\alpha}, \mathbf{w}, \mathbf{t}^{(F)}, \mathbf{t}^{(S)}) \sim \text{Gamma} \left(1 + \sum_{\{i|w_i=1\}} \alpha(t_i^{(S)} | \boldsymbol{\alpha}), 1 + \sum_{\{i|w_i=1\}} (t_i^{(F)} - \mu_1) \right), \tag{13}$$

$$f(\sigma_0^2 | \mathbf{w}, \mathbf{t}^{(F)}) \sim \text{Inverse-Gamma} \left(3 + \frac{1}{2} \sum_{i=1}^m I_{\{w_i=0\}}, 2 + \frac{1}{2} \sum_{\{i|w_i=0\}} (t_i^{(F)})^2 \right). \tag{14}$$

where $\alpha(t_i^{(S)} | \boldsymbol{\alpha}) = \exp(\alpha_0 + \alpha_1 t_i^{(S)})$, $\mu_1 = 0.674$, $I_{\{w_i=0\}} = 1$ if $w_i = 0$ and 0 otherwise.

The posterior sampling for β and σ_0^2 are directly drawn from the full conditional distributions (13) and (14), respectively. For $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, posterior samplings are drawn from a multivariate t distribution with a small number of df ν such as $\nu = 4$, providing three finite moments to approximate the density. Specifically, a multiple-try Metropolis algorithm is employed to increase the step size and acceptance rate (Liu et al., 2000).

For Markov chains, the initial values are set as $\boldsymbol{\alpha}^{(0)} = [0 \ 0]^T$, $\beta^{(0)} = 0.1$, $\sigma_0^{2(0)} = 1$ and $\pi_0^{(0)} = 0.94$. Denote the threshold



that corresponds to $\pi_0^{(0)} = 0.94$ by $\lambda^{(0)}$, $w_i^{(0)} = 1$ if $t_i^{(F)} \geq \lambda^{(0)}$, and 0 otherwise. Then, we can get the estimates of the coefficients based on the logistic regression model on $w_i^{(0)}$ given $t_i^{(S)}$, as the initial values of $\boldsymbol{\gamma}^{(0)} = [\gamma_0^{(0)} \ \gamma_1^{(0)}]^T$. At the k th MCMC iteration, $k = 1, \dots, K$, when $\{\boldsymbol{\alpha}^{(k)}, \boldsymbol{\gamma}^{(k)}, \beta^{(k)}, \sigma_0^{2(k)}\}$ are drawn, the probability of the latent variable $w_i^{(k)} = 1$ given $t_i^{(F)}$, $t_i^{(S)}$ and $\{\boldsymbol{\alpha}^{(k)}, \boldsymbol{\gamma}^{(k)}, \beta^{(k)}, \sigma_0^{2(k)}\}$ denoted by $p_i^{(k)}$, is updated using

$$\begin{aligned}
 p_i^{(k)} &= P(w_i = 1 \mid t_i^{(F)}, t_i^{(S)}, \boldsymbol{\alpha}^{(k)}, \boldsymbol{\gamma}^{(k)}, \beta^{(k)}, \sigma_0^{2(k)}) \\
 &= \left\{ \pi_1(t_i^{(S)} \mid \boldsymbol{\gamma}^{(k)}) f_1(t_i^{(F)} \mid t_i^{(S)}, \boldsymbol{\alpha}^{(k)}, \beta^{(k)}) \right\} \\
 &\times \left\{ \pi_0(t_i^{(S)} \mid \boldsymbol{\gamma}^{(k)}) f_0(t_i^{(F)} \mid \sigma_0^{2(k)}) \right. \\
 &\left. + \pi_1(t_i^{(S)} \mid \boldsymbol{\gamma}^{(k)}) f_1(t_i^{(F)} \mid t_i^{(S)}, \boldsymbol{\alpha}^{(k)}, \beta^{(k)}) \right\}^{-1}. \quad (15)
 \end{aligned}$$

$w_i^{(k)}$ is sampled from the full conditional Bernoulli distribution with probability $p_i^{(k)}$, i.e., $w_i^{(k)} \stackrel{iid}{\sim} \text{Bernoulli}(p_i^{(k)})$.

After obtaining samples for each parameter from their corresponding posterior distributions, we compute the posterior medians as the posterior estimates of the parameters $\{\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}, \hat{\beta}, \hat{\sigma}_0^2\}$. The posterior probability that the i th connectivity link is null, given $t_i^{(F)}$ and $t_i^{(S)}$, denoted by $\widehat{BLfdr}_i(t_i^{(F)}, t_i^{(S)})$. To ensure that the average FDR is controlled at a pre-specified level of q , we then apply the oracle procedure (Sun and Cai, 2007), which regards multiple testing as a compound decision problem and intended to minimize the false non-discovery rate that is subject to a constraint on the $FDR \leq q$: (1) Sort the calculated \widehat{BLfdr}_i , $i = 1, \dots, m$ in ascending order denoted by $\widehat{BLfdr}_{(1)} \leq \widehat{BLfdr}_{(2)} \leq \dots \leq \widehat{BLfdr}_{(m)}$; (2) For a given q ($0 < q < 1$), find k for which $k = \arg \max \left\{ i : \frac{1}{i} \sum_{j=1}^i \widehat{BLfdr}_{(j)} \leq q \right\}$; (3) Reject all $H_{(i)}$, $i = 1, \dots, k$.

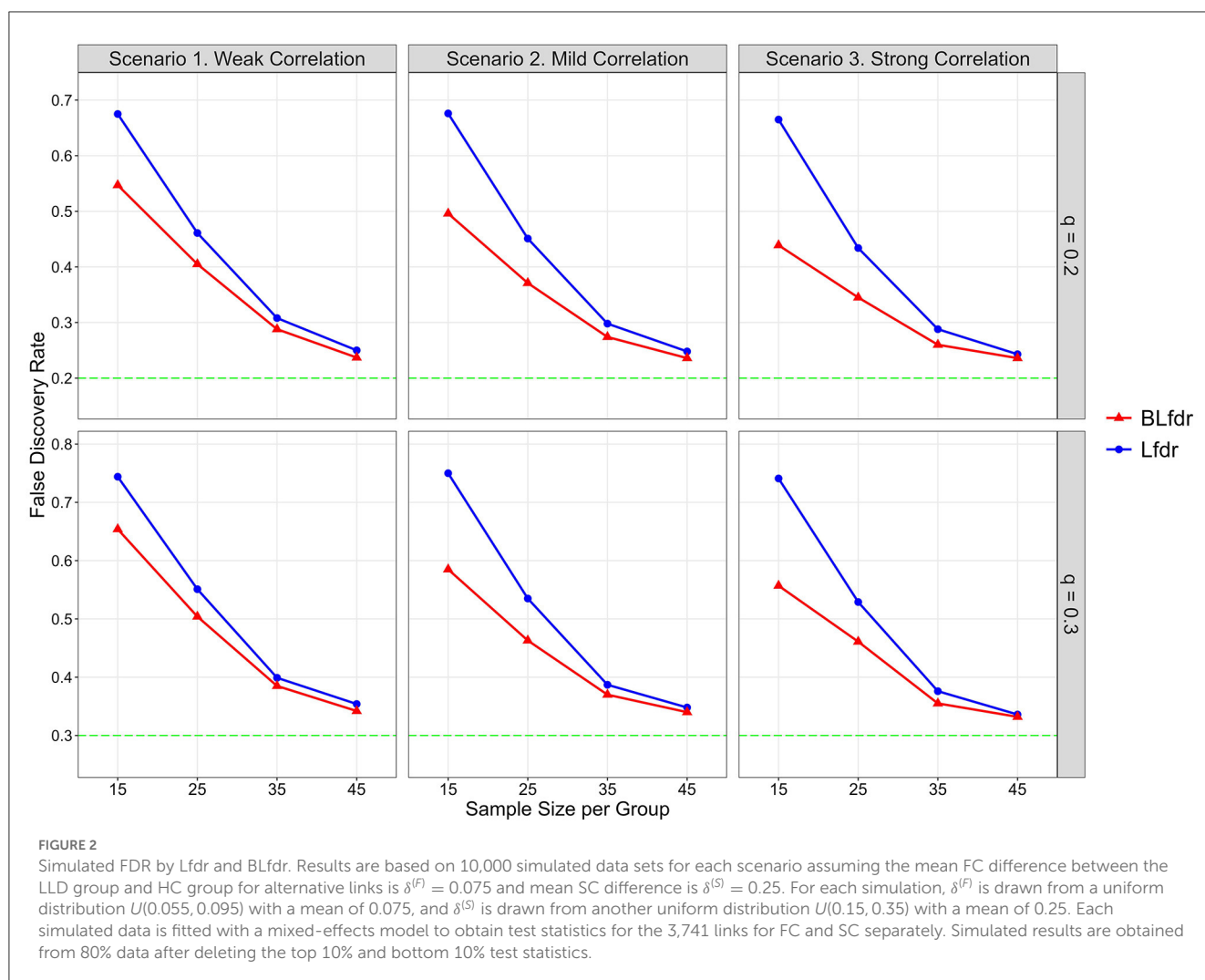
4. Simulation study

We performed a simulation study to see how the proposed Bayesian multimodal approach controls the local false discovery rate. Related algorithms, the combination of parametric space, and corresponding results are given below.

4.1. An algorithm for simulation

Several data sets are generated with varying sample sizes $n = 15, 25, 35$, and 45, following a very similar structure to that of the motivating LLD study. The performance of BLfdr is compared with that of the Lfdr method. To estimate the percentage of null (i.e., no significant difference between two groups) of FC, we computed a t -statistic utilizing the difference of Fisher's Z transformed Pearson correlations of LLD and HC. The corresponding p -value for each link is determined from the t -test statistic. Smaller p -values (< 0.05) are used to screen the first level of the null FC. Similarly, a Poisson distribution is used to detect a null link for SC.

We assume out of a total of 3,741 links, 1% links are non-null in both FC and SC, 1% links are non-null in FC, but those are null in SC, another 1% links are null in FC, but those are non-null in SC, and the rest 97% are null in both FC and SC. The 2% proportion of alternative links in FC aligns with the assumption used in the simulation study with the same LLD neuroimaging data (Song, 2016; Bhaumik et al., 2018b). The between-group difference is measured by the difference between the intercept parameters of a mixed-effects model while comparing LLD with HC. We consider the between-group difference of FC and SC for the alternative links as $\delta^{(F)} = 0.075$ and $\delta^{(S)} = 0.25$, respectively. To account for the potential correlation between FC and SC, we use a bivariate mixed-effects model with heteroscedastic errors to simulate FC



and SC data jointly, where the parameters of the model are set close to the estimated values of those determined from our LLD data discussed before. A detailed description of the simulation strategy with various steps of selection of parameters is given in the [Appendix](#).

Note that our assumption on errors is wide open by employing a gamma distribution to accommodate heterogeneity in links and skewness of the distribution. We make a realistic assumption that alternatives may vary following a gamma distribution, unlike a fixed alternative. Further, we assume that error distributions of FC and SC of HC are different from respective distributions of LLD. As SC and FC are nested within the same subject, it is natural to expect that those are correlated, which is reflected in our assumption.

4.2. Simulated results

We present simulated results in [Figure 2](#) in terms of FDR at pre-specified levels of $q = 0.2$ and 0.3 for a weak correlation = 0.1, mild correlation = 0.4, and strong correlation = 0.9. As mentioned before, neuroimaging studies usually have a small number of subjects; thus, improving the FDR for a small sample is extremely helpful for researchers.

For $n = 15$ (per group) and $q = 0.2$, simulated FDR by BLfdr ranges between 0.439 and 0.547, compared to a range between 0.665 and 0.678 by Lfdr. In [Figure 2](#), we see BLfdr (red color) has a better performance across all three correlations for both $q = 0.20$ and 0.30 compared to Lfdr (blue color). In particular, for $n = 15$, better performance of BLfdr with a wider gap between red and blue curves draws our attention.

As sample size increases, FDRs by both methods are reduced greatly. For $n = 45$, FDRs by the two methods are close (but still BLfdr is slightly better than Lfdr), and both converge to the pre-specified level q , suggesting that even asymptotic results of BLfdr are slightly better than those of Lfdr. FDRs at $q = 0.3$ are larger than those at $q = 0.2$, which is expected. This simulation also provides a guideline to researchers on determining the appropriate sample size ($n = 45$ per group) when the goal is to control the FDR at the desired level (i.e., $q = 0.2$ or 0.3).

The take-home message of this simulation study is that when the goal is to detect disrupted FC, it is better to incorporate additional information from SC (as BLfdr does, but Lfdr does not) for better controlling the FDR. Determination of appropriate sample size is necessary to effectively control the FDR at a desired level q , which this study does. Otherwise, the study will end

up with discovering false biomarkers, defeating the purpose of the study.

5. Application of BLfdr to LLD data

Let us now revisit the motivational example and analyze the study data. Various assumptions were made while developing a statistical model in Section 2. In this section, first, we show how those assumptions are satisfied with the LLD data and why it is appropriate to take $q = 0.2$ or 0.3 . Once the validity of the model is established, we will concentrate on detecting significant disrupted FCs. The details of image acquisition and data processing of the study are described in [Tadayonnejad et al. \(2014\)](#) and [Zhao \(2014\)](#).

As mentioned in the motivating example, the data set has 23 participants; 13 HC and 10 LLD subjects. The adjacency matrix of each of FC and SC has $\binom{87}{2} = 87 \times (87 - 1)/2 = 3741$

unique connectivity measures (or links). The SC was measured using an internal Matlab program by counting the number of fiber tracts found by the tractography algorithm connecting each pair of regions. Some SC data have zero values, suggesting no SC between the corresponding brain regions. Since FC data is measured using Pearson’s correlation coefficients (r), we apply Fisher’s Z transformation to stabilize the variance and get an approximate normal distribution for the transformed data. Moreover, for SC data, cube-root transformation is applied to get an approximate normal distribution ([Bhaumik et al., 2018b](#)).

The p -values from t -tests to compare mean FC measures between LLD and HC are obtained using a mixed-effects model for all 3741 links. Out of 3,741 hypotheses tests, 281 (7.5%) tests have two-sided $p \leq 0.025$. This indicates for the existence of a small proportion of alternative links whose FC measurements differ significantly between LLD and HC and some potential false positives.

We ran Markov chains using the prior distribution for model parameters specified before. The acceptance rates for α and γ using the multiple-try Metropolis within the Gibbs algorithm were satisfactory. We performed model convergence diagnostics to ensure that it converged to the target posterior distribution and reached the stationarity. We took the posterior medians as the posterior estimates of each parameter. The posterior inference, including medians and 95% credible intervals, are presented in [Table 1](#). The results are based on 9,000 posterior samples combined from three parallel Markov chains, where each chain has 325,000 iterations, with the first 25,000 as burn-in and a thinning interval of 100 after burn-in.

TABLE 1 Posterior inference for model parameters.

| Parameter | Median (95% credible interval) |
|--------------|--------------------------------|
| α_0 | 1.120 (0.362, 1.739) |
| α_1 | -0.018 (-0.548, 0.353) |
| β | 2.214 (1.204, 3.359) |
| γ_0 | -2.313 (-3.138, -1.484) |
| γ_1 | -0.692 (-1.680, -0.055) |
| σ_0^2 | 1.100 (0.829, 1.283) |

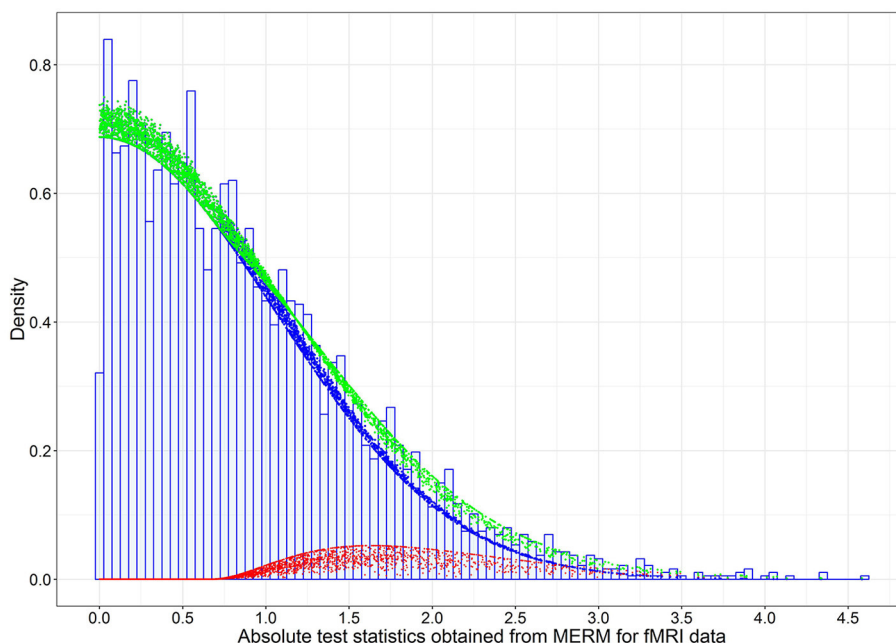


FIGURE 3

The estimated weighted null, weighted alternative, and mixture densities based on the posterior estimates of model parameters for the 3,741 absolute test statistics obtained from the mixed-effects model on FC data in the LLD study. The blue dots are the weighted null density $\pi_0(t_i^{(S)} | \hat{\gamma}) f_0(t_i^{(F)} | \hat{\sigma}_0^2)$, the red dots are the weighted alternative density $\pi_1(t_i^{(S)} | \hat{\gamma}) f_1(t_i^{(F)} | \hat{\alpha}, \hat{\beta}, t_i^{(S)})$, and green dots are the mixture density $\pi_0(t_i^{(S)} | \hat{\gamma}) f_0(t_i^{(F)} | \hat{\sigma}_0^2) + \pi_1(t_i^{(S)} | \hat{\gamma}) f_1(t_i^{(F)} | \hat{\alpha}, \hat{\beta}, t_i^{(S)})$, where $\{\hat{\alpha}, \hat{\gamma}, \hat{\beta}, \hat{\sigma}_0^2\}$ are the posterior estimates ([Table 1](#)).

Assuming these FC links were under the null hypothesis, the estimated variance ($\hat{\sigma}_0^2$) of the null density is 1.01 using absolute test statistics with $p > 0.025$. This is a reasonable fit for the links under the null hypothesis, while the tail of the histogram for the potential alternative links is fitted using a gamma distribution described in the Section of the Bayesian Mixture Model. Thus, we can obtain a conservative estimate of the proportion of null cases (Storey and Tibshirani, 2003), denoted by $\hat{\pi}_0$. In Figure 3, the dashed blue line is the weighted null density $\hat{\pi}_0 \hat{f}_0(t_i^{(F)})$, where $\hat{\pi}_0 = \sum_{i=1}^{3741} I\{p\text{-value}_i > 0.025\} / 3741(1 - 0.025) = 3584 / (3,741 \times 0.975) = 0.983$. The weighted red dots represent the alternative (non-null) density, and the green dots represent mixture density using posterior estimates given in Table 1.

Based on the posterior estimates of the model parameters $\{\hat{\alpha}, \hat{\gamma}, \hat{\beta}, \hat{\sigma}_0^2\}$, we calculated the posterior probability of each link being null given the observed FC and SC test statistics, \widehat{BLfdr}_i , $i = 1, \dots, 3,741$, by (9), and then determined which connectivity links to be rejected using the oracle procedure to control the FDR (Sun and Cai, 2007). The Bayesian multimodal Lfdr method detected 21 connectivity links, significant FC difference between the LLD group ($n_1 = 10$) and HC group ($n_2 = 13$) at $q = 0.2$ is presented in Tables 2, 3, where 15 connectivity links have increased FC significantly (hyperconnectivities), and six connectivity links have decreased FCs significantly (hypoconnectivities). The primary hub, the right caudal middle frontal (RCMF), is identified by both BLfdr and Lfdr methods.

To assess the consistency of our results between the BLfdr and Lfdr methods, we compared the significant FC links identified using these two methods. Using Lfdr by R locfdr package, we get the central matching estimates of the null distribution as $f_0 \sim Normal(\hat{\mu}_0 = 0.038, \hat{\sigma}_0^2 = 1.250)$ and the null proportion as $\hat{p} = 0.988$. This large value of \hat{p} justifies using the FDR $q = 0.2$ or 0.3 . The estimated null density has a variance of 1.250, which is larger than 1.100 using BLfdr. The Lfdr method detects 12 significant FC links at $q = 0.2$, among which 11 links overlap with those identified by BLfdr. Matching with the simulated results, we would like to say that borrowing strength from SC, BLfdr controls the FDR better than Lfdr. Further, to compare our results, we used a semi-parametric Bayesian approach utilizing a nonparametric Dirichlet process mixture model (DPM) (Ghosal, 2019) and Tract-Based Spatial Statistics (TBSS) as discussed in Tadayonnejad et al. (2014). It should be noted that the primary purpose of developing the TBSS was not for classification. Significant FC links detected by various methods are presented in Tables 2, 3. Significant SC links detected by the Univariate Linear Mixed-Effects Model (ULMM) at q level of 0.3 are presented in Tables 2, 3.

5.1. Accuracy in classification

We examine the performance of our approach while classifying subjects with selected links presented in Tables 2, 3. Random Forest, an embedded learning regression method with a leave-one-out cross-validation strategy, is used to classify subjects with selected features presented in Tables 2, 3. Performance of each approach is measured by sensitivity, specificity, accuracy, and area under the curve (AUC), and results are presented in Table 4. BLfdr has a very

high sensitivity (89.05%), specificity (87.63%), accuracy (88.54%), and AUC (98.46%) and performs much better compared to all other methods. One reason for the better performance of BLfdr is the additional strength it borrows from SC. The high accuracy rate of classification of BLfdr encourages us to use it for prediction of early prevention.

6. Model comparison with multimodality

Multimodal neuroimaging data are used to better understand the neurological condition of brain with diseases. Neuroimaging studies with multimodality mainly focus on correlations of modalities or detection of the main components of such correlations. Sui et al. (2013) employed the independent component analysis to estimate the number of independent components using three modalities such as fMRI, DTI, and structural MRI while studying the abnormal structure underlying schizophrenia relative to healthy controls. Zhao (2014) used a bivariate linear mixed-effects model (BLMM) to detect disrupted links of LLD compared to HC. Honey et al. (2009) investigated the correlation between FC (using resting-state fMRI) with SC (using DTI).

Our proposed Bayesian model, i.e., BLfdr, utilizes SC as an auxiliary information to model FC. To enable comparisons, we used the BLMM to model FC and SC jointly. Controlling the FDR at a level of 0.2, we identified 40 significant links for both FC and SC, only two links (left ventral diencephalon (LVD)—right caudal anterior cingulate (RCAC) and left thalamus proper (LTP)—left posterior cingulate (LPC) overlapped with those detected by the BLfdr model. However, considering brain regions as a whole, we found that 15 regions of interest (ROIs) identified by the BLMM overlapped with those detected by the BLfdr model (refer to Table 5). Moreover, BLMM requires subjective decisions regarding the covariance matrix between SC and FC, and it takes a long computational time due to the setting of bivariate outcomes. In contrast, our proposed Bayesian approach is straightforward and properly leverages multimodal information to detect disrupted functional connectivity.

6.1. Model-based analysis results

The traditional method for comparing LLD with HC using FC or SC data typically involves (i) t -tests that ignore between-link correlations and (ii) Hotelling's T-square tests. In addition, we compared the two groups by (iii) Fisher's Z transformation on Pearson correlation of SC and FC. Using (i), we observed 59 significant FCs and 49 significant SCs; (ii) produced 55 significant links for joint FC and SC, and (iii) identified 72 significant links for FC-SC correlations. However, none of these analyses remained significant after controlling for the multiplicity problem and type I error rate using an FDR level of 0.30. We observed almost no overlapping results between FC and SC in (i); and between (i) and (iii). As expected, (i) and (ii) share about 50% of common significant links.

TABLE 2 FC links identified by BLfdr, Efron's Lfdr, DPM, and TBSS at q = 0.2 for LLD study.

| Region 1 ¹ | Region 2 ¹ | BLfdr ^{1,2} | Lfdr ^{1,2} | DPM ^{1,2} | TBSS ^{1,2} | ULMM ^{1,2} |
|------------------------------------|-------------------------------------|----------------------|---------------------|--------------------|---------------------|---------------------|
| Hypoconnectivities | | | | | | |
| R caudal middle frontal (RCMF) | L rostral middle frontal (LRMF) | X | X | X | | |
| R caudal middle frontal (RCMF) | R pallidum (RP) | X | X | | | |
| R caudal middle frontal (RCMF) | R caudal anterior cingulate (RCAC) | X | X | | | |
| R caudal middle frontal (RCMF) | R isthmus cingulate (RIC) | X | X | | | |
| R caudal middle frontal (RCMF) | R thalamus proper (RTP) | X | X | X | | |
| R caudal middle frontal (RCMF) | R posterior cingulate (RPC) | X | | X | | |
| R caudal middle frontal (RCMF) | L isthmus cingulate (LIC) | X | X | X | | |
| R caudal middle frontal (RCMF) | R precentral (RPrC) | | | | | X |
| R accumbens area (RAA) | L isthmus cingulate (LIC) | X | X | | | |
| R accumbens area (RAA) | R medial orbitofrontal (RMO) | | | | X | |
| R accumbens area (RAA) | L caudate (LCau) | | | | | X |
| R posterior cingulate (RPC) | R supramarginal (RS) | X | | X | | |
| L posterior cingulate (LPC) | R supramarginal (RS) | X | | X | | |
| L posterior cingulate (LPC) | R rostral middle frontal (RRMF) | X | X | | | |
| R ventral diencephalon (RVD) | R fusiform (RF) | X | X | X | | |
| R amygdala (RAMY) | R fusiform (RF) | | | | | X |
| R pallidum (RP) | L inferior parietal (LIP) | X | X | | | |
| L pallidum (LP) | L inferior parietal (LIP) | | | | | X |
| L pallidum (LP) | L cerebellum (LC) | | | | | X |
| L insula (LI) | L inferior parietal (LIP) | | | | | X |
| L pallidum (LP) | Brain Stem (BS) | | | | | X |
| L inferior parietal(LIP) | L ventral diencephalon(LVD) | | | | | X |
| R amygdala (RAMY) | R ventral diencephalon(RVD) | | | | | X |
| R amygdala (RAMY) | Brain Stem (BS) | | | | | X |
| R superior frontal (RSF) | R caudal middle frontal (RCMF) | | | X | | |
| R pars triangularis (RPT) | R rostral anterior cingulate (RRAC) | X | | | | |
| R superior frontal (RSF) | R rostral anterior cingulate (RRAC) | | | | X | |
| L superior frontal (LSF) | R rostral anterior cingulate (RRAC) | | | | X | |
| L superior frontal (LSF) | R isthmus cingulate (RIC) | | X | | | |
| R superior parietal (RSP) | R paracentral (RPA) | | | | | X |
| L superior parietal (LSP) | R superior parietal (RSP) | | | X | | |
| L superior parietal (LSP) | R pars opercularis (RPO) | X | | X | | |
| L superior parietal (LSP) | R supramarginal (RS) | | | X | | |
| L superior parietal (LSP) | R insula (RI) | | | | | X |
| L superior parietal (LSP) | L lateral occipital(LLO) | | | | | X |
| L superior parietal (LSP) | L parahippocampal (LPH) | | | | | X |
| L isthmus cingulate (LIC) | L parahippocampal (LPH) | | | | | X |
| L precuneus (LPR) | L parahippocampal (LPH) | | | | | X |
| L inferior parietal (LIP) | R posterior cingulate (RPC) | | | X | | |
| L caudal anteriorc ingulate (LCAC) | R posterior cingulate (RPC) | | | | | X |

(Continued)

TABLE 2 (Continued)

| Region 1 ¹ | Region 2 ¹ | BLfdr ^{1,2} | Lfdr ^{1,2} | DPM ^{1,2} | TBSS ^{1,2} | ULMM ^{1,2} |
|-------------------------------------|------------------------------|----------------------|---------------------|--------------------|---------------------|---------------------|
| L rostral anterior cingulate (LCAC) | R medial orbitofrontal (RMO) | | | | | X |
| L rostral middle frontal (LRMF) | R medial orbitofrontal (RMO) | | | | | X |
| L lateral orbitofrontal (LLO) | R medial orbitofrontal (RMO) | | | X | | |
| L lateral orbitofrontal (LLO) | L medial orbitofrontal (LMO) | | | X | | |
| L frontal pole (LFP) | L medial orbitofrontal (LMO) | | | | | X |
| L precentral (LPrC) | R precentral (RPrC) | | | X | | |
| L precentral (LPrC) | R postcentral (RPoC) | | | X | | |
| L postcentral (LPoC) | L precentral (LPrC) | | | X | | |
| R insula (RI) | R bankssts (RB) | | | | | X |
| R insula (RI) | L lateral occipital (LLO) | | | | | X |
| L insula (LI) | R insula (RI) | | | | | X |
| L insula (LI) | Brain Stem (BS) | | | | | X |
| L insula (LI) | L thalamus proper (LTP) | | | | | X |
| L superior temporal (LST) | L thalamus proper (LTP) | | | | | X |
| R putamen (RPU) | R precuneus (RPR) | | | | | X |
| R putamen (RPU) | R cerebellum (RC) | | | | | X |
| R putamen (RPU) | Brain Stem (BS) | | | | | X |
| Brain Stem (BS) | R temporal pole (RT) | | | | | X |
| Brain Stem (BS) | L superior Temporal (LST) | | | | | X |

¹L, left; R, right; LLD, late-life depression; HC, health control; BLfdr, Bayesian Lfdr model; Lfdr, Efron’s local discovery rate model; DPM, Dirichlet process mixture; TBSS, tract-based spatial statistics; ULMM, univariate linear mixed-effects model for SC links. ²The connectivity links identified by each method are indicated using X.

TABLE 3 FC links identified by BLfdr, Efron’s Lfdr, DPM, and TBSS at q = 0.2 for LLD study.

| Region 1 ¹ | Region 2 ¹ | BLfdr ^{1,2} | Lfdr ^{1,2} | DPM ^{1,2} | TBSS ^{1,2} | ULMM ^{1,2} |
|------------------------------|------------------------------------|----------------------|---------------------|--------------------|---------------------|---------------------|
| Hypoconnectivities | | | | | | |
| R caudate (RCau) | R cuneus (RCun) | | | X | | |
| R caudate (RCau) | L cuneus (LCun) | X | | | | |
| L fusiform (LF) | L pars triangularis (LPT) | X | | | | |
| L fusiform (LF) | L supramarginal (LS) | X | X | X | | |
| L entorhinal (LE) | L supramarginal (LS) | X | | | | |
| L thalamus proper (LTP) | L posterior cingulate (LPC) | X | | | | |
| L ventral diencephalon (LVD) | R caudal anterior cingulate (RCAC) | X | | | | |

¹L, left; R, right; LLD, late-life depression; HC, health control; BLfdr, Bayesian Lfdr model; Lfdr, Efron’s local discovery rate model; DPM, Dirichlet process mixture; TBSS, tract-based spatial statistics; ULMM, univariate linear mixed-effects model for SC links.

²The connectivity links identified by each method are indicated using X.

Using the univariate linear mixed-effects model (ULMM) and controlling the FDR level of 0.3, we detected 28 significant FCs and 31 significant SCs. SC results of ULMM are presented in the last column of Tables 2, 3. The bivariate mixed-effects model detected 40 and 90 significant links when the FDR level was controlled at 0.2 and 0.3, respectively. An interesting finding of this analysis is that (a) the proportion of false discoveries (FDP) for FC with within-subject independence was 0.52, and with within-subject covariance was 0.06, (b) the FDP for SC with within-subject independence was 0.36, and with within-subject covariance was

0.04. These results suggest that when an appropriate covariance structure is implemented, the FDP is close to the desired value of 0.05. However, when the correlation structure was ignored, FDP exceeded the desired level of 0.50.

One possible explanation for these findings is that when tests are correlated, an appropriate dependency structure must be implemented in the computation of the FDP. Highly correlated hypotheses that favor the alternative are likely to be rejected together, while hypotheses that favor the null are likely to be retained together because of their dependency. If the FDP value is

TABLE 4 Classification performance of Lfdr, BLfdr, DPM, and TBSS at $q = 0.2$ for LLD study.

| Model | Accuracy | Sensitivity | Specificity | AUC |
|-------|----------|-------------|-------------|--------|
| BLfdr | 0.8854 | 0.8905 | 0.8763 | 0.9846 |
| Lfdr | 0.7964 | 0.8668 | 0.7043 | 0.9000 |
| DPM | 0.7905 | 0.8027 | 0.7705 | 0.9154 |
| TBSS | 0.6996 | 0.8074 | 0.5560 | 0.6731 |

TABLE 5 Fifteen ROIs detected by BLMM that overlapped with those found by BLfdr.

| Network | ROIs |
|--|------------------------------------|
| Default mode network (DMN) | L inferior parietal (LIP) |
| | L/R posterior cingulate (LPC/RPC) |
| | L/R thalamus proper (LTP/RTP) |
| Saliency network (SN) | R caudal anterior cingulate (RCAC) |
| Central executive network (CEN) | L superior parietal (LSP) |
| | R caudal middle frontal (RCMF) |
| | R rostral middle frontal (RRMF) |
| Visual network/Fusiform face area (VN/FFA) | L cuneus (LCun) |
| | L fusiform (LF) |
| Reward network (RN) | L ventral diencephalon(LVD) |
| | R accumbens area (RAA) |
| | R caudate (RCau) |
| | R pallidum (RP) |

small, it indicates that the number of false discoveries is not large, and any hypothesis that is not rejected is likely not to be significant. However, under the assumption of an uncorrelated structure, if the number of rejections remains the same, the rejected hypotheses are more likely to be randomly selected from all hypotheses, leading to an inflated FDP. These results suggest that the underlying correlation structure plays a vital role in controlling the FDP when dealing with multiple testing problem, and ignoring such correlations increases the likelihood of false discoveries.

Using a bivariate mixed-effects model with FC and SC, we identified the left inferior parietal as the hub of impaired regions. Bayesian results are comparable with those described above by the bivariate mixed-effects model. However, SC links detected by ULMM using SC are not comparable with FC links detected by the bivariate mixed-effects model or BLfdr using FC and SC. **Figure 4** displays the 31 disrupted SC links detected by the ULMM.

7. Neurophysiological importance of our discovery

BLfdr with $q = 0.2$ detects 21 significantly disrupted links on cortical and subcortical gray matter regions in the left and right hemispheres of the brain. **Figure 5** displays all FC connectivity links significantly different between the LLD and HC groups. The primary hub—the right caudal middle frontal, also known

as the right dorsolateral prefrontal cortex (dlPFC), is identified. In LLD patients, we find significantly increased FCs in seven other brain regions, including the bilateral isthmus cingulate (LIC/RIC), left rostral middle frontal (LRMF), right thalamus proper (RTP), right posterior cingulate (RPC), right pallidum (RP), and right caudal anterior cingulate (RCAC). The dlPFC area has been determined as the critical neural substrate for MDD. This area controls cognitive and complex mental processes, including emotion modulation, selective attention, and working memory (**Hamilton et al., 2012**). Some literature indicates that healthy subjects showed bilateral activation during a working memory task. In contrast, depressed patients exhibited asymmetric activity, with the left dlPFC showing an increased activation (**Perrin et al., 2012**). In 2008, the US Food and Drug Administration (FDA) approved a treatment for MDD. This treatment (FDA approval K061053) utilizes the repetitive transcranial magnetic stimulation (rTMS) system, which delivers transcranial repetitive pulsed magnetic fields of sufficient magnitude that induces neural action potentials in the left dlPFC area. Additionally, randomized clinical trials for depressed subjects established the clinical effectiveness of rTMS on the left dlPFC (**Blumberger et al., 2018**).

Our analysis shows a strong evidence of the laterization of right dlPFC with significant increased FCs; i.e., a specific and distinctive FC activation path via the right dlPFC was found in LLD patients. The right dlPFC with an increased FC also served as a critical hub involved in cognition impairment and LLD pain. **Ihara et al. (2019)** found altered FCs in the right dlPFC of patients with chronic neck pain at high risk for depression. Chronic pain is one of the most common comorbid conditions among patients with LLD (**Aziz and Steffens, 2013**). Thus, the right dlPFC with an increased FC is the crucial region for cognition impairment and pain for LLD.

We present the remaining results of our study in **Table 6**, where detected connectivity is given in Column 2, and the corresponding network(s) to which it belongs to is stated in Column 1. The functional activity of detected links and corresponding references are presented in Column 3.

Thus, our findings match most of those in the literature. Our results of increased FC in the bilateral PCC left IPL, and left supramarginal (LS) within the DMN of LLD patients, together with those in the literature, suggest that interventions targeting to decrease FCs within the DMN may have beneficial clinical effects for LLD patients. **Menon (2015)** suggested that the caudal anterior cingulate (CAC) and anterior insula are two major cortical structures of the SN. Moreover, electroconvulsive therapy can help reduce FCs between the right anterior cingulate cortex and the right dlPFC in patients with severe depression (**Perrin et al., 2012**). Our results consistently show increased FC between the right dACC within the SN, the right dlPFC within the CEN, and the right PCC within the DMN, across the major large-scale neurocognitive brain networks, including the DMN, CEN, and SN of LLD patients. These critical brain networks may contribute to a deterioration of memory and cognitive functions in elderly patients with LLD.

8. Conclusions and discussions

Based on the maximum number of disrupted links and their cognitive importance, we would like to declare the right caudal middle frontal (RCMF), also known as dlPFC, as a possible

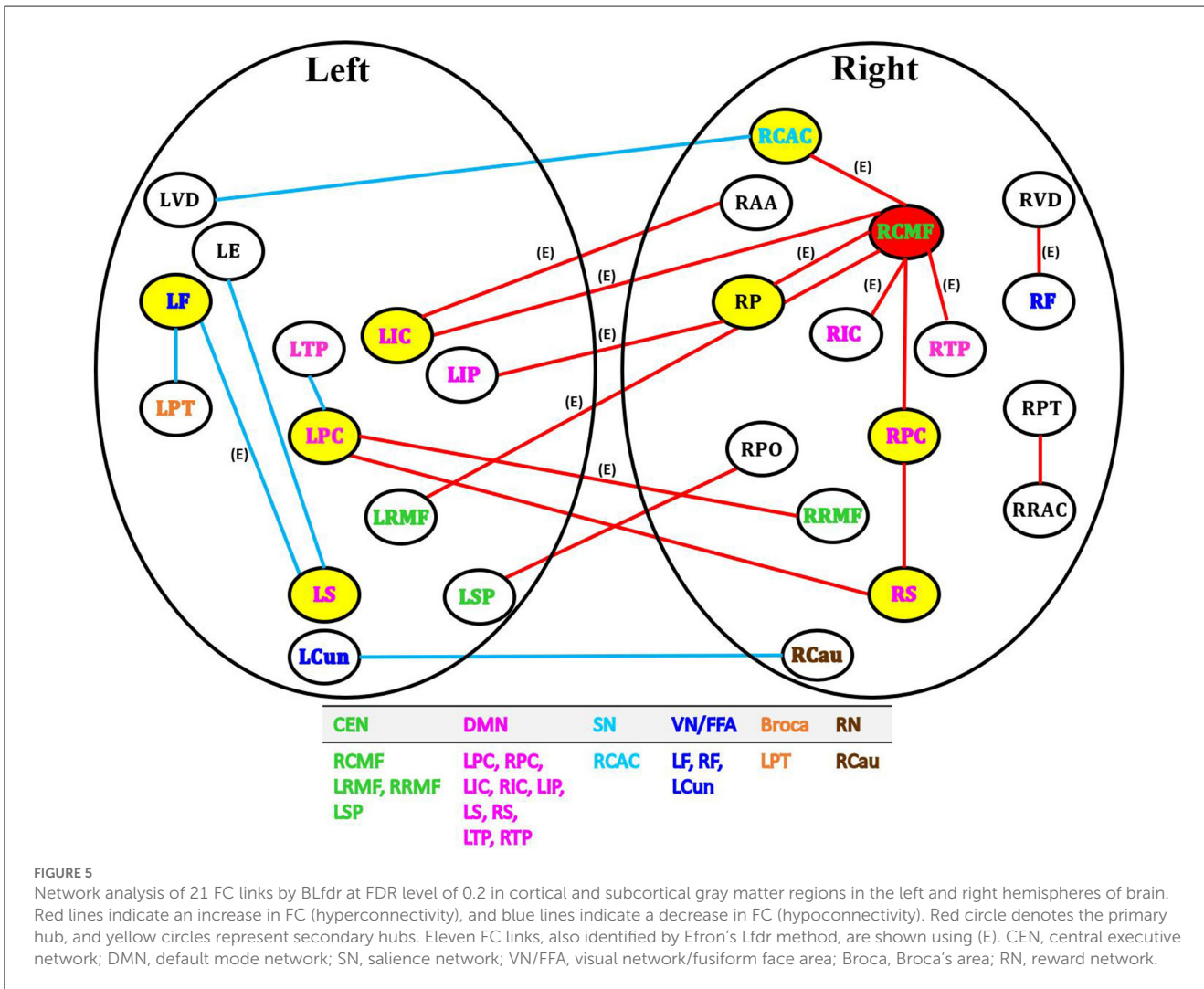
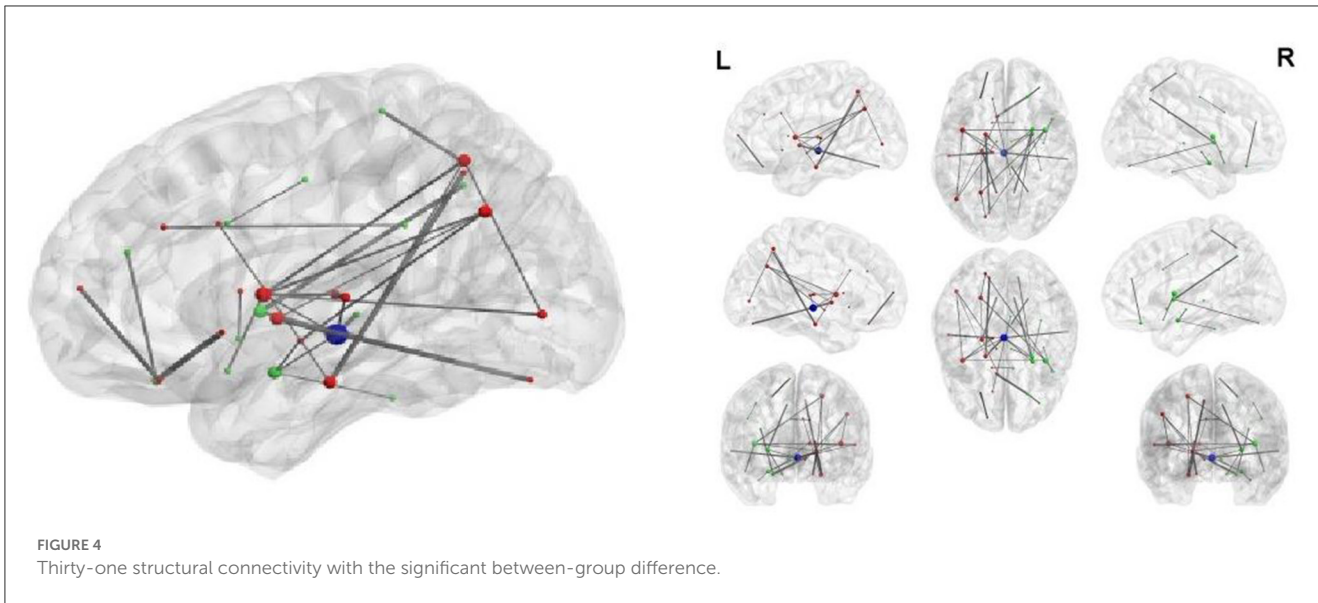


TABLE 6 FC links within CEN, SN, RN, DMN, Broca, and VN/FFA identified using BLfdr for LLD study.

| Network | [+/-] ROIs ¹ | Function of the ROIs/Network |
|--|---|--|
| Central executive network (CEN) | RCMF (CEN) vs. LIC/RIC (DMN) | (1) CEN controls cognitive and mental processes, including emotion modulation, selective attention, and working memory. (2) RCMF is also called dlPFC, which served as the key neural substrates for MDD (Hamilton et al., 2012). (3) Depressed patients exhibit asymmetric activity in dlPFC (Perrin et al., 2012). The left dlPFC shows increased activation. The right dlPFC is a critical hub of the altered FCs in chronic net pain patients at high risk of experiencing depression (Aziz and Steffens, 2013; Ihara et al., 2019). (4) The left dlPFC has been the target site of repetitive transcranial magnetic stimulation (rTMS) as a treatment for medication-resistant MDD (refer to FDA approval K061053; Blumberger et al., 2018). |
| | RCMF (CEN) vs. LRMF/RRMF (DMN) | (1) RRMF is a converging site of the dorsal and ventral attention networks by playing a role in the reorientation of attention. |
| | LSP (CEN) vs. RPO (IFG) | (1) RPO belongs to the right inferior frontal gyrus (IFG), which is associated with cognitive functions, including speech, attention, motor inhibition, imagery. (2) Hyperconnectivity of RPO was observed in adolescents with MDD (Tang et al., 2018). |
| Saliency network (SN) | RCAC (SN) vs. RCMF (dlPFC) | (1) SN works in detecting, filtering, and integrating salient external stimuli with internal states to orchestrate brain network dynamics in the service of goal-directed behaviors and motivated behaviors. (2) SN plays a mediating role in switching between activation and deactivation of internally directed cognition of DMN and externally directed cognition of CEN. (3) SN is associated with MDD/LLD (Yuen et al., 2014). (4) The caudal anterior cingulate (CAC, aka. dACC) and anterior insula are two major cortical structures of the SN. |
| Reward network (RN) | [-] RCau (RN) vs. LCun (VN/FEA) | (1) RN includes reinforcement learning, action monitoring, novelty processing, learning, decision making, and addiction. (2) RCau is a key region with enhanced neural response to positive emotions in MDD (Keren et al., 2018). |
| Default mode network (DMN) | RPC (DMN) vs. RCMF (CEN) and [+] LPC (DMN) vs. RRMF (CEN) | (1) DMN, the default mode of brain functions for internally directed self-referential cognition activated at rest and deactivated during cognitive and mental tasks requiring attention and response. (2) DMN is associated with MDD (Kaiser et al., 2015). (3) The posterior cingulate cortex (PCC)—is considered the posterior hub of the DMN, controlling emotion, cognition, awareness, arousal, and regulatory modulation. PCC exhibits higher activities when the brain is at rest. (4) PCC is associated with MDD (Zhou et al., 2010; Zhao et al., 2019) and LLD (Alexopoulos et al., 2012). |
| | [-] LS (DMN) vs. LF (VN/FFA) and [+] RS (DMN) vs. LPC/RPC (DMN) | (1) Supramarginal gyrus (LS/RS) is part of the inferior parietal lobe that functions in sensory-motor and cognitive domains. (2) Inferior parietal lobe is associated with depression symptoms (Roh et al., 2020). |
| | [-] LTP (DMN) vs. LPC (DMN) and [+] RTP (DMN) vs. RCMF (CEN) | (1) The bilateral thalamus proper (LTP/RTP), also called Dorsal Thalamus (DT). (2) DT is a hub for relaying sensory and motor signals from the senses to the cerebral cortex and receiving feedback from the cortex. DT plays a mediator role in-between cortico-cortical communication processing (Sherman and Guillery, 2002). (3) DT is associated with MDD (Greicius et al., 2007). |
| Broca's area (Broca) | [-] LPT (Broca) vs. LF (VN/FFA) | (1) Broca is essential for the speech process by integrating and coordinating information. (2) Pars Triangularis (PT) is considered a hub of the language-control network (Elmer, 2016). (3) LPT has been observed to activate semantic processing (Friederici et al., 2000). |
| Visual Network / Fusiform Face Area (VN/FFA) | RF (VN/FFA) vs. RVD | (1) The bilateral fusiform gyrus (FG) involves in visual cognition, especially facial cognition. (2) The antidepressants have improved the neural response to positive emotion in the LF and the right dlPFC (Ma, 2015). (3) RVD is associated with LLD (Lebedeva et al., 2017). |

¹[+] indicates increased FCs and [-] indicates decreased FCs.

biomarker. However, it needs larger studies to confirm this result. The long-term goal of this article is to treat medicine-resistant LLD patients who need alternative treatments. Exploring the association of RCMF with the central executive network's function that controls cognitive and complex mental processes, including emotion modulation and selective attention working memory, neuropsychotherapies (e.g., cognitive behavioral therapy) can be developed. Pinpointing the RCMF in the broad area of the central executive network is extremely helpful to use as a target for applying neurotherapeutic intervention (e.g., transcranial

magnetic stimulus, TMS) to get optimal stimulus results. The high accuracy rate of classification encourages us to use RCMF to predict prospective subjects at high risk for early prevention. The team is currently investigating the cognitive gain from applying neurotherapeutic intervention on RCMF, the detected biomarker. Our approach can be extended to traumatic brain injury studies to select the target for TMS or PTSD study for treatment selection. It is worth to extend our methodology for longitudinal neuroimaging studies when detected biomarker needed to be consistent over time.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by UIC Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

Author contributions

DB and YW conceived the novel model. DB, YW, and P-SY contributed equally to revising the initial draft. YW conducted the initial analyses and P-SY assisted with the classification performance evaluation. OA provided the LLD data and helped with the interpretation of neurophysiological discovery results. All authors contributed to the article and approved the submitted version.

References

- Alexopoulos, G. S. (2019). Mechanisms and treatment of late-life depression. *Transl. Psychiatry* 9, 1–16. doi: 10.1038/s41398-019-0514-6
- Alexopoulos, G. S., Hoptman, M. J., Kanellopoulos, D., Murphy, C. F., Lim, K. O., and Gunning, F. M. (2012). Functional connectivity in the cognitive control network and the default mode network in late-life depression. *J. Affect. Disord.* 139, 56–65. doi: 10.1016/j.jad.2011.12.002
- Aziz, R., and Steffens, D. C. (2013). What are the causes of late-life depression? *Psychiatr. Clin. N. Am.* 36, 497–516. doi: 10.1016/j.psc.2013.08.001
- Bartlett, R. (1993). Linear modelling of Pearson's product moment correlation coefficient: an application of fisher's z-transformation. *J. R. Stat. Soc. Ser. D* 42, 45–53. doi: 10.2307/2348110
- Benjamini, Y., and Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Educ. Behav. Stat.* 25, 60–83. doi: 10.3102/10769986025001060
- Bhaumik, D. K., Jie, F., Nordgren, R., Bhaumik, R., and Sinha, B. K. (2018a). A mixed-effects model for detecting disrupted connectivities in heterogeneous data. *IEEE Trans. Med. Imaging* 37, 2381–2389. doi: 10.1109/TMI.2018.2821655
- Bhaumik, D. K., Song, Y., Dan, Z., and Ajilore, O. A. (2018b). Controlling false disrupted connectivities in neuroimaging studies. *J. Biostat. Biometr.* 1, 1–10. doi: 10.29011/JBSB-108.100008
- Biswal, B., Zerrin Yetkin, F., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409
- Blumberger, D. M., Vila-Rodriguez, F., Thorpe, K. E., Feffer, K., Noda, Y., Giacobbe, P., et al. (2018). Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (three-d): a randomised non-inferiority trial. *Lancet* 391, 1683–1692. doi: 10.1016/S0140-6736(18)30295-2
- Chiang, S. (2016). *Hierarchical Bayesian models for multimodal neuroimaging data* (Ph.D. thesis). Rice University, Houston, TX, United States.
- Dai, Z., Yan, C., Li, K., Wang, Z., Wang, J., Cao, M., et al. (2015). Identifying and mapping connectivity patterns of brain network hubs in Alzheimer's disease. *Cereb. Cortex* 25, 3723–3742. doi: 10.1093/cercor/bhu246
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021
- Dunlop, B. W., Binder, E. B., Cubells, J. F., Goodman, M. M., Kelley, M. E., Kinkead, B., et al. (2012). Predictors of remission in depression to individual and combined treatments (predict): study protocol for a randomized controlled trial. *Trials* 13, 1–18. doi: 10.1186/1745-6215-13-106
- Efron, B. (2004). Large-scale simultaneous hypothesis testing: the choice of a null hypothesis. *J. Am. Stat. Assoc.* 99, 96–104. doi: 10.1198/016214504000000089
- Efron, B. (2007). Size, power and false discovery rates. *Ann. Stat.* 35, 1351–1377. doi: 10.1214/009053606000001460
- Efron, B. (2008). Microarrays, empirical Bayes and the two-groups model. *Stat. Sci.* 23, 1–22. doi: 10.1214/07-STS236
- Elmer, S. (2016). Broca pars triangularis constitutes a “hub” of the language-control network during simultaneous language translation. *Front. Hum. Neurosci.* 10:491. doi: 10.3389/fnhum.2016.00491
- Friederici, A. D., Opitz, B., and Von Cramon, D. Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: an fMRI investigation of different word types. *Cereb. Cortex* 10, 698–705. doi: 10.1093/cercor/10.7.698
- Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. (1993). Functional connectivity: the principal-component analysis of large (pet) data sets. *J. Cereb. Blood Flow Metab.* 13, 5–14. doi: 10.1038/jcbfm.1993.4
- Gelman, A., Jakulin, A., Pittau, M. G., and Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Statist.* 2, 1360–1383. doi: 10.1214/08-AOAS191
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). doi: 10.1214/06-BA117A
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. Boca Raton, FL: CRC Press. doi: 10.1201/b16018
- Ghosal, N. (2019). *Clustered-temporal Bayesian models for brain connectivity in neuroimaging data* (Ph.D. thesis). University of Illinois at Chicago, Chicago, IL, United States.
- Grecius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., et al. (2007). Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol. Psychiatry* 62, 429–437. doi: 10.1016/j.biopsych.2006.09.020
- Hamilton, J. P., Etkin, A., Furman, D. J., Lemus, M. G., Johnson, R. F., and Gotlib, I. H. (2012). Functional neuroimaging of major depressive disorder: a meta-analysis and

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnimg.2023.1147508/full#supplementary-material>

- new integration of baseline activation and neural response data. *Am. J. Psychiatry* 169, 693–703. doi: 10.1176/appi.ajp.2012.11071105
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.-P., Meuli, R., et al. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2035–2040. doi: 10.1073/pnas.0811168106
- Ihara, N., Wakaizumi, K., Nishimura, D., Kato, J., Yamada, T., Suzuki, T., et al. (2019). Aberrant resting-state functional connectivity of the dorsolateral prefrontal cortex to the anterior insula and its association with fear avoidance belief in chronic neck pain patients. *PLoS ONE* 14:e0221023. doi: 10.1371/journal.pone.0221023
- Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., and Pizzagalli, D. A. (2015). Large-scale network dysfunction in major depressive disorder: a meta-analysis of resting-state functional connectivity. *JAMA Psychiatry* 72, 603–611. doi: 10.1001/jamapsychiatry.2015.0071
- Keren, H., O'Callaghan, G., Vidal-Ribas, P., Buzzell, G. A., Brotman, M. A., Leibenluft, E., et al. (2018). Reward processing in depression: a conceptual and meta-analytic review across fMRI and EEG studies. *Am. J. Psychiatry* 175, 1111–1120. doi: 10.1176/appi.ajp.2018.17101124
- Kim, S., Dahl, D. B., and Vannucci, M. (2009). Spiked Dirichlet process prior for Bayesian multiple hypothesis testing in random effects models. *Bayesian Anal.* 4:707. doi: 10.1214/09-BA426
- Lebedeva, A. K., Westman, E., Borza, T., Beyer, M. K., Engedal, K., Aarsland, D., et al. (2017). MRI-based classification models in prediction of mild cognitive impairment and dementia in late-life depression. *Front. Aging Neurosci.* 9:13. doi: 10.3389/fnagi.2017.00013
- Liu, J. S., Liang, F., and Wong, W. H. (2000). The multiple-try method and local optimization in metropolis sampling. *J. Am. Stat. Assoc.* 95, 121–134. doi: 10.1080/01621459.2000.10473908
- Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The bugs project: evolution, critique and future directions. *Stat. Med.* 28, 3049–3067. doi: 10.1002/sim.3680
- Ma, Y. (2015). Neuropsychological mechanism underlying antidepressant effect: a systematic meta-analysis. *Mol. Psychiatry* 20, 311–319. doi: 10.1038/mp.2014.24
- Menon, V. (2015). "Large-scale functional brain organization," in *Brain Mapping: An Encyclopedic Reference*, Vol. 2, ed A. W. Toga (San Diego, CA: Academic Press; Elsevier), 449–459.
- Perrin, J. S., Merz, S., Bennett, D. M., Currie, J., Steele, D. J., Reid, I. C., et al. (2012). Electroconvulsive therapy reduces frontal cortical connectivity in severe depressive disorder. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5464–5468. doi: 10.1073/pnas.1117206109
- Ramachandram, D., and Taylor, G. W. (2017). Deep multimodal learning: a survey on recent advances and trends. *IEEE Signal Process. Mag.* 34, 96–108. doi: 10.1109/MSP.2017.2738401
- Robert, C. P., Casella, G., and Casella, G. (1999). *Monte Carlo Statistical Methods*, Vol. 2. New York, NY: Springer. doi: 10.1007/978-1-4757-3071-5
- Roh, H. W., Hong, C. H., Lim, H. K., Chang, K. J., Kim, H., Kim, N.-R., et al. (2020). A 12-week multidomain intervention for late-life depression: a community-based randomized controlled trial. *J. Affect. Disord.* 263, 437–444. doi: 10.1016/j.jad.2019.12.013
- Rykhlevskaia, E., Gratton, G., and Fabiani, M. (2008). Combining structural and functional neuroimaging data for studying brain connectivity: a review. *Psychophysiology* 45, 173–187. doi: 10.1111/j.1469-8986.2007.00621.x
- Sherman, S. M., and Guillery, R. (2002). The role of the thalamus in the flow of information to the cortex. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 357, 1695–1708. doi: 10.1098/rstb.2002.1161
- Song, Y. (2016). *Sample size determination for high-dimensional neuroimaging studies controlling false discovery rate* (Ph.D. thesis). University of Illinois at Chicago, Chicago, IL, United States.
- Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9440–9445. doi: 10.1073/pnas.1530509100
- Sui, J., He, H., Yu, Q., Chen, J., Rogers, J., Pearlson, G. D., et al. (2013). Combination of resting state fmri, DTI, and SMRI data to discriminate schizophrenia by n-way MCCA+ JICA. *Front. Hum. Neurosci.* 7:235. doi: 10.3389/fnhum.2013.00235
- Sun, W., and Cai, T. T. (2007). Oracle and adaptive compound decision rules for false discovery rate control. *J. Am. Stat. Assoc.* 102, 901–912. doi: 10.1198/01621450700000545
- Szucs, D., and Ioannidis, J. P. (2019). Sample size evolution in neuroimaging research: an evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals. *bioRxiv*. 111674. doi: 10.1101/809715
- Tadayonnejad, R., Yang, S., Kumar, A., and Ajilore, O. (2014). Multimodal brain connectivity analysis in unmedicated late-life depression. *PLoS ONE* 9:e96033. doi: 10.1371/journal.pone.0096033
- Tang, S., Lu, L., Zhang, L., Hu, X., Bu, X., Li, H., et al. (2018). Abnormal amygdala resting-state functional connectivity in adults and adolescents with major depressive disorder: a comparative meta-analysis. *EBioMedicine* 36, 436–445. doi: 10.1016/j.ebiom.2018.09.010
- Uddin, L. Q., Kelly, A., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2009). Functional connectivity of default mode network components: correlation, anticorrelation, and causality. *Hum. Brain Mapp.* 30, 625–637. doi: 10.1002/hbm.20531
- Whitfield-Gabrieli, S., and Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect.* 2, 125–141. doi: 10.1089/brain.2012.0073
- WHO (2021). *Depression*. World Health Organization. Available online at: <https://www.who.int/news-room/fact-sheets/detail/depression>
- Xue, W., Bowman, F. D., Pileggi, A. V., and Mayer, A. R. (2015). A multimodal approach for determining brain networks by jointly modeling functional and structural connectivity. *Front. Comput. Neurosci.* 9:22. doi: 10.3389/fncom.2015.00022
- Yuen, G. S., Gunning-Dixon, F. M., Hoptman, M. J., AbdelMalak, B., McGovern, A. R., Seirup, J. K., et al. (2014). The salience network in the apathy of late-life depression. *Int. J. Geriatr. Psychiatry* 29, 1116–1124. doi: 10.1002/gps.4171
- Zablocki, R. W., Schork, A. J., Levine, R. A., Andreassen, O. A., Dale, A. M., and Thompson, W. K. (2014). Covariate-modulated local false discovery rate for genome-wide association studies. *Bioinformatics* 30, 2098–2104. doi: 10.1093/bioinformatics/btu145
- Zhang, M. J., Xia, F., and Zou, J. (2019). Fast and covariate-adaptive method amplifies detection power in large-scale multiple hypothesis testing. *Nat. Commun.* 10:3433. doi: 10.1038/s41467-019-11247-0
- Zhao, Q., Swati, Z. N., Metmer, H., Sang, X., and Lu, J. (2019). Investigating executive control network and default mode network dysfunction in major depressive disorder. *Neurosci. Lett.* 701, 154–161. doi: 10.1016/j.neulet.2019.02.045
- Zhao, W. (2014). *Statistical methodologies for group comparisons of brain connectivity using multimodal neuroimaging data* (Ph.D. thesis). University of Illinois at Chicago, Chicago, IL, United States.
- Zhou, Y., Yu, C., Zheng, H., Liu, Y., Song, M., Qin, W., et al. (2010). Increased neural resources recruitment in the intrinsic organization in major depression. *J. Affect. Disord.* 121, 220–230. doi: 10.1016/j.jad.2009.05.029