

UC Irvine

UC Irvine Previously Published Works

Title

A Bayesian Double Fusion Model for Resting-State Brain Connectivity Using Joint Functional and Structural Data

Permalink

<https://escholarship.org/uc/item/9c63p29r>

Journal

Brain Connectivity, 7(4)

ISSN

2158-0014

Authors

Kang, Hakmook
Ombao, Hernando
Fonnesbeck, Christopher
[et al.](#)

Publication Date

2017-05-01

DOI

10.1089/brain.2016.0447

Peer reviewed

A Bayesian Double Fusion Model for Resting-State Brain Connectivity Using Joint Functional and Structural Data

Hakmook Kang,^{1,2} Hernando Ombao,^{3,4} Christopher Fonnesebeck,^{1,2}
Zhaohua Ding,^{5,6} and Victoria L. Morgan^{5,6}

Abstract

Current approaches separately analyze concurrently acquired diffusion tensor imaging (DTI) and functional magnetic resonance imaging (fMRI) data. The primary limitation of these approaches is that they do not take advantage of the information from DTI that could potentially enhance estimation of resting-state functional connectivity (FC) between brain regions. To overcome this limitation, we develop a Bayesian hierarchical spatio-temporal model that incorporates structural connectivity (SC) into estimating FC. In our proposed approach, SC based on DTI data is used to construct an informative prior for FC based on resting-state fMRI data through the Cholesky decomposition. Simulation studies showed that incorporating the two data produced significantly reduced mean squared errors compared to the standard approach of separately analyzing the two data from different modalities. We applied our model to analyze the resting state DTI and fMRI data collected to estimate FC between the brain regions that were hypothetically important in the origination and spread of temporal lobe epilepsy seizures. Our analysis concludes that the proposed model achieves smaller false positive rates and is much robust to data decimation compared to the conventional approach.

Keywords: diffusion tensor image; functional connectivity; functional magnetic resonance imaging; space-time structure; structural connectivity

Introduction

NEUROIMAGING TECHNIQUES CONTINUE to be developed for investigating the underlying brain functions related to human cognition, emotions, and behaviors. Applications of functional magnetic resonance imaging (fMRI) have shed light on brain research, including studies of the functional and effective connectivity between two regions of the brain providing insights into understanding the functional networks in the brain. Functional connectivity (FC) has been used as a biomarker for neurological and psychiatric disorders, for example, Alzheimer's disease (Joo et al., 2016; Wang et al., 2007; Zippo et al., 2015) and bipolar disorder (Altinay et al., 2015; Dickstein et al., 2010).

To study brain functional and effective connectivity, it is important to understand the extent to which this is facilitated by structural connectivity (SC) that can be estimated by multiple-tensor model or by probabilistic diffusion tractogra-

phy (Behrens et al., 2007) using diffusion tensor imaging (DTI) data. It has been demonstrated in van den Heuvel et al. (2009) the existence of anatomical white matter tracts, SC, and interconnect regions associated with resting-state networks.

A significant amount of work has been done to develop statistical models and procedures to analyze fMRI data (Bowman, 2007; Poldrack et al., 2011; Worsley et al., 1996), fMRI data with a Bayesian approach (Yu et al., 2016; Zhang et al., 2014, 2015), or DTI data (Basser et al., 1994; Behrens et al., 2003, 2007), separately. Recently, Harrison et al. (2015) used an independent component analysis approach to handle spatiotemporal modes in large-scale resting-state fMRI (rs-fMRI), combined with Bayesian approximation.

Although there has not been much effort to develop statistical models that use these two types of data jointly, Olesen et al. (2003) proposed a combined analysis of fMRI and DTI to explore the association between maturation of white

¹Department of Biostatistics, Vanderbilt University, Nashville, Tennessee.

²Center for Quantitative Sciences, Vanderbilt University, Nashville, Tennessee.

³Applied Mathematics and Computational Science, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

⁴Department of Statistics, University of California, Irvine, California.

⁵Institute of Imaging Science, Vanderbilt University, Nashville, Tennessee.

⁶Department of Radiology and Radiological Science, Vanderbilt University, Nashville, Tennessee.

matter and changes in brain activity during childhood. They used structural information as a covariate in a typical multiple regression model for fMRI data analysis. There are other studies that address the advantages of using both fMRI and DTI (Bowman et al., 2012; Greicius et al., 2009; Hending et al., 2003; Honey et al., 2009; Morgan et al., 2009; Werring et al., 1998; Wieshmann et al., 2001; Xue et al., 2015).

Recently, Schmittmann et al. (2015) proposed to distinguish indirect FC from direct FC. On the same note, direct SC can be separated from indirect SC where we define direct SC as structural connection between two regions by white matter tracts, whereas indirect SC as any structural connection beside direct SC between two regions. To capture dynamics in FC, Cribben et al. (2012) proposed dynamic connectivity regression model. In 2014, Deco et al. proposed a method in which SC estimation was iteratively enhanced using empirical FC at bifurcation and Hansen et al. (2015) constructed FC dynamics while SC was fused into the generative model. These two approaches were designed to capture nonstationary nature of FC.

To assess the relationship between FC and SC, Huang and Ding (2016) investigated association between conditional Granger Causality, as a measure of FC, and edge weight as a measure of SC proposed by Hagmann et al. (2008). The mutual relationship between FC and SC was confirmed by Meier et al. (2016), although the relationship was not one to one. More methods and applications of fusing DTI and fMRI can be found in Zhu et al. (2014). However, using the SC acquired by analyzing DTI data to guide the estimates of FC has not been extensively explored.

To properly take into account underlying spatiotemporal correlation in rs-fMRI data while improving precision and accuracy of FC estimation using multimodal MRI data, that is, rs-fMRI and DTI, we developed a novel Bayesian hierarchical modeling framework. Before describing the model in detail, let's define a term "naive FC": The naive FC is defined as the correlation between two averaged time series across voxels within each region of interest without taking into account temporal correlation, which has been also widely used in practice.

Our model takes into account the intrinsic spatial and temporal correlation in rs-fMRI data while using the weighted average of SC from DTI data and the naive FC as a source of prior information for FC. Moreover, we fused two distinct contributions of SC to FC estimation, that is, the effect of indirect SC on FC estimation was treated differently from that of direct SC using a mixture model. Given predefined regions of interests (ROIs), our approach naturally enables us to estimate FC by incorporating both functional and structural data, that is, rs-fMRI and DTI data. As a result, our method is expected to improve estimation accuracy and lead to more reliable inference.

To validate our novel method, we conducted several numerical experiments where we generated data with realistic spatial and temporal correlation at each voxel within a ROI and imposed correlation (i.e., structural and FC) between a pair of ROIs, to appropriately represent the behavior of rs-fMRI data. Then we analyzed the data by applying our Bayesian hierarchical spatiotemporal model. In addition, we applied our model to a data set consisting of rs-fMRI time series and DTI acquired at voxels from predetermined ROIs that were hypothetically important in the origination

and spread of temporal lobe epilepsy seizures (Holmes et al., 2013): left hippocampus (HL), right hippocampus (HR), left thalamus (TL), right thalamus (TR), precuneus (PC), left insula (IL), right insula (IR), and cingulate gyrus (CG). This data set provides us with an ideal opportunity to study the relationships between SC and FC and how SC affects FC estimation through our double-fusion process.

With this data set, we assess which approach is more robust to data decimation: assessing change in inference results over different sample sizes enables us to compare our approach with the most common approach without knowing the ground truth. With the most common approach (denoted as AVG-FC) to estimate FC between a pair of ROIs, all the time series within an ROI are summarized into one time course by simply taking the average of the time series across voxels within that ROI. Then after taking into account temporal correlation in each time series, we compute the correlation between each pair of time series to estimate ROI-level FC.

Materials and Methods

Spatiotemporal hierarchical model

Let C be the number of ROIs and V_c be the number of voxels within the c -th ROI. Denote the time series at voxel v in ROI c to be $Y_{cv}(t)$, $t=1, \dots, T$, which would be a typical structure of observed data in an rs-fMRI study. We define a kernel function $\psi_c(\cdot)$ that generates a valid covariance matrix for local spatial (within ROI) covariance. We assume that $\psi_c(\cdot)$ is a function of Euclidean distance. Using the model described below, we can take into account the distant-dependent spatial correlation between voxels within the same ROI and the temporal correlation within a voxel.

Consider the following spatiotemporal model for rs-fMRI time series:

$$Y_{cv}(t) = \beta_c + b_c(v) + d_c + \epsilon_{cv}(t) \quad (1)$$

- β_c is the grand mean in ROI c ;
- $b_c(v)$ is a zero-mean voxel-specific random effect within an ROI c ; this component captures the local spatial dependency between voxels within the ROI. In this study, we specify the covariance structure to be

$$\text{Cov}(b_c(v), b_{c'}(v')) = \begin{cases} \psi_c(\|v - v'\|), & \text{when } c = c' \\ 0, & \text{when } c \neq c'. \end{cases} \quad (2)$$

Note that when two voxels belong to different ROIs then their corresponding b values are uncorrelated as illustrated in the second condition in equation (2). The function $\psi_c(\cdot)$ can be any valid spatial covariance function, for example, exponential, Gaussian, or Matérn family (Cressie, 1993) in which the correlation between two voxels is inversely proportional to the distance between them. In this article, the exponential function is used as illustrated in equation (3), where

$$\text{Cov}(b_c(v), b_{c'}(v')) = \sigma_c^2 \exp(-\|v - v'\| / \phi_c), \quad (3)$$

and σ_c^2 is defined as the spatial variance at each voxel in ROI c , $\|v - v'\|$ denotes the distance between two voxels, v and v' , and ϕ_c is ROI-specific decaying parameter;

- d_c is a zero-mean ROI-specific random effect with a covariance structure $\text{Cov}(d_c, d_{c'})$ that is used to model FC

between ROIs. Both the SC resulting from DTI and naive FC play a role of prior information when estimating the covariance matrix of $\mathbf{d}=[d_1, d_2, \dots, d_C]^T$; and

- $\epsilon_{cv}(t)$ is the noise that takes into account voxel-specific temporal correlation that is assumed to follow an autoregressive temporal process with order one, that is, AR(1). That is, $\epsilon_{cv}(t) = \phi \epsilon_{cv}(t-1) + w(t)$, where ϕ is the AR(1) coefficient and $w(t)$ denotes Gaussian random noise.

Hierarchical structure

Our main goal is to determine if using SC through a double fusion process improves the accuracy and precision in functional and effective connectivity estimation and enhances power to detect statistically meaningful FC, that is, *important* FC is defined as the probability of FC being >0.4 is larger than 0.5. As no standard threshold has been established, we choose $\Pr(\text{FC} > 0.4) > 0.5$ as *important*, which has been used to reflect moderate functional coherence by Xue et al. (2015). Note that a posterior distribution of each FC can be used as a descriptive tool to characterize the FC, for example, median and interquartile range.

Let's define $\mathbf{Y}_c(t) = [\mathbf{Y}_{c1}(t), \dots, \mathbf{Y}_{cV_c}(t)]^T$. Let $\mathbf{1}_m$ and $\mathbf{I}_{m'}$ denote an all-ones vector with length m and an $m' \times m'$ identity matrix, respectively. Then the model (1) can be rewritten as an ROI-level model,

$$\mathbf{Y}_c(t) = \beta_c + \mathbf{b}_c + \mathbf{d}_c + \epsilon_c(t) \quad (4)$$

where $\beta_c = \beta_c \mathbf{1}_{(V_c)}$, $\mathbf{b}_c = [b_c(1), \dots, b_c(V_c)]^T$, $\mathbf{d}_c = d_c \mathbf{1}_{(V_c)}$, and $\epsilon_c(t) = [\epsilon_{c1}(t), \dots, \epsilon_{cV_c}(t)]^T$. Then this model has the following hierarchical structure:

$$\begin{aligned} \beta_c &\sim N(0, \sigma_{\beta_c}^2) \\ \mathbf{b}_c &\sim N(0, \Sigma_{b_c}) \\ \mathbf{d} &\sim N(0, \Sigma_d) \\ \epsilon_{cv}(t) &\sim N(0, \sigma_{\epsilon_{cv}}^2 / (1 - \phi_{cv}^2)) \end{aligned} \quad (5)$$

where some of terms can be further explained in detail as follows. Each term β_c follows a Gaussian distribution with mean zero and variance $\sigma_{\beta_c}^2$ and β_c is independent of $\beta_{c'}$ where $c \neq c'$. The term $\mathbf{d}=[d_1, \dots, d_C]^T$ is also assumed to follow a Gaussian distribution $N(0, \Sigma_d)$, where the correlation matrix of \mathbf{d} denotes the FC among ROIs. For temporal correlation, we assume that a time series at each voxel v follows an autoregressive model of order one (AR(1)), that is, the covariance of ϵ_{cv} shows an AR(1) structure with AR(1) parameter ϕ_{cv} .

Note that model (4) can take into account (1) distance-dependent spatial correlation between voxels inside the same ROI using $\text{Cov}(\mathbf{b}_c)$ and (2) temporal correlation at each voxel.

Double fusion of structural and FC

We develop a rigorous framework that incorporates structural information in estimating FC. Our approach is a Bayesian hierarchical model (4) and allows us to capture the resting state FC (rs-FC) between brain regions, which is defined as the correlation of $\mathbf{d}=[d_1, \dots, d_C]^T$ in the model (Friston et al., 1993).

Our model captures the FC between ROIs while adjusting for the effect of *local* ROI-specific spatial correlation. Each entry of the correlation matrix of \mathbf{d} corresponds to the FC

of the appropriate pair of ROIs. The prior distribution of the correlation matrix is a function of the structural and naive FC of the ROIs in two distinct steps. The first fusion of SC and naive FC is performed given that the effect of direct SC on FC estimation is different from that of indirect SC on FC estimation. For example, lower values for SC suggest that there is no direct structural coupling between the two ROIs, but there may be indirect structural connection between the two ROIs, possibly leading to high functional coupling. If there is truly no structural connection, then the low SC would push the corresponding FC toward zero. However, indirect SC that cannot be measured in practice should be treated differently from direct SC when fused with naive FC.

Consider the prior distribution of the covariance matrix $\text{Cov}(d_c, d_{c'})$. Let Σ_d denote the covariance matrix of \mathbf{d} that is assumed to be a function of structural and naive FC matrix. Also, let L_{sc} , L_{nfc} , and L_{Σ_d} denote a lower triangular matrix resulting from the Cholesky decomposition of structural covariance matrix, naive functional covariance matrix, and functional covariance matrix, respectively. To distinguish the effects of direct SC from indirect SC, let LD_{Σ_d} and LID_{Σ_d} denote L_{Σ_d} corresponding to the lower triangular matrix resulted from "direct" SC and "indirect" SC, respectively.

Then, we assume that

$$LD_{\Sigma_d}(i, j) = \lambda_k L_{sc}(i, j) + (1 - \lambda_k) L_{nfc}(i, j), \quad (6)$$

$$LID_{\Sigma_d}(i, j) = SC(i, j) \lambda_k L_{sc}(i, j) + (1 - SC(i, j) \lambda_k) L_{nfc}(i, j), \quad (7)$$

$$L_{\Sigma_d}(i, j) = \omega_k LD_{\Sigma_d}(i, j) + (1 - \omega_k) LID_{\Sigma_d}(i, j), \quad (8)$$

where λ_k and ω_k are mixture weights on $[0, 1]$ sampled from a noninformative beta distribution with both the shape and scale parameters being one; $SC(i, j)$ denotes SC measure between i th and j th ROI, where $k=1, \dots, K$ (=the number of nonzero elements in L_{Σ_d}), $i=1, \dots, C$, and $j=1, \dots, C$. In Equation (6), the contribution of direct SC to LD_{Σ_d} is regulated by λ_k that is determined by data, whereas that of indirect SC to LID_{Σ_d} is also regulated by λ_k and the corresponding SC as shown in Equation (7). Therefore, small SC (i, j) derives $LID_{\Sigma_d}(i, j)$ to be a function of only $L_{nfc}(i, j)$ where λ_k controls the effect of $L_{nfc}(i, j)$ on $LID_{\Sigma_d}(i, j)$. The second mixture between $LD_{\Sigma_d}(i, j)$ and $LID_{\Sigma_d}(i, j)$ is regulated by ω_k that is determined by data and results in $L_{\Sigma_d}(i, j)$. In this way, we can accommodate the contribution of both direct and indirect SC for FC estimation. Then, L_{Σ_d} , a hyperprior of Σ_d , can be said as the element-wise weighted average of L_{sc} and L_{nfc} while taking into account both direct and indirect contribution of SC.

Then, the covariance matrix Σ_d is reconstructed as $L_{\Sigma_d} \times L_{\Sigma_d}^T$, where $L_{\Sigma_d}^T$ denotes the transpose of L_{Σ_d} . Simple normalization of Σ_d results in the correlation matrix of \mathbf{d} , rs-FC. This approach can guarantee that the estimated covariance (or correlation) matrix is positive semidefinite, which is one of most important characteristics of a valid covariance (correlation) matrix.

Prior distributions

We used Markov Chain Monte Carlo (MCMC) methods for estimation using Metropolis–Hastings sampler implemented

in PyMC 2.3.4 (Patil et al., 2010). The standard deviation terms σ_c , σ_{b_c} , and σ_{ϵ_c} , as well as the decaying parameter of exponential covariance function in equation (3) ϕ_c , were assumed to follow a uniform distribution, which is a common noninformative prior for a variance parameter: $\sigma_c, \sigma_{\epsilon_c}, \sigma_{b_c} \sim \text{unif}(0, 100)$, and $\phi_c \sim \text{unif}(0, 20)$ where $\text{unif}(a, b)$ denotes a uniform distribution between a and b . For temporal correlation, an AR(1) parameter at each ROI was also assumed to follow a uniform distribution, $\text{unif}(0, 1)$, imposing positive temporal correlation over time. The standard deviation term σ_{d_c} in logarithmic scale was also assumed to follow a uniform distribution, $\text{unif}(0, 20)$, which was used to convert a correlation matrix to the corresponding covariance matrix and vice versa. Since we had no prior information regarding the values of our parameters, we used uninformative priors.

Simulation study: data generation

We generated time series with a length of $T=128$ scans using autoregressive model with order one (AR(1)) at five ROIs in which there are 100 voxels. We assumed positive temporal correlation with the AR(1) coefficient of 0.6. Then we imposed correlation between ROIs using a multivariate normal distribution with zero mean and the correlation matrix shown in matrix (9), assuming that the underlying SC was fixed at the matrix (9) but FC was randomly sampled from a Wishart distribution with mean covariance matrix based on matrix (9) and six degrees of freedom.

Spatial correlation within a ROI was added by applying spatial smoothing kernel with (1) FWHM=0 (no spatial correlation), (2) FWHM=3.53 (moderate spatial correlation), (3) FWHM=8.24 (strong spatial correlation), and (4) FWHM= ∞ (extreme spatial correlation where all the observations are the same within an ROI) across voxels at each time point. In this way, we could explore the effect of the degree of spatial correlation on mean squared error (MSE).

We also generated the data using a t -distribution with three degrees of freedom to investigate the robustness of Gaussian noise assumption at FWHM=3.53. Moreover, at FWHM=3.53, we examined the effect of violating the assumption of AR(1) process: we simulated the data using AR(2) process with AR(1) coefficient of 0.6 and AR(2) coefficient of 0.3 and then analyzed the simulated data using only AR(1) process.

$$\begin{pmatrix} 1 & 0.6 & 0 & 0.5 & 0 \\ & 1 & 0.2 & 0.1 & 0 \\ & & 1 & 0 & 0.1 \\ & & & 1 & 0.2 \\ & & & & 1 \end{pmatrix} \quad (9)$$

Simulation study: estimation

We analyzed the simulated data with two different priors for FC, in particular, an informative prior based on true SC and another prior based on structural independence assumption (i.e., an identity matrix as prior); we computed the MSEs of estimated rs-FC for each case. Moreover, we computed how often three different strength of FCs, that is, zero FC, low FC (0.1 or 0.2), and strong FC (0.5 or 0.6) were claimed *important* at different sample sizes, $n=10, 8, 5$, and 4. For comparison, we also used the conventional approach (AVG-FC). Although it is not necessary to infer FC by finding *important* pairs and

our approach can fully characterize FC in a descriptive way, we used the notion of *important* FC to make direct comparisons between our approach and AVG-FC.

Subjects

To examine the impact of the SC between brain regions associated with the origination of and spread of temporal lobe epilepsy seizures, on the rs-FC, we utilized our proposed hierarchical spatiotemporal model. A total of 7 healthy subjects (37.3 ± 16.5 years, 2 men, right-handed), without a history of neurological, psychiatric, or medical conditions, participated in this study. All subjects gave written informed consent.

Magnetic resonance imaging

All MRI was performed using a Philips Achieva 3T MRI scanner (Philips Healthcare, Inc., Best, Netherlands) using a 32-channel head coil. Informed consent was obtained before scanning as per institutional review board guidelines. The imaging protocol for each subject included a three-dimensional, T1-weighted high-resolution image series across the whole brain for intersubject normalization ($1 \times 1 \times 1 \text{ mm}^3$) and FreeSurfer segmentation, and an fMRI T2* weighted gradient echo, echo-planar image series at rest with eyes closed—matrix 80×80 , FOV=240 mm, 34 axial slices, TE=35 ms, TR=2 sec, slice thickness=3.5 mm with 0.5 mm gap, 300 volumes. Physiological monitoring of cardiac and respiratory fluctuations was performed at 500 Hz using the MRI scanner integrated pulse oximeter and the respiratory belt.

Diffusion weighted images were acquired to quantify white matter integrity parameters using a single shot, spin-echo, echo-planar sequence— $b=1600 \text{ s/mm}^2$, 92 diffusion sensitizing directions, TR=8.5 s, TE=65 ms, matrix= 96×96 , 50 slices, $2.5 \times 2.5 \times 2.5 \text{ mm}^3$, no gap, 3 averages.

Imaging processing

The fMRI data were corrected for slice timing effects and motion occurring between the fMRI scans using SPM8 software (www.fil.ion.ucl.ac.uk/spm/software/spm8). The x , y , and z translations and rotations of each volume acquisition were saved as potential confounds in the FC computation. The images were then corrected for physiological noise using a RETROICOR protocol (Glover et al., 2000) using the measured cardiac and respiratory time series. The corrected fMRI images were spatially normalized to the Montreal Neurological Institute template. The normalized fMRI time series were then low pass filtered at a cutoff frequency of 0.1 Hz (Cordes et al., 2001).

Probabilistic tractography

The probabilistic tractography was created by analyzing the DTI using the algorithm proposed by Behrens et al. (2003) implemented in FSL software (Behrens et al., 2007). More details about the algorithm are well described in Behrens et al. (2003). From each voxel in the first ROI (ROI 1, seed ROI), we created 5000 streamlines which could travel in constrained directions due to the anisotropic characteristic of diffusion coefficient of water molecule. Therefore, some of the streamlines can reach to any voxels in the second ROI (ROI 2, target ROI). We count the number of voxels in ROI 1 that contain at least one streamline

TABLE 1. TOTAL MSE FOR EACH MODEL UNDER FOUR DIFFERENT STRENGTHS OF SPATIAL CORRELATION, FWHM=0, 3.53, 8.24, AND ∞ : CORRECTSC, INDEPENDENCE, AND AVG-GLM

<i>Spatial correlation</i>	<i>MSE by type of SC</i>	<i>Bayesian correct SC</i>	<i>Bayesian Independence</i>	<i>AVG-FC</i>
No correlation FWHM=0	MSE for low SC	0.551	0.455 ^a	1.100
	MSE for high SC	0.048 ^a	0.118	0.130
Moderate correlation FWHM=3.53	MSE for low SC	0.555	0.454 ^a	1.112
	MSE for high SC	0.044 ^a	0.106	0.123
Strong correlation FWHM=8.24	MSE for low SC	0.573	0.473 ^a	1.128
	MSE for high SC	0.054 ^a	0.116	0.138
Extreme correlation FWHM= ∞	MSE for low SC	0.533	0.453 ^a	1.087
	MSE for high SC	0.051 ^a	0.115	0.135

Total MSE is divided into MSE for low SC (i.e., 0, 0.1, and 0.2) and MSE for high SC (i.e., 0.5 and 0.6).

For each case, ^adenotes the smallest MSE among CorrectSC, Independence, and AVG-GLM.

FC, functional connectivity; MSE, mean squared error; SC, structural connectivity.

reached to ROI 2. The proportion of those voxels to the total number of voxels in ROI 1 is considered as the probability of the SC between the two ROIs. The same procedure is done by letting ROI 2 be a seed ROI. Then, finally the maximum of the two probabilities is claimed as the SC between ROI 1 and ROI 2.

Results

Simulation study: results

We evaluated each approach in terms of MSE of the FC based on 400 Monte Carlo simulations. The results are summarized in Table 1. In the Table 1, our proposed approach with informative prior (true structural correlation) and independence assumption (no structural correlation) is denoted as ‘‘CorrectSC’’ and ‘‘Independence,’’ respectively. The conventional approach based on averaged time series across voxels in ROIs is denoted as ‘‘AVG-FC.’’ Because, we only consider positive FC, all the estimated values corresponding to zero true connectivity are slightly overestimated, whereas the overestimation tapers out with bigger true connectivity values.

It is noteworthy that our approach, that is, CorrectSC and Independence, outperforms the conventional approach (AVG-FC) in terms of MSEs: the MSE value resulted from CorrectSC, Independence, and AVG-FC is 0.599, 0.560, and 1.235, respectively (see Table 1 when FWHM=3.53, sum of MSEs for high and low SC, e.g., 0.599 [total MSE]=0.555 [low SC]+0.044 [high SC] for CorrectSC). Although using the underlying true SC is expected to achieve the lowest MSE, Independence SC produces the lowest MSE because the true SC consists of many lower values, that is, 0, 0.1, and 0.2 where Independence SC is a quite informative prior. However, for higher values, that is, 0.5 and 0.6, Independence SC does not play a role of informative prior.

In Table 1, the MSE for each approach is reported along with MSE for low SC (i.e., 0, 0.1, and 0.2) and MSE for high SC (i.e., 0.5 and 0.6) under four different strengths of underlying spatial correlation. The MSE for high SC with CorrectSC approach (0.044) is the lowest while the MSE for low SC with Independence (0.454) outperforms that with CorrectSC (0.555) at FWHM=3.53. However, both CorrectSC and Independence always outperform AVG-FC. Our Bayesian spatiotemporal models with double fusion

reduce almost 50% of the MSE resulted from AVG-FC approach. Given that a true underlying SC based on probabilistic tractography is somewhere between unknown truth and complete disconnection (i.e., no structural connection between two regions), then we would expect to achieve significant reduction in MSE of FC by taking into account the underlying distant-dependent spatial correlation and using an extra information regarding SC.

The similar results were observed when the data were generated without spatial dependence (FWHM=0), with strong spatial correlation (FWHM=8.24), and with extreme spatial correlation (FWHM= ∞), when the Gaussian noise assumption was violated, and when the temporal correlation structure was misspecified.

Moreover, we assessed which model was most robust to data decimation. That is, 400 simulated data sets were divided into 40, 50, 80, and 100 group-level data sets in which there were 10, 8, 5, and 4 subjects per group, respectively. For each sample size, the rate of positive findings, that is, number of times rejecting the null (FC=0) divided by a total number of repetitions (=number of group-level data sets), was computed for each FC. For the Bayesian approach, the number of times rejecting the null is equivalent to the number of times claiming FC *important*.

In Figure 1, the rates of positive findings were summarized following three categories of FC: (1) zero FC (i.e., FC=0, blue) where positive findings were all erroneous, (2) low FC (i.e., FC=0.1 or 0.2, green) where positive findings may or may not be erroneous, and (3) strong FC (i.e., FC=0.5 or 0.6, red) where positive findings are desired. As shown in Figure 1, two Bayesian models denoted by solid line (CorrectSC) and dashed line (Independence) are quite robust against data decimation, compared to AVG-FC (dotted line) showing significant decrease in rate of positive findings as decrease in sample size. Moreover, note that high false positive rate is observed with AVG-FC when the underlying true FC=0, denoted by dotted blue line. When the underlying FC is strong, CorrectSC (solid red line) outperforms the other two approaches but Independence (dashed red) also seems to be very robust to data decimation. With very large sample size, it is expected that the performance of Independence converges to that of CorrectSC, while AVG-FC is more likely to claim all FCs statistically significant.

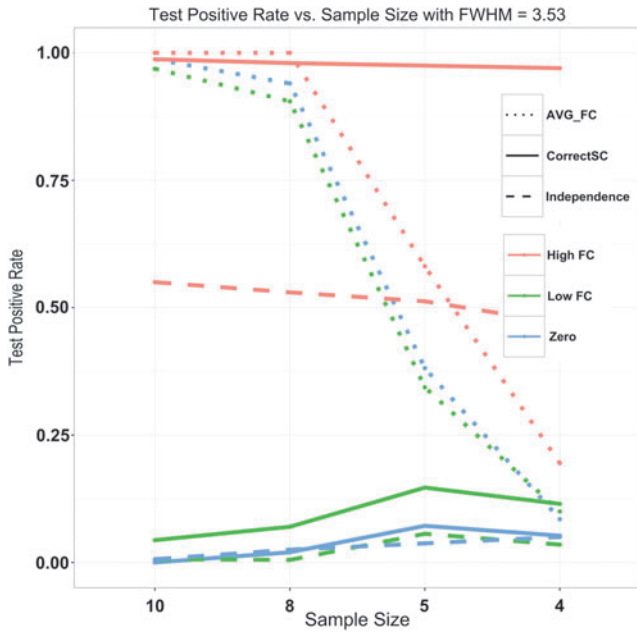


FIG. 1. Test positive rate at different sample sizes with moderate spatial correlation (FWHM=3.53). Three methods were denoted by different types of line: AVG-FC (dotted line), Bayesian with CorrectSC prior (solid line), and Bayesian with Independence prior (dashed line). Three different colors denote the underlying true strength of FC: red denotes high FC (i.e., FC=0.5 and 0.6), green denotes low FC (i.e., 0.1 and 0.2), and blue denotes zero FC. FC, functional connectivity.

Data analysis: results

We assessed FC by looking into the following probability: $\Pr(\text{FC} > 0.4)$, where FC denotes the FC estimated by the median value of posterior distribution of each FC. Instead of computing p values based on t -statistics as done for the AVG-FC approach, we computed the posterior probability and used it as an inference tool, that is, any ROI pair with $\Pr(\text{FC} > 0.4) > 0.5$ would be considered *important* as Xue et al. (2015) used 0.4 as a threshold for moderate or above functional coherence. Although it seems that the threshold values are arbitrary, by generating a series of probabilistic statements with different thresholds such as $\Pr(\text{FC} > 0.3)$, $\Pr(\text{FC} > 0.4)$, and $\Pr(\text{FC} > 0.5)$, we would be able to understand how those probabilities change with different thresholds, indicating which pairs of ROIs are of great interest and importance in terms of rs-FC. The resulting 13 *important* FC estimates are all above the reference line at $\Pr(\text{FC} > 0.4)$ (Fig. 2 and Table 2).

It is noteworthy that the FC between TL and TR shows the smallest change over the three thresholds, indicating that the FC is quite strong and highly likely to be > 0.5 . The second strong FC is observed between HL and HR. In contrast, a significant change in probability that the FC between HR and TR, that is, $\Pr(\text{FC} > 0.3) = 0.7876$ to $\Pr(\text{FC} > 0.5) = 0.2633$, indicates that the strength of FC between the two ROIs would lie between 0.3 and 0.4 with high probability.

With the AVG-FC approach, all of 28 FC values were statistically significant at FDR=0.05. As shown in the previous

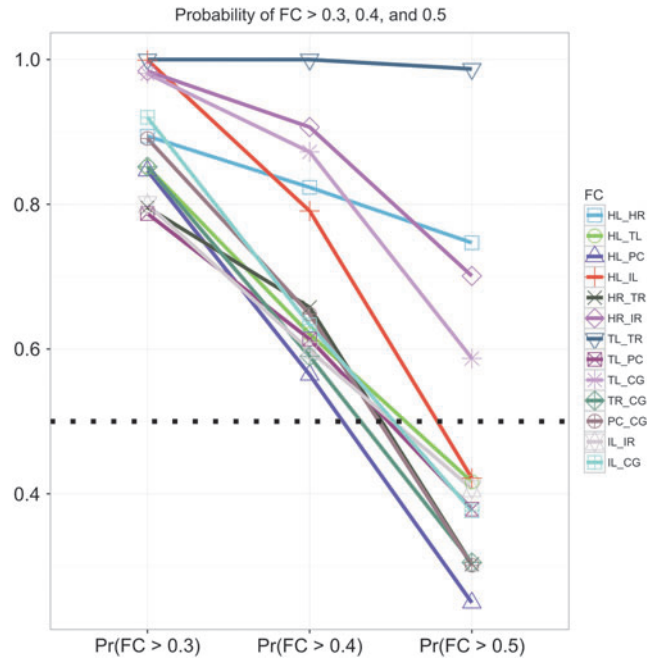


FIG. 2. A graphical summary of 13 *important* functional connectivity estimates: Probability of FC >0.3, 0.4, and 0.5. The middle dotted line denotes a probability of 0.5. The 13 *important* pairs are: HL and HR, HL and TL, HL and PC, HL and IL, HR and TR, HR and IR, TL and TR, TL and PC, TL and CG, TR and CG, PC and CG, IL and IR, and IL and CG. CG, cingulate gyrus; HL, left hippocampus; HR, right hippocampus; IL, left insula; IR, right insula; PC, precuneus; TL, left thalamus; TR, right thalamus.

simulation study, the conventional approach tends to result in high false positives because ignoring spatial correlation within a ROI causes to underestimate the variance associated with FC, leading to smaller p values and much tight 95% CI on each FC. In Table 2, 13 *important* FC values are summarized along with 95% credible intervals and statistics resulted from AVG-FC. Note that 95% CI is always tighter than the corresponding 95% credible interval, because AVG-FC underestimates the variance associated with FC. Probabilistic SC values corresponding to those 13 pairs are also reported in Table 2. Note that HL-HR and TL-TR pairs show the two strongest functional coupling detected by both the conventional approach and our spatiotemporal model. Although the SC for HL-HR (0.093) is much smaller than that for TL-TR (0.382), both pairs are claimed *important* by our Bayesian approach, indicating that our model can well accommodate both direct and indirect effect of SC on FC estimation.

One way to compare methods without knowing the ground truth is to assess how sensitive each method is to data decimation (Yang et al., 2014). If one method is very robust and reliable, then the method is expected to show certain degree of robustness to data decimation, that is, with smaller sample size the method gives rise to the similar results to the original inference result. To this end, we reduced the sample size ($n=6, 5$, and 4) with all possible combinations and then analyzed the data with both our Bayesian approach and AVG-FC. With each sample size, we counted how many times the

TABLE 2. SUMMARY OF FC VALUES WITH 95% CREDIBLE INTERVALS AND $\Pr(\text{FC} > 0.4)$ FOR 13 *IMPORTANT* ROI PAIRS BASED ON THE BAYESIAN SPATIO-TEMPORAL APPROACH

ROI pairs	FC by Bayesian (95% credible intervals)	$\Pr(\text{FC} > 0.4)$	FC by AVG-FC (95% confidence intervals)	p values for AVG-FC	Probabilistic SC
HL-HR	0.605 (0.255–0.734)	0.921	0.726 (0.632–0.823)	<0.001	0.093
HL-TL	0.449 (0.150–0.650)	0.594	0.437 (0.235–0.639)	0.003	0.452
HL-PC	0.495 (0.335–0.730)	0.763	0.588 (0.477–0.700)	<0.001	0.105
HL-IL	0.486 (0.366–0.770)	0.820	0.604 (0.461–0.747)	<0.001	0.299
HR-TR	0.466 (0.163–0.637)	0.655	0.472 (0.286–0.656)	0.002	0.427
HR-IR	0.504 (0.336–0.722)	0.786	0.677 (0.574–0.780)	<0.001	0.320
TL-TR	0.674 (0.531–0.809)	1.000	0.842 (0.776–0.907)	<0.001	0.382
TL-PC	0.503 (0.147–0.769)	0.687	0.577 (0.370–0.781)	0.002	0.117
TL-CG	0.521 (0.327–0.797)	0.910	0.662 (0.554–0.768)	<0.001	0.096
TR-CG	0.458 (0.107–0.775)	0.659	0.611 (0.449–0.770)	<0.001	0.112
PC-CG	0.455 (0.322–0.701)	0.832	0.641 (0.513–0.767)	<0.001	0.148
IL-IR	0.429 (0.075–0.755)	0.579	0.678 (0.524–0.833)	<0.001	0.085
IL-CG	0.442 (0.189–0.734)	0.648	0.693 (0.564–0.820)	<0.001	0.028

For comparison, the corresponding FC estimates with 95% confidence intervals and p -values based on the AVG-FC are listed. In the last column, probabilistic structural connectivity values are reported.

CG, cingulate gyrus; HL, left hippocampus; HR, right hippocampus; IL, left insula; IR, right insula; PC, precuneus; ROI, regions of interest; TL, left thalamus; TR, right thalamus.

important or statistically significant FC pairs at $\text{FDR} = 0.05$ shown in Table 2 were correctly found by the corresponding method. The results are summarized in Figure 3. It is obvious that our Bayesian spatiotemporal model with double fusion is much more robust to data decimation than AVG-FC. Even with $n=4$, on average 89% of the time our approach can result in the same result as $n=7$ while only 76% of the time AVG-FC results in the same result as $n=7$.

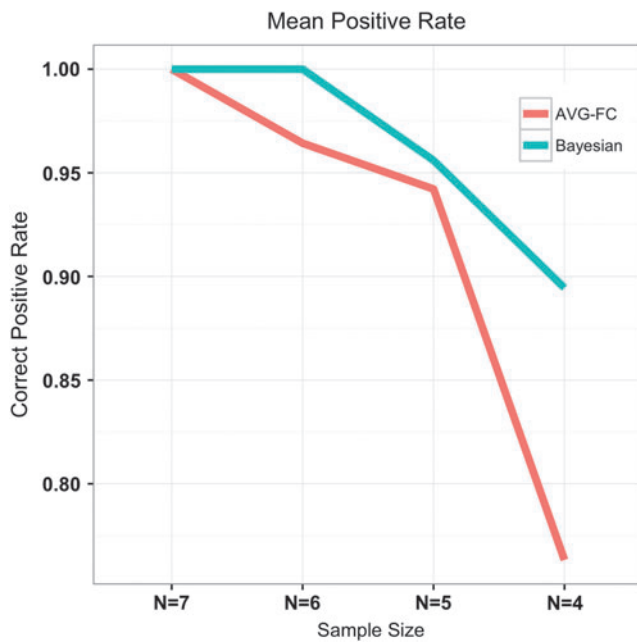


FIG. 3. Mean correct positive rates at different sample sizes, $n=7, 6, 5,$ and 4 . “Correct” positive rate denotes how often the 13 FC values summarized in Table 2 were “correctly” detected *important* by Bayesian spatiotemporal model with double fusion (cyan) and all of 28 FC values were “correctly” classified as significant by AVG-FC (red).

Discussion

We proposed a Bayesian hierarchical spatiotemporal model with double fusion to analyze multimodal neuroimaging data, that is, concurrently acquired rs-fMRI and DTI data, which could properly take into account the intrinsic spatial correlation among voxels within an ROI, and temporal correlation at each voxel. Accounting for those underlying correlation enables us to draw a valid inference about between-ROI functional couplings. Using a Bayesian analytic approach, it is very natural to combine prior knowledge about the SC based on DTI data with rs-fMRI data for updating posterior information about the FC between ROIs. Moreover, accounting for the role of direct and indirect association between FC and SC through double fusion provides a unified framework for incorporating multimodal MRI data. Our model is flexible enough to be modified to accommodate longitudinal or task-induced fMRI data.

We validated our approach using stimulation study and also applied our model to analyze multimodal MRI data to assess the impact of the SC between brain regions on rs-FC. The simulation study showed that an informative prior based on the SC reduced the MSE of FC estimates, compared to using a prior based on structural independence. In addition, it showed that failing to take into account the underlying spatial correlation within an ROI highly inflated false positive rates, leading to erroneous conclusions, for example, tight 95% confidence intervals. The similar finding was reported with task-induced fMRI analysis (Kang et al., 2012).

In data analysis, 13 ROI pairs show *important* functional couplings in the brain network through our Bayesian approach, whereas all 28 ROI pairs were claimed statistically significant through AVG-FC, although it might be false positive findings. When we considered more conservative threshold such as $\Pr(\text{FC} > 0.5) > 0.5$, then only four ROI pairs remained as *important*: HL-HR, HR-IR, TL-TR, and TL-CG. We were also able to show that our Bayesian spatiotemporal model was much robust to data decimation, which would be one of the most desirable properties of statistical

models. Although the data decimation approach would be more meaningful when we had a large sample size, it would be still valid to show that our approach was more robust than AVG-FC with such small sample size.

Our model requires estimating parameters using MCMC methods, which in general are more computationally demanding than frequentist methods, for example, AVG-FC and require the convergence of the MCMC chains that is always challenging. In this study, we confirmed the convergence by the Gelman–Rubin diagnostic (Gelman and Rubin, 1992): the diagnostic values associated with our simulation studies and data analysis were <1.2 . However, the Bayesian framework allows us to easily combine two distinct sources of information to estimate FC with higher precision and accuracy as measured by MSE. Moreover, we were able to obtain a smaller MSE using the informative prior based on DTI data.

Then one can argue that our approach heavily relies on the assumption that we have good estimates of the SC. As shown in the simulation study, reliable prior structural information that is somewhere between true underlying SC and structural independence SC can reduce the MSE of estimates. However, if the prior information is incorrect but we have a large sample size, which can be true in fMRI data analysis, then the point estimates of interest would not be heavily affected using an incorrect prior. But the variance of the estimates will increase and the time to attain the convergence of Markov chains may also increase.

Conclusion

In summary, we proposed a novel Bayesian spatiotemporal model with double fusion to better estimate FC using both rs-fMRI and DTI data. The main advantages of using our model relative to current standard methodologies are: (1) It is a novel approach that combines multimodal data producing estimates of FC with lower MSEs in simulation. (2) It gives valid inference about ROI-level FC while properly accounting for underlying spatiotemporal correlations. (3) It characterizes a probability distribution of each FC estimate, which is a unique feature of a Bayesian approach. (4) It can incorporate both direct and indirect effects of SC on FC estimation. (5) It provides an inference tool that is much more robust to data decimation compared to the conventional approach.

Acknowledgments

The authors thank Drs. David Degras (University of Massachusetts), Mark Fiecas (University of Minnesota), and Baxter Rogers (Vanderbilt University) for helpful discussions. This work was, in part, supported by NIH (Grant No. R01 NS075270).

Author Disclosure Statement

No competing financial interests exist.

References

Altinay MI, Hulvershorn LA, Karne H, Beall EB, Anand A. 2016. Differential Resting-State Functional Connectivity of Striatal Subregions in Bipolar Depression and Hypomania. *Brain Connect* 6:255–265.

Basser PJ, Mattiello J, LeBihan D. 1994. MR diffusion tensor spectroscopy and imaging. *Biophys J* 66:259–267.

Behrens TE, Berg HJ, Jbabdi S, Rushworth MF, Woolrich MW. 2007. Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *Neuroimage* 34:144–155.

Behrens TE, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, et al. 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn Reson Med* 50:1077–1088.

Bowman FD. 2007. Spatiotemporal models for region of interest analyses of functional neuroimaging data. *J Am Stat Assoc* 102:442–453.

Bowman FD, Zhang L, Derado G, Chen S. 2012. Determining functional connectivity using fMRI data with diffusion-based anatomical weighting. *NeuroImage* 62:1769–1779.

Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, et al. 2001. Frequencies contributing to functional connectivity in the cerebral cortex in “resting-state” data. *AJNR Am J Neuroradiol* 22:1326–1333.

Cressie NAC. 1993. *Statistics for Spatial Data*. New York: Wiley.

Cribben I, Haraldsdottir R, Atlas LY, Wager TD, Lindquist MA. 2012. Dynamic connectivity regression: determining state-related changes in brain connectivity. *NeuroImage* 61:907–920.

Deco G, McIntosh AR, Shen K, Hutchison RM, Menon RS, Everling S, et al. 2014. Identification of optimal structural connectivity using functional connectivity and neural modeling. *J Neurosci* 34:7910–7916.

Dickstein DP, Gorrostieta C, Ombao H, Goldberg LD, Brazel AC, Gable CJ, et al. 2010. Fronto-temporal spontaneous resting state functional connectivity in pediatric bipolar disorder. *Biol Psychiatry* 68:839–846.

Friston KJ, Frackowiak RSJ, Liddle PF, Frith CD. 1993. Functional connectivity: the principal-component analysis of large (PET) data sets. *J Cereb Blood Flow Metab* 13:5–14.

Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. *Statist Sci* 7:457–511.

Glover GH, Li TQ, Ress D. 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med* 44:162–167.

Greicius MD, Supekar K, Menon V, Dougherty RF. 2009. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb Cortex* 19:72–78.

Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, et al. 2008. Mapping the structural core of human cerebral cortex. *PLoS Biol* 6:e159.

Hansen EC, Battaglia D, Spiegler A, Deco G, Jirsa VK. 2015. Functional connectivity dynamics: modeling the switching behavior of the resting state. *Neuroimage* 105:525–535.

Harrison SJ, Woolrich MW, Robinson EC, Glasser MF, Beckmann CF, Jenkinson M, et al. 2015. Large-scale probabilistic functional modes from resting state fMRI. *NeuroImage* 109:217–231.

Hendler T, Pianka P, Sigal M, Kafri M, Ben-Bashat D, Constantini S, et al., 2003. Two are better than one: combining fMRI and DTI based fiber tracking for effective pre-surgical mapping. *Proc Int Soc Magn Reson Med*, p. 394.

Holmes MJ, Yang X, Landman BA, Ding Z, Kang H, Abou-Khalil B, et al. 2013. Functional networks in temporal-lobe epilepsy: a voxel-wise study of resting-state functional connectivity and gray-matter concentration. *Brain Connect* 3:22–30.

Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, et al. 2009. Predicting human resting-state functional

- connectivity from structural connectivity. *Proc Natl Acad Sci USA* 106:2035–2040.
- Huang H, Ding M. 2016. Linking functional connectivity and structural connectivity quantitatively: a comparison of methods. *Brain Connect* 6:99–108.
- Joo SH, Lim HK, Lee CU. 2016. Three large-scale functional brain networks from resting-state functional MRI in subjects with different levels of cognitive impairment. *Psych Invest* 13:1–7.
- Kang H, Ombao H, Linkletter C, Long N, Badre D. 2012. Spatio-spectral mixed effects model for functional magnetic resonance imaging data. *J Am Stat Assoc* 107:568–577.
- Meier J, Tewarie P, Hillebrand A, Douw L, van Dijk BW, Stufflebeam SM, et al. 2016. A mapping between structural and functional brain networks. *Brain Connect* 6:298–311.
- Morgan VL, Mishra A, Newton AT, Gore JC, Ding Z. 2009. Integrating functional and diffusion magnetic resonance imaging for analysis of structure-function relationship in the human language network. *PLoS One* 4:e6660.
- Olesen P, Nagy Z, Westerberg H, Klingberg T. 2003. Combined analysis of DTI and fMRI data reveals a joint maturation of white and grey matter in a fronto-parietal network. *Cogn Brain Res* 8:48–57.
- Patil A, Huard D, Fonnesebeck CJ. 2010. PyMC: Bayesian stochastic modelling in python. *J Stat Softw* 35:1–81.
- Poldrack RA, Mumford JA, Nichols TE. 2011. *Handbook of Functional MRI Data Analysis*. New York, NY: Cambridge University Press.
- Schmittmann VD, Jahfari S, Borsboom D, Savi AO, Waldorp LJ. 2015. Making large-scale networks from fMRI data. *PLoS One* 10:e0129074.
- van den Heuvel MP, Mandl RC, Kahn RS, Hulshoff Pol HE. 2009. Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Hum Brain Mapp* 30:3127–3141.
- Wang K, Liang M, Wang L, Tian L, Zhang X, Li K, et al. 2007. Altered functional connectivity in early Alzheimer's disease: a resting-state fMRI study. *Hum Brain Mapp* 28:967–978.
- Werring DJ, Clark CA, Barker GJ, Miller DH, Parker GJ, Brammer MJ, et al. 1998. The structural and functional mechanisms of motor recovery: complementary use of diffusion tensor and functional magnetic resonance imaging in a traumatic injury of the internal capsule. *J Neurol Neurosurg Psychiatry* 65:863–869.
- Wieshmann UC, Krakow K, Symms MR, Parker GJ, Clark CA, Barker GJ, et al. 2001. Combined functional magnetic resonance imaging and diffusion tensor imaging demonstrate widespread modified organisation in malformation of cortical development. *J Neurol Neurosurg Psychiatry* 70:521–523.
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. 1996. A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp* 4:58–73.
- Xue W, Bowman FD, Pileggi AV, Mayer AR. 2015. A multi-modal approach for determining brain networks by jointly modeling functional and structural connectivity. *Front Comput Neurosci* 9:1–11.
- Yang X, Kang H, Newton AT, Landman BA. 2014. Evaluation of statistical inference on empirical resting state fMRI. *IEEE Trans Biomed Eng* 61:1091–1099.
- Yu Z, Prado R, Quinlan EB, Cramer SC, Ombao H. 2016. Understanding the impact of stroke on brain motor function: a hierarchical Bayesian approach. *J Am Stat Assoc* 111:549–563.
- Zhang L, Guindani M, Vannucci M. 2015. Bayesian models for fMRI data analysis. *Wiley Interdiscip Rev Comput Stat* 7:21–41.
- Zhang L, Guindani M, Versace F, Vannucci M. 2014. A spatio-temporal nonparametric Bayesian variable selection model of fMRI data for clustering correlated time courses. *NeuroImage* 95:162–175.
- Zhu D, Zhang T, Jiang X, Hu X, Chen H, Yang N, et al. 2014. Fusing DTI and fMRI data: a survey of methods and applications. *Neuroimage* 102:184–191.
- Zippo AG, Castiglioni I, Borsa VM, Biella GE. 2015. The compression flow as a measure to estimate the brain connectivity changes in resting state fMRI and 18FDG-PET Alzheimer's disease connectomes. *Front Comput Neurosci* 9:148.

Address correspondence to:
 Hakmook Kang
 Department of Biostatistics
 Vanderbilt University
 2525 West End Avenue
 Suite #1100
 Nashville, TN 37203

E-mail: hakmook.kang@vanderbilt.edu