Received: 12 December 2021 | Revised: 7 May 2022 | Accepted: 6 June 2022

DOI: 10.1002/hbm.26000

RESEARCH ARTICLE

WILEY

On predictability of individual functional connectivity networks from clinical characteristics

Emily L. Morris[1] | Stephan F. Taylor[2] | Jian Kang[1]

[1]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA

[2]Department of Psychiatry, University of Michigan, Ann Arbor, Michigan, USA

Correspondence
Jian Kang, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA.
Email: jiankang@umich.edu

Funding information
National Institutes of Health, Grant/Award Numbers: R01 DA048993, R01 GM124061, R01 MH105561

Abstract

In recent years, understanding functional brain connectivity has become increasingly important as a scientific tool with potential clinical implications. Statistical methods, such as graphical models and network analysis, have been adopted to construct functional connectivity networks for single subjects. Here we focus on studying the association between functional connectivity networks and clinical characteristics such as psychiatric symptoms and diagnoses. Utilizing machine learning algorithms, we propose a method to examine predictability of functional connectivity networks from clinical characteristics. Our methods can identify salient clinical characteristics predictive of the whole brain network or specific subnetworks. We illustrate our methods on the analysis of fMRI data in the Philadelphia Neurodevelopmental Cohort study, demonstrating clinically meaningful results.

1 | INTRODUCTION

It is widely acknowledged that the integrated behavior of the approximately 100 billion neurons of the human brain in connected networks provides the substrate for complex behavior (Schultz et al., 1997). Growing interest in understanding the association between abnormal brain development and vulnerability to psychiatric disorders has motivated recent research in functional brain connectivity, utilizing functional magnetic resonance imaging (fMRI) (Greicius et al., 2004; Satterthwaite et al., 2016; Sheline et al., 2010; Sripada et al., 2012). Correlated activity of low-frequency fluctuations of the blood-oxygenation level dependent (BOLD) signal provides a putative marker of large-scale networks, which may be used to find patterns among patients linked to specific clinical states. Identifying connectivity patterns associated with patient characteristics is clinically relevant for classifying high risk patients or identifying disease markers. These neuroimaging datasets, often obtained while the brain is "at rest," that is, not engaged in any particular task, capture dynamic activity across the brain with complex spatial and temporal covariance patterns. Extracting clinically meaningful information from these patterns represents important progress in the analysis of fMRI images.

1.1 | Motivating data: The PNC study

The motivating data set for this analysis is from the Philadelphia Neurodevelopmental Cohort (PNC) study (Satterthwaite et al., 2014). The PNC is a community sample of 9500 young persons from an urban hospital who presented for care for a wide range of physical and brain illnesses. Among this large sample, over 1400 underwent neuroimaging, from which subjects for the current study were derived. The open-source dataset provides the opportunity to study the relationship between brain development and psychiatric symptoms.

The PNC study has been used to investigate abnormal brain development with an aim to identify youth at risk of developing psychiatric disorders (Kessler et al., 2016; Satterthwaite et al., 2016; Xia et al., 2018). Kessler et al. (2016) used independent component analysis to generate "growth charts" for functional brain network, linking this maturation to predict task outcomes. Xia et al. (2018) identified

functional connectivity patterns associated with four categories of psychopathology using canonical correlation analysis. These important findings assist in furthering knowledge of the link between brain development and psychopathology. In the present work, we use novel methods to identify clinical characteristics predictive of functional connectivity, enabling a more granular level of both connectivity and patient symptoms.

1.2 | Existing methods

Several statistical methods have been proposed to estimate networks from these fMRI data, graphical models being one popular choice; see Smith et al. (2013) for a recent overview of the functional connectivity methods. It has been shown that partial correlation, inverse covariance estimation, and Bayes net methods, can capture accurate connectivity estimation but computational challenges limit these methods in many cases. When using undirected network estimation, there are often limitations regarding choice of link strength (Solo et al., 2018). Although the drawbacks of various methods may be known, it is often difficult to compare approaches or decide on the best method given multiple options. Most approaches for comparison of networks involve simple summary characteristics of a network, but these summary metrics may not capture the true differences or similarities across metrics. Here the proposed method aims to utilize connectivity in a different way, by not relying on network summaries or dimension reduction approaches, than existing methods to establish associations with clinical symptoms (Solo et al., 2018).

Understanding the predictive utility of functional connectivity is of critical interest to clinicians. Diseases like depression and Alzheimer's disease may be associated with unique connectivity patterns, but research to demonstrate true predictive power of these networks for clinical use is still in its infancy (Greicius et al., 2004; Sheline et al., 2010). On the other hand, predicting connectivity using clinical characteristics may be more accurate than the reverse. Because the connectivity matrix is difficult to estimate, current methods that rely on the connectivity estimate to model phenotype may be unreliable or infeasible due to computational limitations. While clinicians are ultimately interested in using connectivity to diagnosis or identify patients at greater risk of symptom emergence or worsening, researchers seeking to understand mechanisms of disease would like to know which cluster of symptoms predict, or map to, which patterns of connectivity. Using clinical characteristics as predictors can identify associations between multiple phenotypic characteristics and functional brain connectivity, without relying on the accuracy of estimating a potentially noisy connectivity matrix. Furthermore, nosological uncertainty about the validity of existing disease classifications in psychiatry has led to calls for developing new categories of disorders based on underlying biological mechanisms (Insel et al., 2010); thus, using the clinical phenotype to predict new associations with underlying connectivity patterns is an important step in the direction of finding more biologically driven categories of psychopathology.

1.3 | Contribution

This work aims to provide a broad framework for constructing a reliable functional connectivity network and analyzing relationships with clinical characteristics using machine learning methods. Specifically, we propose a three-level model to specify the association between the voxel-level fMRI times series, the region-level brain networks and the clinical characteristics. We develop a fast computing method to estimate the model parameters and make predictions on brain functional connectivity by integrating different machine learning methods. A joint modeling framework is then used to improve connectivity estimation from the results of modeling connectivity with clinical characteristics. See Figure 1 for an illustration of the model and the estimation method. We illustrate the proposed method on evaluating the predictability of individual functional connectivity networks from the clinical characteristics in the PNC study, identifying the important functional subnetworks that are highly associated with psychiatric syndromes such as Post-Traumatic Stress Disorder (PTSD) and Psychosis. We also perform a simulation study to evaluate the performance of the proposed framework in terms of selection and prediction accuracy. In addition, we have developed an R package that implements the proposed method and provides user-friendly software to study the association between functional brain networks and clinical characteristics, including the fMRI voxel-level time series prepossessing and the graphical presentation of the model fitting results. The R package will be freely available online after the paper is published and it is now available upon request and on GitHub.

2 | METHODS

In this work, we propose a general modeling framework for analysis of brain functional connectivity and clinical characteristics. It consists of two major steps: (1) individual functional connectivity network construction and (2) covariate feature screening and network prediction. To improve the selection and estimation accuracy, we propose a joint estimation method which adopts the alternating direction method of multipliers (ADMM) algorithm to update the sparse precision matrices in the model.

Suppose we collect data from n subjects. We collect the resting-state fMRI signal with T scans, where the whole brain regions consist of V voxels and R regions. Let $i\,(i=1,\ldots,n)$ index the subject, $v\,(v=1,\ldots,V)$ index the voxels, $t\,(t=1,\ldots,T)$ index the time scans. Let $r_v \in \{1,\ldots,R\}$ be the region index for voxel v. Let $y_{i,v,t}$ represent the observed resting state fMRI signals for subject i at voxel v and time t. For each subject, we also collect p covariates of clinical characteristics. Let $j\,(j=1,\ldots,p)$ index the covariates and let $x_{i,j}$ denote the measurements of covariate j for subject i. Write $\mathbf{x}_i = \left(x_{i,1},\ldots,x_{i,p}\right)^\top$.

2.1 | A generative modeling framework

We consider a generative modeling framework to specify the associations between the voxel-level fMRI time series, the region-level brain

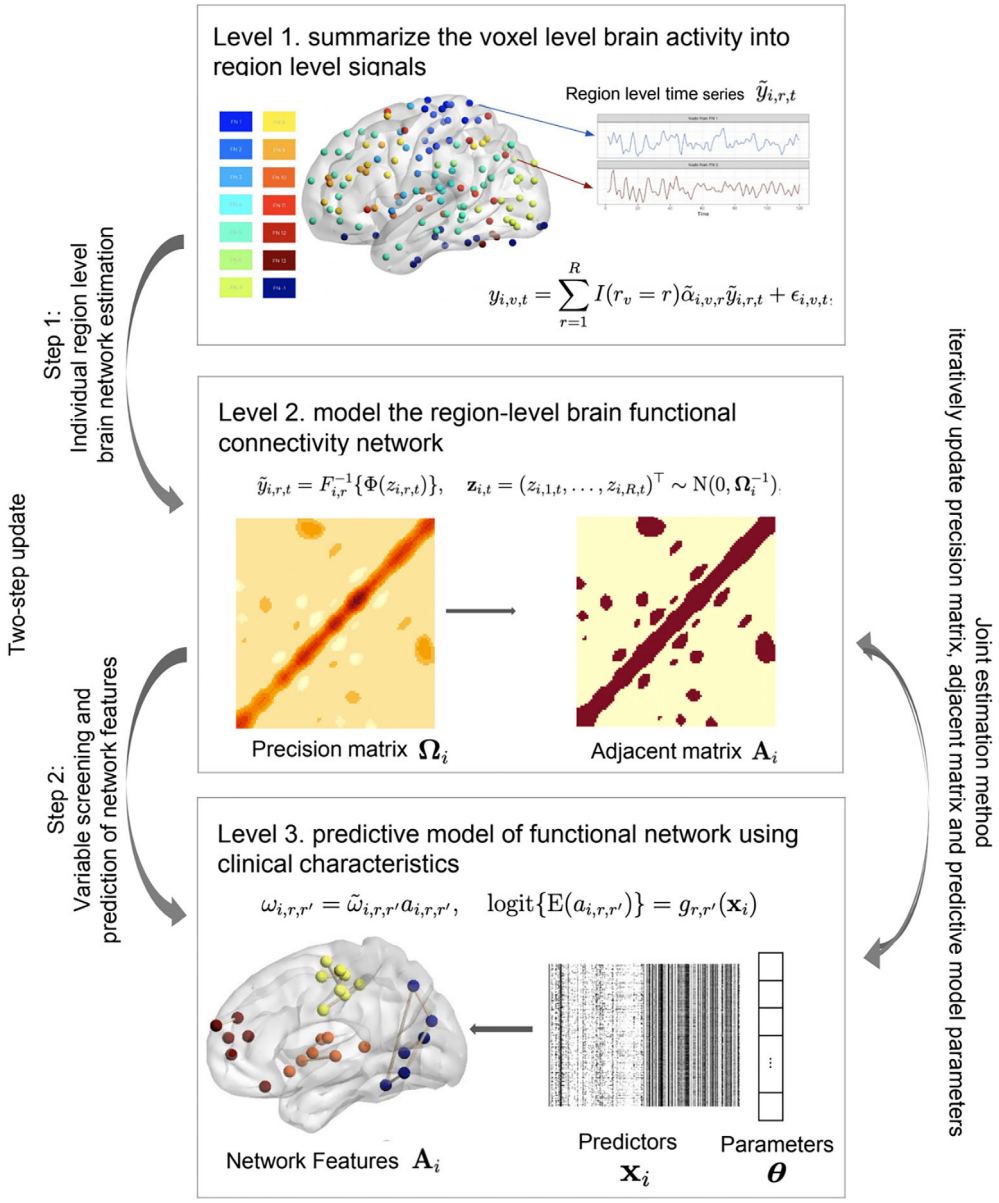

FIGURE 1 Illustrations of the proposed three-level hierarchical model along with the two-setup update and the joint estimation method

networks and the clinical characteristics. We consider a three-level hierarchical model.

At Level 1, we summarize the voxel level brain activity into region level signals:

$$y_{i,v,t} = \sum_{r=1}^{R} I(r_v = r)\tilde{\alpha}_{i,v,r}\tilde{y}_{i,r,t} + \epsilon_{i,v,t}, \tag{1}$$

where $\tilde{y}_{i,r,t}$ represents the summarized neural activity in region r at time t for subject i and $\alpha_{i,v,r}$'s are the weight coefficients that represent the contribution of voxel v to region r for subject i. We assume the random error $\epsilon_{i,v,t}$ with mean zero and constant variance.

At Level 2, we model the region-level brain functional connectivity network. We consider a Gaussian copula graphical model. We transform each region specific signal $\tilde{y}_{i,r,t}$ into a latent variable $z_{i,r,t}$ according to the marginal distribution. We assume those latent variables follow a multivariate normal distribution. In particular, we have

$$\tilde{y}_{i,r,t} = F_{i,r}^{-1}\{\Phi(z_{i,r,t})\}, \quad \mathbf{z}_{i,t} = (z_{i,1,t}, \ldots, z_{i,R,t})^\top \sim N(0, \mathbf{\Omega}_i^{-1}), \tag{2}$$

where $\mathbf{\Omega}_i = \{\omega_{i,r,r'}\}$ is an $R \times R$ precision matrix. The function $F_{i,r}(\cdot)$ is the cumulative distribution function of $\tilde{y}_{i,r,t}$ and $\Phi(\cdot)$ is the cumulative distribution function of the standard Gaussian distribution.

At Level 3, we impose sparsity on the precision matrix $\mathbf{\Omega}$. We introduce a latent selection indicator $a_{i,r,r'} \in \{0,1\}$ for each region pair (r,r') to indicate whether the region r and r' are function connected for subject i; and for each region pair, we model the conditional distribution of $a_{i,r,r'}$ given the clinical characteristics $\mathbf{x}_i$ through a logistic regression model:

$$\omega_{i,r,r'} = \widetilde{\omega}_{i,r,r'} a_{i,r,r'}, \quad \text{logit}\{\mathrm{E}(a_{i,r,r'})\} = g_{r,r'}(\mathbf{x}_i), \tag{3}$$

where $g_{r,r'}(\cdot)$ is an unknown function representing the log odds of region r and r' being functional connected for subject i with clinical characteristics $\mathbf{x}_i$.

2.2 | Individual functional connectivity network estimation

We may consider some fast computing methods for Level 1 estimation: for example, the averaging voxel level signals within each region or using principal component analysis (PCA) to summarize the region level signal. Specifically in the first case, taking a simple average of voxels uses $\widetilde{\alpha}_{i,v,r} = 1$. We assume equal weight across all voxels and average within regions:

$$\widetilde{y}_{i,r,t} = \frac{1}{\sum_v I(r_v = r)} \sum_{v: r_v = r} y_{i,v,t}, \tag{4}$$

For the PCA method, we can use the first principal component (DuBois Bowman et al., 2012) to estimates $\widetilde{\alpha}_{i,v,r}\widetilde{y}_{i,v,t}$ and construct the weights. However, the PCA method may lose power to detect the weak signals in the region-level function connectivity (Jian Kang et al., 2016). In this paper, we choose the method of averaging voxel level signals.

Our network estimation for Level 2 implements the Meinshausen and Bühlmann method of estimating a sparse graphical model (Zhao et al., 2015). We have selected this method in part due to the flexibility to relax the normality assumption often imposed on observations in graphical models. Through fitting semi-parametric Gaussian copula models, this approach aims to better recover the true underlying undirected graph structure (Liu et al., 2009).

In particular, let $\Phi^{-1}(\cdot)$ be the quantile function of the standard Gaussian distribution. We have $z_{i,r,t} = \Phi^{-1}\{F_{i,r}(\widetilde{y}_{i,r,t})\}$ and the connectivity matrix is estimated as

$$\widehat{\boldsymbol{\Omega}}_i = \text{argmin}_{\boldsymbol{\Omega}_i}[\text{tr}\{\boldsymbol{\Omega}_i S_T(\mathbf{Z}_i)\} - \text{logdet}(\boldsymbol{\Omega}_i) + \lambda \| \boldsymbol{\Omega}_i \|_1], \tag{5}$$

where $S_T(\mathbf{Z}_i)$ is the sample covariance of the transformed region-level connectivity signals $\mathbf{Z}_i = (\mathbf{z}_{i,1}, \ldots, \mathbf{z}_{i,R})$, $\det(\boldsymbol{\Omega})$ is the determinant of $\boldsymbol{\Omega} = \{\omega_{r,r'}\}$, λ is a regularization parameter and $\| \boldsymbol{\Omega} \|_1 = \sum_{r,r'} |\omega_{r,r'}|$ is the entry-wise L_1 norm. The solution of Equation (5) enjoys the sparsity, thus we estimate the latent connectivity indicator to obtain the functional connectivity matrix for the entire brain $\mathbf{A}_i = (a_{i,r,r'})$ with

$$a_{i,r,r'} = I(\omega_{i,r,r'} \neq 0),$$

We use this network construction as the connectivity outcome when fitting the prediction models in the Level 3 estimation.

2.3 | Prediction of network features

We can estimate the relationship between the covariates and connectivity of regions r and r', $g_{r,r'}(\cdot)$ using the initial estimate of the functional connectivity network. Although the framework presented in Section 2.1 is flexible and may incorporate more general models, we start from a linear model for simplicity:

$$\text{logit}\{E(a_{i,r,r'})\} = \sum_{j=1}^{p} \theta_{r,r'j} x_{i,j}, \tag{6}$$

where x_{ij} is the measured value for covariate j of subject i.

With a large number of region pairs and covariates to consider, it may be difficult to perform model estimation. A more efficient approach is to carry out a screening mechanism before fitting edgewise prediction models. We proposed to implement the elastic net regression to screen out clinical variables that are not associated with a given edge connection. We only consider predicted edges where at least 5% of subjects have a connection.

Machine learning methods are increasingly popular for predictive modeling. We consider two common machine learning methods to predict edgewise connectivity: support vector machine (SVM) and random forests (Breiman, 2001; Cortes & Vapnik, 1995).

When performing SVM we aim to minimize the following loss function for each region pair r, r':

$$\begin{aligned} &\min_{\boldsymbol{\theta}_{r,r'}} \left\{ \frac{1}{2} \mathbf{w}_{r,r'}^\top \mathbf{w}_{r,r'} + c_{r,r'} \sum_{i=1}^{n} \xi_{i,r,r'} \right\}, \\ &\text{such that} \quad (2a_{i,r,r'} - 1)\{\mathbf{w}_{r,r'}^\top \boldsymbol{\phi}(\mathbf{x}_i) + b\} \geq 1 - \xi_{i,r,r'}, \end{aligned} \tag{7}$$

where $\boldsymbol{\theta}_{r,r'} = \{\mathbf{w}_{r,r'}, b_{r,r'}, \{\xi_{i,r,r'}\}_{i=1}^n\}$ and $\boldsymbol{\phi}(\mathbf{x})$ is a vector of features in the transformed feature space derived from the kernel $k(\mathbf{x}, \mathbf{x}')$, such that $k(\mathbf{x}, \mathbf{x}') = \boldsymbol{\phi}(\mathbf{x}) \cdot \boldsymbol{\phi}(\mathbf{x}')$. In this setup $\mathbf{w}_{r,r'}$ denotes the weight vector used to maximize the margin around the hyperplane separating subjects with and without a connected edge between region pairs r and r' in the network. The penalty term $c_{r,r'} \sum_{i=1}^{n} \xi_{i,r,r'}$ is used to penalize for observations that are misclassified.

Random forest (Breiman, 2001) implements a series of decision trees, where individual trees form based on minimizing the residual sum of squares. In particular, for each region pair r, r', we aim to predict the functional connectivity $a_{i,r,r'}$ using clinical characteristics $\mathbf{x}_i$ using classification probability $m_{r,r'}(\mathbf{x}) = \Pr(a_{i,r,r'} = 1 \mid \mathbf{x}_i = \mathbf{x})$. We model $m_{r,r'}(\mathbf{x})$ as an ensemble of M randomized regression trees, that is,

$$m_{r,r'}(\mathbf{x}) = \frac{1}{M} \sum_{j=1}^{M} m(\mathbf{x}, \mathcal{T}_{r,r'j}), \quad m(\mathbf{x}, \mathcal{T}_{r,r'j}) = \sum_{s=1}^{S} p_{r,r'j,s} I(\mathbf{x} \in \mathcal{A}_{r,r'j,s}), \tag{8}$$

where $m(\mathbf{x}, \mathcal{T}_{r,r'j})$ is the classification probability given $\mathbf{x}_i = \mathbf{x}$ by the jth tree for region pairs r, r'. Each tree $\mathcal{T}_{r,r'j}$ consists of a tree-based partition $\{\mathcal{A}_{r,r'j,s}\}_{s=1}^S$ of the sample space $\mathcal{X}$ with $\mathcal{X} = \bigcup_{s=1}^S \mathcal{A}_{r,r'j,s}$ and $\mathcal{A}_{r,r'j,s} \cap \mathcal{A}_{r,r'j,s'} = \emptyset$ for $s \neq s'$ and the corresponding classification probability $p_{r,r'j,s}$ for partition $\mathcal{A}_{r,r'j,s}$. To construct each tree $\mathcal{T}_{r,r'j}$, we draw a subsample of the training set with replacement and grow a tree on the

sub training set using random feature selection, that is, randomly select a subset of predictors to split on and grow the tree using the classification and regression tree (CART) method (Breiman et al., 2017) to maximum size without pruning. The out-of-bag estimates are used to monitor the classification errors. The subsampling and random feature selection provide a mechanism to help decorrelate trees.

For the tuning parameters, we use the standard settings for SVM, using the normal kernel function and soft margin classification, and tune the number of trees and number of candidate variables at each split for random forest. Cross-validated AUC is used to evaluate the predictive performance of edge classification.

The procedure of estimating the functional connectivity network and then fitting prediction models to each edge will be referred to as the two-step update. These estimates will be used as the initial values in the iterative algorithm detailed in the next section.

2.4 | Joint estimation method

We develop a method to jointly estimate the individual functional connectivity networks and the predictive models of network features. We focus on illustrating our method by using logistic regression as the predictive model, while SVM and RF can be derived in a similar fashion. Let $\mathbf{S}_i = S_T(\mathbf{Z}_i)$, $\mathbf{G}(\mathbf{x}_i;\boldsymbol{\theta}) = \{g_{r,r'}(\mathbf{x}_i;\boldsymbol{\theta})\}_{R\times R}$ and $\mathbf{L}(\mathbf{x}_i;\boldsymbol{\theta}) = \{l_{r,r'}(\mathbf{x}_i;\boldsymbol{\theta})\}$ be three $R\times R$ matrices, where $l_{r,r'}(\mathbf{x}_i;\boldsymbol{\theta}) = \log\left[1+\exp\{g_{r,r'}(\mathbf{x}_i;\boldsymbol{\theta})\}\right]$. Let $\boldsymbol{\Theta} = \left\{\{\boldsymbol{\Omega}_i\}_{i=1}^n, \{\mathbf{A}_i\}_{i=1}^n, \boldsymbol{\theta}\right\}$ represent all the unknown parameters in our problem, where $\boldsymbol{\Omega}_i$ is an $R\times R$ symmetric positive definite matrix as the precision matrix of the transformed region-level fMRI time series for subject i, $\mathbf{A}_i \in \{0,1\}^{R\times R}$ is a $R\times R$ binary matrix indicating functional connectivity patterns for subject i, and $\boldsymbol{\theta}$ is the parameter in the logistic regression. We estimate $\boldsymbol{\Theta}$ by solving the following constrained optimization problem

$$\widehat{\boldsymbol{\Theta}} = \text{argmin}_{\boldsymbol{\Theta}} \sum_{i=1}^{n}\left[-\log\det(\boldsymbol{\Omega}_i) + tr(\boldsymbol{\Omega}_i\mathbf{S}_i) + \lambda\|\boldsymbol{\Omega}_i\|_1 - tr\{\mathbf{A}_i\mathbf{G}(\mathbf{x}_i;\boldsymbol{\theta})\} + \mathbf{1}_R^\top \mathbf{L}(\mathbf{x}_i;\boldsymbol{\theta})\mathbf{1}_R\right]$$
$$\text{subject to } a_{i,r,r'} = I(\omega_{i,r,r'} \neq 0), \text{for all } i,r,r', \tag{9}$$

where $\mathbf{1}_R$ is a column vector of R ones. In the objective function, $\{\mathbf{S}_i\}_{i=1}^n$ are observed data, the term $\sum_{j=1}^{n}[-\log|\boldsymbol{\Omega}_i| + \text{tr}(\boldsymbol{\Omega}_i\mathbf{S}_i) + \lambda\|\boldsymbol{\Omega}_i\|_1]$ is the penalized loss functions for estimating the sparse $\{\boldsymbol{\Omega}_i\}_{i=1}^n$ over all the subjects, where λ is a tuning parameter. If we consider $\{\mathbf{A}_i\}_{i=1}^n$ as a collection of binary response variables and $\{\mathbf{x}_i\}_{i=1}^n$ as the predictors, then the term $\sum_{i=1}^{n}\left[-\text{tr}\{\mathbf{A}_i\mathbf{G}(\mathbf{x}_i;\boldsymbol{\theta})\} + \mathbf{1}_R^\top \mathbf{L}(\mathbf{x}_i;\boldsymbol{\theta})\mathbf{1}_R\right]$ becomes to the summation of all the cross-entropy loss functions for logistic regression over all the region pairs. The sparsity of $\boldsymbol{\Omega}_i$ and the constraints between all the elements of $\boldsymbol{\Omega}_i$ and $\mathbf{A}_i$ define the connections between the terms.

It is challenging to directly solve (9). We propose an approximating objective function by removing the constraint in (9) but adding another term "$-\gamma\text{tr}\{(2\mathbf{A}_i-1)|\boldsymbol{\Omega}_i|\}$", where $\gamma \in (0,\lambda)$ is a tuning parameter and $|\boldsymbol{\Omega}_i| = (|\omega_{i,r,r'}|)$. Thus, the constrained optimization (9) is approximated by an unconstrained optimization problem

$$\widehat{\boldsymbol{\Theta}} = \text{argmin}_{\boldsymbol{\Theta}} \sum_{i=1}^{n}\Big[-\text{logdet}(\boldsymbol{\Omega}_i) + tr(\boldsymbol{\Omega}_i\mathbf{S}_i) + \lambda\|\boldsymbol{\Omega}_i\|_1 - \gamma tr\{(2\mathbf{A}_i-1)|\boldsymbol{\Omega}_i|\} - tr\{\mathbf{A}_i\mathbf{G}(\mathbf{x}_i;\boldsymbol{\theta})\} + \mathbf{1}_R^\top \boldsymbol{L}(\mathbf{x}_i;\boldsymbol{\theta})\mathbf{1}_R\Big]. \tag{10}$$

In (10), the term $\lambda\|\boldsymbol{\Omega}_i\|_1 - \gamma\text{tr}\{(2\mathbf{A}_i-1)|\boldsymbol{\Omega}_i|\} = \sum_{r,r'}\{\lambda - \gamma(2a_{i,r,r'}-1)\}|\omega_{i,r,r'}|$. When $|\omega_{i,r,r'}|$ is large, taking $a_{i,r,r'} = 1$ leads to a smaller objective function compared to taking $a_{i,r,r'} = 0$. On the other hand, when $a_{i,r,r'} = 1$, the penalty term for $\omega_{i,r,r'}$ reduces to $(\lambda-\gamma)|\omega_{i,r,r'}|$ from $\lambda|\omega_{i,r,r'}|$. Then the solution to $\omega_{i,r,r'}$ is more likely to be nonzero. In contrast, when $a_{i,r,r'} = 0$, the penalty term for $\omega_{i,r,r'}$ increases to $(\lambda+\gamma)|\omega_{i,r,r'}|$, then the solution to $\omega_{i,r,r'}$ is more likely to be shrunk toward to zero.

We use the solution of (10) to approximate the solution of (9) by iteratively updating $\boldsymbol{\Omega}_i$, $\boldsymbol{\theta}$, and $\mathbf{A}_i$ until the algorithm converges. To specify the sparse initial values, that is, $\boldsymbol{\Omega}_i^{(0)}$, $\mathbf{A}_i^{(0)}$, and $\boldsymbol{\theta}^{(0)}$, we reduce the dimension of the candidate predictors based on the two-step variable screening result. To reduce the computational complexity, over iterations of the joint estimation approach, we do not perform variable screening but fit the logistic regression, SVM, or random forest in the reduced parameter space.

2.4.1 | Update $\boldsymbol{\Omega}_i$ by ADMM

In the k th iteration ($k = 1,2,\ldots$), we first update $\boldsymbol{\Omega}_i^{(k)}$ by minimizing the objective function with respect to $\boldsymbol{\Omega}_i$ and fixing $\mathbf{A}_i$ at the previous iteration, that is, $\mathbf{A}_i^{(k-1)}$, for $i = 1,\ldots,n$,

$$\widehat{\boldsymbol{\Omega}}_i = \text{argmin}_{\boldsymbol{\Omega}_i}\Big[-\log\{\det(\boldsymbol{\Omega}_i)\} + \text{tr}(\boldsymbol{\Omega}_i\mathbf{S}_i) + \lambda\|\boldsymbol{\Omega}_i\|_1 - \gamma tr\left\{\left(2\mathbf{A}_i^{(k-1)}-1\right)|\boldsymbol{\Omega}_i|\right\}\Big], \tag{11}$$

where λ and γ are tuning parameters.

To implement the ADMM algorithm, we introduce $\mathbf{Y}_i^{(k)}$ and let $\mathbf{Z}_i^{(k)} = \boldsymbol{\Omega}_i^{(k)} - \mathbf{Y}_i^{(k)}$, we now minimize the following objective function with respect to $\boldsymbol{\Omega}_i$

$$\boldsymbol{\Omega}_i^{(k)} = \underset{\boldsymbol{\Omega}_i}{\text{argmin}}\left\{-\log\left(\det(\boldsymbol{\Omega}_i)\right) + \frac{\mu}{2}\left\|\boldsymbol{\Omega}_i + \left(\mathbf{Z}_i^{(k-1)} - \mathbf{Y}_i^{(k-1)} + \frac{1}{\mu}\mathbf{S}_i - \frac{\gamma}{\mu}\left(2\mathbf{A}_i^{(k-1)}-1\right)\right)\right\|^2\right\}.$$

Taking the derivative with respect to $\boldsymbol{\Omega}_i$, and defining $\mathbf{Y}_i - \mathbf{Z}_i^{(k-1)} - \frac{1}{\mu}\mathbf{S}_i + \frac{\gamma}{\mu}(2\mathbf{A}_i-1) = \mathbf{U}_i\boldsymbol{\Lambda}_i\mathbf{U}_i^\top$ with $\boldsymbol{\Lambda}_i = \text{diag}(\lambda_1,\ldots,\lambda_R)$ results in the following equations to solve $\boldsymbol{\Omega}_i^{(k)}$.

$$\begin{aligned}\mathbf{0} &= -\boldsymbol{\Omega}_i^{-1} + \mu\boldsymbol{\Omega}_i - \mu\left\{\mathbf{Y}_i^{(k-1)} - \mathbf{Z}_i^{(k-1)} + \frac{1}{\mu}\mathbf{S}_i + \frac{\gamma}{\mu}(2\mathbf{A}_i-1)\right\}.\\ \mathbf{0} &= -F_\mu^{-1}(\boldsymbol{\Lambda}_i) + \mu F_\mu(\boldsymbol{\Lambda}_i) - \mu\boldsymbol{\Lambda}_i\end{aligned}$$

The solution takes the form:

$$\boldsymbol{\Omega}_i^{(k)} = \boldsymbol{U}_i F_\mu(\boldsymbol{\Lambda}_i)\boldsymbol{U}_i^\top = \frac{1}{2}\boldsymbol{U}_i\left\{\text{diag}\left(\lambda_{i1}+\sqrt{\lambda_{i1}^2+\frac{4}{\mu}},\ldots,\lambda_{iR}+\sqrt{\lambda_{iR}^2+\frac{4}{\mu}}\right)\right\}\boldsymbol{U}_i^\top, \tag{12}$$

where $F_\mu(\boldsymbol{\Lambda}_i) = \text{diag}\{f_{i1},\ldots,f_{iR}\}$ with $f_{ir} = \frac{1}{2}\left(\lambda_{ir}+\sqrt{\lambda_{ir}^2+\frac{4}{\mu}}\right)$ for $r = 1,\ldots,R$.

Update $\boldsymbol{\theta}$

Next to update $\boldsymbol{\theta}^{(k)}$ we minimize the objective function (10) with respect to $\boldsymbol{\theta}$, given the current estimates $\boldsymbol{\Omega}_i^{(k)}$ and $\mathbf{A}_i^{(k-1)}$. This update is the same as solving for $\boldsymbol{\theta}$ in logistic regression using the negative log likelihood.

$$\boldsymbol{\theta}^{(k)} = \underset{\boldsymbol{\theta}}{\text{argmin}} \sum_{i=1}^{n}\left[-tr\left\{\boldsymbol{A}_i^{(k-1)}\boldsymbol{G}(\boldsymbol{x}_i;\boldsymbol{\theta})\right\} + \boldsymbol{1}_R^\top \boldsymbol{L}(\boldsymbol{x}_i;\boldsymbol{\theta})\boldsymbol{1}_R\right]. \tag{13}$$

For each node pair (r,r') this becomes:

$$\theta_{r,r'}^{(k)} = \underset{\theta_{r,r'}}{\text{argmin}} \sum_{i=1}^{n}\left[-a_{r,r',i}^{(k-1)}\, g_{r,r'}(\boldsymbol{x}_i;\boldsymbol{\theta}) + log\left(1+exp\left\{g_{r,r'}(\boldsymbol{x}_i;\boldsymbol{\theta})\right\}\right)\right]. \tag{14}$$

If we suppose $g_{r,r'}(\mathbf{x}_i;\boldsymbol{\theta}) = x_i\theta$, such as in the case of logistic regression, we can obtain the logistic regression estimates for $\boldsymbol{\theta}$. Similarly the log-odds can be estimated from SVM or random forest instead of logistic regression. Please refer to models (7) and (8) along with the description of methods in Section 2.3.

Update $\mathbf{A}_i$.

Next to update $\mathbf{A}_i^{(k)}$ we minimize the objective function (10) with respect to $\mathbf{A}_i$, fixing estimates of $\boldsymbol{\Omega}_i^{(k)}$ and $\boldsymbol{\theta}^{(k)}$.

$$\boldsymbol{A}_i^{(k)} = \underset{\boldsymbol{A}_i}{\text{argmin}} \sum_{i=1}^{n}\left[\gamma\,\text{tr}\left\{(2\boldsymbol{A}_i-1)\boldsymbol{\Omega}_i^{(k)}\right\} - tr\left\{\boldsymbol{A}_i\boldsymbol{G}\left(\boldsymbol{x}_i;\boldsymbol{\theta}^{(k)}\right)\right\} + \boldsymbol{1}_R^\top \boldsymbol{L}\left(\boldsymbol{x}_i;\boldsymbol{\theta}^{(k)}\right)\boldsymbol{1}_R\right]. \tag{15}$$

Note that $\mathbf{A}_i \in \{0,1\}^{R\times R}$. For each region pair (r,r'), we minimize this function by comparing the objective function for $a_{i,r,r'} = 0$ and $a_{i,r,r'} = 1$ given the values in $\mathbf{A}_i$ at all other pairs. Equivalently, we can minimize the following for each subject i at each pair (r,r'):

$$a_{i,r,r'}^{(k)} = \underset{a_{i,r,r'}\in\{0,1\}}{\text{argmin}}\left[a_{i,r,r'}\left\{2\gamma\omega_{i,r,r'} - g_{r,r'}\left(\boldsymbol{x}_i,\boldsymbol{\theta}^{(k)}\right)\right\} - \omega_{i,r,r'}\right]. \tag{16}$$

Combining these steps, the full algorithm is presented in Algorithm 1.

Iterate until the estimates converge. The choices of initial values for $\boldsymbol{\Omega}$ and $\mathbf{A}$ and $\boldsymbol{\theta}$ are described in Sections 2.2 and 2.3. Iteratively updating these estimates could improve the predictive performance and power to detect true associations compared to the proposed framework without the joint estimation method.

wThe choice of λ impacts the level of sparsity when estimating $\boldsymbol{\Omega}_i$. Individual level λ_i is used to control the sparsity of the initial estimates of $\boldsymbol{\Omega}_i$ in Section 2.2. We have opted to use the same subject specific λ_i in the joint estimation algorithm, though the results do not change significantly in simulation for a common population level λ. Similarly the ratio between λ and μ impacts the sparsity of the estimate for $\boldsymbol{\Omega}_i$, so μ is chosen to satisfy the desired level of sparsity. γ is used to control the similarity between $\mathbf{A}$ and $\boldsymbol{\Omega}$. For this reason we increase γ over iterations of k, presumably as the two estimates converge towards the same sparsity pattern this parameter enforces that relationship.

ALGORITHM 1 **The joint estimation algorithm to update $\boldsymbol{\Omega}_i$ and $\mathbf{A}_i$, and $\boldsymbol{\theta}$**

Data: $\{\boldsymbol{\Omega}_i^{(1)}, \mathbf{A}_i^{(1)}, \mathbf{x}_i\}_{i=1}^n$; Number of iterations K; Updating rate μ; Penalty term λ and tuning parameter γ.

Result: $\{\boldsymbol{\Omega}_i^{(K)}, \mathbf{A}_i^{(K)}\}_{i=1}^n$

begin

Initialize $\boldsymbol{\Omega}_i^{(1)}$ and $\mathbf{A}_i^{(1)}$

for $k = 1,\ldots,K$ **do**

for $i = 1,\ldots,n$ **do**

$\mathbf{Y}_i^{(k)} = UF_\mu(\Lambda_i)U^\top = \frac{1}{2}U\left(diag\left(\lambda_{ir}+\sqrt{\lambda_{ir}^2+\frac{4}{\mu}}\right)\right)U^\top$

$\boldsymbol{\Omega}_i^{(k)} = \begin{cases} Y_i^{(k)} + Z_i^{(k)} - \frac{\lambda}{\mu} & \text{if } Y_i^{(k)} + Z_i^{(k)} \geq \frac{\lambda}{\mu} \\ Y_i^{(k)} + Z_i^{(k)} + \frac{\lambda}{\mu} & \text{if } Y_i^{(k)} + Z_i^{(k)} \leq \frac{\lambda}{\mu} \\ 0 & \text{otherwise} \end{cases}$

Update $Z_i^{(k)} = Z_i^{(k)} + \mu\,(\boldsymbol{\Omega}_i^{(k)} - Y_i^{(k)})$.

end

for *each pair* (r,r') **do**

Estimate $g_{r,r'}(\mathbf{x};\boldsymbol{\theta}^{(k-1)})$ using logistic regression, SVM, or random forest.

end

for $i = 1,\ldots,n$ **do**

$F_k = \text{tr}\{\gamma\,(2\mathbf{A}_i^{(k)} - 1)\boldsymbol{\Omega}_i^{(k)} - \mathbf{A}_i^{(k)}\mathbf{G}(\mathbf{x}_i;\boldsymbol{\theta}^{(k)})\}$

for *each* (r,r') **do**

if $a_{i,r,r'}^{(k)} = 0$ **then**

if $F_k + \gamma\,\omega_{(r,r')}^{(k)} - \mathbf{x}_i\boldsymbol{\theta}_{r,r'}^{(k)} < F_k$ **then**

$a_{i,r,r'} = 1$

end

else

if $F_k - \gamma\,\omega_{(r,r')}^{(k)} < F_k$ **then**

$a_{i,r,r'} = 0$

end

end

end

end

end

end

3 | SIMULATIONS

3.1 | Performance of the two-step update

The prediction procedure, detailed in Section 2, was evaluated using simulated data based on the real data application. True signals in the clinical variables were simulated by generating a $\boldsymbol{\theta}$ vector and using the observed clinical covariates to generate the corresponding time series data for 346 subjects. The ability to recover the true signals was evaluated by comparing the subset of variables selected by elastic net to those with true nonzero signals in $\boldsymbol{\theta}$. This performance was summarized using sensitivity, specificity, and false discovery rate (FDR). FDR is defined as the proportion of true signals identified out of all selected variables. The performance of the prediction procedure was evaluated using AUC, comparing the ability to correctly classify edges as connected or not.

The data generation process was as follows: simulate $\boldsymbol{\theta}$, use observed clinical variables and simulated $\boldsymbol{\beta}$ to assign connectivity to each edge for each subject, simulate precision matrix $\boldsymbol{\Omega}_i$ from a mixture of normal distributions with mean 3 or -3 and standard deviation 1 for connected edges, ensure $\boldsymbol{\Omega}_i$ is positive definite, and finally simulate T time points for each node of each subject from a multivariate normal distribution with mean zero and covariance $\boldsymbol{\Omega}_i$.

Simulation was also used to understand how characteristics of the fMRI data impacts the Gaussian copula network estimation, implemented using the R package *huge* (Zhao et al., 2015). We considered how a differing number of time points in the fMRI and a different network size affected the ability to recover the connectivity matrix. Table 1 presents the results for each of the settings considered. We compared results using *huge* to other network construction methods, *clime* and *tiger*, and found that *huge* outperformed other existing methods for estimating sparse graphical models (Li et al., 2019). Due to space limitations, we did not report the detailed results here.

Though we do not achieve high power with this process, we consistently see good control in the FDR; edges that are assigned to be connected are typically correctly labeled, and more frequently those incorrectly labeled are among edges that are connected in truth and not connected in the network estimate. This leads to the conclusion that signals identified from this procedure are likely to be true associations, though some true associations may be missed.

In Table 2a–c we present the number of subjects that are correctly and incorrectly identified as having a connection at a given edge, using SVM to predict connectivity. We compare the results to both the true simulated network and the network estimate. This is an example using one edge across 346 subjects.

Table 2 presents the accuracy of the network estimate compared to the true network connectivity in addition to the predicted network compared to the estimated and true network. Table 2a shows that the network estimate has a very low FDR, with no edges incorrectly identified as connected. Using the network estimate to perform the prediction, we can evaluate the performance of the prediction results (Table 2c). This simulation provides evidence that we can be confident in the network estimate and prediction results compared to the true underlying network structure. Though some signals of connected edges are missed, those identified as connected most often are connected in the underlying network.

Simulation was also used to evaluate the ability to recover true signals among the clinical characteristics. When simulating 3814 edges we observed an average FDR of about 11%. Average sensitivity is only 5% and specificity is 99%. The true signal is sparse, among 286 variables there are 20 true nonzero signals randomly selected for each edge. Again we conclude that there is low power in the procedure but good false discovery control. The variables selected are likely to be true associations, although many true signals will be missed. We expect this to be the case when using a relatively large network structure compared to a small number of subjects.

TABLE 2 Simulation results for network estimation using the prediction procedure and *huge*

	True network	
Connectivity matrix estimate	**Number connected**	**Number not connected**
(a) Comparison of network estimated through graphical modeling to the true simulated network		
Number connected	15	0
Number not connected	0	321
(b) Comparison of true simulated network to the predicted connectivity using SVM		
Number connected	7	4
Number not connected	18	317
(c) Comparison of network estimated through graphical modeling to the predicted connectivity using SVM		
Number connected	6	5
Number not connected	9	326

TABLE 1 Results of simulation evaluating how well the connectivity estimation can recover the simulated connectivity network

Number of nodes	Number of time points	% of connected edges recovered	% of edges not connected recovered
50	120	47%	99%
	1000	54%	98%
264	120	12%	99%
	1000	51%	99%
	5000	73%	99%

TABLE 3 Performance of the joint estimation algorithm using different methods to perform the prediction: Logistic regression, SVM, and random forest (RF)

Setting			A estimation			θ estimation		
N	*p*	Method	Sens	Spec	FDR	Sens	Spec	FDR
100	3	Logistic	0.61	0.86	0.50	0.52	0.68	0.21
		SVM	0.57	0.80	0.60	0.23	0.92	0.12
		RF	0.33	0.97	0.37	0.18	0.93	0.12
		two-step	0.20	0.73	0.63	0.00	1.00	0.00
100	5	Logistic	0.73	0.86	0.50	0.64	0.63	0.20
		SVM	0.67	0.77	0.60	0.32	0.88	0.12
		RF	0.43	0.97	0.32	0.24	0.91	0.12
		two-step	0.20	0.74	0.63	0.00	1.00	0.00
300	5	Logistic	0.87	0.86	0.42	0.97	0.48	0.21
		SVM	0.77	0.74	0.60	0.48	0.77	0.19
		RF	0.49	0.98	0.28	0.41	0.87	0.11
		two-step	0.69	0.98	0.12	0.00	1.00	0.00

Notes: Several simulation settings are presented with a different sample size (N) and average effect size of nonzero $\boldsymbol{\Omega}$ (β). Estimation of the adjacency matrix ($\boldsymbol{A}$) is evaluated as an average of the following metrics over 100 iterations: Sensitivity (Sens), specificity (spec), and FDR. Accuracy of the selected clinical characteristics ($\boldsymbol{\theta}$) is measured using sensitivity, specificity, and FDR.

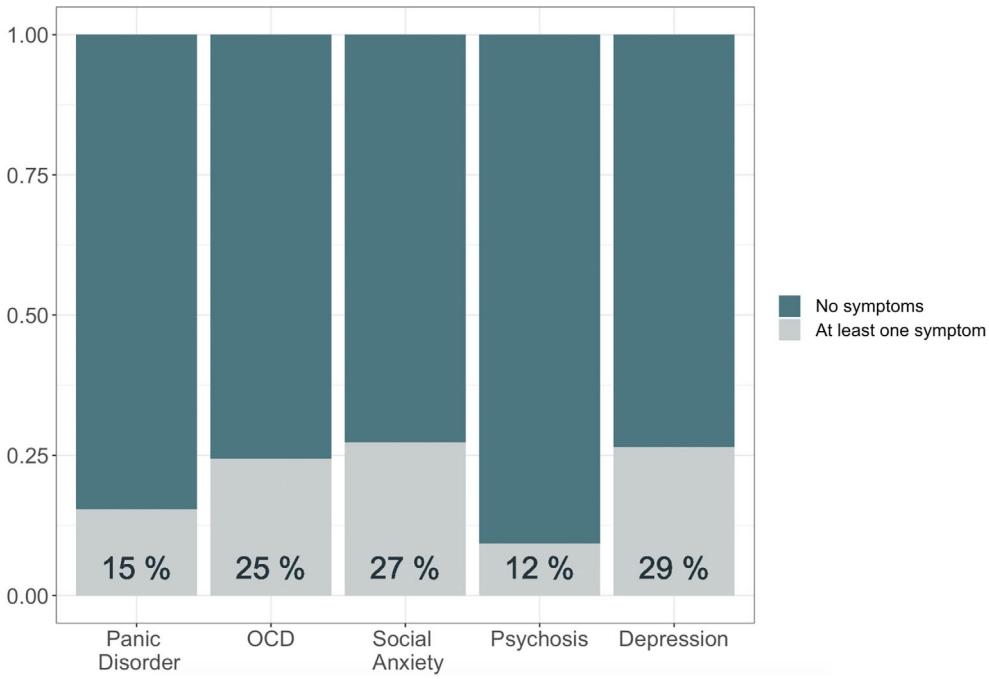

FIGURE 2 Proportion of children in the Philadelphia Neurodevelopmental Cohort experiencing symptoms of psychiatric disorders

3.2 | Performance of the joint estimation algorithm

Potential improvement due to using a joint estimation procedure was quantified through simulation. We considered a toy example for demonstration purposes with a small network with 10 nodes and a sample of 100 subjects. Because the estimates from the two-step update are the initial values for the joint estimation algorithm, the performance depends on how well the two-step estimation performs, detailed in Section 3.1.

In this simulation setting the estimation of both **A** and **θ** are evaluated. Estimation of **A** is evaluated using sensitivity, specificity, and FDR in terms of correctly identifying the connected and not connected edges. Similarly estimation of **θ** is evaluated in terms of sensitivity, specificity, and FDR, but defined in terms of the accuracy of clinical characteristics with nonzero effect estimates.

Table 3 compares three methods using the joint estimation algorithm: logistic regression, SVM, and random forest. Each is also compared to the initial values of the algorithm, obtained via the two-step update using graphical lasso and SVM. The results in Table 3 are the average of 100 iterations of the simulated setting. Increasing the effect size of nonzero signal in the simulation leads to higher sensitivity across all the methods, though it does not reduce FDR of **A** estimation in most cases.

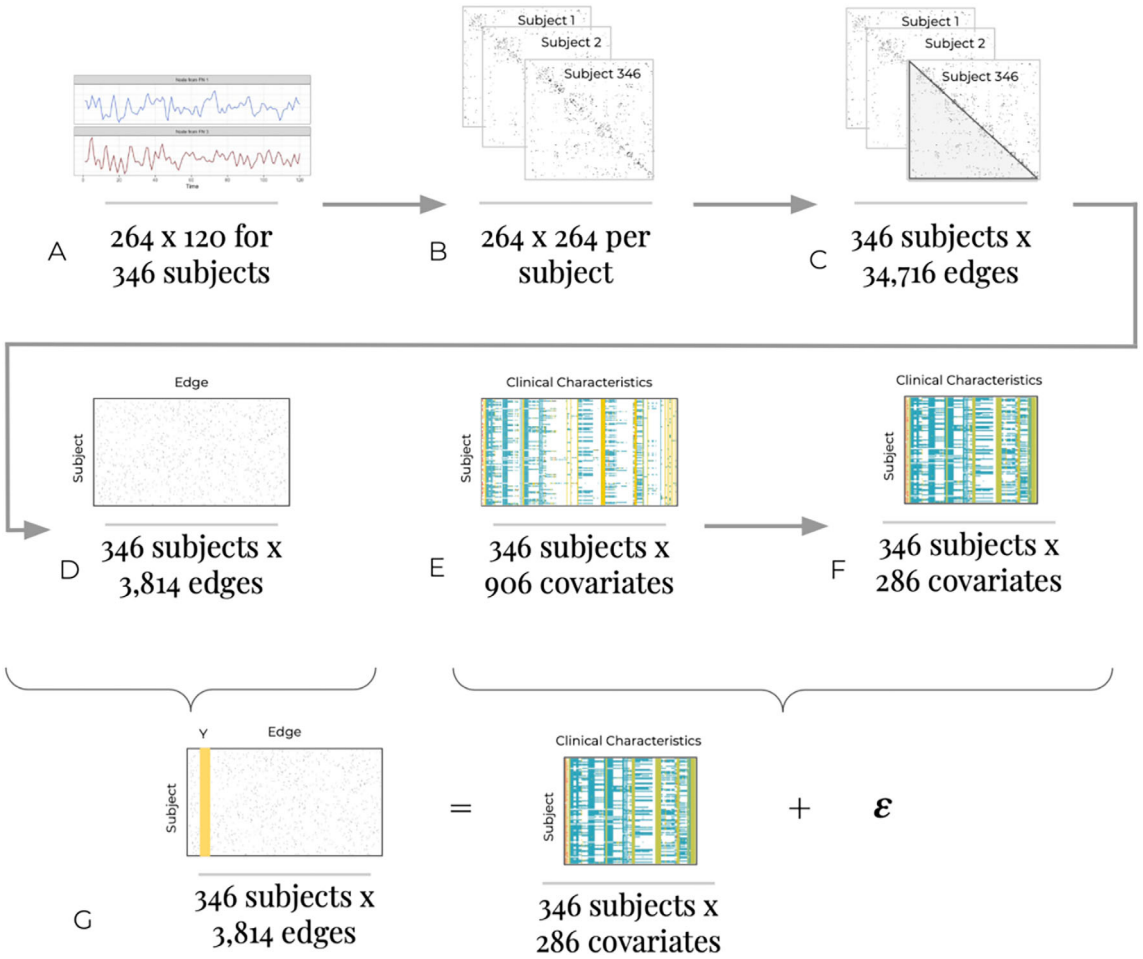

FIGURE 3 Procedure for Philadelphia Neurodevelopmental Cohort data analysis

Increasing the sample size also seems to lead to improved sensitivity and lower FDR in some cases (logistic regression and random forest). In terms of identifying covariates that are truly associated with connectivity patterns ($\boldsymbol{\theta}$ Estimation in Table 3), increasing sample size leads to much higher power with an increase in sensitivity from 0.52 with 100 subjects to 0.97 using 300 subjects when applying logistic regression in the joint estimation algorithm. In terms of the FDR among these covariates, it is not too large with a maximum of about 20% across all methods. SVM and random forest seem to perform better than logistic regression in terms of controlling FDR, which leads to slightly lower power in terms of lower sensitivity. With these simulation settings the two-step update does not identify any variables to be included using elastic net, which demonstrates that incorporating the joint estimation algorithm provides a more powerful approach.

4 | DATA APPLICATION

4.1 | PNC data

Extensive assessment of behavior, life events, demographics and neuropsychological performance was obtained on all subjects in the PNC, in addition to performing resting state fMRIs. The sample has children with mental illnesses as well as healthy individuals, and all subjects underwent a structured neuropsychiatric interview to establish the presence, duration, and effect of multiple psychiatric symptoms (if present) on functioning. The broad range of psychiatric disorders assessed included: depression, mania, simple phobia, social phobia, generalized anxiety, separation anxiety, social anxiety, panic disorder, obsessive–compulsive disorder, posttraumatic stress disorder (PTSD), eating disorders, and psychosis. Figure 2 demonstrates how frequently characteristics of these disorders are observed in the PNC, by presenting the proportion of subjects experiencing at least one symptom of a panic disorder, OCD, social anxiety, psychosis, or depression.

MRI data for the resting state scans were obtained with BOLD-sensitive image acquisitions over 6 minutes, with 120 frames, each 2 s in duration. Voxel size was 3×3 in the transverse plane and 3 in the axial plane, yielding approximately 100,000 voxels in the brain (Satterthwaite et al., 2014). Each voxel constituted a time series. The number of measurements in the time series and the temporal sampling rate are both fixed by the type of MRI performed. Data preprocessing was done to correct for timing differences in the acquisition, realign individual subject scans, and map the images to a common anatomical space so that the images could be combined and analyzed

TABLE 4 Power brain functional modules and associated brain functions (Power et al., 2011)

#	Function	#	Function
1	Sensory/somatomotor Hand	8	Fronto-parietal Task Control
2	Sensory/somatomotor Mouth	9	Salience
3	Cingulo-opercular Task Control	10	Subcortical
4	Auditory	11	Ventral attention
5	Default Mode	12	Dorsal attention
6	Memory Retrieval	13	Cerebellar
7	Visual	−1	Uncertain

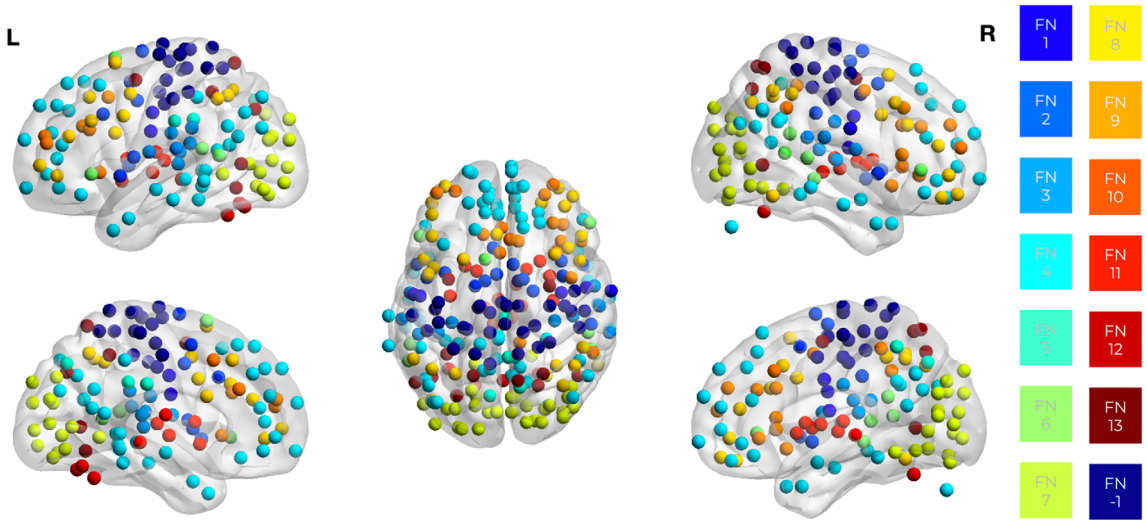

FIGURE 4 Power 264 node spatial parcellation. Each color represents one of the 13 functional brain networks of interest. Generated using BrainNet viewer (Xia et al., 2013)

across subjects. This preprocessing occurred with the pipeline used for the 1000 Functional Connectomes Project.

Subjects with excessive movement (defined as subjects with greater than 0.25 mm volume-to-volume displacement) or poor scan quality were excluded from the analysis, which yielded 500 scans from the original 1442 subjects scanned. Images were also spatially filtered to reduce residual anatomic variability, and for each time voxel, a time series was extracted and bandpass filtered (0.1–0.01 hz) to remove physiological artifact from respirations and heart rate. In addition, regressors for white matter and cerebral spinal fluid were obtained, and variability from these additional sources of noise was removed. From a set of a priori nodes (see Figure 4), a time series from each node (10 mm sphere) was extracted, and a cross-correlation matrix of Pearson *r*-values was obtained and *Z*-transformed for each of the 264 nodes with every other node. The PNC data was obtained from the NCBI database of Genotypes and Phenotypes (dbGaP), a publicly available database. Informed consent was obtained for all subjects who participated in the PNC study; the original publication states "Participants had been previously enrolled in a genomics study at CAG and they and/or their parents had provided informed consent (assent) to be re-contacted for participation in additional studies such as this one. The institutional review boards of both the University of Pennsylvania and the Children's Hospital of Philadelphia approved all study procedures."

4.2 | Analysis pipeline for PNC data analysis

Preprocessing of the PNC imaging data was done to reduce bias from motion and other known confounders. Additional steps were taken to reduce the sample and network size based on missingness and variability. The entire procedure using the PNC data is summarized in Figure 3.

Figure 3 shows the steps performed to manipulate the data into a workable form, as well as steps for data reduction. For each subject, the rfMRI data was converted from region-level time series (Panel A) into binary connectivity indicators for all region pairs using the method described in Section 2.2 (Panel B), that is, the subject-level adjacency matrices to define the edges in the brain networks. Then the connectivity estimates were stacked across subjects (Panel C). We reduced the brain network to edges with sufficient variability for modeling, removing those connected in fewer than 5% of subjects (Panel D). For the clinical characteristics we reduced the subset of potential variables based on missingness; if a given variable was missing in more than 5% of subjects it was removed from the analysis (Panels E and F). This conveniently reduced the number of variables (p) from 906 to 286 which is fewer than the sample size (n), reducing to an $n > p$ problem. Lastly Panel G depicts the final model fitted associating the connectivity across subjects at a given edge and the clinical characteristics.

4.3 | Brain networks of interest

For the analysis that follows, we used an a priori anatomic parcellation of the brain, which utilized 264 nodes organized into 13 different functional modules (FM, Table 4), identified by Power et al. in the 2011 paper *Functional Network Organization of the Human Brain* (Power et al., 2011).

The networks, made up of several regions, identified in this paper align well with others proposed, such as the default mode network (Greicius et al., 2003), dorsal and ventral attention (Fox et al., 2006), and fronto-parietal task control. The authors classified the remaining subnetworks by associated functions, including visual processing, memory, sensory and motor control, auditory, and somatosensory. Compared to voxel-based approaches to connectivity, these networks should minimize connectivity contributions from image smoothness, which causes adjacent voxels in an image to have very high correlation coefficients, irrespective of functional connections. Compared to anatomically defined nodes, such as automated anatomic labeling (Tzourio-Mazoyer et al., 2002) these units more likely reflect intrinsic functional organization in the brain and may be more meaningful probes of functional brain networks.

Figure 4 shows the location of the nodes in the power parcellation, and colors indicate membership to the 13 identified functional networks.

4.4 | Predictability of network

The ability to predict binary connectivity of the entire functional brain network using clinical characteristics was evaluated in addition to the ability to predict connectivity within the subnetworks of interest. We found that the ability to capture the entire brain network was very limited using the machine learning methods tested, SVM and random forest. Fivefold cross-validation was used to evaluate the predictive performance, meaning 80% of subjects were used for estimation and to obtain predictions for the remaining 20% of subjects. The average performance of predicting a connection across the entire brain was no better than choosing at random, with an average AUC of 0.5 across the five cross-validation folds. However, when looking within specific regions we do see fairly high predictive performance for some. In the final iteration of the joint estimation algorithm, only 480 edges had any clinical variables selected for inclusion in a model, so the following results reflect the performance within that subset of the network. Note that the "truth" for computation of AUC was defined as the estimate of **A** from the joint estimation algorithm before updating with random forest or SVM results, since we do not know the true underlying connectivity as we did in the simulations.

We observed the greatest ability to detect connected edges on average in the subcortical network (10) using random forest. Table 5 contains the average AUC, range of AUC from first to the third quartile, and the maximum AUC for a given edge within each network for both methods, random forest, and SVM. Though on average the AUC within some networks was not high, we did have good performance (AUC > 0.95) for some edges in the following functional networks: somatomotor hand, auditory, default mode, visual, fronto-parietal task control, salience, and subcortical. These networks had the highest maximum AUC observed, but the most edges with good performance were in the somatomotor hand, default mode, and visual networks. The variables selected in the model with high predictive performance (highest AUC) for connectivity in the sensory and somatomotor hand include: race, indicator of liver disease, indicator of infectious disease, four questions from the Structured Interview of Psychosis-risk Syndromes survey, an indicator of social anxiety, and questions from the Penn Age Differentiation Test, Penn Emotion Differentiation Test, and Visual Object Learning Test.

One model with AUC >0.95 predicting connectivity for an edge from the visual network includes the following variables: indicator of vision problems and two components of the 1-Back trials. These

TABLE 5 Mean AUC for edges contained within each functional module from two methods, SVM and random forest

	SVM		Random forest	
Functional brain module	**Mean (Q1 – Q3)**	**Max**	**Mean (Q1 – Q3)**	**Max**
1 Sensory/somatomotor Hand	0.45 (0.45, 0.52)	0.59	0.67 (0.52, 0.89)	1.00
2 Sensory/somatomotor Mouth	0.34 (0.32, 0.40)	0.40	0.54 (0.38, 0.62)	0.86
3 Cingulo-opercular Task Control	0.44 (0.44, 0.48)	0.60	0.60 (0.55, 0.73)	0.76
4 Auditory	0.47 (0.46, 0.50)	0.69	0.69 (0.57, 0.80)	1.00
5 Default Mode	0.47 (0.45, 0.53)	0.59	0.63 (0.48, 0.77)	1.00
6 Memory retrieval	0.49 (0.49, 0.49)	0.49	0.45 (0.45, 0.45)	0.45
7 Visual	0.45 (0.41, 0.51)	0.57	0.64 (0.52, 0.70)	1.00
8 Fronto-parietal Task Control	0.49 (0.46, 0.51)	0.60	0.54 (0.43, 0.55)	1.00
9 Salience	0.41 (0.44, 0.50)	0.57	0.69 (0.49, 0.85)	1.00
10 Subcortical	0.57 (0.55, 0.59)	0.59	**0.83 (0.66, 1.00)**	1.00
11 Ventral Attention	0.40 (0.35, 0.46)	0.49	0.66 (0.55, 0.76)	0.79
12 Dorsal Attention	0.49 (0.44, 0.54)	0.56	0.60 (0.41, 0.71)	0.83
13 Cerebellar	0.54 (0.54, 0.54)	0.54	0.44 (0.44, 0.44)	0.44

TABLE 6 Variables selected for the model with AUC = 1.00 (using random forest) from network 1, associated with sensory somatomotor hand control

Variable	Description	Estimate
Race	Self-reported ethnicity of participant (EA, AI)	−0.63
MED807	Liver disease	0.33
MED809	Infectious disease	−4.25
SIP015	SIPS feeling odd things going on	−0.09
SIP016	SIPS feeling able to predict the future	−0.07
SIP018	SIPS Feeling different due to superstitions	−0.10
SOC001	Feeling afraid in social settings	−0.04
PADT	Penn Age Differentiation Test	−0.34
PADT	Number of correct responses with no age difference	−0.01
PEDT	Penn Emotion Differentiation Test	−1.30
VOLT	Visual Object Learning Test	0.18

Abbreviations: PADT, Penn age differentiation test; PEDT, Penn emotion discrimination test; SIPS, structured interview of psychosis-risk syndromes; VOLT, visual object learning test.

TABLE 7 Variables selected for the model with AUC = 1.00 (using random forest) from network 7, associated with visual processing

Variable	Description	Estimate
MED622	Vision problems	−0.75
PFMT	Penn Face Memory Test	−0.18
PEIT	Penn Emotion Identification	−1.09
LNB	Number of correct responses to 1-back trials	0.12
LNB	Number of incorrect responses to 1-back trials	−0.05

Abbreviations: PFMT, Penn face memory test; PEIT, Penn emotion identification test; LNB, Penn letter N-Back test which tests working memory.

variables are expected to be associated with activation in the visual processing network and confirm these known associations.

The variables selected for two models with the best predictive performance are presented in Tables 6 and 7.

Several variables were frequently selected across models with good performance of classifying connected nodes, these include: indicator of ear/nose/throat problems, indicator of metabolic disease, separation anxiety, having thoughts of suicide, and results from the Penn Conditional Exclusion Test. The Penn Conditional Exclusion Test is designed to assess executive functioning ability.

Figure 5 presents the proportion of models for each functional module that selected at least one of the clinical characteristics in the groups. The groups and abbreviations are as follows: demographics, overall medical metrics (health), Attention Deficit Disorder (ADD), Conduct Disorder (CDD), Depression (DEP), Eating Disorder (EAT), Generalized Anxiety Disorder (GAD), Children's Global Assessment Scale (GAF), Mania/Hypomania (MAN), other medical conditions (MED), Obsessive Compulsive Disorder (OCD), Oppositional Defiant Disorder (ODD), Panic Disorder (PAN), Specific Phobia (PHB), Psychosis (PSY), Post-Traumatic Stress (PTD), general probes about counseling and emotions (SCR), Separation Anxiety (SEP), Structured Interview for Prodromal Symptoms (SIP), Social Anxiety (SOC), Suicide (SUI), Penn Age Differentiation Test (PADT), Penn Facial Memory Test (PFMT), Penn Emotion Identification Test (PEIT), Penn Word Memory Test (PWMT), Penn Verbal Reasoning Test (PVRT), Penn Emotion Differentiation Test (PEDT), Penn Matrix Reasoning Test (PMAT), Tap hand trials (TAP), Visual Object Learning Test (VOLT), Letter N-Back test (LNB), Penn Conditional Exclusion Test (PCET), Penn Continuous Performance Test (PCPT), Penn Line Orientation Test (PLOT), Wide Range Assessment Test (WRAT), and Penn Motor Praxis Test (MP). These variable groupings were defined by the surveys and tools used to collect the data in the PNC study.

4.5 | Association analysis

In addition to predicting connectivity in the network, our method can also be applied to the more common goal of understanding characteristics or symptoms associated with levels of connectivity in the resting state brain. This is the diagnostic task of, given a label, what patterns of connectivity distinguish those with a label from those without a label. One of the main goals of the PNC study was to establish associations between brain development and psychiatric diseases.

We can build a brain network associated with one disease or symptom of interest. For example, the network built of edges associated with psychosis spans all 13 functional networks; for specific diseases we may be interested in how connectivity of a specific network, rather than the entire brain, relates to the disorder of interest. Figure 6 shows the edges within the fronto-parietal task control network that are associated with psychosis variables (variables from the SIPS and psychosis surveys).

When we consider PTSD, associated connections span four behavioral networks. Connections associated with PTSD span four behavioral networks: cingulo-opercular task control, memory retrieval, visual, and fronto-parietal task control. Figure 7 presents these connections across the four functional networks associated with PTSD.

5 | DISCUSSION

In this work, we present the framework for predicting functional connectivity with clinical characteristics and demonstrate that it is feasible to predict some subnetworks in the brain. Many clinical characteristics identified in the PNC application are consistent with previous findings, as the simulation results suggest the ability to identify some true signals among clinical characteristics. Other findings suggest novel associations between behavioral measures and brain networks. By taking an atheoretical approach and searching across a large clinical parameter space, the analysis turned up both expected and unexpected findings.

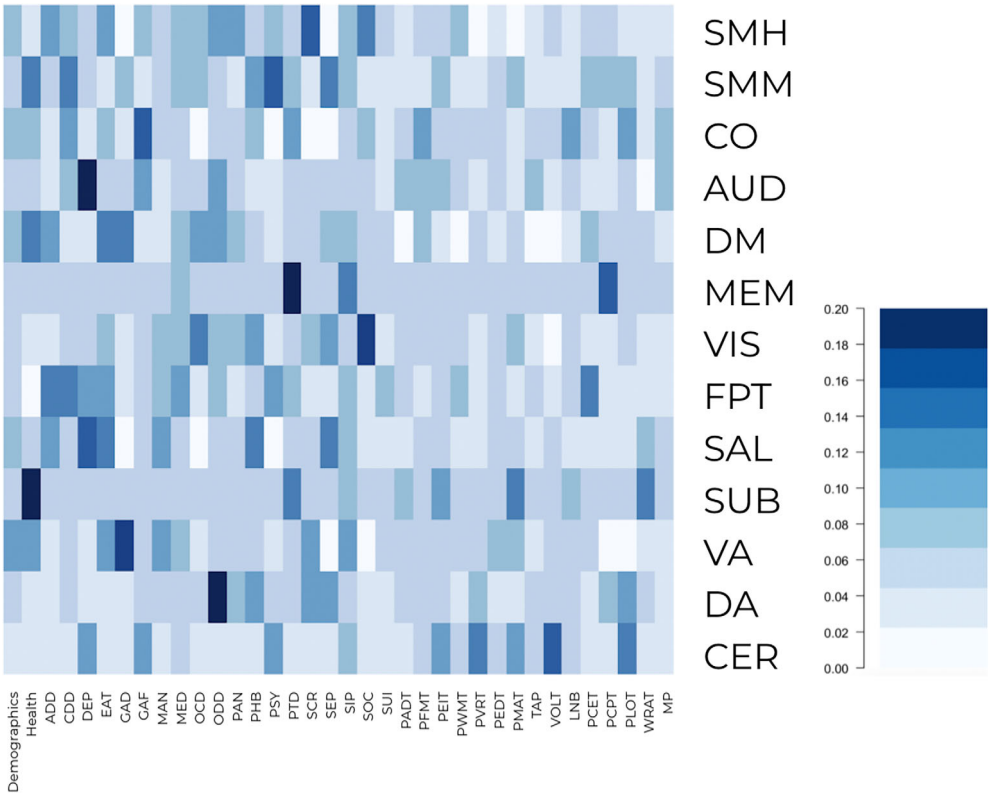

FIGURE 5 Proportion of models which a variable in the group was selected across edges in each functional module

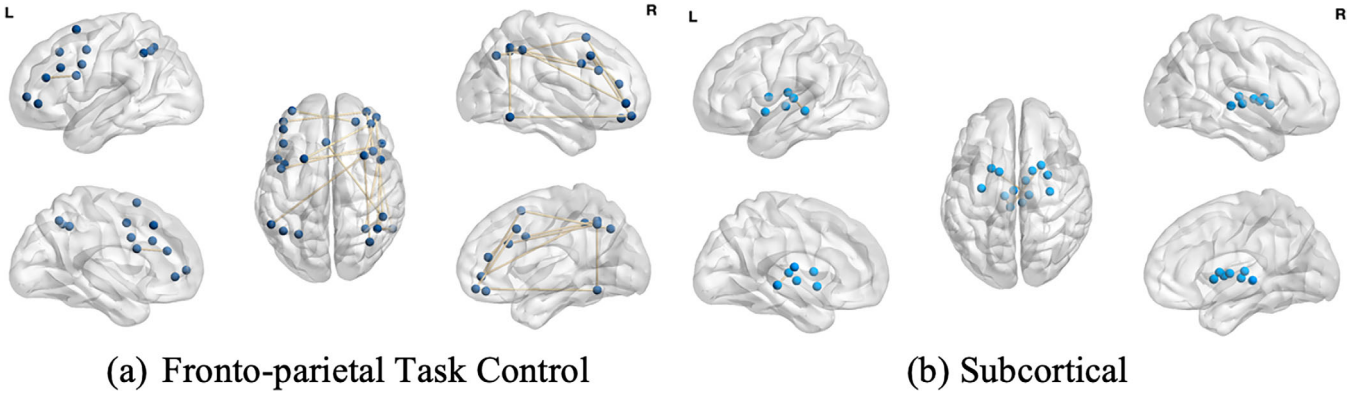

FIGURE 6 Edges associated with SIPS or psychosis variables within fronto-parietal task control and subcortical functional networks

The PNC, as a large community cohort, contains a range of psychopathology, measured in the clinical assessment tools. Most of the participants were not engaged in active treatment for psychiatric conditions, although, like any large, community sample, significant amounts of psychopathology exists within this sample. Thus, while the results are not entirely comparable with selected cohorts of psychiatric syndromes, other analyses of the PNC data indicates that similar patterns of aberrant connectivity are found in PNC youth who meet criteria of psychosis-risk conditions as those with psychosis (Satterthwaite et al., 2015). Accordingly, the networks associated with psychosis spectrum, as in Figure 6 in our analysis, identified some of the commonly reported areas of dysfunction in psychosis, such as fronto-parietal and subcortical networks. The fronto-parietal networks, linked with executive control functions and working memory, are reported to be deficient in psychotic disorders (Minzenberg et al., 2009; Sheffield et al., 2015). The subcortical networks, specifically the striatum, heavily enervated by dopaminergic projections from the midbrain, have also been implicated in the pathophysiology of schizophrenia (Clinton & Meador-Woodruff, 2004; Howes & Kapur, 2009; Weinstein et al., 2017). In addition to these expected findings, we also found several predictive clinical variables in the sensorimotor hand subnetwork, such as subclinical psychosis symptoms and cognitive performance measures. Although the sensorimotor hand network is not commonly associated with psychopathology, emerging data have begun to report that serious mental disorders, such as schizophrenia and bipolar disorder, show imbalances in this network (Anticevic et al., 2014; Martino et al., 2020). This finding provides an example

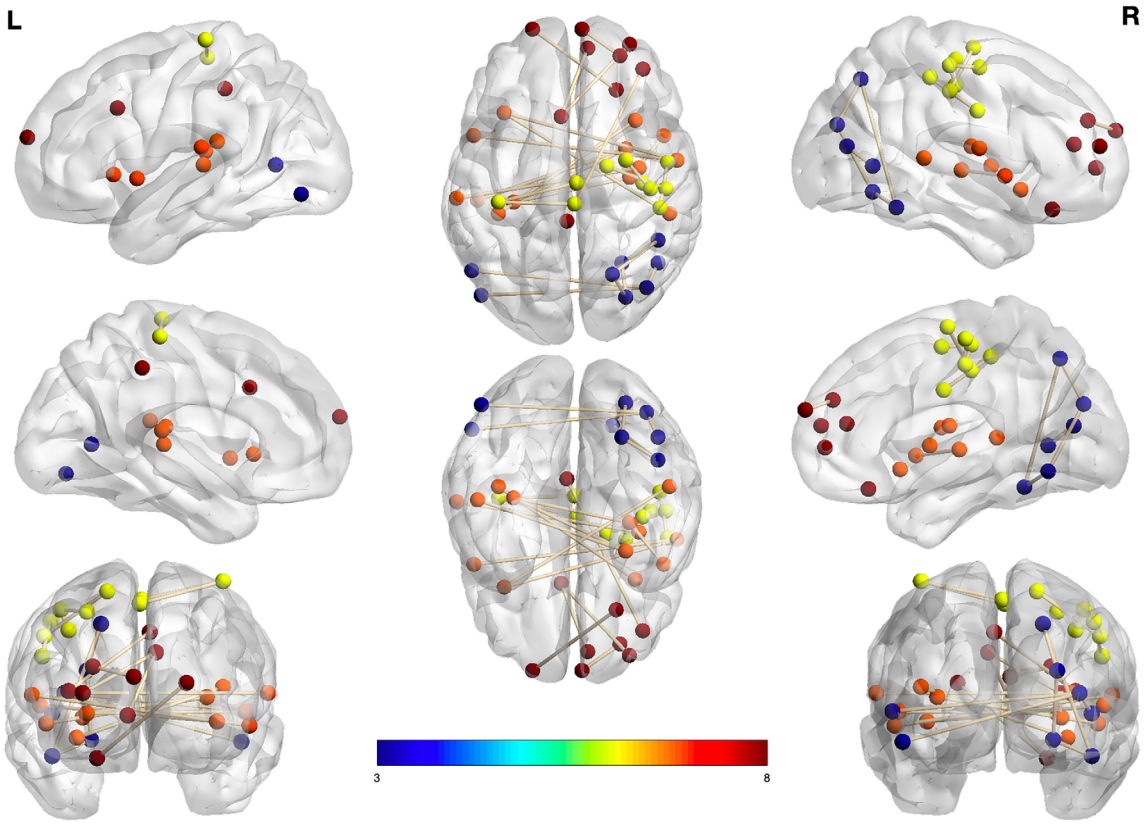

FIGURE 7 Network for indicator of posttraumatic stress disorder. Nodes colored by corresponding functional networks

where this technique might identify new associations between psychopathology and network dysfunction.

In the association of edges with the PTSD symptoms, we found a pattern consistent with previous research showing altered connectivity within and between subnetworks of the brain when compared to control patients (Bao et al., 2021; Schmalfuss, 2009; Sripada et al., 2012). In the prediction analysis, the MEM subnetwork was particularly well-represented when the PTSD variables were used to predict the edges. In the Power parcellation (Power et al., 2011) we used to extract the connectome for the analysis, this MEM, or memory retrieval, subnetwork corresponds to the posterior DMN in parcellations, strongly implicated in PTSD, possibly linked to aberrant memory formation around traumatic events (Bao et al., 2021; Schmalfuss, 2009; Sripada et al., 2012).

One observation about our findings is the highly distributed nature of the relationships uncovered between clinical and neural variables. This observation accords with reports that dysconnectivity in psychopathology encompasses multiple networks, although there are some subnetworks that may have diagnostic specificity (Brandl et al., 2019; Doucet et al., 2020). We also used very broad categories, for example, demographics and medical problems, opting for a broad search of the parameter space. This generated some associations that might not be directly relevant to psychopathology. Our simulations showed that the FDR was relatively well-controlled, so the associations we found can be considered relatively reliable, but the stimulation also showed that sensitivity was low, so that it is likely that we missed an unknown proportion of associations. Hence, more subtle connections that may have been more specific to a diagnosis may have been missed.

While the field of psychiatry has traditionally relied on diagnostic categories, dating back to clinical observations made in the late 19th century, there has been increasing recognition for new methods of classification, such as theoretical constructs derived from neuroscience research (Insel et al., 2010). The framework illustrated here is an atheoretical, purely data-driven approach that, lacking constraining assumptions (within the subnetwork parcellation we chose to test), has the potential to provide new insights into brain and behavior correlations. Though it is difficult to detect associations with relatively few subjects compared to network size and number of clinical characteristics, this work provides a way to identify subnetworks that are predictive given clinical characteristics of interest.

Future extensions of this work may involve making valid statistical inferences on parameters and predictions. In addition, changing the graph estimation procedure or voxel level summary could influence the ability to predict connectivity. These extensions may provide further insight into the relationship between functional brain connectivity and behavioral phenotypes, relevant for psychopathology.

ACKNOWLEDGMENTS

This work was partially supported by the NIH grants R01 MH105561 (Kang, Morris), R01 DA048993 (Kang, Morris) and R01 GM124061 (Kang).

CONFLICT OF INTEREST

The authors have declared no conflicts of interest.

DATA AVAILABILITY STATEMENT

PNC data were obtained from the NCBI database of Genotypes and Phenotypes (dbGaP) and are available to researchers through the repository. Code for this analysis is available as an R package at https://github.com/EmilyLMorris/ImagingPred.

ORCID

Emily L. Morris https://orcid.org/0000-0001-7357-0483

REFERENCES

Anticevic, A., Cole, M. W., Repovs, G., Murray, J. D., Brumbaugh, M. S., Winkler, A. M., Savic, A., Krystal, J. H., Pearlson, G. D., & Glahn, D. C. (2014). Characterizing thalamo-cortical disturbances in schizophrenia and bipolar illness. *Cerebral Cortex*, *24*(12), 3116–3130. https://doi.org/10.1093/cercor/bht165

Bao, W., Gao, Y., Cao, L., Li, H., Liu, J., Liang, K., Xinyue, H., Zhang, L., Xinyu, H., Gong, Q., & Huang, X. (2021). Alterations in large-scale functional networks in adult posttraumatic stress disorder: A systematic review and meta-analysis of resting-state functional connectivity studies. *Neuroscience & Biobehavioral Reviews*, *131*, 1027–1036.

Brandl, F., Avram, M., Weise, B., Shang, J., Simoes, B., Bertram, T., Ayala, D. H., Penzel, N., Gursel, D. A., Bauml, J., Wohlschlager, A. M., Vukadinovi, Z., Koutsouleris, N., Leucht, S., & Sorg, C. (2019). Specific substantial dysconnectivity in schizophrenia: A transdiagnostic multimodal meta-analysis of resting-state functional and structural magnetic resonance imaging studies. *Biological Psychiatry*, *85*(7), 573–583.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). *Classification and regression trees*. Routledge.

Clinton, S. M., & Meador-Woodruff, J. H. (2004). Thalamic dysfunction in schizophrenia: Neurochemical, neuropathological, and in vivo imaging abnormalities. *Schizophrenia Research*, *69*(2–3), 237–253. https://doi.org/10.1016/j.schres.2003.09.017

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Doucet, G. E., Janiri, D., Howard, R., O'Brien, M., Andrews-Hanna, J. R., & Frangou, S. (2020). Transdiagnostic and disease-specific abnormalities in the default-mode network hubs in psychiatric disorders: A meta-analysis of resting-state functional imaging studies. *European Psychiatry*, *63*(1), E57. https://doi.org/10.1192/j.eurpsy.2020.57

DuBois Bowman, F., Zhang, L., Derado, G., & Chen, S. (2012). Determining functional connectivity using fmri data with diffusion-based anatomical weighting. *NeuroImage*, *62*(3), 1769–1779.

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, *103*(26), 10046–10051. https://doi.org/10.1073/pnas.0604187103

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, *100*(1), 253–258. https://doi.org/10.1073/pnas.0135058100

Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(13), 4637–4642. https://doi.org/10.1073/pnas.0308627101

Howes, O. D., & Kapur, S. (2009). The dopamine hypothesis of schizophrenia: Version III – Tthe final common pathway. *Schizophrenia Bulletin*, *35*(3), 549–562. https://doi.org/10.1093/schbul/sbp006

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (Eds.). (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*(7), 748–751.

Jian Kang, F., Bowman, D. B., Mayberg, H., & Liu, H. (2016). A depression network of functionally connected regions discovered via multi-attribute canonical correlation graphs. *NeuroImage*, *141*, 431–441.

Kessler, D., Angstadt, M., & Sripada, C. (2016). Growth charting of brain connectivity networks and the identification of attention impairment in youth. *JAMA Psychiatry*, *73*(5), 481–489. https://doi.org/10.1001/jamapsychiatry.2016.0088

Li, X., Zhao, T., Wang, L., Yuan, X., & Liu, H. (2019). *flare: Family of lasso regression*. R package version 1.6.0.2. https://CRAN.R-project.org/package=flare

Liu, H., Lafferty, J., & Wasserman, L. (2009). The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. *Journal of Machine Learning Research*, *10*(10), 2295–2328.

Martino, M., Magioncalda, P., Conio, B., Capobianco, L., Russo, D., Adavastro, G., Tumati, S., Tan, Z., Lee, H.-C., Lane, T. J., Amore, M., Inglese, M., & Northoff, G. (2020). Abnormal functional relationship of sensorimotor network with neurotransmitter-related nuclei via subcortical-cortical loops in manic and depressive phases of bipolar disorder. *Schizophrenia Bulletin*, *46*(1), 163–174.

Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S., & Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of General Psychiatry*, *66*(8), 811–822.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L., & Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, *72*(4), 665–678. https://doi.org/10.1016/j.neuron.2011.09.006

Satterthwaite, T. D., Connolly, J. J., Ruparel, K., Calkins, M. E., Jackson, C., Elliott, M. A., Roalf, D. R., Hopson, R., Prabhakaran, K., Behr, M., Qiu, H., Mentch, F. D., Chiavacci, R., Sleiman, P. M. A., Gur, R. C., Hakonarson, H., & Gur, R. E. (2016). The Philadelphia neurodevelopmental cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage*, *124*, 1115–1119. https://doi.org/10.1016/j.neuroimage.2015.03.056

Satterthwaite, T. D., Elliott, M. A., Ruparel, K., Loughead, J., Prabhakaran, K., Calkins, M. E., Hopson, R., Jackson, C., Keefe, J., Riley, M., Mentch, F. D., Sleiman, P., Verma, R., Davatzikos, C., Hakonarson, H., Gur, R. C., & Gur, R. E. (2014). Neuroimaging of the philadelphia neurodevelopmental cohort. *Neuroimage*, *86*, 544–553. https://doi.org/10.1016/j.neuroimage.2013.07.064

Satterthwaite, T. D., Vandekar, S. N., Wolf, D. H., Bassett, D. S., Ruparel, K., Shehzad, Z., Craddock, R. C., Shinohara, R. T., Moore, T. M., Gennatas, E. D., Jackson, C., Roalf, D. R., Milham, M. P., Calkins, M. E., Hakonarson, H., Gur, R. C., & Gur, R. E. (2015). Connectome-wide network analysis of youth with psychosis-spectrum symptoms. *Molecular Psychiatry*, *20*(12), 1508–1515.

Schmalfuss, I. M. (2009). Petrous apex. *Neuroimaging Clinics*, *19*(3), 367–391.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Sheffield, J. M., Repovs, G., Harms, M. P., Carter, C. S., Gold, J. M., MacDonald, A. W., III, Ragland, J. D., Silverstein, S. M., Godwin, D., & Barch, D. M. (2015). Fronto-parietal and cingulo-opercular network integrity and cognition in health and schizophrenia. *Neuropsychologia*, *73*, 82–93. https://doi.org/10.1016/j.neuropsychologia.2015.05.006

Sheline, Y. I., Price, J. L., Yan, Z., & Mintun, M. A. (2010). Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. *Proceedings of the National Academy of Sciences*, *107*(24), 11020–11025. https://doi.org/10.1073/pnas.1000446107

Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., Nichols, T. E., Robinson, E. C., Salimi-Khorshidi, G., Woolrich, M. W., Barch, D. M., Ugurbil, K., & Van Essen, D. C. (2013). Functional connectomics from resting-state fMRI. *Trends in Cognitive Sciences*, *17*(12), 666–682. https://doi.org/10.1016/j.tics.2013.09.016

Solo, V., Poline, J.-B., Lindquist, M. A., Simpson, S. L., Bowman, F. D. B., Chung, M., & Cassidy, B. (2018). Connectivity in fMRI: Blind spots and breakthroughs. *IEEE Transactions on Medical Imaging*, *37*(7), 1537–1550. https://doi.org/10.1109/TMI.2018.2831261

Sripada, R. K., King, A. P., Garfinkel, S. N., Wang, X., Sripada, C. S., & Robert, C. (2012). Altered resting-state amygdala functional connectivity in men with posttraumatic stress disorder. *Journal of Psychiatry & Neuroscience*, *37*(4), 241–249. https://doi.org/10.1503/jpn.110069

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Weinstein, J. J., Chohan, M. O., Slifstein, M., Kegeles, L. S., Moore, H., & Abi-Dargham, A. (2017). Pathway-specific dopamine abnormalities in schizophrenia. *Biological Psychiatry*, *81*(1), 31–42. https://doi.org/10.1016/j.biopsych.2016.03.2104

Xia, C. H., Ma, Z., Ciric, R., Gu, S., Betzel, R. F., Kaczkurkin, A. N., Calkins, M. E., Cook, P. A., de la Garza, A. G., Vandekar, S. N., Cui, Z., Moore, T. M., Roalf, D. R., Ruparel, K., Wolf, D. H., Davatzikos, C., Gur, R. C., Gur, R. E., Shinohara, R. T., … Satterthwaite, T. D. (2018). Linked dimensions of psychopathology and connectivity in functional brain networks. *Nature Communications*, *9*(3003), 1–14. https://doi.org/10.1038/s41467-018-05317-y

Xia, M., Wang, J., & He, Y. (2013). BrainNet viewer: A network visualization tool for human brain connectomics. *PLoS One*, *8*(7), e68910.

Zhao, T., Li, X., Liu, H., Roeder, K., Lafferty, J., & Wasserman, L. (2015). *Huge: High-dimensional undirected graph estimation*. R package version 1.2.7. https://CRAN.R-project.org/package=huge

AUTHOR BIOGRAPHIES

Emily L. Morris (emorrisl@umich.edu) is a PhD candidate in the Department of Biostatistics at the University of Michigan Ann Arbor, MI.

Stephan Taylor, MD (sftaylor@med.umich.edu) is Professor and Associate Chair for Research and Research Regulatory Affairs in the Department of Psychiatry.

Jian Kang (jiankang@umich.edu) is Professor in the Department of Biostatistics at the University of Michigan, Ann Arbor, MI.

How to cite this article: Morris, E. L., Taylor, S. F., & Kang, J. (2022). On predictability of individual functional connectivity networks from clinical characteristics. *Human Brain Mapping*, *43*(17), 5250–5265. https://doi.org/10.1002/hbm.26000