



Published in final edited form as:

Neuroimage. 2013 July 15; 75: 165–175. doi:10.1016/j.neuroimage.2013.02.054.

Inferring functional connectivity in MRI using Bayesian network structure learning with a modified PC algorithm

Swathi Iyer¹, Izhak Shafran⁴, David Grayson^{1,5}, Kathleen Gates⁶, Joel Nigg², and Damien Fair^{1,2,3,*}

¹Department of Behavioral Neuroscience, Oregon Health & Science University, Portland

²Department of Psychiatry, Oregon Health & Science University, Portland ³Advanced Imaging

Research Center, Oregon Health & Science University, Portland ⁴Department of Biomedical Engineering, Oregon Health & Science University, Portland ⁵Center for Neuroscience, University of California – Davis, Davis, CA ⁶Virginia Polytechnic Institute and State University Psychology Department and Arlington Innovation Center

Abstract

Resting state functional connectivity MRI (rs-fcMRI) is a popular technique used to gauge the functional relatedness between regions in the brain for typical and special populations. Most of the work to date determines this relationship by using Pearson's correlation on BOLD fMRI timeseries. However, it has been recognized that there are at least two key limitations to this method. First, it is not possible to resolve the direct and indirect connections/influences. Second, the direction of information flow between the regions cannot be differentiated. In the current paper, we follow-up on recent work by Smith et al (2011), and apply a Bayesian approach called the PC algorithm to both simulated data and empirical data to determine whether these two factors can be discerned with group average, as opposed to single subject, functional connectivity data. When applied on simulated individual subjects, the algorithm performs well determining indirect and direct connection but fails in determining directionality. However, when applied at group level, PC algorithm gives strong results for both indirect and direct connections and the direction of information flow. Applying the algorithm on empirical data, using a diffusion-weighted imaging (DWI) structural connectivity matrix as the baseline, the PC algorithm outperformed the direct correlations. We conclude that, under certain conditions, the PC algorithm leads to an improved estimate of brain network structure compared to the traditional connectivity analysis based on correlations.

Keywords

fMRI; Bayesian Network; PC algorithm; directed functional connectivity; effective connectivity

© 2013 Elsevier Inc. All rights reserved

*Corresponding authors: Damien A. Fair, PA-C, Ph.D. Oregon Health and Science University Behavioral Neuroscience Department 3181 SW Sam Jackson Park Road HRC5D30 Portland, Oregon 97239 Office: 503-418-0995 faird@ohsu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Resting-state functional connectivity MRI (rs-fcMRI), measures intrinsic, temporally correlated low-frequency BOLD signals in subjects at rest (not performing an explicit task) (Biswal et al., 1995). The measurement has shown increased popularity in recent years to examine the functional relatedness of independent brain regions and various forms of network phenomena due to its relatively straightforward implementation in clinical and non-clinical samples.

However, while rs-fcMRI has become a useful tool to examine brain organization there are at least two important considerations regarding typical approaches to operationalize functional connectivity that indicates need for improvement. These stem from the fact that, many investigations to date simply apply Pearson's correlations to determine whether two regions are “functionally connected”. While this measurement of functional relatedness has offered powerful insights into brain organization (Fox and Raichle 2007; Van Dijk et al., 2010), there are two clear limitations. One limitation relates to the ability to differentiate direct and indirect influences. The second limitation pertains to its inability to clarify the direction of information flow (i.e. directed functional connectivity (Stephan and Friston 2010)).

As an illustration of indirect relationships between two nodes we use an example of Flu and Hay fever. Because flu and hay fever often co-occur during the same seasons of the year, they are correlated, leading a naïve observer to potentially conclude that hay-fever causes flu, or that flu causes hay-fever. In other words, the two separate entities are directly linked (or in our case functionally connected) as illustrated in Figure 1. However, these constructs are not directly linked. Hay fever doesn't cause the flu, nor does the flu cause Hay fever. It's only through a shared cause, in this case seasonal changes, that Hay fever and the Flu are correlated or connected statistically (Koller and Friedman 2009).

It is likely, and indeed has already been demonstrated in both human (Habeck and Moeller 2011; Honey et al., 2009; Zalesky, Fornito, Bullmore 2012), and animal work (Vincent et al., 2007) that these sorts of indirect influences account for some of the functional connectivity measurements determined with traditional fcMRI methods. For example, Vincent et al (2007) demonstrated how segments of the left and right primary visual cortex, known to lack direct anatomical connections (Van Essen, Newsome, Bixby 1981), show strong functional connectivity. The authors show that this relationship is likely mediated through polysynaptic pathways in the brain. As Hay fever and Flu co-occur seasonally, they might be thought to be directly linked to each other as shown in figure (a). However, the reason why they co-occur, or correlate is because of a common third factor, which is seasonal change in this case as shown in figure (b). Traditional correlation methods when examining functional connectivity fail account for similar dynamics in the brain as diagrammed in figure 2.

Figure 2 provides a schematic of the likely phenomenon.

With regard to directed functional connectivity, (i.e. determining who is directing who), traditional correlation methods cannot decipher the direction of the association. Many attempts to deduce this information via other methods have been noted, but the accuracy of these methodologies has recently been challenged (Smith et al 2011, Friston 2011).

In examining different approaches to determine directed functional connectivity, Smith et al (2011) used simulated BOLD data to determine the quality of several common methods that attempt to identify A) true structure of a network (direct and indirect relationship), and B) causality within the network (who is directing who). The methods examined were

correlation, partial correlation, inverse covariance (ICOV), Patel's conditional dependence approach and selected Bayesian network (Bayes Net) methods including CCD (Cyclic Coordinate Descent), CPC (Conservative PC), FCI (Fast Casual Inference), PC and GES (Greedy Equivalence Search) (Smith et al., 2011). The partial correlation, ICOV and the Bayes Net methods performed well at identifying the structure of the system and Patel's τ was the best at identifying the direction but there was no method that did well in identifying both the structure and the directionality.

Importantly, much of the work in Smith et al (2011) focused on determining how well structure and causality could be detected across single subject simulated data. However, the noise inherent in a single subject's data may lead to reduced performance of the measurements versus the more typical group-based analyses that many investigators apply. While there are certain drawbacks to examining group averaged matrices, where one first generates one composite matrix by averaging across the single subjects, it may assist performance on some of these measurements.

In this paper, we focus on re-examining one of the prior approaches highlighted in Smith et al (2011), using two sets of simulated data but on group average matrices (average of all the subjects' connection matrices). We also examine performance of the approach on empirical data. In this latter case, structural connectivity matrices based on diffusion-weighted imaging (DWI) served as our baseline. While the ability to ascertain causal relationships is still evolving, Bayesian inference can provide a reasonable estimate based on observed statistical evidence as indicated above. These methods are particularly powerful in certain networks where the ambiguities like feedback and forward connections are limited. This perspective is adopted here and we exploit conditional independence test to infer probabilistic dependence between observed variables.

The Bayesian approach has been successful in applications like finding causal relation in protein signalling, in cellular networks, to drug design and more. (Friedman 2004; Gertrudes et al., 2012; Sachs et al., 2005) suggesting it warrants closer scrutiny for brain imaging data as well. We used the PC algorithm developed by Spirtes et al (2000) to determine the direct and indirect relationship and the directed functional connectivity in the examined systems.

2. Methods

We used two simulation data sets and one empirical data set as follows

2.1. Simulation 1

Simulation 1 includes the data set ($n=50$) designed by Smith et al (2011). The neural network model in the simulation was based on the dynamic causal modelling (DCM) which was then fed through a non-linear balloon model for the vascular dynamics. The neural network model was defined as

$$z' = \sigma Az + Cu$$

where z is the neural timeseries, z' is its rate of change, u is the external inputs and C the weights controlling how the external inputs feed into the network, A has the network connections between nodes and σ is the neural lag. This neural signal was fed into the balloon model for which the parameters are set in DCM.

The neural lag was set to 50 ms as a realistic estimate of the timelag in neural signalling. Twenty eight data sets were created with different number of nodes, different TR and

different percent of noise. Table 1 shows the parameters for the data we used from the Smith et al (2011) data set.

Here we chose network simulations of varying size (i.e. 5, 10, and 50 nodes) like the other groups (Ramsey, Hanson, Glymour 2011; Smith et al., 2011), except that Gates et al (2012) did not use data with 50 nodes. We also chose a subset of simulations that varied session duration as there is significant variability in the duration of resting state sessions across connectivity studies. We also decided to examine a network with cyclic connections. Cyclic connections present links that start and end on the same vertices, and as mentioned below are often present in biological systems. The set, which includes cyclic connections, may be particularly challenging for structure learning because Bayesian networks are not optimized to support cyclic relationships. Last we decided to include a set that has feedback or backward connections in the system, as most biological links examined with rs-fcMRI are likely to have both feedforward and feedback connections.

2.2. Simulation 2

Another data set was created that explicitly models contemporaneous relations separately from the effects which occur at a lag of one TR (Gates & Molenaar, 2012). Since fMRI data is known to contain sequential dependencies (i.e., cross-correlation and autoregressive effects), it is important to test the present method's ability to accurately recover the true relations and not be biased by these lagged influences. The data was created using Monte Carlo simulation for the following model

$$\eta_t = A\eta_t + \sum_{i=1}^q \phi_i \eta_{t-1} + \zeta_t$$

where η is the ROI time series, A is the contemporaneous relation among ROI, Φ is the lagged relation, and ζ is the error/white noise with variance 1. (Fnaiech and Ljung 1987; Gates et al., 2011; Kim et al., 2007; Mathews and Sicuranza 2000)

We used three data sets for a 10 node network. Each had the same structure depicted in the 10 node graphic in Figure 4. One network featured contemporaneous relations between the nodes, second network featured lagged relations and third network featured contemporaneous relations among the nodes and the autoregressive relation for each node (i.e., how the signal at a given time point predicts the signal of the same node at the next time point).

2.3. Empirical Data

To obtain empirical data we used data from 8 healthy subjects, seven females and one male, with average age of 26.5. Imaging was performed during a single session for each subject on a 3T Siemens Tim Trio scanner with a 12-channel head coil. Data acquisition included a T1-weighted image for anatomical reference, a functional MRI scan, a T2-weighted image, and a high-angular resolution diffusion weighted image (HARDI).

2.3.1. T1-weighted structural MRI—First, a whole-brain, high-resolution T1-weighted magnetization-prepared gradient-echo image (MP-RAGE) was acquired with the following parameters: repetition time (TR) = 2,300 ms, inversion time (TI) = 900 ms, echo time (TE) = 3.58 ms, flip angle (FA) = 10°, 1 mm³ voxels, 160 slices, FOV = 240×256 mm). Tissue segmentation into white and gray matter was performed on the T1 image using Freesurfer software (<http://surfer.nmr.mgn.harvard.edu>). Regions of interest were then defined using the boundaries of each contiguous region within a given network identified in (Yeo et al.,

2011). Importantly, ROIs were defined for each subject using surface registration in Freesurfer, and were then applied to both structural and functional data.

2.3.2. Diffusion-weighted imaging—A HARDI scan was performed using an EPI sequence consisting of 72 gradient directions with $b\text{-value} = 3,000 \text{ mm}^2/\text{s}^2$ along with 10 unweighted B0 images. Acquisition parameters for the scan included the following: TR = 7100 ms, TE = 112 ms, 2.5 mm^3 voxels, 48 slices, FOV = $230 \times 230 \text{ mm}$. Diffusion data processing was carried out by connectomemapper (<http://www.connectomics.org/connectomemapper/>), and consists of four stages: coregistration of the T1-weighted image and B0 image, diffusion data reconstruction, tractography and identification of connections.

2.3.2.1. Coregistration of T1-weighted and B0 image: To facilitate accurate registration of the T1-weighted anatomical image onto the B0 image of the diffusion-weighted data, a T2-weighted image was also acquired (TR = 3200 ms, TE = 497 ms; 1 mm^3 voxels, 160 slices, FOV = $256 \times 256 \text{ mm}$) as an intermediary. We performed a rigid-body transformation of the T1-weighted image onto the T2-weighted image, and then nonlinear registration of the T2-weighted image onto the B0 image. The additional step of nonlinear registration allowed us to account for image distortion common in diffusion-weighted data, such as susceptibility artifact, B0 field inhomogeneity, and eddy-current distortions. Skull-stripping was also performed on the T2-weighted and B0 images prior to this step to ensure robustness of the nonlinear algorithm. Every subject was then manually inspected to ensure high-quality accuracy for each step in the registration procedure.

2.3.2.2. Diffusion Reconstruction: Diffusion data was resampled into 2 mm^3 voxel size and reconstructed using a Q-BALL scheme (Tuch et al., 2003) into an orientation distribution function (ODF) at each voxel. The ODF was defined on a tessellated sphere of 181 vertices, and represents the estimated diffusion intensity in each direction. At each voxel, we defined up to 3 directions of maximum diffusion as defined by the local maxima of the ODF. This step is analogous to computing the principal eigenvector when using the Diffusion Tensor Imaging (DTI) model.

2.3.2.3. Tractography: At each voxel of white matter, we initiated 32 evenly-spaced fibers for every direction of maximum diffusion. Each fiber was propagated in opposite directions, and upon reaching a new voxel, continued in the direction of whichever maximal diffusion direction was closest to its current direction. The growth process of a fiber was stopped whenever this resulted in a change of direction sharper than 60° , or when its ends left the white matter mask. Additionally, fibers shorter than 20 mm in length were considered potentially spurious and were removed. This resulted in a large sample of reconstructed white-matter fibers across the whole brain.

2.3.2.4. Connection Identification: Structural connections between cortical ROIs were identified by combining the results of the tractography with the cortical parcellation. For example, two ROIs i and j were said to be structurally connected if there existed a fiber with endpoints in i and in j . Connections were weighted by the total number of fibers between two ROIs (<http://www.connectomics.org/connectomemapper/>).

2.3.3. Resting-state functional connectivity (rs-fcMRI)—Functional data was acquired using a gradient-echo echo-planar imaging (EPI) sequence with the following parameters: TR = 2500 ms, TE = 30 ms, FA = 90° , 3.8 mm^3 voxels, 36 slices with interleaved acquisition, FOV = $240 \times 240 \text{ mm}$). Subjects were instructed to remain still and passively fixate on a crosshair for 10–20 minutes.

2.3.3.1 Processing: The raw fMRI data underwent standard fMRI preprocessing including slice-time correction, debanding, motion-correction, registration onto the T1 image, and resampling into 3 mm³ voxel size. Several additional steps were also taken to prepare the data for connectivity analyses (Fox et al., 2005), including temporal bandpass filtering (0.009 Hz < f < 0.08 Hz), spatial smoothing (6 mm full-width at half-maximum), and regression of nuisance signals. The latter includes the whole-brain signal, signals from ventricular matter and white matter, and the six parameters related to rigid-body motion correction.

2.3.3.2 Motion Censoring: Subjects underwent several rigorous steps to correct for head motion during scanning. First, frame-to-frame displacement (FD) was calculated for every time point. FD was calculated as a scalar quantity using a formula that sums the values for framewise displacement in the six rigid body parameters ($FD_i = |\Delta d_{ix}| + |\Delta d_{iy}| + |\Delta d_{iz}| + |\Delta \alpha_i| + |\Delta \beta_i| + |\Delta \gamma_i|$, where $\Delta d_{ix} = d(i-1)_x - d_{ix}$, and similarly for the other five rigid body parameters) (Power et al., 2012). At each time point, if the FD was greater than 0.2mm, the frame was excluded from the subject's time series, along with one preceding frame and the two following frames (Power et al., 2012).

2.3.3.3 Connectivity: Time series were then computed for each of the cortical ROIs by averaging the signal intensity across all voxels within the ROI for each time point. Cross-correlations were computed between the time series of all ROI pairs.

2.3.4. Modularity Analysis—To divide our data set (see section 2.3.1) into specific brain networks we applied a community detection algorithm, as introduced by Newman (Newman 2006), to the structural matrix generated from our tractography procedures. This method uses modularity, a quantitative measure of the observed versus expected intra-community connections, as a means to guide assignments of nodes into specific communities or networks. The method by Newman was selected based on prior evidence that it is efficient, accurate, and commonly applied to both structural and functional brain data (Fair et al., 2009; Power et al., 2011; van den Heuvel et al., 2009; Van Dijk et al., 2010). The community detection algorithm split our large scale system of structural data (DWI data) into 5 networks: A presumptive default network, fronto-parietal network, cingulo-opercular network, visual network and motor network. We examined the first 3 of these networks here due to the distributed nature of their connectivity and the nodes that defined them (See Figure 8). The functional correlation and PC algorithm results were then compared to the network formed by the modularity analysis. That is, we clarified whether the functionally connected nodes (by both Pearson's correlation and PC algorithm) matched with the structurally connected nodes within a given network using the structural DWI result as the baseline.

2.4. Conditional Independence and PC algorithm

The PC algorithm is based on a Bayesian principles; a graphical model that shows the joint probability distribution. The graphical model represents a directed acyclic graph (G) with the nodes or vertices (V) and the edges (E). The nodes represent the random variables and the edges represent the relation or causal dependency between two variables. Figure 1 shows an example of Bayesian network where season, hay fever and flu are three variables/nodes and the connection between season and the two illnesses represents the edges/ causal relationship.

A joint probability distribution is the combined probability of all possible events to occur. A Bayesian network pairs each variable x_i in the network to a conditioned distribution $P(x_i | pa_i)$, where pa_i represents the parent of x_i which is given as

The conditioned probability is given as when two events X and Y are conditionally independent given the third event Z that is; when the occurrence of X and Y are independent events in their conditional probability distribution given Z. For the network shown in figure 1, the joint probability distribution for hay fever would be $P(H|S)$ (Pearl 1995; Sachs et al., 2005). PC algorithm (Sprites and Glymour C. and Scheines, R. 2000) incorporates this notion of conditional independence. The algorithm begins with Step1: a complete undirected graph Step2: Edges between the nodes are deleted based on the conditional independence test. As illustrated in the example given node S, Node F and Node H are conditionally independent and hence the edge between F and H is deleted. Node S is saved as a separation set (Sepset) $S[F,H]$ and $S[H,F]$. In Step3, the v-structures are oriented first and in Step4: the remaining edges are orientated using the rules mentioned in Sprites et al (2000).

The function $pc()$ in R (Kalisch et al., 2012) implements the four steps mentioned above.

2.5. Algorithm

For simulation 1, the correlations between nodes were created for every subject individually. Using the Schmidt-Hunter method for meta analyses, the correlation coefficient (r) between the nodes were averaged across all the subjects which was used as an input to the PC algorithm in R (Kalisch et al., 2012).

The procedure generates a matrix called $pMax$ with a dimension of $N \times N$ (N equals the number of nodes/vertices) for a given alpha value. Alpha is the tuning parameter for the PC algorithm. $pMax$ is the maximum p-value of all the conditional independence test between two given nodes. At any given alpha level, if the $pMax$ for any two given nodes is less than the alpha value, there exists a connection between those two nodes. Based on this, the $pMax$ matrix was converted into binary form. Importantly, the PC algorithm is threshold (alpha) dependent (a similar problem as with traditional correlation analyses). As such, using optimum threshold is very important as the results can be biased. So we focused here on developing a method that would not be based on a single threshold. We start this procedure with the binary matrix being created for varying alpha levels starting from 0.0 to 1 with an increment of 0.01. The intuition behind our approach is based on the notion that Bayesian average increase the robustness of the model and reduces the pitfall of overtraining or being overly sensitive to a specific parameter setting.

The binary form of the $pMax$ was used in the iteration as shown in the Figure 3. The notations used in the procedure were: P_n , the $pMax$ matrix for alpha value n ; B_n , the binary form of P_n ; $A_{0.1-m}$, the average of $B_{0.1}$ to B_m (the average of $B_{0.1}, B_{0.2}, B_{0.3} \dots B_m$). Any value greater than 0 in the average matrix $A_{0.1-m}$ was treated as a connection. The iteration always starts with m which is the max alpha that we use in our iteration equals 0.9.

To illustrate how the iterations work let us take an example. The iteration begins with $m = 0.9$

Step 1: $T = m - 0.05 = 0.85$ and $C = 0$, $A_{0.1-0.9}$ is calculated

Step 2: Check if the structure (connections) in $A_{0.1-0.9}$ is same as $B_T = B_{0.85}$.

Step 3: (A) If the structure is same find the lowest n for which the structure of B_n is still same as the average. The structure and the directions of the connection for this B_n are treated as the true ones. There is a range of alpha levels where the structure remains constant. So the lowest alpha value where the structure remains consistent is considered to be the most stable for the structure and the directions.

(B) If the structure is not the same, $C = C + 1$;

- (i) If $C = 1$, $T = T - 0.5 = 0.85 - 0.05 = 0.8$ and go to step 2
- (ii) If $C = 2$, $m = m - 0.1 = 0.9 - 0.1 = 0.8$ and go to step 1

To compare the results with Smith et al (2011), we ran `pc()` on individual subjects for logarithmically spaced alpha between 0.01 and 1 on data set 2. We assigned connection strengths for all the connections as $-\log(\alpha)$ where alpha is the lowest value at which the connection first appears (Smith et al., 2011). Then we calculated the number of true connection whose strength was greater than the 95% percentile of the false positive connection strength. This number was divided by the total number of true connections. This fraction is defined as c-sensitivity in Smith et al (2011). C-sensitivity was calculated for each individual subject separately and averaged across the subjects. For the directionality, we found the most consistent alpha for each subject and found the probability of the true connections to have true directions.

For simulation 2, the averaged maximum correlation was considered as input for the contemporaneous plus lagged relation model and lagged relation model to generate the pMax matrix. For the contemporaneous network, the input was average correlation as was in Simulation 1 from Smith et al (2011). As explained earlier, the binarized pMax was generated and rest of the iteration was carried out.

For the empirical data, we used the three networks namely default network, fronto-parietal network and cingulo-opercular network, which we found with modularity analysis (see figure 8). The structural data contains the information about the number of fibres connecting two nodes. This data was combined for each connection across all subjects and recoded as binary (connected versus non-connected) as follows. If the connection appears in more than 50% of the subject, then the connection was assigned a 1 or else 0. This binarized structural data was treated as the gold standard (baseline). For each network, the correlation coefficients for the functional data were averaged for the 8 subjects. The connections from this functional connectivity and `pc` algorithm were then compared for 100 different thresholds to the baseline and the true positive rate (TPR) and false positive rate (FPR) were found.

3. Results

3.1. Simulation 1

As mentioned earlier, eight sets of simulations were used from the Smith et al (2011) study. Figure 5a shows the structure and direction of the information flow recovered by the algorithm for the data set with 5 nodes. The algorithm was able to recover all the connection and directions accurately. Similarly the structure and the causality for 10 nodes data and 50 nodes data were accurately recovered. Figure 5b shows the result for the data set with 10 nodes. The algorithm even worked excellently for varied session duration (Data set 4, 5 and 6). For the cyclic connection, the structure was accurately recovered but the direction indicated in Figure 5c in red was identified incorrectly. Figure 5d shows the result from the algorithm for backward connection. The ground truth by Smith et al (2011) for the backward connection is shown in SI figure 1.

As explained earlier, we calculated the c-sensitivity for each subject. Figure 6a shows the distribution of c-sensitivity for all the subjects. The white dot in the figure shows the mean of the distribution which is above 0.9 i.e. 90%. This finding replicates prior findings by Smith et al (2011). Figure 6b shows the probability distribution for the true direction and the mean as shown in the figure is around 0.5 which is 50%.

3.2. Simulation 2

For Simulation 2, the algorithm was successful in finding the structure for all three data sets (Figure 7a and 7b). The algorithm was also somewhat successful at identifying directionality. As shown in Figure 7a, for the contemporaneous data, four out of eleven directions were incorrect. Out of the four two were bi-directional meaning the algorithm was not able to decide how to orient these connections. As shown in Figure 7b, for the mixture of contemporaneous and lagged data and only lagged data, five out of eleven connections were oriented incorrectly.

3.3. Empirical data

3.3.1. Modularity results—As mentioned in section 2.3.3, modularity analysis was applied to our empirical structural matrix to detect networks within the structural data. This data driven approach splits the whole brain into 5 different networks namely default network, fronto-parietal network cingulo-opercular network, visual network, and motor network. We use default mode network, fronto-parietal network and cingulo-opercular network as shown in figure 8 to examine how functionally deduced connections with our algorithm and traditional correlation methods relate to the structural connections within these identified networks.

3.3.2. Comparison—As noted above, here we compared the structure based on the Pearson's correlation and the structure that we obtain from `pc()` with 3 networks identified with community detection using the structural data (Newman 2006; Fair 2009). The three networks included the default network (pink), the cingulo-opercular network (red) and the fronto-parietal network (green).

Figure 9 shows the radar graph indicating the difference between the true positive rate (TPR) and false positive rate (FPR) for the two methods. The value on the circumference of the circle indicates the different level of thresholds. The values on the radar axis (vertical axis) indicate the difference. As one moves towards the edge of the radar graphs the better the functional connectivity agrees the structural connectivity baseline. For all three networks we can see that PC algorithm was more accurate than the straight correlation. We also conducted a t-test between the PC algorithm and straight correlation for the difference of TPR and FPR. We found that for all the three networks the p-value was <0.0001 .

4. Discussion

In recent years, resting-state functional connectivity MRI (rs-fcMRI) has become popular tool that measures temporally correlated low-frequency BOLD signals in subjects at rest (Biswal et al., 1995). This measurement of functional relatedness has offered significant insight into brain organization. However, most investigators examine the functional relatedness of regions using Pearson's correlation coefficient, which has the advantage of ease of application, but is not able to differentiate between direct and indirect influences or identify the direction of information flow (i.e. directed functional connectivity). In this report we examine the ability of the PC algorithm, based on conditional independence to elucidate these aspects of connectivity on two simulated data sets and an empirical data set.

We find that, as previously reported by Smith et al (2011), based on individual subject data, the approach yields relatively strong results with regard to identifying direct versus indirect connectivity; however, the algorithm does poorly with regard to directionality on individual subject. In contrast, applying the algorithm to averaged connection matrices yields strong results for both indirect/direct connections and directionality for some simulated models (Smith et al (2011)). The method provides modest results with regard to these outcomes for

the other models (Simulation 2 (Gates et al (2012))) as well. In the latter case, the simulated data contains lagged information in two of the data sets, which reduces the efficacy somewhat. Note, probabilistic inference of direction of influence has certain inherent limitations. Of the three elemental structures, only one of the V-structure can be inferred unambiguously. In many applications where Bayes network can be successfully inferred, the V-structures over all the three variable combinations strongly constraints the directions of the influence amongst all the variables. In both of our simulations, the underlying structures had links that cannot be unambiguously resolved; however, we thought it is important to begin our inquiry with data structures that have been examined in prior reports ((Gates 2012; Ramsey, Hanson, Glymour 2011; Smith et al., 2011)). Importantly, using empirical data to compare the correlation based method and the PC algorithm to the structural data; we see that the PC algorithm performs better than the traditional correlation methods.

We note that the work by Smith et al (2011) was followed up by at least two studies beside ours. First, Ramsey et al (2011) (Ramsey, Hanson, Glymour 2011) applied two algorithms ImaGES (modification of Greedy Equivalence Search) and LOFS (modification of Linear, Non-Gaussian, Acyclic Causal Model) on the Smith et al (2011) data. Looking at the results from data set with 10 nodes, the two methods correctly identified nearly all the connections and 90% of the directionality on group of subjects. They used 10 subjects out of the 50 subjects with replacement to calculate the connection and orientation precision. Notably, when using the algorithm on single subjects, the algorithm correctly identified 90% of the connections right but only 63% of the causal directions correct. Second, Gates et al (Gates 2012) used Group Iterative Multiple Model Estimation (GIMME) which is an extension of unified structural equation modeling (uSEM) (Gates et al., 2010). Individual and group level analysis was carried out with this algorithm re-examining the 10 node data set from Smith et al (2011). For Smith et al. (2011) the 10 node data set, the individual-level uSEM obtained a c-sensitivity of 90% but only detected 50% of the directions. For the GIMME analysis (50 subjects), all of the connections were recovered and 90% of the directions were accurate for Smith's 10 node data set. The algorithm becomes highly complex for larger number of nodes and may be difficult to apply to large networks even for group level analysis. It will be important for long-term that the methods applied to these types of data are able to handle large datasets, as whole brain systems could include large number of nodes even more than 50. Thus, while these methods all show some promise, the results all highlight the need for further methods development. Furthermore, all of these tests examined thus far have looked explicitly at the same single simulated dataset, which may not reflect all of the complexities of empirical data. The work in this report begins to fill some of this void.

4.1. Threshold related problems and our solution

One parameter that has been difficult to deal with when using correlation based measures is what threshold to signify connected versus non connected nodes. Although it has been shown that thresholding is a necessary step to remove spurious connections while generating a graph (Rubinov and Sporns 2010), the choice of threshold is a critical decision point in analytical processes. A choice of R approaching 1.0 will generate very sparse graphs, with a limited number of edges (i.e., few connections). On the other hand a choice of R approaching 0.0 will generate densely connected graphs. The PC algorithm has a similar problem; alpha is the tuning parameter for PC algorithm (Kalisch and Buhlmann 2007). The structure, number and the orientation of the edges depend on the alpha level. In other words, as with correlation based methods, network changes depend upon the alpha chosen.

As an example we applied three differing thresholds using the 10 node model (data set 2) from Smith et al (2011) – each of which provides a unique result. At one alpha the true network structure is revealed, and at the others the structure is lost (see Figure 10). Because of these discrepancies we attempted to generate an algorithm that could identify the structure

of the system independent of a specific alpha, but rather a system that would rely on the most consistent level of alpha. Our method is inspired by ideas of Bayesian smoothing where the models are smoothed across different parameter settings using a prior. In our case, there is no information prior hence a uniform prior implies averaging the results from all the settings. The current report using both simulated and empirical data suggest that this approach provides consistent and accurate findings with this regard.

4.2. Comparing the PC algorithm implementation in the current report with the Smith et al (2011) implementation

In Smith et al (2011), 38 different methods potentially capable of determining true network structure and in some instances directionality of information flow in BOLD fMRI data were examined. Partial correlation, ICOV and various Bayes method performed well in finding the true network structure but none of the methods were capable of determining the direction of information flow when using single subject data. In the current manuscript we focused on one of the Bayes methods tested by Smith et al (2011) – the PC algorithm (Spirtes et al. 2000). The PC algorithm functions by building a Bayesian Network from the data based on conditional independence tests on subsets of three variables. Although the algorithm used here produced only a slightly better result than Smith et al (2011) in its ability to detect the true network structure in individual subject data, when data was averaged across participants, the results were largely superior. In the case of averaged data, the algorithm did better job at detecting directionality, as well. The analysis on the individual level is likely more prone to inherit noise, whereby using group averaged matrices likely improved the signal to noise ratio, thus allowing for the superior result. This is important support for the PC algorithm because for many research questions, averaged data are sufficient.

The differences in the varying results are also likely secondary to the two program's (R and Tetrad) implementation of the PC algorithm. The PC algorithms used in the two papers were slightly unique. Smith et al (2011) used the tetrad software, while here we used an algorithm implemented in R. While the implementations are largely similar, there are some differences. These differences appear to be nested in the implementation of the conditional independence tests (or Meek's rules <http://uai.sis.pitt.edu/papers/95/p403-meek.pdf>) as shown in SI figure 2 and the post-processing of the Bayes network.

The PC algorithm in R uses the orientation rules described in the book (Spirites and Glymour C. and Scheines, R. 2000), p 85. The algorithm uses some but not all the four Meek's rule for orientation of pattern where as the pc algorithm in tetrad uses all four Meek's rule (http://www.phil.cmu.edu/projects/tetrad_download/files/manual.pdf). These rules are illustrated in SI figure 2 and are also explained below.

R (Rule) 1: if $A \rightarrow B - C$ and if A and C are not adjacent then orient $B \rightarrow C$.

R2: if $A \rightarrow B \rightarrow C$ and $A - C$, orient $A \rightarrow C$.

R3: if $A \rightarrow B \leftarrow C$, A is not adjacent to C and $A - D - C$, then orient $D \rightarrow B$

R4: if $A \leftarrow B \leftarrow C$, A is not adjacent to C and $A - D - C$, then orient $D \rightarrow A$

Also tetrad stops its processing after these rules are implemented. With the R implementation one can include an option which runs the conditional independence test along with 100 different combinations for an ambiguous direction. The orientation of the edge is overwritten at later point if a new orientation is found.

4.3. Backward and Cyclic connections

Of note, when one considers feedforward and feedback connections together for any given circuit in the simulation, the method performs poorly on both respects. While simulating

feedback connections for fMRI data is not straightforward (Smith et al 2011), the findings based on how they were modelled in the data by Smith et al (2011) are potentially informative and an important consideration. The large majority of links in the brain contain both feedforward and backward connections. In the neocortex, most of the feedforward connections arise from supragranular layers, i.e. layers II and III, whereas feedback connections arise from the infragranular layers, i.e. layer V and VI. Another factor to consider is inhibitory and excitatory connections, which can also be difficult to model, but that change considerably in the life span and are likely to reduce the accuracy of methods aimed at identifying direct and indirect links in fMRI data. In addition, other aspects of the model, such as the presence or absence of cyclic connections add complexity to the discovery (Friston 2011). In other words, the relationship of feedforward and feedback connections, inhibitory and excitatory connections, and the presence or absence of cyclic connections in the brain is significantly complex and unlikely to be fully captured by any given model considering the relatively sluggish fMRI BOLD response. It is likely that current models (for full list see Smith et al (2011)) aimed at identifying true connection structure and directionality are unable to incorporate all these true biological dynamics.

Learning the links of influence between variables in a Bayesian network is a challenging problem with combinatorial complexity. The PC algorithm simplifies the task by breaking up the problem into independent subtasks. At the most elemental level, the subtasks involves testing conditional independence between three variables. Among the three possible directed influence structures amongst three variables, one of them (often referred to as the V structure) is unambiguous. The ambiguous links of influence are subsequently resolved by finding an explanation that is consistent with results from all the other three variable subtasks. In certain cases, there may not be a consistent explanation or the explanation may not be sufficient to resolve all ambiguities in the network. This is the trade-offs made by PC algorithm to simplify the structure learning problem. So the accuracy of the algorithm depends on the underlying true structure that generated the data and hence is dependent on the application. Thus, while apparently better than traditional correlation analyses, the PC algorithm may not perform optimally in “real world” systems (Friston 2011). Our examination of empirical structural and functional data using the PC algorithm supports this notion. We note that Dynamic Causal Modelling (DCM) was developed with some of the above noted limitations in mind and might be used to improve on such findings - for details see Friston (Friston 2011).

4.4. Homogeneity assumption

The simulated data networks examples used here were homogeneous across individuals. Testing novel methods against homogeneous data sets reflects the current standard. However, emerging evidence suggests that such simulations may not be entirely representative of data sets found in empirical studies because heterogeneity likely exists across individual brain maps (Fair et al., 2012) see also (Hillary et al., 2011; Kherif et al., 2003; Kherif et al., 2009; Seghier et al., 2008). In the presence of heterogeneity, empirical (Miller et al., 2002; Miller and Van Horn 2007) and simulated (Molenaar, P. C. M., Huizenga, H. M., & Nesselroade, J. R. 2002) data examples have frequently pointed out that traditional group-level analysis results in findings that fail to describe any individual comprising the group. The degree to which the aggregated result differs from individual results relates to the degree of heterogeneity found across individuals (Gates 2012). Thus, heterogeneity may present a problem only in extreme cases where individuals vary greatly. Since aggregation in some form has been shown to pick out signal from noise (e.g., Ramsey, Hanson, Glymour 2011), the benefit may outweigh the negatives in some cases. More work is needed to ascertain the degree of heterogeneity that exists in specific populations and how

this may affect the reliability of methods, such as the present one, which aggregates across individuals using averaging.

4.5. Empirical data

As noted above, while the simulations by Smith et al (2011) and the additional simulations (Gates et al (2012)) included here are a good start, there are limitations to these simulations. One of the problems with these models is that they are not small-world i.e., clustering coefficient = 0. This reduces their biological plausibility. Hence to test the procedure in a more biological plausible context, in the present report we attempted to generate an empirical test of our method compared to the traditional correlation analyses by examining how close the methods relate to the underlying structural connectivity based on high resolution tractography. While we recognize that high resolution tractography has its own limitations (Dell'acqua and Catani 2012; Masutani et al., 2003), we believe it likely reflects an estimate of the underlying structural matrix and should serve as a good baseline to compare methods. In this instance, the method applied to the fMRI data best able to identify the true structure of the matrix should recapitulate the average matrix and network structure of structural connectivity data.

Despite limitations in the method's ability to clarify structure in data that includes feedforward and feedback connections, the PC algorithm outperformed the traditional correlation based method. However, while the Bayesian networks appear to outperform traditional correlations, in identifying structure neither are significantly strong in these biologically plausible scenarios as explained above. This lack of correspondence is potentially related to several factors. The imprecise nature of the methods as noted in our simulations is likely leading to reduce efficacy in identifying the true structure of the matrix. However, it is possible that the low performance obtained by both methods is at least partially reflective of the imprecise nature of the baseline, or "gold standard," used here, being high resolution DWI.

Nonetheless, in all instances it appears that the PC algorithm, as implemented here, outperformed traditional correlation methods suggesting that it is modestly superior at reflecting the true underlying connectivity structure.

5. Conclusion

In conclusion, these findings reveal that 1) Bayesian methods such as the PC algorithm may be superior at matching functional connectivity matrices to structural connectivity, 2) in the simplest, non-biological, cases these methods can reveal both structure and directionality, but 3) when considering empirical biological data, or more biologically plausible simulations, such as feedforward and backward connections (a phenomena quite ubiquitous in the brain), both methods perform poorly at identifying directionality.

We believe these results suggest a potentially alternative way to examine connectivity using BOLD fMRI data. But this method cannot be applied to find the direction of information flow. Future work will need to build on approaches, as examined here, that provide a baseline for comparisons purposes and additional neuronal and BOLD models that better capture various biological phenomena.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by R00 MH091238 (Fair), R01 MH096773 (Fair), R01 MH086654 (Nigg), Oregon and Translational Research Institute (Fair), McDonnell Foundation (Petersen/Sporns/Fair). We thank Corinne Stevens, Sam Carpenter and Taciana Costa Dias for their important discussions and help with data collection.

References

- Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*. 1995; 34(4):537–541. [PubMed: 8524021]
- Dell'acqua F, Catani M. Structural human brain networks: Hot topics in diffusion tractography. *Current Opinion in Neurology*. 2012; 25(4):375–383. [PubMed: 22766720]
- Fair DA, Bathula D, Nikolas MA, Nigg JT. Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109(17):6769–6774. [PubMed: 22474392]
- Fair DA, Cohen AL, Power JD, Dosenbach NU, Church JA, Miezin FM, Schlaggar BL, Petersen SE. Functional brain networks develop from a “local to distributed” organization. *PLoS Computational Biology*. 2009; 5(5):e1000381. [PubMed: 19412534]
- Fnaiech F, Ljung L. Recursive identification of bilinear systems. *International Journal of Control*. 1987; 45(2):453–470.
- Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews.Neuroscience*. 2007; 8(9):700–711.
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(27):9673–9678. [PubMed: 15976020]
- Friedman N. Inferring cellular networks using probabilistic graphical models. *Science (New York, N.Y.)*. 2004; 303(5659):799–805.
- Friston KJ. Functional and effective connectivity: A review. *Brain Connectivity*. 2011; 1(1):13–36. [PubMed: 22432952]
- Gates KM, Molenaar PC, Hillary FG, Slobounov S. Extended unified SEM approach for modeling event-related fMRI data. *NeuroImage*. 2011; 54(2):1151–1158. [PubMed: 20804852]
- Gates KM, Molenaar PC, Hillary FG, Ram N, Rovine MJ. Automatic search for fMRI connectivity mapping: An alternative to granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM. *NeuroImage*. 2010; 50(3):1118–1125. [PubMed: 20060050]
- Gates, KM; P.C.. Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*. 2012
- Gertrudes JC, Maltarollo VG, Silva RA, Oliveira PR, Honorio KM, da Silva AB. Machine learning techniques and drug design. *Current Medicinal Chemistry*. 2012
- Habeck C, Moeller JR. Intrinsic functional-connectivity networks for diagnosis: Just beautiful pictures? *Brain Connectivity*. 2011; 1(2):99–103. [PubMed: 22433005]
- Hillary FG, Medaglia JD, Gates K, Molenaar PC, Slocomb J, Peechatka A, Good DC. Examining working memory task acquisition in a disrupted neural network. *Brain : A Journal of Neurology*. 2011; 134(Pt 5):1555–1570. [PubMed: 21571783]
- Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, Hagmann P. Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106(6):2035–2040. [PubMed: 19188601]
- Kalisch M, Buhlmann P. Estimating high-dimensional directed acyclic graphs with PC-algorithm. *Journal of Machine Learning Research*. 2007; 8:613–636.
- Kalisch M, Machler M, Colombo D, Maathuis MH, Buhlmann P. Causal inference using graphical models with the R package pcalg. *Journal of Statistical Software*. 2012 in Press.

- Kherif F, Josse G, Seghier ML, Price CJ. The main sources of intersubject variability in neuronal activation for reading aloud. *Journal of Cognitive Neuroscience*. 2009; 21(4):654–668. [PubMed: 18702580]
- Kherif F, Poline JB, Meriaux S, Benali H, Flandin G, Brett M. Group analysis in functional neuroimaging: Selecting subjects using similarity measures. *NeuroImage*. 2003; 20(4):2197–2208. [PubMed: 14683722]
- Kim J, Zhu W, Chang L, Bentler PM, Ernst T. Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping*. 2007; 28(2): 85–93. [PubMed: 16718669]
- Koller, D.; Friedman, N. Probabilistic graphical models: Principles and techniques. MIT press; 2009.
- Masutani Y, Aoki S, Abe O, Hayashi N, Otomo K. MR diffusion tensor imaging: Recent advance and new techniques for diffusion tensor visualization. *European Journal of Radiology*. 2003; 46(1):53–66. [PubMed: 12648802]
- Mathews, VJ.; Sicuranza, GL. Polynomial signal processing. John Wiley & Sons; Chichester, England: 2000.
- Miller MB, Van Horn JD. Individual variability in brain activations associated with episodic retrieval: A role for large-scale databases. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*. 2007; 63(2):205–213. [PubMed: 16806546]
- Miller MB, Van Horn JD, Wolford GL, Handy TC, Valsangkar-Smyth M, Inati S, Grafton S, Gazzaniga MS. Extensive individual differences in brain activations associated with episodic retrieval are reliable over time. *Journal of Cognitive Neuroscience*. 2002; 14(8):1200–1214. [PubMed: 12495526]
- Molenaar, PCM.; Huizenga, HM.; Nesselrode, JR. The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In: Staudinger, UM.; Lindenberger, U., editors. Anonymous Understanding human development. Kluwer; Dordrecht, Netherlands: 2002. p. 339-360.
- Newman ME. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(23):8577–8582. [PubMed: 16723398]
- Pearl J. Bayesian network. 1995
- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*. 2012; 59(3):2142–2154. [PubMed: 22019881]
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, et al. Functional network organization of the human brain. *Neuron*. 2011; 72(4):665–678. [PubMed: 22099467]
- Ramsey JD, Hanson SJ, Glymour C. Multi-subject search correctly identifies causal connections and most causal directions in the DCM models of the smith et al. simulation study. *NeuroImage*. 2011; 58(3):838–848. [PubMed: 21745580]
- Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*. 2010; 52(3):1059–1069. [PubMed: 19819337]
- Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP. Causal protein-signaling networks derived from multiparameter single-cell data. *Science (New York, N.Y.)*. 2005; 308(5721):523–529.
- Seghier ML, Lazeyras F, Pegna AJ, Annoni JM, Khateb A. Group analysis and the subject factor in functional magnetic resonance imaging: Analysis of fifty right-handed healthy subjects in a semantic language task. *Human Brain Mapping*. 2008; 29(4):461–477. [PubMed: 17538950]
- Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, Ramsey JD, Woolrich MW. Network modelling methods for FMRI. *NeuroImage*. 2011; 54(2):875–891. [PubMed: 20817103]
- Sprites, P.; Glymour, C.; Scheines, R. Causation, prediction and search. 2nd ed. MIT Press; 2000.
- Stephan KE, Friston KJ. Analyzing effective connectivity with fMRI. *Wiley Interdisciplinary Reviews. Cognitive Science*. 2010; 1(3):446–459. [PubMed: 21209846]
- Tuch DS, Reese TG, Wiegell MR, Wedeen VJ. Diffusion MRI of complex neural architecture. *Neuron*. 2003; 40(5):885–895. [PubMed: 14659088]

- van den Heuvel MP, Mandl RC, Kahn RS, Hulshoff Pol HE. Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Human Brain Mapping*. 2009; 30(10):3127–3141. [PubMed: 19235882]
- Van Dijk KR, Hedden T, Venkataraman A, Evans KC, Lazar SW, Buckner RL. Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization. *Journal of Neurophysiology*. 2010; 103(1):297–321. [PubMed: 19889849]
- Van Essen DC, Newsome WT, Bixby JL. The pattern of interhemispheric connections and its relationship to extrastriate visual areas in the macaque monkey. *The Journal of Neuroscience*. 1981; 2(3):265–283. [PubMed: 7062108]
- Vincent JL, Patel GH, Fox MD, Snyder AZ, Baker JT, Van Essen DC, Zempel JM, Snyder LH, Corbetta M, Raichle ME. Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*. 2007; 447(7140):83–86. [PubMed: 17476267]
- Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zollei L, Polimeni JR, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*. 2011; 106(3):1125–1165. [PubMed: 21653723]
- Zalesky A, Fornito A, Bullmore E. On the use of correlation as a measure of network connectivity. *NeuroImage*. 2012; 60(4):2096–2106. [PubMed: 22343126]

Highlights

- We use a Bayesian approach called PC algorithm to find effective connectivity.
- The algorithm finds true connections and directions in Smith et al (2011) data.
- The algorithm performs better than traditional correlation method with the empirical data.

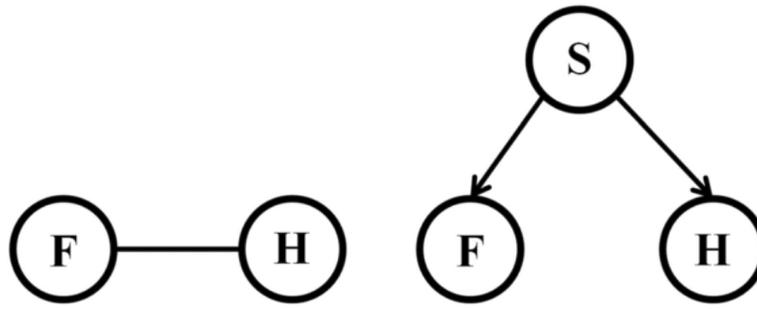


Figure 1. Example illustrating indirect and direct connections and the direction of information flow

As Hay fever and Flu co-occur seasonally, they might be thought to be directly linked to each other as shown in figure (a). However, the reason why they co-occur, or correlate is because of a common third factor, which is seasonal change in this case as shown in figure (b). Traditional correlation methods when examining functional connectivity fail account for similar dynamics in the brain as diagrammed in figure 2.

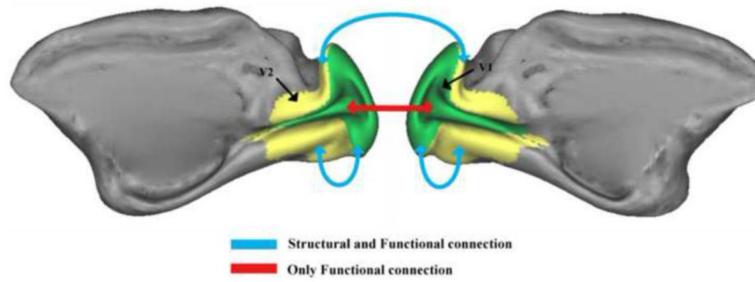


Figure 2. Example illustrating the findings of Vincent et al (2007)

Vincent and colleagues demonstrated how segments of the left and right primary visual cortex, known to lack direct anatomical connections, show strong functional connectivity. This functional connectivity between non-connected regions appeared to be mediated via a secondary brain region.

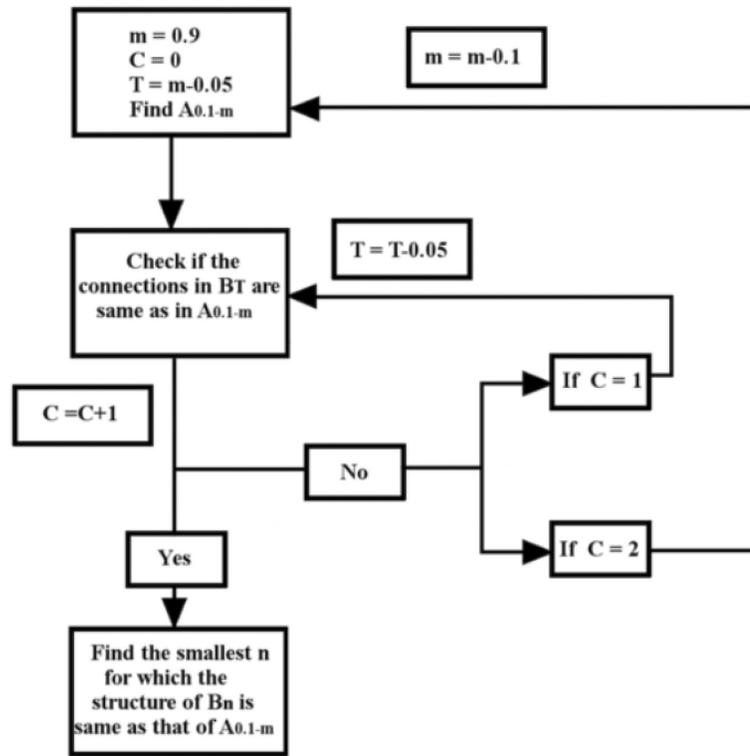


Figure 3. Flowchart of the PC algorithm iteration applied here

The iteration applied here is inspired by ideas of Bayesian smoothing where the models are smoothed across different parameter settings. Starting with the highest alpha value, the most consistent network structure is found with the iterations outlined in the figure

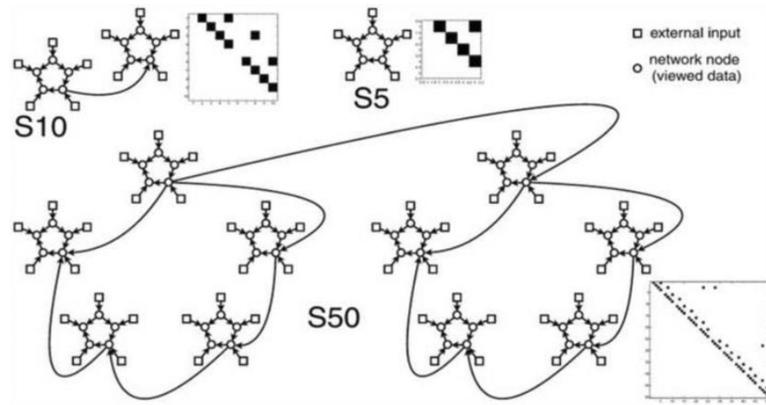


Figure 4. The structure and causality of the Smith et al data (2011)

Three network structures were chosen from Smith et al (2011) to test the capability of the algorithm. The first one has 5 nodes, second one has 10 nodes and the third network has 50 nodes.

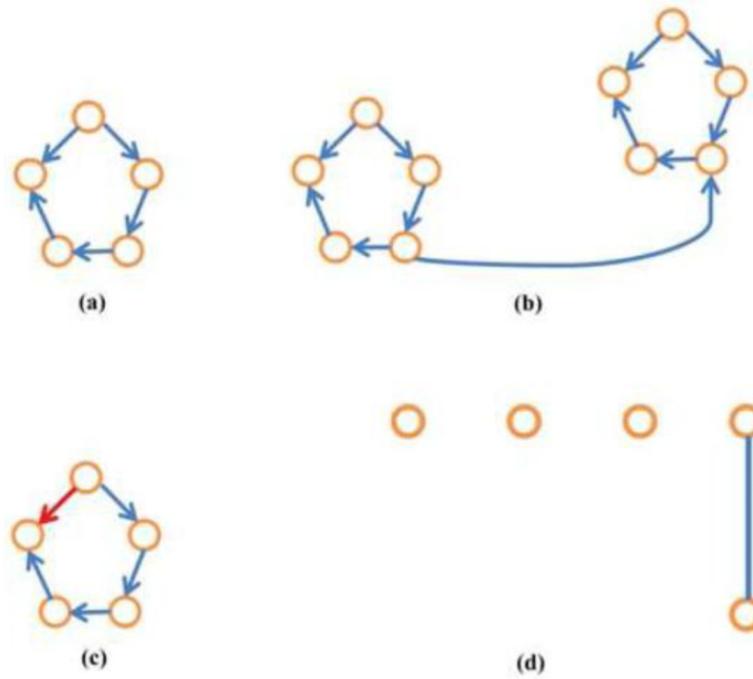


Figure 5. Results from simulation 1

(a) Estimated model for data set 1 with 5 nodes, (b) Estimated model for data set 2 with 10 nodes, (c) Estimated model for the data set 7: cyclic connection, (d) Estimated model for data set 8: feedforward and feedback connections. (For information about the network structure refer Smith et al (2011)).

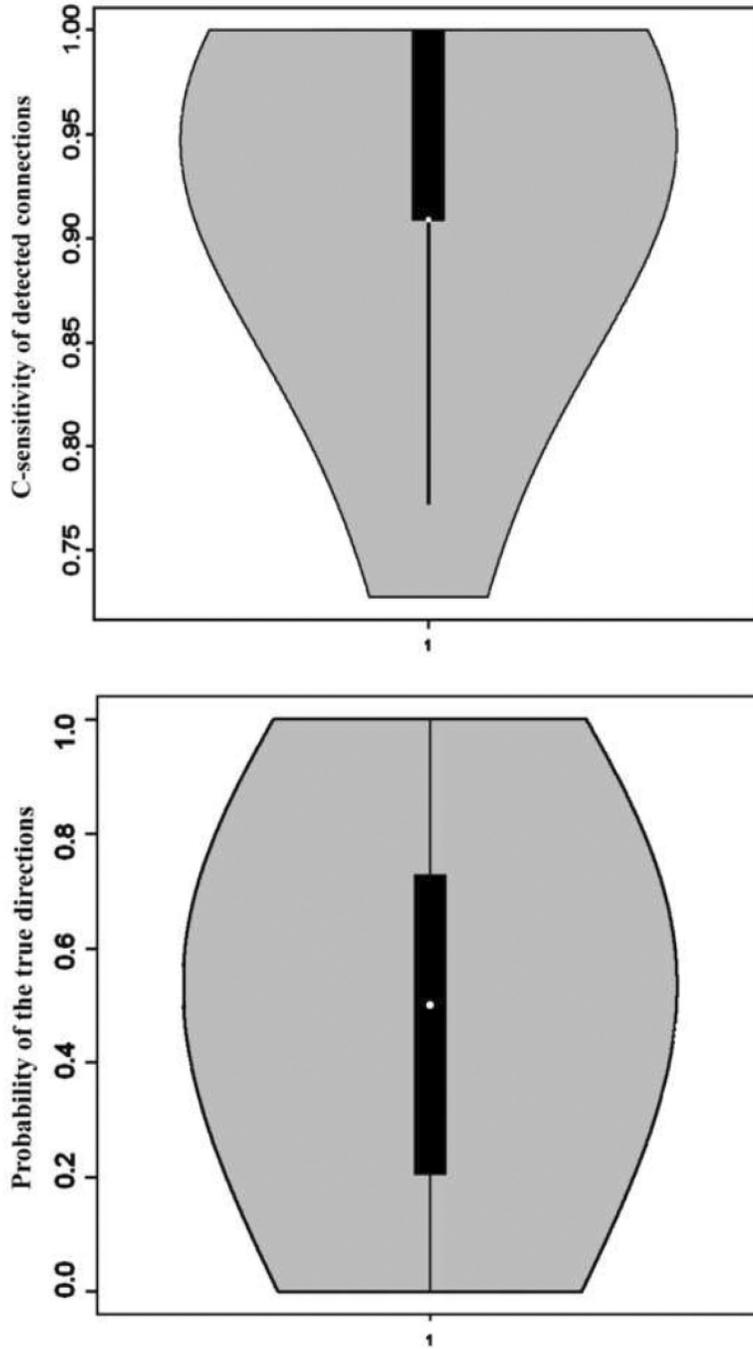


Figure 6. Results from simulation 1, 10 node network: To compare our results with the previous work (Smith et al (2011)) we applied the algorithm on individual subjects. (a) shows the C-sensitivity for connection/edges detected. (The connection strengths for every connection were assigned as $-\log(\alpha)$ where α is the lowest value at which the connection first appears (Smith et al., 2011). Then we calculated the number of true connection whose strength was greater than the 95% percentile of the false positive connection strength. This number was divided by the total number of true connections. This fraction is defined as c-sensitivity in Smith et al (2011)). The white dot in the figure shows the mean of the

distribution which is above 0.9 i.e. 90%. This finding replicates prior findings by Smith et al (2011) and (b) shows the probability of correct direction. The mean as shown in the figure is around 0.5 which is 50%.

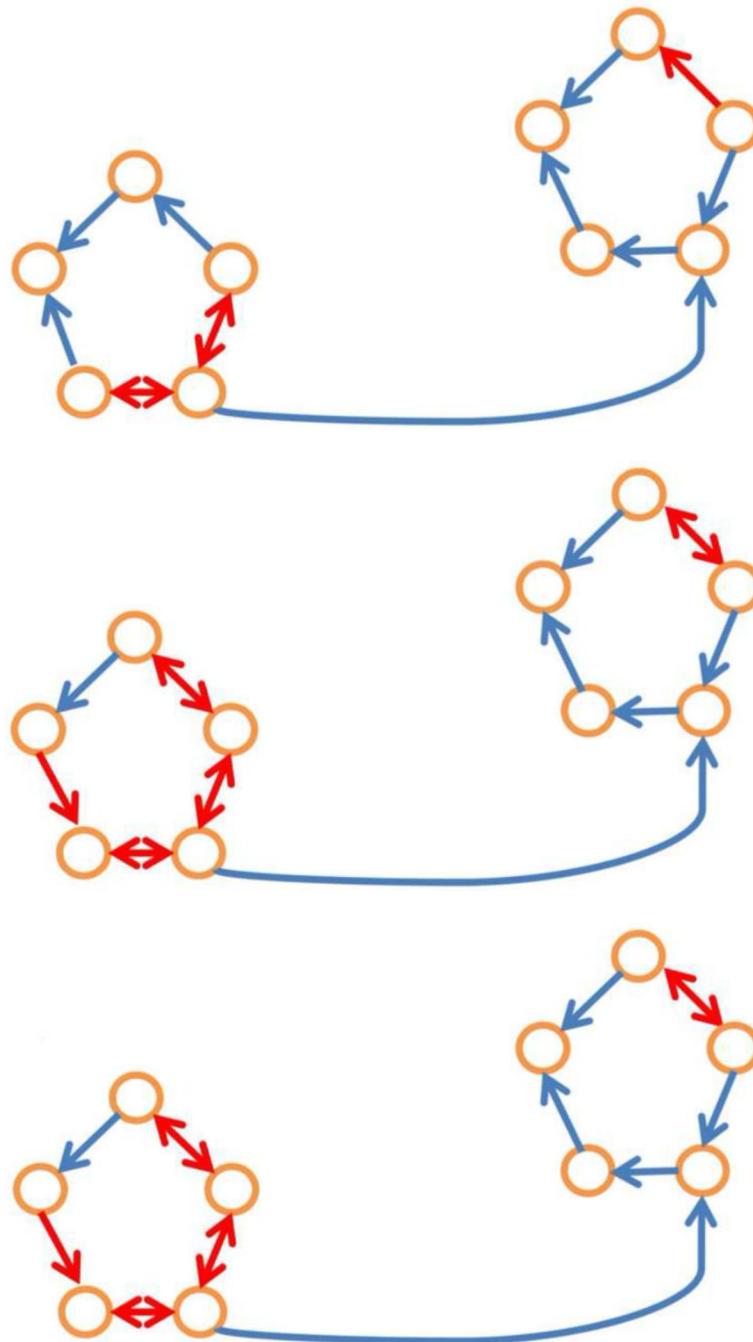


Figure 7. Results from simulation 2

(a) Estimated model for Data set 1, Contemporaneous (b) Estimated model for Data set 2, Contemporaneous and lagged, and Data set 3, lagged only. Structure links are indicated with blue or red arrows. Blue arrows indicate correct direction, and red arrows indicate incorrect direction.

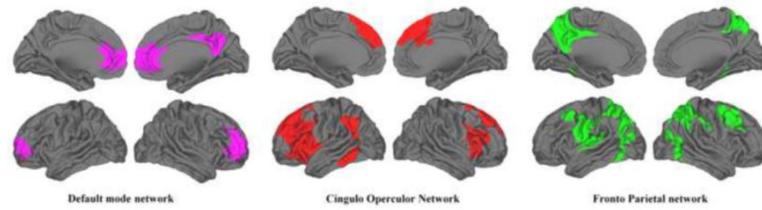


Figure 8. Three network modules

Figure shows the three distributed networks that we obtain from community detection and diffusion weighted imaging. The Default network is indicated with pink color, the cingulo-opercular network is indicated with red color and fronto-parietal network with green color.

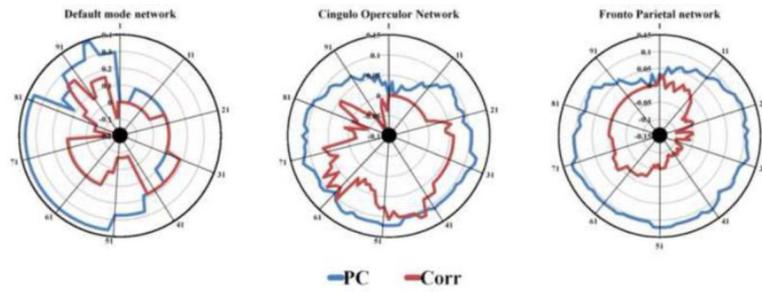


Figure 9. Radar graph

The graph shows the difference between True positive rate and the False positive rate when comparing PC algorithm and straight correlation with the structural data for each network (a) default mode network, (b) cingulo-opercular network and (c) fronto parietal network.

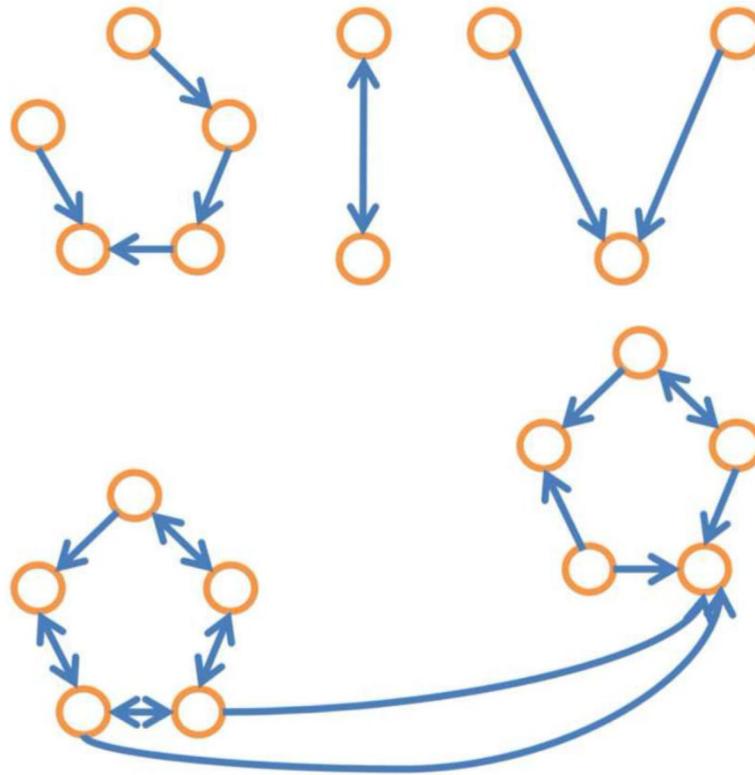


Figure 10. Different thresholds provide different network structures

The implementation of our algorithm (Figure 3) is inspired by ideas of Bayesian smoothing and accounts for this problem where the models are smoothed across different parameter settings using a prior. In our case, there is no information prior hence a uniform prior implies averaging the results from all the settings. (a) Network structure at 0.48 alpha value (b) network structure at 0.93 alpha value.

Table 1

Parameters for the simulation 1

Data Set	Node	Session Duration (mins)	TR (s)	Noise (%)	Condition
1	5	10	3.00	1.0	
2	10	10	3.00	1.0	
3	50	10	3.00	1.0	
4	5	60	3.00	1.0	
5	10	60	3.00	1.0	
6	5	5	3.00	1.0	
7	5	10	3.00	1.0	Cyclic connection
8	5	10	3.00	1.0	Backward connection

Table 2

Details about simulation 2

Data Set	Nodes	Relation
1	10	Contemporaneous
2	10	Lagged
3	10	Contemporaneous & Lagged