Causal search procedures for fMRI: review and suggestions

Teague Henry & Kathleen Gates

Behaviormetrika

ISSN 0385-7417

Behaviormetrika
DOI 10.1007/s41237-016-0010-8

Springer

Behaviormetrika
DOI 10.1007/s41237-016-0010-8

INVITED PAPER

Causal search procedures for fMRI: review and suggestions

Teague Henry[1] · **Kathleen Gates**[1]

Received: 30 August 2016 / Accepted: 1 December 2016

Abstract In this article, the most commonly used algorithms for causal search on fMRI data are reviewed and discussed, with particular attention paid to aspects of the algorithms useful for substantive neuroimaging researchers. Classic algorithms, such as PC and GES, as well as more contemporary algorithms, such as IMaGES–LOFS and GIMME, are compared and contrasted. One major difference between algorithms, the use of lagged variables to infer direction vs. the assumption of non-normality to infer direction is discussed, with eye on the impact those choices have on substantive findings. The algorithms share some commonalities as well as differences. For instance, some use lagged variables to assist in ascertaining directionality, whereas others rely on aspects of non-normality in the data to infer directionality. Attention is given to the impact of these choices on the reliability of results.

Keywords fMRI · Causal search · Greedy search · DAGs · GIMME · IMaGES · LiNGAM · DCM

1 Introduction

Functional magnetic resonance imaging, or fMRI, is a neuroimaging technique that allows researchers to examine the neurological functioning of the brain as measured by the proxy of blood oxygenation (Huettel et al. 2014). In recent years, there has been an increasing interest in functional connectivity analysis, or how different

Communicated by Shohei Shimizu.

✉ Teague Henry
trhenry@email.unc.edu

Kathleen Gates
gateskm@email.unc.edu

[1] University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Published online: 20 December 2016

regions of the brain functionally connect to each other (Sporns et al. 2004). One of the main end goals of functional connectivity analysis is to determine the set of causal relations that describe the functional connections within a set of brain regions, so that researchers can say, for example, that Region A causally influences Region B. The difficulty with inferring causal relations is twofold. The first difficulty is that of orientation of a potential causal relation. Having identified that Region A is associated with Region B, the question then is, does Region A causally effect Region B, or does Region B causally effect Region A. Second, it is difficult to infer causal relations between two regions without taking into account the causal relations between all other regions within a given network. To account for these difficulties, many methodologists have developed what are termed *causal search procedures*, which attempt to infer the network of causal relations between brain regions. These causal search procedures are what we focus our review on here.

One seminal paper on causal search procedures for fMRI was that of Smith et al. (2011). This article provided one of the first reviews and evaluations of a variety of causal search procedures on simulated fMRI data. Therefore, prominent is this article that the majority of causal search procedures validate themselves on the Smith et al. (2011) data simulations. The authors simulated fMRI-type data using a feed-forward balloon model (Friston et al. 2003), which allowed for the explicit modeling of the BOLD response signal. They then evaluated a variety of causal search procedures, such as Granger causality (Granger 1969), the PC algorithm (Spirtes and Glymour 1991), LiNGAM (Shimizu et al. 2006), and GES (Meek 1997). Smith et al. (2011) found that these algorithms were fairly efficient at determining the presence of an association between variables, but less effective at orienting that association into a causal relation.

A second prominent review of causal search procedures is that of Mumford and Ramsey (2014). In their review, the authors provide detailed information as to the specifics of each algorithm and outline various additional results for several newer algorithms, such as the IMaGES algorithm (Ramsey et al. 2010) and the GIMME algorithm (Gates et al. 2010).

Due to the previous thorough reviews of various causal search methods, the present review balances a mix of technical detail, advice for use on fMRI data, and descriptions of how these approaches have been used in practice. To that end, we will not provide detailed descriptions of each of the algorithms, but rather summarize important aspects of each algorithm that inform practical concerns. Furthermore, we point out previous uses of each algorithm on fMRI data, and provide some discussion on the use of each algorithm. This paper provides a practical guide for selecting algorithms given the qualities of the data and the research questions at hand.

The structure of this review is as follows: we begin by briefly reviewing necessary concepts and terminology from fMRI practice, followed by a more through review of necessary concepts in causal inference. Once these are complete, we move onto reviewing each of our selected methods in turn. Within each section, we provide an overview, several technical details, and a discussion as to how each method can be best used on fMRI data. Finally, we summarize our findings.

1.1 An overview of fMRI concepts

While an in-depth review of fMRI practice is beyond the scope of this paper, it is important to understand precisely what type of data fMRI is capable of collecting. To begin, we need to make a distinction between resting state, block, and event-related designs. Resting state data are gathered, while a subject is resting within a scanner, and not doing anything in particular (Biswal et al. 1995; Huettel et al. 2014). By contrast, block and event-related designs have the participant performing a task within the scanner by responding to some type of stimuli or simply observe some stimuli. In block designs, the participant is doing something rather consistent throughout a given time, such as watching a video or doing the same task in rapid succession. Here, changes that relate to a task are not directly modeled using the causal search. Researchers may sometimes conduct separate DAG analysis for different blocks, or compare results from resting state with models obtained from while an individual was performing a task in a given block. Event-related designs attempt to identify changes occurring within a given frame of time, and may have multiple tasks interleaved.

Data obtained during resting state and block designs are the most commonly used data types when using causal search procedures. Unless otherwise stated, all methods reviewed here were developed in a resting state or block design context. Of note, two methods we review, dynamic causal modeling (Friston et al. 2011) and extended unified SEM in GIMME (Gates and Molenaar 2012; Gates et al. 2011), are immediately capable of modeling event-related task data directly by modeling how functional connectivity among brain regions can change in the presence of stimuli that varies across time. In many cases, all the DAG methods presented here can be extended to task-based paradigms to examine causal relations within a task session. For the purposes of this review, we focus on the use of these methods on resting-state data or block designs, where we do not have to consider the presentation of tasks.

fMRI data, fundamentally, are a multivariate time series that has undergone a series of preprocessing steps. We do not concern ourselves with the specifics of preprocessing, and instead assume that all preprocessing has already been done. However, it should be noted that choices made during preprocessing can have significant impacts on the results of a causal search procedure. Specifically, preprocessing steps that alter the temporal relations between timepoints will change the associated causal structure. One prime example off this is high-pass and low-pass filtering (Huettel et al. 2014). These steps in preprocessing remove low-frequency components that correspond with physiological phenomena, such as respiration and heartbeat, as well as high-frequency components that correspond to noise. Obviously, if these steps are not performed, the causal relations between regions of the brain could be confounded by the systematic physiological noise or by the higher frequency random noise. However, if the filters are too strict and remove skewness in the data, then some causal search procedures would miss any number of causal relations. The Butterworth filter and SPM low-pass filter remove non-normality that reduces the ability of some approaches to detect connections (Ramsey et al. 2014). The key concern in preprocessing fMRI data for later use with

causal search algorithms is that one must remove irrelevant sources of noise, but preserve the functional relations between regions.

The fMRI time series is initially at the voxel level. These voxels are three-dimensional volumes that form the smallest unit of resolution the scanner can obtain (Huettel et al. 2014), typically on the order of 2–5 mm^3. In any given fMRI scan, there are usually several thousand voxels. Most analytic techniques are incapable of handling that number of dimensions, so what is typically done is combining voxels into brain regions of interest, or ROIs. There are several methods of doing this, from using predefined anatomical atlases, such as the canonical Brodmann atlas (Brodmann 1909), to determining ROIs statistically (Craddock et al. 2012; Sanchez-Romero 2012). It is these ROIs that are the usual unit of analysis for causal search procedures. Selection of ROIs and other preprocessing options, such as motion correction methods, have been shown to have influence on results for any functional time of fMRI analysis (Power et al. 2011; Weissenbacher et al. 2009). Given the large scope of this topic, the reader is suggested to evaluate emerging preprocessing steps and critically consider pipelines prior to conducing DAG analysis. One clear suggestion of preprocessing steps suggested for use with functional connectivity can be found in (Power et al. 2014).

Finally, it is worthwhile to note that fMRI does not collect information about the neurological functioning of brain regions directly, rather fMRI collects the blood oxygenation-level-dependent signal, or BOLD signal (Logothetis 2008; Huettel et al. 2014). This signal represents the flow of deoxygenated blood in a particular region of the brain, and can be considered to be measuring neurological activity with error. Most causal search procedures are agnostic to what the data are, be it BOLD signal, or a financial time series, and it is important to keep in mind the actual unit of analysis when interpreting the results of a causal search procedure. With fMRI practice being briefly reviewed, we can now turn our attention to concepts relating to causal search.

1.2 Concepts in causal search

In this review, we will use a network centered approach to discussing causality. As such we describe the system of causal relations between a set of variables as a *directed graph*, an example of which is shown in Fig. 1:

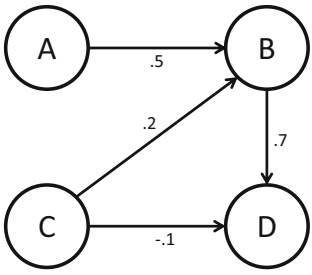

Fig. 1 Example of a directed graph representing a causal structure. *Numbers* indicate edge strength on some arbitrary metric

A graph such as the one in Fig. 1 is constructed by a set of *nodes* (here, brain regions) and a set of directed *edges* (i.e., connections). These edges represent causal relations between the variables represented by the nodes. The *weights* on each edge, here using an arbitrary scale for illustration purposes, represent the strength of that causal effect. In this graph, variable A causes variable B, but has no direct causal relation to variable D. However, due to variable B's causal relation with variable D, A will not be marginally independent from variable D.

The algorithms discussed here seek to identify not just the presence of the edge, but also the direction and the weight. More specifically, the types of causal structures that are primarily dealt with in causal search are *directed acyclic graphs*, which simply mean that there are no cycles within the graph. A cycle makes interpreting a causal structure difficult, and the causal search algorithms that are reviewed here typically restrict the causal structure to directed acyclic graphs. Several of the methods reviewed here, chief of which is the LOFs algorithm (Ramsey et al. 2011), can also infer undirected connections. Typically, these undirected connections are inferred when the directionality of the causal relation is unclear, but it is clear that the variables in question are associated some how.

Directed acyclic graphs that suppose to represent causal structure have also been called *Bayesian networks* (Pearl 1986) due to the decomposition of the conditional probabilities. One important quality of these networks is that of the Causal Markov condition (Spirtes et al. 1993), which indicates that variables are conditionally independent of variables not directly or indirectly connected to them. This is to say that if variable X causally influences variables Y and Z, Causal Markov Independence states that conditions on X, Y and Z are independent of each other.

Causal DAGs have several other implicit assumptions that are instantiations of the assumption of causality itself. Key among these for fMRI data is the temporal nature of causality. Simply put, for Cause A to have Effect B, Cause A must happen before Effect B (Holland 1986). However, in fMRI causal search procedures, a great deal of attention is paid to so-called contemporaneous or "instantaneous" causation, where a causal connection can be oriented between two regions measured at the same time. When data are sampled at rates far slower than the rate of the actual process under study, effects often surface as contemporaneous (Granger 1969). This is the case in fMRI data, where the BOLD signal occurs and is sampled at a temporal resolution that is far slower than the neuronal activity of interest. In Smith et al. (2011), the only approaches that correctly recovered the presence of a relation among brain regions were contemporaneous modeling approaches. For this reason, only a few of the causal search procedures presented in this review use the information from previous timepoints, with GIMME (Gates et al. 2010) being the only one that explicitly uses lagged effects.

One final note of import is that these DAG causal structures make no distributional assumptions in of themselves. However, certain graphical search algorithms do require distributional assumptions to be correct (statistically consistent). For instance, a path diagram from a structural equation model (Bollen 1989) implies a resulting multivariate normal distribution (and could indeed have a DAG structure), and the LiNGAM algorithms utilize non-normality of the data to identify directionality (Shimizu et al. 2006). Nonetheless, a basic Bayesian network

needs not imply any particular family of distributions. We now provide an overview of the approaches used in the fMRI literature.

2 GES

2.1 Overview

Greedy equivalence search (GES) was developed as a general search procedure (Meek 1997) for determining causal structure. The algorithm makes modifications to the graph one edge at a time, an approach that has been found to be simple to implement, computationally efficient, and one that works well in practice (Chickering and Meek 2002).

2.2 Details

GES is a score-based algorithm in that it utilizes the Bayesian information criterion (BIC; Schwarz 1978) to guide model building. The BIC is a model selection index that approximates a Bayes Factor, and is formed using the log-likelihood of the model along with a penalty term that reflects the number of parameters in the model. In GES, improvement of a BIC score with the addition of an edge indicates that the edge should be included. The algorithm rests heavily on the knowledge that the BIC is locally consistent, meaning that the BIC will favor the addition of edges that suggest dependence among variables only if that dependence is in the generative model (Chickering 2002). GES is a forward selection and backward deletion approach. It begins by adding edges to an empty graph. Then, GES deletes edges in a backward search also utilizing the BIC. Note that the removal of edges is only done after the addition of all edges using the forward search, as opposed to performing pruning after each individual edge is added (Chickering and Meek 2002).

As specified in Chickering and Meek (2002), the GES algorithm returns a partially directed acyclic graph (PDAG). These PDAGs consist of oriented causal relations and undirected associations. The reason why some edges are oriented, while others are undirected lie in the notion of *equivalence classes*. Directed Acyclic Graphs within the same equivalence class are indistinguishable from each other in terms of their fit to the data, but could have several edges oriented in different directions (Pearl 2004). This problem with equivalence classes is met head on with the GES algorithm, in that it orients edges only if they are determined to have a unique orientation. This can only occur if the two nodes connected by an edge do not share the same parents (Chickering and Meek 2002). Using this fact, the GES algorithm returns a set of directed causal relations, and a set of undirected associations where the direction of any potential causal relation is unclear.

2.2.1 Use on fMRI data

Since the algorithm does not inherently include temporal information, it may be applied to graphs that are static in nature, such as diffusion tensor imaging (DTI). However, to date, the GES algorithm has only been applied on fMRI data. Studies have used GES to investigate differences in brain processes among individuals diagnosed with Autism spectrum disorder (Hanson et al. 2013) and those with traumatic brain injury (Dobryakova et al. 2015). One study compared a few Bayesian network search approaches described in the present paper (GES, PC, Granger causality, and IMaGES) on fMRI data obtained, while participants watched a video. They found consistency in patterns from the GES and PC algorithms, concluding that this convergence suggested that they are arriving at some true underlying pattern of effects (Sun et al. 2012).

The GES algorithm is available from the Tetrad project: http://www.phil.cmu.edu/tetrad/current.html.

2.2.2 Drawbacks and Limitations

GES requires a large number of timepoints to appropriately recall the true edges and their directions. Results for lengths of time series typically seen in fMRI might be unreliable (Mumford and Ramsey 2014). This limits the utility for individual-level analysis using fMRI data. An approach that utilizes information across individuals (IMaGES; Ramsey et al. (2010), described below) has been shown to greatly improve upon this situation (Perez et al. 2010). For this reason, IMaGES is more commonly used and recommended for fMRI researchers. It should also be noted that the BIC typically assumes a linear Gaussian likelihood, and the consistency of selection methods using BIC is not guaranteed when the data are non-normally distributed.

Finally, as originally proposed GES could only reasonably handle 5–15 variables, which limits the number of ROIs that can be analyzed in a causal structure. In recent work, Ramsey (2015), demonstrated that GES can be scaled up to handle millions of variables in a reasonable amount of time. This would allow GES to comfortably analyze complete parcellations of the brain.

3 IMaGES

3.1 Overview

IMaGES [Independent Multiple-Sample Greedy Equivalent Search; (Ramsey et al. 2010)] is an extension of the Greedy Equivalent Search Algorithm (Meek 1997) for multiple subjects. IMaGES pools information from multiple subjects to estimate a single DAG structure that best fits the data from the multiple subjects. This method differs from simply concatenating the data from multiple subjects as IMaGES allows the weights on each edge of the DAG to differ between samples.

As was mentioned above, the GES algorithm has trouble determining the directionality of the edges within a given DAG. Ramsey et al. (2010) proposes a further modification of the IMaGES algorithm to address this issue. In their modification, they propose systematically backshifting variables in each data set by a certain lag and then running IMaGES on the resulting set of backshifted data. The analysis then selects the IMaGES run that demonstrates the best fit. This results in a DAG, where some variables have been backshifted by a certain amount, while other variables are unmodified. By performing this backshifting, the analyst can ameliorate timing issues caused by the BOLD signal, where the timing of one ROI's signal might be offset from the rest.

3.2 Details

The IMaGES algorithm follows closely its parent algorithm, GES. As noted above, GES begins with an empty graph and adds edges based on BIC (Schwarz 1978). IMaGES takes advantage of having multiple data sets, and instead of concatenating data add edges based on an averaged BIC-type measure, defined as follows:

$$\text{IMaGESBIC} = -\frac{2}{m}\sum_{i=1}^{m}\ell(D_i, G) + ck\ln(\text{n}) \tag{1}$$

where m is the number of data sets, $\ell(D_i, G)$ is the maximum log-likelihood of the model described by the DAG G applied to the data set D_i, c is a integer penalty term set by the analyst, k is the number of free parameters in the model described by DAG G, and n is the number of timepoints in any data set. Note that this expression works only if there are equal numbers of timepoints in all data sets, else the expression will be weighted to account for differing numbers of timepoints for each subject. It is also important to note that the IMaGES algorithm, such as the GES algorithm, is a feed-forward then backward prune procedure. Specifically, both begin with an empty graph, add edges until no improvement can be made, and then removes all edges that do not result in a decrease in fit.

The penalty term c acts as a penalty to the complexity of the DAG G, and is used to filter out spurious triangle structures. These triangle structures occur when X causes Y and Y causes Z, but due to lack of information, a contemporaneous relation between X and Z emerges. These can be viewed as spurious as they can result in cycles, which violate causal assumptions. Ramsey et al. (2010) suggest incrementing c up until triangle structures stop appearing in the graph.

IMaGES has been shown to be a consistent estimator of the graph structure when the GES would also be consistent. However, conditions for this consistency are quite strict. Ideally, all subjects would have identical connectivity patterns, but can have differing strengths of those causal connections. That being said, Ramsey et al. (2010) notes that IMaGES is robust to mild variation in subject's connectivity patterns.

3.2.1 Use on MRI data

The IMaGES algorithm is best used on ROI-level data, with multiple individuals' data sets. Use on voxel-level data is not advised for several reasons. The first is that

the IMaGES algorithm needs to estimate a maximum likelihood solution for each graph evaluated for each data set. As Ramsey et al. (2010) points out, this is typically only feasible for Gaussian distributions, as maximum likelihood estimates can be obtained quickly. The second reason why IMaGES is best used for ROI-level data is that the number of causal relations increases exponentially with the number of variables being evaluated. While IMaGES can quickly evaluate a set of 20 ROIs, the speed of the algorithm would be several orders of magnitude slower if used on a voxel-level data set. Finally, voxel-level data have considerably more noise than ROI-level data, and as such would be much more difficult to evaluate for causal relations.

There have been several recent uses of the IMaGES algorithm in high impact journals. One particular example is that of Manelis et al. (2016). They examined differences in the functional connectivity of the occipital sub-network and a fronto-parietal-temporo-striatal sub-network for a sample of typical controls, individuals with bipolar disorder, and individuals with major depressive disorder, under conditions of win anticipation and loss anticipation. The authors shows significant differences in overall levels of connection between the diagnostic groups during the different task settings. This study demonstrates the utility of IMaGES for task- and group-level analysis.

3.2.2 Drawbacks and limitations

As IMaGES is a greedy search procedure, it is limited in the number of ROIs as well as the number of subjects it can handle. As originally tested in Ramsey et al. (2010), IMaGES comfortably handled 10 ROIs. Increasing the number of ROIs past this will simply be a matter of computation time, rather than some innate weakness of the algorithm. As IMaGES is based on the GES algorithm, the new scaling of GES to over a million variables (Ramsey 2015) could be extended to IMaGES, but as of writing, this has not occurred.

A more pressing limitation is the fact that IMaGES produces a single graph structure for every subject in the sample. While it does allow different weights on the edges for each subject, the fact that it hypothesizes the same set of causal relations for each subject may or may not be a tenable assumption based on the research question.

IMaGES is available from the TETRAD project (http://www.phil.cmu.edu/tetrad/).

4 LiNGAM, pairwise LiNGAM, ParceLingam, pooled LiNGAM and LV-LiNGAM

Linear non-gaussian acyclic models (LiNGAM; Shimizu et al. 2006) takes advantage of the assumption of non-normal error terms to determine the best fitting directed acyclic graphical model for a single data set. Many methods for discovering DAG structures use Gaussian error assumptions, which make it difficult if not impossible to determine the direction of a causal link if no other information is

present, due to the symmetry of a covariance matrix which uniquely defines a multivariate Gaussian distribution (this results in an equivalence class of DAGs, where causal direction of edges are indeterminable Spirtes et al. 1993). The LiNGAM approach uses the assumption of non-normality to determine the unique set of directed paths that best describe the given subject's data set. There have been several improvements to the original LiNGAM algorithm (pairwise, parceled, and pooled) that we discuss here.

Pairwise LiNGAM (Hyvärinen et al. 2010) seeks to orient the edges in an already established undirected graph. This differs from the original LiNGAM approach, where both the graph structure as well as the edge orientation are estimated by the same algorithm. Pairwise LiNGAM uses the same assumptions as LiNGAM, that being the linear non-Gaussian error terms, and orients edges one at a time. Hyvarinen and Smith (2013) show using the same data from Smith et al. (2011) that pairwise LiNGAM orients the edges correctly 75% of the time, and that for analysis of concatenated data, the orientation accuracy was 100%. It is important to note that this method requires an already established graph structure; however, any algorithm that can provide an undirected graph structure, such as IMaGES or GES, can be used in tandem with pairwise LiNGAM.

ParceLiNGAM (Tashiro et al. 2014) is a alteration of the original LiNGAM algorithm to be robust to latent or unobserved confounders. In simulations, ParceLiNGAM outperformed both pairwise LiNGAM as well as the original LiNGAM algorithm for the majority of the simulated conditions.

Pooled LiNGAM (Xu et al. 2014) is a more recent development. The method constructs a set of aggregate "subjects" by randomly selecting subjects and then concatenating the data sets in a random order. The original LiNGAM method is then run on these aggregate data sets and the results compared. As the authors note, the Pooled LiNGAM method still makes the same assumption that the original LiNGAM method does, that of homogeneity in both graph structure as well as edge weights. Pooled LiNGAM showed performance nearly equivalent to other commonly used Bayes net approaches (Xu et al. 2014).

Finally, latent variable LiNGAM (Hoyer et al. 2008) is a variant of LiNGAM that accounts for the presence of latent confounding variables. Unobserved variables can potentially have causal effects on observed variables, and if left unmodeled this association can be incorrectly expressed as a causal connection between observed variables. LV-LiNGAM proposes a set of causal structures, where associations between observed variables can be due to the presence of an unobserved variable. The algorithm then simplifies this set of causal structures into a canonical model, which includes, in some sense, a minimum number of latent variables and causal connections needed to optimally model the observed data. One key feature of this method is that it allows for non-Gaussian latent variables to take advantage of the LiNGAM framework.

4.1 Details

Here, we will review the computational aspects of the original LiNGAM algorithm, as well as mention the contribution of the pairwise LiNGAM.

Shimizu et al. (2006) begins by proposing that the data generating model for an observed set of multivariate data X is

$$\mathbf{X} = \mathbf{BX} + \epsilon \tag{2}$$

where ϵ is a set of non-Gaussian error terms, and $\mathbf{B}$ is a square matrix that can be permuted to strict lower triangularity. This is true given the assumption that the data generating graph is indeed a DAG, with the attendant acyclity. (Shimizu et al. 2006) then note that if one solves for $\mathbf{X}$, the resulting equation is identical to that of independent component analysis (Comon 1994):

$$\mathbf{X} = \mathbf{A}\epsilon. \tag{3}$$

LiNGAM's innovation lies in determining the correct order of the independent components that corresponds to the matrix $\mathbf{B}$ that is as close as possible to strict lower triangularity.

From (Shimizu et al. 2006), their algorithm in brief follows the steps as follows:

1. Mean center the data, $\mathbf{X}$, and apply the ICA algorithm to get an estimate of $\mathbf{A}^{-\mathbf{1}}$.
2. Determine the permutation of $\mathbf{A}^{-\mathbf{1}}$ that results in a matrix with no zeroes on the diagonal. Normalize that permuted matrix.
3. Estimate $\mathbf{B}$ with $\mathbf{I} - \mathbf{A}^{-\mathbf{1}*}$, where $\mathbf{A}^{-\mathbf{1}*}$ is the permuted and normalized matrix from the previous step.
4. Find the permutation of $\mathbf{B}$ that is closest to strict lower triangularity.

Algorithms to find the permutations in steps 2 and 5 are found in Shimizu et al. (2006).

Pairwise LiNGAM (Hyvärinen et al. 2010) relies on the assumption of non-normality to determine the direction of the causal relationship between two variables at a time. Instead of using ICA, however, they instead use a likelihood ratio test to determine if the relation is $X \rightarrow Y$ or $Y \rightarrow X$. As was noted above, pairwise LiNGAM can only be used to orient edges that are already determined to be present, necessitating the use of a separate algorithm to determine the overall graph structure.

4.2 Uses on fMRI data

While there has been a considerable amount of literature discussing the use of the LiNGAM family of algorithms on fMRI data, there has been considerably less literature using any of the LiNGAM algorithms on fMRI data. In one article, Liu et al. (2015) uses the original LiNGAM algorithm to identify connectivity differences in the default-mode network between individuals with bipolar disorder and individuals with major depressive disorder, and showed several differences in connectivity between brain regions.

The LiNGAM family of algorithms is well suited for analyzing fMRI data at an ROI level for several reasons. One is that the assumption of non-normality is in many ways less restrictive than the normal assumption when applied to fMRI data.

The magnitude of an fMRI signal is not normally distributed, but rather follows a Rice distribution (Wink and Roerdink, 2006). This Rice distribution can approximate a normal distribution under certain circumstances; however, it typically has a non-negligible amount of skewness. The LiNGAM family of algorithms uses this non-normality to orient the edges of the causal graph. However, it is important that the researcher not use a Butterworth filter (as was done in Smith et al. (2011)) as this can remove the non-normalities of the data. In that paper, LiNGAM could not recover the majority of the directed paths, with performance being better only at longer session times of 1–4 h (Smith et al. 2011). This suboptimal performance is likely due to LiNGAM's single subject nature (i.e., uses relatively short time series and no shared information across individuals). However, another contributing factor was the use of the Btterworth filter which removed the skewness. Using filters that maintain the skewness greatly improves LiNGAM-type algorithms ability to recover directionality (Ramsey et al. 2014).

Code for the original LiNGAM and pairwise LiNGAM is available at https://sites.google.com/site/sshimizu06/lingam.

ParceLiNGAM and pooled LiNGAM are not available on any accessible site.

4.3 Drawbacks and limitations

The primary drawback of the LiNGAM family of algorithms as applied to fMRI data is the assumption of homogeneity across individuals in both graph structure as well as edge weights when data are concatenated to arrive at one data set describing all individuals. While these assumptions of homogeneous functional connectivity could be correct when one has a sample of fairly homogeneous subjects, with a more complex sample, the LiNGAM family of algorithms runs the risk of committing the ecological fallacy, and presenting a connectivity network that represents no-one in particular in the sample (Molenaar 2004). However, this is not necessarily a deeply embedded feature of this family of algorithms, and indeed, the pooled LiNGAM (Xu et al. 2014) moves one step closer in its testing of multiple aggregate subjects and determining the number of agreed upon edges.

It is also worthwhile to note that the LiNGAM family of algorithms can be applied to individual data sets without the assumption of homogeneity. This instead results in a weaker assumption of homogeneity, as it amounts to fitting a LiNGAM to each subject in isolation. However, since performance in recovery of direction decreases with the number of timepoints, this does not always provide reliable models. LiNGAM family methods require a fairly large number of timepoints for any degree of accuracy, and also are limited to the number of ROIs that they can handle, both being limitations to using these methods on fMRI data.

5 LOFS

An orientation algorithm that deserves its own section apart from the LiNGAM category is that of LOFS, or LiNG Orientation, Fixed Structure (Ramsey et al. 2011). LOFS is a LiNGAM-type algorithm most related to the pairwise approach of

Hyvärinen et al. (2010). What makes it somewhat unique is its development as an edge orientation post-processor for the IMaGES algorithm (Ramsey et al. 2011). As discussed above, the IMaGES algorithm has a difficult time correctly orienting the edges of a graph, but typically is quite successful in determining the undirected graph structure. The LOFS post-processor is capable of taking that undirected graph structure and orienting the edges to produce a directed graph.

5.1 Details

The LOFS algorithm, such as the closely related LiNGAM family of algorithms, relies on the assumption of non-normality of the error terms. Specifically, LOFS takes advantage of the fact that the distribution of the residuals from the correctly specified non-Gaussian linear model will be more non-Gaussian than the distribution of the residuals from a incorrectly specified model (Ramsey et al. 2011).

To use this fact, the LOFS algorithm tests non-normality using the Anderson–Darling statistic (Anderson and Darling 1952), and to avoid having to test all possible directed graphs given a particular graph structure, implements a localized search algorithm to orient edges. Ramsey et al. (2011) proposes 2 search rules, where Rule 1 iterates through each node in the graph and orients the edges of X's neighbors, whereas Rule 2 does the same, but takes into account the local neighborhood of X and its neighbors. Rule 2 can be applied after an application of Rule 1, and results in marginally improved recovery of edge orientation (Ramsey et al. 2011).

In both cases, the LOFS algorithm iterates through each node in a graph and tests the orientation of the edges by determining which set of oriented edges maximizes the non-normality of the residuals, in this case measured by the Anderson–Darling statistic. Furthermore, the LOFS algorithm, as an extension of IMaGES, is designed to work with multiple data sets, and when orienting edges allows for different edge weights for each subject data set. This is one of the differences between this algorithm and the pairwise LiNGAM algorithm of Hyvärinen et al. (2010).

One final feature of the LOFS algorithm is its ability to retain undirected edges. If an edge cannot be oriented by the LOFS algorithm, it is still returned un-oriented. This allows the analyst to identify edges that are potentially impacted by Gaussian noise or latent confounders enough to not uniquely determine the orientation of the edge.

5.2 Uses on fMRI data

IMaGES with the LOFS post-processor has been used to examine the functional connectivity of individuals with autism (Hanson et al. 2013), as the functional connectivity of cocaine smokers (Ray et al. 2015) and the functional connectivity of bilingual individuals (Boukrina et al. 2014).

In the simulation studies of Ramsey et al. (2011), wherein they replicated the standard Smith et al. (2011) simulations and evaluated IMaGES–LOFS, the combined algorithm outperformed GES, and the original IMaGES algorithm by a

wide margin. Indeed, IMAGES-LOFS had nearly perfect recall and precision in a number of the simulation conditions, making this algorithm one of the top performers. However, in using LOFS, there are several aspects to consider.

Like pairwise LiNGAM, LOFS is an edge orientation technique, and requires a companion algorithm to determine the graph structure. In this case, that algorithm is readily found in the IMaGES procedure, which is designed to work with ROI-level fMRI data. This carries all the attendant drawbacks of the IMaGES procedures noted above.

The LOFS algorithm as part of the IMaGES–LOFS is available from the TETRAD project (http://www.phil.cmu.edu/tetrad/).

5.3 Drawbacks and limitations

The LOFS algorithm assumes and indeed requires non-normality of the residuals. This is potentially an issue if the fMRI data are normally distributed; however, as in the case of the LiNGAM algorithms, this feature is advantageous given that fMRI data are likely not perfectly normally distributed. That being said, certain preprocessing steps can force normality on fMRI data sets (Huettel et al. 2014), and would render LOFS, and other LiNGAM methods unreliable.

6 Dynamic causal modeling

Dynamic Causal modeling, or DCM (Friston et al. 2003), is a method that can be used to infer a causal structure from fMRI data. It differs from the other discussed methods in both its complexity and the properties of the resulting graph structure. As originally developed in Friston et al. (2003), DCM is not a causal search procedure. Instead, it is a method to estimate non-linear relations between brain regions that could be impacted by different tasks or stimuli during scan. Immediately, this distinguishes DCM from other methods discussed thus far. Both (IMa)GES and the LiNGAM families of methods are linear in structure, meaning that each causal relation is additive, while DCM allows for the approximation of bilinear effects. In addition, DCM can be used with task-related fMRI data, and can explicitly model the effects of task and stimuli on the causal relations between brain regions. This is different than other methods discussed hitherto, which are made to be used with resting-state data or data from a block design. Finally, unlike the other methods discussed here, which use the BOLD signal time series as is, DCM explicitly attempts to model the underlying neural signal by estimating a hemodynamic response function (Friston et al. 2003).

DCM was adapted into a causal search method by Friston et al. (2011). This causal search procedure is unique among the other methods discussed here. Instead of restricting the graph structure to directed acyclic graphs, Friston et al. (2011) allows for directed cyclic graphs which can reflect reciprocal relations between ROIs. Indeed, the causal search procedure outlined in Friston et al. (2011) estimates all edges to be directed and reciprocal, which the authors claim is a reasonable assumption based on anatomical observations.

6.1 Details

Causal search with DCM consists of three components. The first two components, the hemodynamic response model (relating the neuronal activity to the BOLD response) and the coupling function (relating the neuoronal activity of regions to each other) are aspects of the original DCM algorithm, while the final component, the model search, was developed in Friston et al. (2011).

To start, DCM utilizes the Balloon–Windkessel (Buxton et al. 1998; Mandeville et al. 1999) model to decompose the observed BOLD signal to neuronal activity. Briefly, this model uses a system of equations with parameters informed by biophysiological information regarding brain blood flow. This includes an increase in blood flow following activity-dependent vasodilatory signal, which in turn increases the volume and dilutes deoxyhemoglobin (i.e., the BOLD signal). The exact details of this function are beyond the scope of this review and can be found in previously published work (Stephan et al. 2007). The coupling function is a complex system of differential equations that describe how ROIs relate to each other at very short timescales. Friston et al. (2011) eventually reduces down to the following:

$$\dot{x} = Ax + \boldsymbol{\omega} \tag{4}$$

where x is a vector of the estimated neural signal obtained from the Balloon model, $\dot{x}$ is the first derivative of x, effectively modeling the rate of change in x, A is a square matrix of parameters that can be thought of as the functional connectivity matrix, and $\boldsymbol{\omega}$ is a vector of random errors. If the structure of the functional connectivity matrix is known, then the parameters can be estimated using expectation maximization (Friston et al. 2003). Notably, this equation is different than that presented in Friston et al. (2003). There, the authors assume a known causal structure, and allow for bilinear effects as well as exogenous inputs, making the equation:

$$\dot{x} = \left(A + \sum_j u_j B_j\right) x + Cu \tag{5}$$

where B is a matrix containing the effects of the inputs u on the relations between xs and C is a matrix containing constant effects of the exogenous inputs u. The key difference between the DCM for network search, and the regular DCM is that DCM for network search uses a set of stochastic differential equations, that as specified in (Friston et al. 2011), do not contain bilinear terms or exogenous inputs. Friston et al. (2011) does note that there is a deterministic setup that does use exogenous inputs for network search in DCM. In essence, the DCM model for network search describes set of bidirectional causal relations, and provides a means of evaluating the fit of this causal structure to the fMRI data. The relevant detail is that these models restrict the edge orientation to be bidirectional, a somewhat unique feature of DCM.

When the causal structure is not known, the causal search procedure is used. As Friston et al. (2011) rightly points out, enumerating over all possible graphs while fitting the model to each quickly becomes impossible as one increases the number of ROIs. As such the authors present the following method. Beginning with a full model M_F, they iterate through the matrix A, setting subsets of the A equal to 0. They then score each model using $\ln p(y|m_i) \approx \ln q(A_i = 0|M_F) - \ln p(A_i = 0|M_F)$, where $q(A_i = 0|M_F)$ is the marginal conditional density and $p(A_i = 0|M_F)$ is the prior probability that the subset A_i is 0. This scheme allows for quick scoring of all possible causal structures. In addition, the assumption that all edges are reciprocal leads to a reduced number of models to evaluate.

6.2 Uses on fMRI

Dynamic causal modeling is used extensively in fMRI research, and here, we will focus on applications of DCM with causal search. Since its publication in 2011, the network search with DCM procedure has seen few uses. One application uses a network search with DCM to determine the connectivity structure between two ROIs, the right posterior STS and the left cerebellar Crus 1, as well as the modulation of these connections by stimuli (Sokolov et al. 2012). Another article examined the propagation of epileptic seizures, and utilized a network search over a set of theoretically plausible networks. Finally, and more recently, researchers used a network search to determine which DCM best described the default-mode network during resting-state data (Di and Biswal 2014). It should be noted that in all three of these articles, DCM was evaluated over a set of predefined network structures, and not in a full causal search.

DCM with network search is in many ways uniquely suited for the analysis of fMRI data. DCM itself is designed to take into account the BOLD signal response, and to extract an estimated neural signal. In simulation studies, DCM with network search successfully recovered network structure with a high degree of accuracy, though the simulations were done with a generative DCM model (Friston et al. 2011). In addition, DCM is uniquely flexible in its ability to estimate reciprocal relations as well as model the effect of inputs, such as tasks and stimuli.

DCM with network search is available in the SPM package for MATLAB (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/).

6.3 Drawbacks and limitations

One of the largest limitations of DCM is its inability to handle large numbers of ROIs. The causal search procedure outlined in Friston et al. (2011) appears to be able to iterate through possible causal structures; however, it does not appear that it can be applied to more than six ROIs at a time. This is consistent with other applications of DCM in the neuroimaging literature.

In addition to the low numbers of ROIs being handled by DCM, the assumption of reciprocal relations between ROIs is problematic. While the authors argue that this is not proposing an instantaneous causality between two ROIs, and indeed, that the underlying dependency graph is acyclic, the issue here is that there are other

potential causes of apparent reciprocal relations, such as unobserved mutual causes or simple measurement error. The forced assumption of reciprocal relations is not as ideal as an algorithm with the ability to estimate edges as undirected, such as the IMaGES–LOFS algorithm, particularly at the timescales that fMRI data operates at, as the interval between timepoints can be several seconds long.

These two limitations, the low number of ROIs and the forced assumption of reciprocity makes network search with DCM, in our view, less suitable for causal search than other algorithms in this review. However, the flexibility and sophistication of the DCM approach make it an excellent choice for confirmatory research, and one potential option would be to use DCM estimation on a causal structure that has been discovered by a different algorithm.

7 PC algorithm

7.1 Overview

The Peter Spirtes and Clark Glymour algorithm (PC algorithm; Spirtes and Glymour 1991) was introduced in the early 1990s. It has been used on a variety of time series data, such as those seen in studies of economics (Hoover 2008), which made the transition into functional MRI data rather seamless. The PC algorithm performs in two stages. First, it identifies the *presence* of an edge. Next, it identifies the *direction* of the edge (Meek 1995). In some cases, the algorithm may be unable to determine the direction of an edge, and so simply outputs an undirected edge much like IMaGES.

7.2 Details

The algorithm begins with all connections possible being in the model. It then removes edges if two nodes are independent or conditionally independent. The tests for independence condition on all the possible combinations of nodes, starting with zero nodes, then two, and so on until either all possible combinations are exhausted or the nodes being tested are found to be conditionally independent. Corrections for multiple testing may be applied to these series of comparisons, to appropriately adjust the false positive rate.

The direction of the edge is considered once the undirected edges are identified. This may be done in many ways (Meek 1995), most of which consider the set of nodes that caused the conditional independence decisions as well as the common neighbors shared by two given nodes (Mumford and Ramsey 2014). For instance, if two conditionally independent variables, X and Z are found to be conditionally dependent when conditioned on a third variable Y (that was not in the set used to identify conditional independence), then Y would be a collider. From this, we know the following: $X - Y - Z$. The PC algorithm would then identify that the edges originate from the conditionally dependent nodes X and Z and end at the collider: $X \rightarrow Y \leftarrow Z$. In some cases, after considering colliders, the directionality of other

edges becomes apparent. In this way, directionality may be assessed (Spirtes et al. 1993).

7.2.1 Use on fMRI data

Much like GES, the PC algorithm does not inherently include effects regarding temporal dependencies and can be used on either structural or functional MRI data. The PC algorithm has been used on structural data across multiple participants to investigate cortical connectivity (Joshi et al. 2010). That paper used the PC algorithm to arrive at the presence of edges and then used structural equation modeling (SEM) to estimate the weight of these effects.

For fMRI data, the PC algorithm is usually considered at contemporaneous timepoints. This choice is informed by the findings from Smith et al. (Smith et al. 2011), which found that lagged approaches failed to arrive at the presence and direction of effects used to generate the data. If desired, dynamics can be considered by including a lagged set of variables (Mumford and Ramsey 2014), but the algorithm does not prevent backwards directionality. In this case, backwards directionality should be excluded a priori by background knowledge input into the algorithm by the user. Alternatively, the residuals obtained from a VAR model can be considered as a series of data with its temporal dependencies (i.e., dynamics) removed (Swanson and Granger 1997) and these residuals can then be used in PC analysis.

Studies using both structural and functional MRI data have found that results from PC algorithm run on fMRI data corresponds with inferences obtained from structural data (Dawson et al. 2013; Iyer et al. 2013). As noted above in the GES section, one empirical investigation using fMRI found consistency between subjects in patterns of edges obtained from GES and PC (Sun et al. 2012). Thus, the use of PC for fMRI data to determine the presence of edges has been promising, at least for arriving at the presence of an edge between nodes.

A number of extensions make the algorithm even more attractive to fMRI researchers. In particular, approaches have been developed to control for false discovery rate (Li and Wang 2009; Liu et al. 2012; Strobl et al. 2016) and detection of change points, or shifts in the brain process (Lian et al. 2014). These developments have been motivated specifically by issues seen in fMRI data and have largely been validated using designs commonly seen in this literature.

The PC algorithm is available from the Tetrad project: http://www.phil.cmu.edu/tetrad/current.html.

7.2.2 Drawbacks and limitations

Results from the Smith simulations indicated that, at the individual level, the PC algorithm can reliably detect the presence of an edge but not the direction (Smith et al. 2011). This is an important finding, since directionality of edges is often of interest in fMRI studies. However, if the data are aggregated across individuals, reliable networks may be obtained on this data (Iyer et al. 2013). This requires the assumption of homogeneity across individuals in terms of their brain processes, an

assumption that may not always be met (Finn et al. 2015; Smith et al. 2012). These undirected networks could then be used as the input for edge orientation algorithms, such as LOFs or pairwise LiNGAM.

8 Granger causality analysis

8.1 Overview

Granger causality builds models from within a vector autoregression (VAR) framework to arrive at lagged and/or contemporaneous relations, depending on the search space provided by the researcher. Originally introduced within the economics community, Granger causality has been used on fMRI data for some time. Typically for brain studies into brain sciences, only lagged relations are explored, with exceptions existing (see description of GIMME below). Since temporal information is needed, this approach requires data that are in the form of time series. As such it has been used solely for functional MRI.

Granger causality identifies edges to add after considering the autoregressive influence for each region (Granger 1969). In this way, directed edges between brain regions are retained only if a given region predicts subsequent values for a target region after controlling for the target region's ability to predict subsequent values in itself using the values at a prior timepoint. A brain region is said to "Granger cause" another region if these conditions are met.

8.2 Details

Granger causality can be defined in terms of VAR models, which can be defined as follows for up to a lag of L:

$$y_t = \upsilon + \sum_{l=1}^{L} (\Phi_{t-l} + \Phi_{t-l}^{g}) y_{t-l} + \zeta_t \tag{6}$$

where υ is intercept, y is a vector of observed multivariate time series data at a given time t, and ζ is the residual vector for each point in time t. Φ matrices are the $p \times p$ matrices of directed edges and edge weights for p variables. The diagonal at a lag of zero (i.e., contemporaneous effects) must be zero, whereas the diagonals of lagged Φ matrices represent the AR effects. Granger causality begins by estimating the AR effects for all variables. These weights indicate the degree to which a given variable can be predicted by itself using previous timepoints. After accounting for these effects, one can identify if a variable Granger causes another variable by looking at the significance levels of the off-diagonal elements of the Φ matrices.

Granger causality can be additionally defined on the frequency domain of a signal (Kamiński et al. 2001), which can be useful in examining neurological phenomena. This type of Granger causality is related to the notion of signal *coherence* (or in the multivariate case, multiple coherence) (Chicharro 2011), which broadly speaking, is a measure of the similarity of a signal to another signal within

the frequency domain. Coherence is somewhat analogous to a correlation in the time domain. Chicharro (2011) details how Granger causality in the spectral domain is an alternative method to the VAR framework detailed above for testing the notion of Granger causality between variables, particularly when a multivariate time series might not follow a VAR-type model.

Finally, as an aside, there is a even more general form of Granger causality termed information-theoretic Granger causality (Schreiber 2000; Lungarella et al. 2007). This form of Granger causality defines causation in terms of mutual information, a quantity that reflects non-linear dependencies between random variables. The advantages of this approach lie in its generality, and allow for non-linear causal effects to be detected. In addition, the information-theoretic definition of Granger causality has direct relations with both the VAR definition and the frequency-domain definition, but has theoretical advantages as the information-theoretic definition is based on conditional probabilities, which are somewhat more directly interpretable in causal inference.

8.2.1 Use on fMRI data

The form of Granger causality introduced to the fMRI literature (Roebroeck et al. 2005; Goebel et al. 2003) relies heavily on an approach developed by Geweke (Geweke 1982). Here, temporal precedence is required to identify the direction of edges. Linear dependence between two given brain regions (or voxels), $F_{x_{\text{res}},y_{\text{res}}}$, is arrived at as follows:

$$F_{x_{\text{res}},y_{\text{res}}} = F_{x_{\text{res}}\to y_{\text{res}}} + F_{y_{\text{res}}\to x_{\text{res}}} + F_{x_{\text{res}}\cdot y_{\text{res}}}. \tag{7}$$

The variables x_{res} and y_{res} are the residuals of the original x and y variables after regressing out the AR effects (Goebel et al. 2003). $F_{x_{\text{res}}\to y_{\text{res}}}$ and $F_{y_{\text{res}}\to x_{\text{res}}}$ are measures of the directed linear influence of x_{res} predicting y_{res} and y_{res} predicting x_{res}, respectively. The contemporaneous effects (contained in $F_{x_{\text{res}}\cdot y_{\text{res}}}$ are bidirectional. Total independence of two given variables (be it ROIs or voxels) in this iteration can be obtained by assessing that (1) there is no directed lagged relation among the variables and (2) that there is no undirected contemporaneous relation between variables. If it is seen that the two variables are not independent, the approach identifies if one variable causes the other. Note that they may both Granger cause each other.

Granger causality has been conducted on fMRI data at the ROI level, the voxel level, and hybrid versions of this (as done in Roebroeck et al. 2005). Indeed, it is one of the more highly used approaches for arriving at directed connectivity models. For instance, it has been used in fMRI studies of depression (Hamilton et al. 2011), treatment for ADHD (Peterson et al. 2009), hepatic encepholopathy (Qi et al. 2013), and nicotine addiction (Ding and Lee 2013).

Several toolboxes are available on NITRC for Granger causality specifically for fMRI applications, including: GCCA, which uses a spectral version of Granger causality (Seth 2010), REST-GCA (Zang et al. 2012), and BSMART

(Cui et al. 2008). An extension of the linear approach enables the use of non-linear effects (Marinazzo et al. 2011).

8.2.2 Drawbacks and limitations

Granger causality using only lagged (i.e., no contemporaneous directed effects) has been shown to be unable to detect the true direction of edges in data simulated to emulate fMRI data at the individual level (Smith et al. 2011). It is unclear the degree to which models can be recovered if data are considered in group-level design. In practice, this analytic approach has typically been conducted on data, where each individual's data have been concatenated to each other individual. This requires that the same causal model for brain activity exists for all individuals, an assumption that likely is not met. From a statistical standpoint, this heterogeneity across individuals will likely cause spurious findings if time series are assumed to be homogeneous in terms of the relations among variables but indeed are not (Molenaar 2004).

Granger causality, as commonly implemented in fMRI research, also does not include directed contemporaneous paths. However, when the temporal resolution of data collection is low relative to the process being studied, edges may appear contemporaneously (Granger 1988, 1969). This describes exactly what exists in the collection of fMRI data. Indeed, evaluations using benchmark fMRI data simulations (Smith et al. 2011) reveal that approaches that consider contemporaneous effects are the best at returning the generating data structure of directed functional connectivity (Smith et al. 2011; Ramsey et al. 2014; Gates and Molenaar 2012; Mumford and Ramsey 2014). For this reason and others, the utility of Granger causality with VAR for fMRI connectivity mapping has been under much debate in recent literature (e.g., Wen et al. 2013; Deshpande and Hu 2012). In addition, Granger Causality approaches would be unable to distinguish associations between variables that were due to latent confounders vs. true causal connections.

With those criticisms being voiced, it is worthwhile to note that the LiNGAM family of algorithms can be viewed as an extended version of Granger causality (Hyvärinen et al. 2010) that includes both lagged and instantaneous causal paths. As was discussed above, LiNGAM achieves this using non-normalities within the data to inform the directionality of instantaneous causation. Alternatively, Schiatti et al. (2015) provide a VAR-based extended Granger causality framework that uses the pairwise LiNGAM procedure (Hyvarinen and Smith 2013) to orient contemporaneous causal connections, after candidate contemporaneous associations are identified by a lag-1 VAR model.

9 FCI and GFCI

9.1 Overview

Fast causal inference (FCI; Spirtes et al. 1993, 1995) and greedy fast causal inference (GFCI; Ogarrio et al. 2016) are causal search procedures that account for

the effect of unobserved, or latent, confounders as well as the presence of selection bias. These methods are search procedures that attempt to construct a more general *partially directed acyclic graph* that represents a class of equivalent conditional independence structures, and from this general structure, finds the directed paths that are present in every member of the equivalence class. These paths are the *necessary* causal paths for the overarching causal structure. Latent variables and selection variables are accounted for by including hypothetical latent variables and selection variables in the independence structure. This family of methods bears a resemblance to the GES and IMaGES algorithms, as all these techniques search over *equivalence classes*, or partially directed graphs that represent a set of statistically indistinguishable directed graphs.

9.2 Details

Both the FCI and GFCI attempt to find the partially directed graph that represents the most likely equivalence class of directed causal graphs. These methods differ from the previous algorithms in that they include the possible presence of latent variables, as well as selection bias within the equivalence class. By including the presence of these sources of endogeneity, the FCI and GFCI algorithms return three distinct types of edges:

- *Directed edges* Edges that can only be due to a directed connection. One can consider these edges as potential causal effects.
- *Bidirectional edges* These edges represent the potential for a latent variable being the dual cause of the two target variables. This type of edge represents evidence for some sort of association between the variables, likely due to an unobserved variable.
- *Induced edges* These edges are ambiguous in their direction, in that it is unlikely that they are true bidirectional effects, but there is not enough information to uniquely determine the direction of the connection.

In the large sample limit, i.e., as the number of observed datapoints approaches infinity, the FCI and by extension the GFCI methods have been proven to return the true causal structure (Spirtes et al. 1995; Ogarrio et al. 2016). However, this crucially relies on several assumptions regarding causal structures, as well as the assumption of independent identically distributed random variables. The GFCI is a recent innovation of the FCI that performs better at small sample sizes.

9.3 Use on fMRI data

To our knowledge, FCI nor GFCI has ever been applied to fMRI data. There is one significant advantage to using these methods with fMRI data, and that is in the accounting for of latent confounders. Typically, causal search procedures are performed not on the whole brain, but rather a subset of ROIs. This immediately raises the question of unmodeled ROIs accounting for the association between two

modeled ROIs. The use of FCI or GFCI could potentially reduce the impact of unmodeled ROIs.

9.3.1 Limitations and drawbacks

The key limitation to FCI or GFCI as applied to fMRI data is the need for independent identically distributed random variables. fMRI signal is almost by definition neither independent between timepoints, nor identically distributed. Due to this, the FCI and GFCI are not guaranteed to return the true causal structure as the number of timepoints gets large. This appears to be a fatal flaw in the application of FCI and GFCI to fMRI data, as well as a potentially fruitful line of research, as the ability to account for latent confounders would be extremely useful in the fMRI context.

10 GIMME

10.1 Overview

Group iterative multiple model estimation (GIMME; Gates and Molenaar 2012) is a model-building approach from within a structural equation modeling framework. It is similar to IMaGES in that shared information is used to arrive at a directed connectivity model for all individuals, with the edges in the model estimated separately for each individual. In addition, each individual is considered separately without concatenation. A final similarity is that they both were motivated by problems seen with arriving at data-driven models on fMRI data. A notable difference is that GIMME allows for individual-level edges to exist, whereas IMaGES requires all individuals to have the same pattern of edges. In this way, GIMME attends to heterogeneity in brain processes that is likely to exist in many samples (Finn et al. 2015; Laumann et al. 2015; Gates et al. 2014).

10.2 Details

GIMME works from within a unified SEM framework (Kim et al. 2007), which is related to structural VAR, another approach used in fMRI (Chen et al. 2011). The uSEM contains both contemporaneous and lagged effects to conduct VAR analysis using SEM. In this way, GIMME does not require the researcher to choose one or the other temporal scale–both are used as needed. This can prevent the presence of spurious effects if both contemporaneous and lagged edges exist in the true model yet only one is modeled (Gates et al. 2010).

Much like the GES algorithms, GIMME is a score-based, feed-forward algorithm that also conducts backward pruning search. GIMME works from within a Granger causality framework described above. However, it adheres to the original definition by allowing for directed contemporaneous edges (Granger 1969, 1988). Beginning with an empty graph (i.e., no edges), GIMME iteratively selects which directed edge (either lagged or contemporaneous) would improve the greatest number of

individuals' models. This is done in a feed-forward approach based on modification indices (Sörbom 1989) until no edge would significantly improve the majority of individuals' models. Next, the model is pruned to ensure that all edges in the current group-level model is significant for the majority of individuals. Next, GIMME uses the group-level edges as a prior and searches for edges that may be needed at the individual to arrive at a model that fits excellently for that individual. Both the group- and individual-level edge weights are estimated uniquely for each individual. GIMME does not require that each individual has the same number of timepoints.

The final models obtained from GIMME for up to a lag of L can be defined as follows:

$$\eta_t = (A_i + A_i^g)\eta_{i,t} + \sum_{l=1}^{L}(\Phi_{i,t-l} + \Phi_{i,t-l}^g)\eta_{i,t-l} + \zeta_{i,t} \tag{8}$$

where $\mathbf{A}$ is the $p \times p$ matrix of contemporaneous effects for p variables and contains a zero diagonal. Φ is the matrix of lagged effects with AR effects along the diagonal, η is the manifest time series (either for a group or individual in this general formula), and ζ is the residual for each point in time t. The subscript i indicates individual-level estimates and superscript g denotes group-level patterns of effects. Note that all effects error variances are estimated at the individual level with no constraints regarding the distribution of parameters.

10.2.1 Use on MRI data

Much like VAR and Granger causality, GIMME utilizes temporal information. Hence, it can be used for fMRI data but not DTI or other data that focus solely on static information. GIMME is best used on ROI-level data, because it cannot currently handle the computational demands required for large scale, voxel-level matrices. Like IMaGES, GIMME has been shown to be robust to a number of conditions likely seen in fMRI data (see Smith et al. 2011). As a further benefit, GIMME can reliably recover effects when data are either normally distributed or have a skewed distribution, whereas other approaches require filtering techniques that do not normalize the data (Ramsey et al. 2014). In this way, GIMME is robust to certain preprocessing choices made by the researcher, since it can be used with or without Butterworth and low-pass filters.

GIMME has been used in clinical samples ranging from children with attention-deficit/hyperactivity disorder (Gates et al. 2014) to traumatic brain injured (Hillary et al. 2014), nicotine-addicted (Nichols et al. 2014; Zelle et al. 2016), and clinically depressed (Price et al. 2016) adults. There have been several recent uses of the GIMME on non-clinical samples, such as the study of language acquisition (Yang et al. 2015), alcohol use in college-aged students (Beltz et al. 2013; Beltz and Molenaar 2015) and olfactory processes (Karunanayaka et al. 2015). Even in normative samples, all these applications have found differences in the pattern of effects across individuals. This further highlights the need for approaches that can accommodate individual-level differences in the presence and direction of edges.

Extensions to GIMME have further improved its utility for fMRI researchers. These include: subgrouping individuals based on similar brain processes (Gates et al. 0000); enabling multiple patterns of edge results (Beltz and Molenaar 2016); modeling task effects convolved with the hemodynamic response function (Gates et al. 2011); and a posteriori model validation (Beltz and Molenaar 2015).

GIMME is available from as an R package on CRAN *gimme* and on NITRC as a Matlab toolbox (https://www.nitrc.org/projects/gimme/) that depends on LISREL software. Instructions and additional resources can be found here: http://gimme.web.unc.edu/.

10.2.2 Drawbacks and limitations

A drawback of GIMME is that it cannot handle the number of nodes that some other methods can as it has an upper limit of about 25. In addition, it currently only models linear relations among effects. Finally, the estimation of effects may be biased, since the likelihood function to which the uSEM is fit assumes independence of the rows of observations, an assumption that is violated here due to the temporal nature of the data.

11 MDM–IPA

Multiregression dynamic models with integer programming algorithm (Costa et al. 2015) is a recently developed method for determining causal structure in fMRI data. Fundamentally, this algorithm is a combination of a dynamic model, found in the multiregression dynamic model component, and a causal search procedure, found in the integer programming algorithm. The MDM component models an observed fMRI timeseries as a dynamic system, where the relations between regions can vary over time. The IPA component searches over all possible causal structures using a score function that radically reduces the computation time for the search.

11.1 Details

MDM–IPA is a ROI-level technique that is applied to individual data sets, and does not currently have the capability of integrating multiple subjects data without assuming that the same data generating model applies to all subjects.

The MDM component of MDM–IPA consists of two equations, that of observation equations, of which there are as many as there are ROIs, which maps the relation between ROI time series as such (Costa et al. 2015):

$$Y_t(r) = F_t(r)'\boldsymbol{\theta}_t(r) + \epsilon_t(r) \tag{9}$$

where $Y_t(r)$ is the value the rth ROI at timepoint t, $F_t(r)$ is the data at the tth timepoint for the parent ROIs, or the ROIs that causally predict the rth ROI, $\boldsymbol{\theta}_t(r)$ is the vector of time-varying parameters governing the relation between the ROIs, and $\epsilon_t(r)$ is a normally distributed error term with ROI specific variance.

The second component of the MDM is the system equation that determines the time-varying nature of the parameter set:

$$\boldsymbol{\theta}_t = G_t \boldsymbol{\theta}_{t-1} + w_t \tag{10}$$

where G_t is a block diagonal matrix, where each block corresponds to the relations between parameters governing one particular ROI, and w_t is a normally distributed error term with 0 mean, and a block diagonal covariance matrix. Finally, one must specify the initial conditions, and as this is a Bayesian method, one can do this using priors. The details of these priors are beyond the scope of this review, but can be found in Costa et al. (2015).

The setup of the MDM allows for the calculation of a *model score*, in this case a joint log predictive likelihood. Furthermore, in the case of MDM–IPA, the model score is a sum of the individual log predictive likelihoods for each ROI's equation. This makes evaluating which causal structure best fits the data a simpler and quicker proposition than a more complex model would allow.

The MDM component of the MDM–IPA differs from other approaches described here due to its dynamic parameters. This allows for substantial heterogeneity in the temporal dynamics of the fMRI signal; however, this does raise the specter of overfitting.

To determine the predictors of each ROI, Costa et al. (2015) use an integer programming technique. Effectively, integer programming attempts to directly solve the model selection process, rather than searching through all possible models. In their application, Costa et al. (2015) uses integer programming to attempt to find the optimally fitting causal structure, and if a direct application of integer programming does not yield a specific solution, they resort to a search procedure beginning from the best solution from the integer programming. This process drastically reduces the amount of time needed to find the optimally fitting causal structure.

11.2 Uses on fMRI

As this is an extremely recent method, there have been no uses of MDM–IPA in a substantive fMRI study, so we will restrict our discussion to the advantages of this method over the others discussed here. Costa et al. (2015) evaluated the performance of MDM–IPA on the canonical Smith et al. (2011) simulation data from a single condition which had 5 ROIs. Costa et al. (2015) showed that MDM–IPA performed well, with specificity and sensitively above 75%, and a detection of the correct orientation of an edge at 63%. These values are better than, for example, the GES and PC algorithm, though they do not outperform IMaGES–LOFS or GIMME. When Costa et al. (2015) simulated data that had time-varying connection strengths, MDM–IPA significantly outperformed both GES and PC, both of which are incapable of modeling dynamic functional connectivity.

This result highlights the unique contribution of MDM–IPA, that of its ability to model dynamic functional connectivity. There is currently no other causal search method that allows for time-varying relations between ROIs, and this development

is extremely timely given the current interest in dynamic functional connectivity (Hutchison et al. 2013). In addition, the integer programming approach to causal search is a novel contribution to the causal search literature, and could be applied to other algorithm's search procedures. While these are two major advantages to MDM–IPA, there are several drawbacks and limitations that should be discussed.

11.2.1 Drawbacks and limitations

One of the key limitations of this method is the computation time. Costa et al. (2015) indicate that for 11 ROIs and 100 timepoints, the time to determine the best fitting causal structure is approximately 168 min, and this is for a single subjects data set. This large amount of time is likely due to the dynamic nature of the MDM component, as the authors mention that once the necessary scores are computed, determining the best causal structure with IPA takes around 30 s. This lengthy computation time renders the MDM–IPA less usable for larger numbers of ROIs.

In addition, the dynamic nature of the MDM component, while innovative, runs the risk of overfitting to noise, and care should be taken when interpreting the results of the time-varying parameters. Finally, MDM–IPA is a single subject method, and does not allow for the incorporation of more than one subjects data set so as to infer a group causal structure. This reduces the usability of MDM–IPA to single subject analyses.

12 Summary

The causal search procedure for fMRI data literature is growing at an increasing rate, with new algorithms and constant improvements being published all the time. In our review, we have endeavored to present a mix of historical, but still widely used methods along side of more recent and more accurate methods. We have summarized many relevant details of the presented methods in Table 1.

There are several overall points to make regarding causal search applied to fMRI. The two methods that appear to perform best with regards to accuracy and specificity, the IMAGES-LOFS algorithm (Ramsey et al. 2011) and the GIMME algorithm (Gates et al. 2010) achieve their gains in performance in two very different ways. The IMAGES-LOFS algorithm takes inspiration from the LiNGAM family of algorithms and uses the assumption of non-normality to orient edges with a high degree of accuracy, while the GIMME algorithm uses lagged information to orient the causal relations. Both these approaches have advantages and disadvantages. The assumption of non-normality might not be met within a data set, or the data set might have normality forced upon it by preprocessing. The use of lagged information might not accurately reflect the causal relations due to the interval between each timepoint, which might be in the realm of 2–5 s. However, it is somewhat comforting to see that both of the algorithms have similar performance on the same simulated data sets of Smith et al. (2011).

The next point is that these methods (with the notable exception of GES) can only be used on ROI-level data. This is both for computational reasons, in that if

Table 1 Summary of reviewed methods

Method	Non-normality required	Cyclical relations	Lagged info	Instant causality	Requires graph	# of ROIs
PC	No	No	No	Yes	No	5–15
GES	No	No	No	Yes	No	5–15 (>1000 K)
FCI/GFCI	No	No	No	Yes	No	5–15
Granger	No	Yes	Yes	No	No	5–15
GIMME	No	Yes	Yes	Yes	No	5–20
LiNGAM	Yes	No	No	Yes	No	5–15
ParceLiNGAM	Yes	No	No	Yes	No	5–15
Pooled LiNGAM	Yes	No	No	Yes	No	5–15
Pairwise LiNGAM	Yes	No	No	Yes	Yes	NA
IMaGES	No	No	Yes	Yes	No	5–20
LOFS	Yes	Yes	NA	Yes	Yes	NA
Network search DCM	No	All edges	Yes	Yes	No	4–6

used on several thousand voxels the majority of these methods would take several decades to complete estimation, and for theoretical reasons, in that the signal from a single voxel has a considerable amount of noise in it. The use of ROI-level data allows for an aggregation over the noise of individual voxels, which should result in a more robust set of causal estimates. These methods do differ in how many ROIs they can handle, with the greatest number of ROIs going to the IMAGES-LOFS algorithm. An analyst should in part select a causal search procedure based on how many ROIs they are interested in. The exception to this is found in GES, where recent developments have shown that the algorithm can scale to over a million variables (Ramsey 2015). However, in an fMRI context, overfitting and voxel-level noise are still a primary concern.

It is worthwhile to note that very few algorithms presented here account for latent confounders, with the exceptions being FCI, GFCI, and LV-LiNGAM. The presence of latent variables within the data generating causal structure can have significant impacts on the detected causal structure, typically resulting in spurious edges or incorrect orientations. While a moderate amount of attention has been paid to latent variables in the causal search literature, very little attention has been paid to latent confounders in the fMRI literature. In future work, latent confounders, such as structurally important ROIs that are left unmodeled (take, for example, ROIs from the default-mode network, which have wide ranging effects on other networks), need to be carefully considered.

A final point is about the heterogeneity that each method allows. Many of the methods discussed here are used only on a single subjects data set, and require separate estimations on each individual collected. However, two methods in

particular, that of IMaGES–LOFS and GIMME allow for information to be combined across subjects. IMaGES–LOFS estimates a single causal structure for every individual, but allows the weights on each edge to differ (Ramsey et al. 2011). GIMME estimates group, subgroup, and individual-level causal structures, and also allows each weight to differ between individuals. These distinctions between allowed differences can be important when considering the population under study. In a more heterogeneous population, such as one that has a clinical disorder, it might be best to allow for more heterogeneity, and, therefore, use GIMME. If the population under study are typically developing adults, and one can assume that the causal structure should be approximately the same for each individual, then IMaGES–LOFS would be a better choice. Suffice it to say, it is not enough to choose a causal search procedure solely on the basis of performance, but one also must consider what differences in causal structures are allowed by each algorithm.

Causal search algorithms are a valuable tool for a neuroimaging researcher, one that allows them to uncover hidden structure and better understand the system-wide functioning of the brain. These methods are in their infancy, with continuous development, and the coming years will bring great advances.

References

Anderson TW, Darling DA (1952) Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. Ann Math Stat 23(2):193–212

Beltz AM, Gates KM, Engels AS, Molenaar PCM, Pulido C, Turrisi R, Berenbaum SA, Gilmore RO, Wilson SJ (2013) Changes in alcohol-related brain networks across the first year of college: a prospective pilot study using fMRI effective connectivity mapping. Addict Behav 38(4):2052–2059

Beltz AM, Molenaar PCM (2015) A posteriori model validation for the temporal order of directed functional connectivity maps. Front Neurosci 9:304

Beltz AM, Molenaar PCM (2016) Dealing with multiple solutions in structural vector autoregressive models. Multivar Behav Res 51(2):357–373

Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar mri. Magn Reson Med 9:537–541

Bollen KA (1989) Structural equation models with latent variables, vol 9. Wiley, New York

Boukrina O, Hanson SJ, Hanson C (2014) Modeling activation and effective connectivity of VWFA in same script bilinguals. Hum Brain Mapp 35(6):2543–2560

Brodmann K (1909) Vergleichende Lokalisationslehre der Großhirnrinde: in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Johann Ambrosius Barth, Leipzig

Buxton RB, Wong EC, Frank LR (1998) Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn Reson Med 39(6):855–864

Chen G, Glen DR, Saad ZS, Hamilton JP, Thomason ME, Gotlib IH, Cox RW (2011) Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis. Comput Biol Med 41(12):1142–1155

Chicharro D (2011) On the spectral formulation of Granger causality. Biol Cybern 105(5–6):331–347

Chickering DM (2002) Optimal structure identification with greedy search. J Mach Learn Res 3(3):507–554

Chickering DM, Meek C (2002) Finding optimal bayesian networks. In: Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc, pp 94–102

Comon P (1994) Independent component analysis, a new concept? Signal Process 36(3):287–314

Costa L, Smith J, Nichols T, Cussens J, Duff EP, Makin TR, Duff EP (2015) Searching multiregression dynamic models of resting-state fMRI networks using integer programming. Bayesian Anal 10(2):441–478

Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS (2012) A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp 33(8):1914–1928

Cui J, Xu L, Bressler SL, Ding M, Liang H (2008) BSMART: a Matlab/C toolbox for analysis of multichannel neural time series. Neural Netw 21(8):1094–1104

Dawson DA, Cha K, Lewis LB, Mendola JD, Shmuel A (2013) Evaluation and calibration of functional network modeling methods based on known anatomical connections. Neuroimage 67:331–343

Deshpande G, Hu X (2012) Investigating effective brain connectivity from FMRI data: past findings and current issues with reference to granger causality analysis. Brain Connect 2(5):235–245

Di X, Biswal BB (2014) Identifying the default mode network structure using dynamic causal modeling on resting-state functional magnetic resonance imaging. NeuroImage 86:53–59

Ding X, Lee S-W (2013) Changes of functional and effective connectivity in smoking replenishment on deprived heavy smokers: a resting-state FMRI study. PLoS One 8(3):e59331

Dobryakova E, Boukrina O, Wylie GR (2015) Investigation of information flow during a novel working memory task in individuals with traumatic brain injury. Brain Connect 5(7):433–441

Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT (2015) Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nature Neurosci 18(11):1664–1671

Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. NeuroImage 19(4):1273–1302

Friston KJ, Li B, Daunizeau J, Stephan KE (2011) Network discovery with DCM. NeuroImage 56(3):1202–1221

Gates KM, Lane ST, Varangis E, Giovanello K, Guskiewicz KM (2016) Unsupervised classification during time series model building. Multivar Behav Res (in press)

Gates KM, Molenaar PCM (2012) Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. NeuroImage 63(1):310–319

Gates KM, Molenaar PCM, Hillary FG, Ram N, Rovine MJ (2010) Automatic search for fMRI connectivity mapping: an alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM. Neuroimage 50(3):1118–1125

Gates KM, Molenaar PCM, Hillary FG, Slobounov S (2011) Extended unified SEM approach for modeling event-related fMRI data. NeuroImage 54(2):1151–1158

Gates KM, Molenaar PCM, Iyer SP, Nigg JT, Fair DA (2014) Organizing heterogeneous samples using community detection of GIMME-derived resting state functional networks. PloS One 9(3):e91322

Geweke J (1982) Measurement of linear dependence and feedback between multiple time series. J Am Stat Assoc 77(378):304–313

Goebel R, Roebroeck A, Kim D-S, Formisano E (2003) Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn Reson Imaging 21(10):1251–1261

Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econom J Econom Soc 37(3):424–438

Granger CWJ (1988) Some recent development in a concept of causality. J Econom 39(1):199–211

Hamilton JP, Chen G, Thomason ME, Schwartz ME, Gotlib IH (2011) Investigating neural primacy in major depressive disorder: multivariate Granger causality analysis of resting-state fMRI time-series data. Mol Psychiatry 16(7):763–772

Hanson C, Hanson SJ, Ramsey J, Glymour C (2013) Atypical effective connectivity of social brain networks in individuals with autism. Brain Connect 3(6):578–89

Hillary FG, Medaglia JD, Gates KM, Molenaar PC, Good DC (2014) Examining network dynamics after traumatic brain injury using the extended unified SEM approach. Brain Imaging Behav 8(3):435–445

Holland PW (1986) Statistics and causal inference. J Am Stat Assoc 81(396):968

Hoover KD (2008) Causality in economics and econometrics. The new Palgrave dictionary of economics, vol 2. Palgrave Macmillan, London

Hoyer PO, Shimizu S, Kerminen AJ, Palviainen M (2008) Estimation of causal effects using linear non-Gaussian causal models with hidden variables. Int J Approx Reason 49(2):362–378

Huettel S, Song A, McCarthy G (2014) Functional magnetic resonance imaging, 3rd edn. Sinauer Associates Inc, Sunderland

Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Della Penna S, Duyn JH, Glover GH, Gonzalez-Castillo J, Handwerker DA, Keilholz S, Kiviniemi V, Leopold DA, de Pasquale F, Sporns O, Walter M, Chang C (2013) Dynamic functional connectivity: Promise, issues, and interpretations. NeuroImage 80:360–378

Hyvarinen A, Smith SM (2013) Pairwise likelihood ratios for estimation of non-Gaussian structural equation models. J Mach Learn Res 14:111–152

Hyvärinen A, Sugiyama M, Yang Q (2010) Pairwise measures of causal direction in linear non-Gaussian acyclic models. In: JMLR workshop and conference proceedings. Proc. 2nd Asian Conference on Machine Learning, ACML2010, vol 13, pp 1–16

Iyer SP, Shafran I, Grayson D, Gates K, Nigg JT, Fair DA (2013) Inferring functional connectivity in MRI using Bayesian network structure learning with a modified PC algorithm. Neuroimage 75:165–175

Joshi AA, Joshi SH, Leahy RM, Shattuck DW, Dinov I, Toga AW (2010) Bayesian approach for network modeling of brain structural features. In: SPIE medical imaging. International Society for Optics and Photonics, pp 762607

Kamiński M, Ding M, Truccolo WA, Bressler SL (2001) Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol Cybern 85(2):145–157

Karunanayaka PR, Wilson DA, Vasavada M, Wang J, Martinez B, Tobia MJ, Kong L, Eslinger P, Yang QX (2015) Rapidly acquired multisensory association in the olfactory cortex. Brain Behav 5(11):e00390

Kim J, Zhu W, Chang L, Bentler PM, Ernst T (2007) Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. Hum Brain Mapp 28(2):85–93

Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen MY, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NUF, Schlaggar BL, Mumford JA, Poldrack RA, Petersen SE (2015) Functional system and areal organization of a highly sampled individual human brain. Neuron 87(3):658–671

Li J, Wang ZJ (2009) Controlling the false discovery rate of the association/causality structure learned with the PC algorithm. J Mach Learn Res 10:475–514

Lian Z, Li X, Xing J, Lv J, Jiang X, Zhu D, Zhang S, Xu J, Potenza MN, Liu T et al. (2014) Exploring functional brain dynamics via a Bayesian connectivity change point model. In: 2014 IEEE 11th international symposium on biomedical imaging (ISBI), IEEE, pp 600–603

Liu A, Li J, Wang ZJ, McKeown MJ (2012) A computationally efficient, exploratory approach to brain connectivity incorporating false discovery rate control, a priori knowledge, and group inference. Comput Math Methods Med 2012:967380

Liu Y, Wu X, Zhang J, Guo X, Long Z, Yao L (2015) Altered effective connectivity model in the default mode network between bipolar and unipolar depression based on resting-state fMRI. J Affect Disord 182:8–17

Logothetis NK (2008) What we can do and what we cannot do with fMRI. Nature 453(7197):869–878

Lungarella M, Ishiguro K, Kuniyoshi Y, Otsu N (2007) Methods for quantifying the causal structure of bivariate time series. Int J Bifurc Chaos 17(03):903–921

Mandeville JB, Marota JJ, Ayata C, Zaharchuk G, Moskowitz MA, Rosen BR, Weisskoff RM (1999) Evidence of a cerebrovascular postarteriole windkessel with delayed compliance. J Cereb Blood Flow Metab Off J Int Soc Cereb Blood Flow Metab 19(6):679–689

Manelis A, Almeida JRC, Stiffler R, Lockovich JC, Aslam HA, Phillips ML (2016) Anticipation-related brain connectivity in bipolar and unipolar depression: a graph theory approach. Brain 139(9):2554–2566

Marinazzo D, Liao W, Chen H, Stramaglia S (2011) Nonlinear connectivity by Granger causality. Neuroimage 58(2):330–338

Meek C (1995) Causal inference and causal explanation with background knowledge. In: Proceedings of the Eleventh conference on uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc, pp 403–410

Meek C (1997) Graphical models: selecting causal and statistical models. Ph.D. thesis, Carnegie Mellon University

Molenaar PCM (2004) A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. Meas Interdiscip Res Perspect 2(4):201–218

Mumford JA, Ramsey JD (2014) Bayesian networks for fMRI: a primer. NeuroImage 86:573–582

Nichols TT, Gates KM, Molenaar P, Wilson SJ (2014) Greater BOLD activity but more efficient connectivity is associated with better cognitive performance within a sample of nicotine-deprived smokers. Addict Biol 19(5):931–940

Ogarrio JM, Spirtes P, Ramsey J (2016) A hybrid causal search algorithm for latent variable models. J Mach Learn Res 52:368–379

Pearl J (1986) Fusion, propagation, and structuring in belief networks. Artif Intell 29(3):241–288
Pearl J (2004) Graphical models for probabilistic and causal reasoning. Computer science handbook, 2nd edn. Chapter 70, pp 1–18. CRC Press, Boca Raton, Florida
Perez CA, El-Sheikh EM, Glymour C (2010) Discovering effective connectivity among brain regions from functional MRI data. Int J Comput Healthc 1(1):86–102
Peterson BS, Potenza MN, Wang Z, Zhu H, Martin A, Marsh R, Plessen KJ, Yu S (2009) An FMRI study of the effects of psychostimulants on default-mode processing during Stroop task performance in youths with ADHD. Am J Psychiatry 116(11):1286–1294
Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE (2011) Functional network organization of the human brain. Neuron 72(4):665–678
Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE (2014) Methods to detect, characterize, and remove motion artifact in resting state fMRI. NeuroImage 84:320–341
Price RB, Lane S, Gates K, Kraynak TE, Horner MS, Thase ME, Siegle GJ (2016) Parsing heterogeneity in the brain connectivity of depressed and healthy adults during positive mood. Biol Psychiatry. doi:10.1016/j.biopsych.2016.06.023
Qi R, Zhang LJ, Zhong J, Zhang Z, Ni L, Jiao Q, Liao W, Zheng G, Lu G (2013) Altered effective connectivity network of the basal ganglia in low-grade hepatic encephalopathy: a resting-state fMRI study with Granger causality analysis. PloS One 8(1):e53677
Ramsey JD, Hanson SJ, Glymour C (2011) Multi-subject search correctly identifies causal connections and most causal directions in the DCM models of the Smith et al. simulation study. NeuroImage 58(3):838–848
Ramsey J, Hanson SJ, Hanson C, Halchenko YO, Poldrack RA, Glymour C (2010) Six problems for causal inference from fMRI. NeuroImage 49(2):1545–1558
Ramsey JD (2015) Scaling up greedy causal search for continuous variables (arXiv preprint)
Ramsey JD, Sanchez-Romero R, Glymour C (2014) Non-Gaussian methods and high-pass filters in the estimation of effective connections. NeuroImage 84:986–1006
Ray S, Gohel SR, Biswal BB (2015) Altered functional connectivity strength in abstinent chronic cocaine smokers compared to healthy controls. Brain Connect (150610044446004) (in press)
Roebroeck A, Formisano E, Goebel R (2005) Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage 25(1):230–242
Sanchez-Romero R (2012) Formation of variables for brain connectivity. Ph.D. thesis, Carnegie-Mellon University
Schiatti L, Nollo G, Rossato G, Faes L (2015) Extended Granger causality: a new tool to identify the structure of physiological networks. Physiol Meas 36(4):827–43
Schreiber T (2000) Measuring information transfer. Phys Rev Lett 85(2):461–464
Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464
Seth AK (2010) A MATLAB toolbox for Granger causal connectivity analysis. J Neurosci Methods 186(2):262–273
Shimizu S, Hoyer PO, Hyvärinen A, Kerminen A (2006) A linear non-Gaussian acyclic model for causal discovery. J Mach Learn Res 7:2003–2030
Smith SM, Bandettini PA, Miller KL, Behrens TEJ, Friston KJ, David O, Liu T, Woolrich MW, Nichols TE (2012) The danger of systematic bias in group-level FMRI-lag-based causality estimation. Neuroimage 59(2):1228–1229
Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, Ramsey JD, Woolrich MW (2011) Network modelling methods for FMRI. NeuroImage 54(2):875–891
Sokolov AA, Erb M, Gharabaghi A, Grodd W, Tatagiba MS, Pavlova MA (2012) Biological motion processing: the left cerebellum communicates with the right superior temporal sulcus. NeuroImage 59(3):2824–2830
Sörbom D (1989) Model modification. Psychometrika 54(3):371–384
Spirtes P, Glymour C (1991) An algorithm for fast recovery of sparse causal graphs. Soc Sci Comput Rev 9(1):62–72
Spirtes P, Glymour C, Scheines R (1993) Causation, prediction, and search. MIT press, Boston
Spirtes P, Meek C, Richardson T (1995) Causal Inference in the presence of latent variables and selection bias. In: Proceedings of the Eleventh conference on uncertainty in artificial intelligence, pp 499–506
Sporns O, Chialvo DR, Kaiser M, Hilgetag CC (2004) Organization, development and function of complex brain networks. Trends Cognit Sci 8(9):418–425

Stephan KE, Weiskopf N, Drysdale PM, Robinson PA, Friston KJ (2007) Comparing hemodynamic models with DCM. NeuroImage 38(3):387–401

Strobl EV, Spirtes PL, Visweswaran S (2016) Estimating and controlling the false discovery rate for the PC algorithm using edge-specific P-values. arXiv:1607.03975 (arXiv preprint)

Sun J, Hu X, Huang X, Liu Y, Li K, Li X, Han J, Guo L, Liu T, Zhang J (2012) Inferring consistent functional interaction patterns from natural stimulus FMRI data. NeuroImage 61(4):987–999

Swanson NR, Granger CWJ (1997) Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. J Am Stat Assoc 92(437):357–367

Tashiro T, Shimizu S, Hyvärinen A, Washio T (2014) ParceLiNGAM: a causal ordering method robust against latent confounders. Neural Comput 26(1):57–83

Weissenbacher A, Kasess C, Gerstl F, Lanzenberger R, Moser E, Windischberger C (2009) Correlations and anticorrelations in resting-state functional connectivity mri: a quantitative comparison of preprocessing strategies. Neuroimage 47(4):1408–1416

Wen X, Rangarajan G, Ding M (2013) Is Granger causality a viable technique for analyzing fMRI data? PloS One 8(7):e67428

Wink AM, Roerdink JBTM (2006) BOLD noise assumptions in fMRI. Int J Biomed Imaging 2006:12014

Xu L, Fan T, Wu X, Chen K, Guo X, Zhang J, Yao L (2014) A pooling-LiNGAM algorithm for effective connectivity analysis of fMRI data. Front Comput Neurosci 8(October):125

Yang J, Gates KM, Molenaar P, Li P (2015) Neural changes underlying successful second language word learning: an fMRI study. J Neurolinguistics 33:29–49

Zang Z-X, Yan C-G, Dong Z-Y, Huang J, Zang Y-F (2012) Granger causality analysis implementation on MATLAB: a graphic user interface toolkit for fMRI data processing. J Neurosci Methods 203(2):418–426

Zelle SL, Gates KM, Fiez JA, Sayette MA, Wilson SJ (2016) The first day is always the hardest: functional connectivity during cue exposure and the ability to resist smoking in the initial hours of a quit attempt. NeuroImage. doi:10.1016/j.neuroimage.2016.03.015