

## Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples

Kathleen M. Gates\*, Peter C.M. Molenaar

Department of Human Development and Family Studies, Pennsylvania State University, USA

### ARTICLE INFO

#### Article history:

Accepted 15 June 2012

Available online 23 June 2012

#### Keywords:

Effective connectivity mapping

Networks

SEM

Heterogeneity

Automatic search

### ABSTRACT

At its best, connectivity mapping can offer researchers great insight into how spatially disparate regions of the human brain coordinate activity during brain processing. A recent investigation conducted by Smith and colleagues (2011) on methods for estimating connectivity maps suggested that those which attempt to ascertain the direction of influence among ROIs rarely provide reliable results. Another problem gaining increasing attention is heterogeneity in connectivity maps. Most group-level methods require that the data come from homogeneous samples, and misleading findings may arise from current methods if the connectivity maps for individuals vary across the sample (which is likely the case). The utility of maps resulting from effective connectivity on the individual or group levels is thus diminished because they do not accurately inform researchers. The present paper introduces a novel estimation technique for fMRI researchers, Group Iterative Multiple Model Estimation (GIMME), which demonstrates that using information across individuals assists in the recovery of the existence of connections among ROIs used by Smith and colleagues (2011) and the direction of the influence. Using heterogeneous in-house data, we demonstrate that GIMME offers a unique improvement over current approaches by arriving at reliable group and individual structures even when the data are highly heterogeneous across individuals comprising the group. An added benefit of GIMME is that it obtains reliable connectivity map estimates equally well using the data from resting state, block, or event-related designs. GIMME provides researchers with a powerful, flexible tool for identifying directed connectivity maps at the group and individual levels.

© 2012 Elsevier Inc. All rights reserved.

### Introduction

Brain connectivity maps represent the state-of-the-art methods for understanding brain processing. Despite great advances in the field, the utility and robustness of these methods often fall short (Smith et al., 2011). One routinely overlooked cause of inaccuracies in connectivity maps, also called “networks”, is the aggregation of the data across individuals prior to estimation (Kherif et al., 2003; Ramsey et al., 2010). Statistical and empirical work has demonstrated that in cases where processes are heterogeneous across individuals, aggregation of the data to arrive at a “group” solution may fail to describe any individual in the sample (Miller and van Horn, 2007; Molenaar and Campbell, 2009). While researchers acknowledge that the group model may not describe any one individual, they rarely assess the degree to which the resulting model describes the individuals comprising the group. Hence it typically remains unknown if the published findings derived from data aggregated in this manner relate to how individuals' brains actually function. This is particularly

evident in the case of effective, or directed (Friston, 2007), connectivity mapping approaches.

There are three primary ways that group models may fail to describe individuals in the context of effective connectivity modeling. First, the beta weights (parameters associated with the couplings of two regions) may vary across participants. Evidence exists for beta weights systematically varying across subjects (e.g., Kim et al., 2007), and examination of how variation in beta weight estimates relates to clinical diagnoses or performance is a focal point for many researchers. Only if the influence that other ROIs might have on the target ROI is explicitly accounted for may unbiased estimates be obtained for the beta weight of interest. Hence, acquiring reliable connectivity map structures is a first requirement for analysis on the beta weights.

Second, the network of couplings among ROIs may differ. Emerging evidence suggests that much heterogeneity in brain processes exists across individuals both in terms of the presence of relations among regions (Fair et al., 2010; Hillary et al., 2011) and which regions become active (Kherif et al., 2003, 2009; Miller and van Horn, 2007; Miller et al., 2002; Seghier et al., 2008). When traditional group-level analysis is conducted on processes which differ, spurious findings emerge (e.g., Miller and van Horn, 2007; Miller et al., 2002). To date, the predominate methods for analyzing brain data assume

\* Corresponding author at: Department of Human Development and Family Studies, Pennsylvania State University, S110 Henderson, University Park, PA 16802, USA.  
E-mail address: [kgates@psu.edu](mailto:kgates@psu.edu) (K.M. Gates).

homogeneity across individuals despite these findings that individuals may differ greatly (although one study found evidence for similarities in connectivity maps, see [James et al., 2009](#)). Taken together, it appears that there may be some paths that are common to the majority of individuals comprising the group while some paths may be unique to certain individuals or to a subgroup of individuals. This may be because of different relations among the regions selected for connectivity analysis, or perhaps because for some individuals a region is not part of the connectivity map. If the presence of paths differs among individuals, then the aggregated results may not represent any of the individuals comprising the group ([Molenaar and Campbell, 2009](#)). Currently, no method utilizes shared information during model selection to find paths common to the group while allowing for paths unique to individuals.

Third, connectivity maps may evidence heterogeneity across individuals in terms of the direction of an effect. Much attention has been given to the inadequacy of most individual-level effective connectivity approaches for arriving at models that identify both the presence and direction of connections ([Smith et al., 2011](#)). In this examination of both effective and functional connectivity mapping approaches, methods for arriving at connectivity maps that are agnostic in terms of directionality when detecting the presence of a relation prevailed. Hence little is known about the true degree of heterogeneity that may exist across individuals in terms of the direction of effects between a given pair of ROIs since previous inferences (particularly those drawn from lagged methods) may have relied on spurious findings.

Examination of functional heterogeneity in brain processing across individuals first requires reliable recovery of true effects. This is true for the examination of varying strengths or different patterns of relations among regions. One potential solution would be to first arrive at individual-level maps and then identify subgroups (should they exist) using graphical clustering across individuals ([Van den Heuvel et al., 2008](#)). Given recent evidence that even the best methods fail to recover the structure of effective connectivity maps at the individual level when certain types of noise and conditions are present ([Smith et al., 2011](#)), this approach may lead to spurious findings since the subgroups will be based on unreliable maps. On the other end of the spectrum, [Ramsey et al. \(2011\)](#) have achieved great success when using a multi-sample approach for estimating effective connectivity maps. Building from the Greedy Equivalence Search (GES; [Meek, 1997](#)), an approach which identifies candidate paths to free up using a score function based on the maximum likelihood estimations, the approach capitalizes on commonalities among individuals to arrive at a group model ([Ramsey et al., 2010](#)). Extensions of this approach were recently applied to the [Smith et al. \(2011\)](#) data sets and have surfaced as the first set of methods for identifying true connections as well as the directionality with as few as 10 individuals in a group ([Ramsey et al., 2011](#)). [James et al. \(2009\)](#) developed a promising model selection approach that applies all viable models to data concatenated across individuals and selects the best one. Going one step further than many researchers, [James et al. \(2009\)](#) assessed the degree to which the group-derived model fit individuals. The group-derived model for an empirical data set fit 85% of the individuals excellently. Clearly powerful procedures which offer an improvement upon connectivity models to date, both require concatenation across individuals much like other current methods, suggesting that it may be best for situations in which homogeneity is suspected.

The presence of noise in fMRI data further confounds the issue of heterogeneity. An approximation of neural activity, fMRI data is subject to many sources of noise due to measurement error which may make individuals' processes look different when they are same. The [Smith et al. \(2011\)](#) simulations provided an excellent opportunity to investigate the impact of varied neural-hemodynamic relations on the ability to recover reliable connectivity maps. Their results have become the benchmark for success when evaluating the utility of

novel analytic methods. In the scenario where one neuronal input influences activity across all ROIs, the best of 38 methods tested recovered only 50% of the true connections. In the face of shortened neural lags and shared inputs the best recovery rates neared 70%. In this way, individual nuances outside the control or the measurement of the researcher may produce maps which appear heterogeneous when in fact the neural structure is homogeneous across individuals.

These realities place researchers at a quandary. On the one hand, true heterogeneity in individual processes may make group-level findings erroneous because the resulting map may not apply to any one individual ([Miller and van Horn, 2007](#); [Molenaar, 2004](#); [Ramsey et al., 2010](#)). On the other hand, the signal-to-noise ratio in fMRI data may be too low for reliable recovery at the individual level, making designs which capitalize on shared information across individuals appear to be a promising approach. Recently, a suggestion has been made for the development of a procedure that enables individual-level modeling while arriving at robust similarities across individuals ([Smith, 2012](#)). We present a timely, reliable, and novel approach, the Group Iterative Multiple Model Estimation (GIMME), which addresses the issue of heterogeneity (i.e., the need for individual-level maps) in effective connectivity mapping while capitalizing on shared information to arrive at group inferences. The model estimation procedure incorporates the following improvements. First and foremost, the selection procedure of candidate paths utilized by GIMME appropriately picks up signal from noise to arrive at a common structure among individuals (should one exist) that appropriately describes the majority of individuals. Second, GIMME continues to evaluate candidate paths at the individual level, freeing those paths for that individual which will improve the model fit. Taken together, this approach enables reliable group inferences with a model selection algorithm similar to those based on maximum likelihood improvements ([James et al., 2009](#); [Ramsey et al., 2011](#)) that diverges from these approaches by enabling reliable estimation of paths unique to the individual and not forcing a group model. This paper presents the formal aspects of GIMME and demonstration of GIMME using the homogeneous [Smith et al. \(2011\)](#) data and heterogeneous in-house data.

## Methods: formal explication of GIMME

GIMME estimates both the unified SEM (uSEM; [Kim et al., 2007](#)) and extended unified SEM (euSEM; [Gates et al., 2011](#)). The uSEM assuages problems encountered when solely lagged or contemporaneous relations are modeled with fMRI data ([Gates et al., 2010](#)) and is useful for block designs (when the researcher would like separate maps for each block) and resting-state analysis. The euSEM builds from uSEM by allowing for direct and modulating (i.e., bilinear) experimental manipulation effects (much like the DCM) for block and event-related designs. The logic behind GIMME follows directly from well-established principles. GIMME first identifies a group model (i.e. a connectivity map common to most individuals in the sample) by selecting paths which will improve the majority of individuals' maps in an iterative forward selection procedure. Next, individual-level paths that will optimally improve that model are opened ([Gates et al., 2010](#)).

GIMME may be applied to most connectivity mapping approaches and has been tested for use with effective connectivity methods. We focus here on the euSEM ([Gates et al., 2011](#)) since it contains within it the uSEM ([Kim et al., 2007](#)) and thus offers a more general description. The euSEM is as follows:

$$\eta(t) = \underbrace{A\eta(t)}_{\text{Cont.}} + \underbrace{\sum_{m=1}^q \phi_m \eta(t-m)}_{\text{Lagged}} + \underbrace{\sum_{j=0}^k \gamma_j u(t-j)}_{\text{Input}} + \underbrace{\sum_{m=1}^s \sum_{j=1}^r \tau_{mj} \eta(t-m) u(t-j)}_{\text{Modulating}} + \zeta(t) \quad (1)$$

where  $q, k, s,$  and  $r$  represent the lags,  $\eta(t), t = 1, 2, \dots, T$ , indicates the manifest  $p$ -variate timeseries of ROI activity (where  $t$  ranges across the sequence of scans),  $A$  the  $(p, p)$ -dimension matrix of contemporaneous relations among ROIs,  $\varphi_m$  the  $(p, p)$ -dimension matrix of the associations among ROIs at a lag of  $m$ ,  $u(t-j)$  a univariate input series at lag  $j$  (which may be expanded to include multiple inputs) convolved with a hemodynamic response function which models the lag between neuronal activity and blood response known to exist in fMRI data (Sarty, 2007),  $\Upsilon_j$  a vector of input effects on ROIs,  $\tau_{m,j}$  the  $(p, p)$ -dimension matrix associated with the bilinear term  $\eta(t-m)u(t-j)$ , and  $\zeta(t)$  a  $p$ -vector error series assumed to be a white noise process. The linear combination  $\tau_{m,j}u(t-j)$  represents the modulating effect of the input  $u(t-j)$  on the connections among ROIs at lag  $j$  (see Fig. 1). Please note that  $u(t-j)$  may be an input vector from any design; in the case of a block design a series of ones would indicate when a specific condition was present whereas for an event-related design the ones may be spread out more. The input vector may even indicate the participant's response, such as when a wrong answer is given, to identify processes relating to poor performance. In this last case, the ability to detect an effect may differ across participants (e.g., the power to detect effects among those with few incorrect answers will be lower than those with many incorrect answers). GIMME will help in detecting these effects by looking across individuals.

The euSEM as written above applies to a single multivariate time series (i.e., a series of scans at a given TR) which may be an individual's series of node activity when run on individual data or may be for a series representing the group by aggregating the data into one series. GIMME further extends this modeling approach by identifying a connectivity map structure for the group as well as connections for the individual. The above general equation may be modified as follows to represent this by adding the subscript "i" to indicate the parameters unique to the  $i$ -th individual,  $i = 1, 2, \dots, N$ , and the superscript "g" to indicate parameters in the group model:

$$\eta_i(t) = \underbrace{(A_i + A^g)}_{\text{Contemporaneous}} \eta_i(t) + \underbrace{\sum_{m=1}^q (\phi_{i,m} + \phi_m^g)}_{\text{Lagged}} \eta_i(t-m) + \underbrace{\sum_{j=0}^k (\gamma_{i,j} + \gamma_j^g)}_{\text{Input}} u_i(t=j) + \underbrace{\sum_{m=1}^s \sum_{j=1}^r (\tau_{i,m,j} + \tau_{m,j}^g)}_{\text{Modulating}} \eta_i(t-m) u_i(t-j) + \zeta_i(t). \quad (2)$$

Setting  $q=k=s=r=1$  represents a model for a lag of one. A notable characteristic of the parameter matrices  $A^g, \phi^g, \Upsilon^g,$  and  $\tau^g$  is that they would not be identical to those found in the first equation but rather indicate the mean parameter matrices found for the group using the iterative procedure described above. For connections which do not exist in the group model, the mean will likely be near zero since the parameter is not estimated for everyone. The distribution of those parameters not identified in the group model will thus be bimodal (or even trimodal if there are negative and positive estimates across individuals) rather than normal, with a spike at zero. This represents an important difference from random effects models which induce a normal distribution on parameters under the assumption that one model can explain all individuals. Unlike random (or mixed) effects models, GIMME allows for the structure of the connectivity maps to be unique across individuals (i.e., person-specific).

Model identification via LISREL uses quasi-maximum likelihood estimation (Jöreskog and Sörbom, 1992; Molenaar, 1985) which consists of deriving the covariance matrix of the ROIs at a finite number of lags, and in the case of euSEM estimation, the direct and bilinear effects of experimental manipulation. It capitalizes on the fact that GIMME is linear in the parameter matrices, enabling straightforward derivation of the implied covariance structure and application of the commercially available SEM software to fit the model.

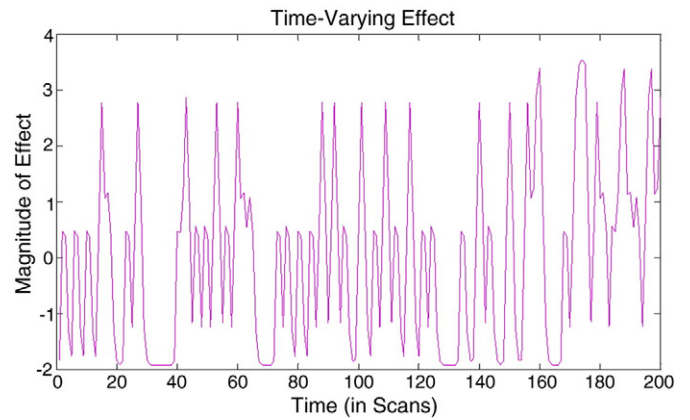


Fig. 1. Time-varying influence of simulated series ROI 4 on ROI 5 as a function of onset vector representing experimental manipulation in an event-related design ( $\tau_{m,j}u(t-j)$ ).

The model selection procedure begins by running the null (empty) model on each participant's data using the proprietary LISREL program. In the case where a priori anatomical or other knowledge suggests certain paths should be freed, the following path-selection procedure may be implemented in a semi-confirmatory manner by having the null model include the desired paths. In either case, running the null model results in a matrix of Lagrange multiplier test equivalents called, "modification indices" (MI; Sörbom, 1989) for each individual. These asymptotically chi-square distributed indices indicate the expected increase in likelihood if that parameter (in this case, path) were freed. These values will be sensitive to the order in which candidate paths are freed since estimating a chosen path will influence how much variation remaining candidate paths may explain. The candidate paths represent lagged or contemporaneous effects if the uSEM is implemented for resting-state or block design data, and lagged, contemporaneous, experimental manipulation, or bilinear effects (i.e., how the relationship between two ROIs is modulated by a given experimental manipulation) if euSEM is implemented for the data from event-related designs.

When conducting analysis on a single data set, the MI indicates which parameter (path) would optimally improve the individual model if freed. The GIMME program follows this rationale by identifying which MIs are significant at the .01 level. The parameter selected to be freed for the group is the one that is significant for the greatest number of individuals. By looking across individual samples for consistency, effects common to most become clear. This signal processing approach is most common in genomics (e.g., Guttman et al., 2007). The goal at this stage is to identify the existence of any paths common to the group. Thus it is necessary to require that model fits would significantly improve for the majority of individuals. Towards this end, each candidate parameter receives a count of how many individuals' models would significantly improve at the .01 level if the parameter were freed:

$$D_j = \sum_{i=1}^N Y_{j,i} / N$$

where

$$\begin{cases} MI_{j,i} \geq \chi_{1,\alpha=.01}^2, & Y = 1 \\ MI_{j,i} < \chi_{1,\alpha=.01}^2 & Y = 0 \end{cases} \quad (3)$$

'j' indicates the candidate parameter, 'i' the individual,  $D_j$  the proportion of individuals for whom a relationship is detected,  $\chi_{1,\alpha=.01}^2$  the nominal value of the chi-squared variate with one degree of freedom at the chosen alpha of .01, and  $Y$  indicates the presence or absence of the detection of a path according to the MI.

The proportion for what constitutes the cutoff, i.e., the expected value of  $D_j$  needed for the parameter to be freed across all individuals, may be set by the researcher. A value of 1.0 would necessitate that this parameter, if freed, would describe each individual's map. For data such as fMRI data which likely contains much noise (as discussed

in the [Introduction](#)) and potential for false negatives, a looser criterion may be appropriate. In the case of GIMME, the detection of a connection is directly related to the MIs, which are generated for each individual independent of other individuals in the sample. [Smith et al. \(2011\)](#) indicate that, across “typical” data sets, the following probability of detection the presence of an effect may be expected for each session length: “60 min: 100%; 10 min: 95%; 5 min: 77%; and 2.5 min: 59%” (p. 887). Probability of detection at the individual level varies across method of choice and length of series.

To recap the group model selection procedure, the program iteratively identifies the next parameter (directed path) that optimally improves the fit of the common model for most individuals according to the  $D_j$  criterion described above until no further parameters can be found. Then, in a next step, the obtained common model is pruned by eliminating those paths which, because of the freeing up of paths at later iterations, no longer are acceptable according to the  $D_j$  criterion. This procedure is consistent with the steps in the individual-level model search within LISREL ([Jöreskog and Sörbom, 1992](#)).

GIMME identifies models on the individual level in a semi-confirmatory manner. The search for the optimal model on the individual level does not begin with the null (empty) model as for the group. Rather, the first iteration estimates the structure of paths found in the group search. Then, the automatic search procedure within LISREL identifies iteratively which individual-specific parameter according to the MI would optimally improve the model. At the final stage the nonsignificant paths in the subject-specific part of each individual's model are removed (providing they do not exist in the group model) and the resulting GIMME model is checked in a final confirmatory fit. In this manner, individual-level parameter estimates are obtained for both the connections that were identified in the group-level analysis as well as those that surfaced only for the individual. The final model for each individual must meet the criteria below on two of the following four fit indices which demonstrated reliability in simulation studies ([Brown, 2006](#)): root mean square error of approximation (RMSEA), non-normed fit index (NNFI), comparative fit index (CFI), and standardized root mean square residual (SRMR). For RMSEA, values less than .05 indicate an acceptable fit, for SRMR values less than .05, and for NNFI and CFI values greater than .95 indicate an excellent fit. [Fig. 2](#) offers a schematic diagram of the process. The final model for each individual consists of a partial connectivity map which is common for all individuals added to a partial individual-specific connectivity map. Below, the ability for GIMME to

correctly identify the existence of a relation and the direction is evaluated across individuals exactly as in the [Smith et al. \(2011\)](#) paper.

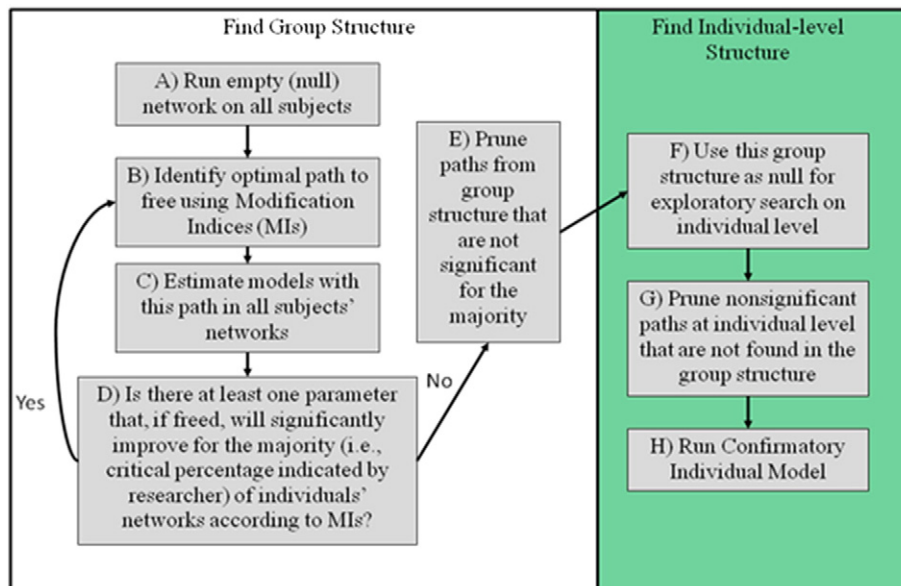
## Data and results

### Homogenous data: [Smith et al. \(2011\)](#) simulation 2

The reader is referred to [Smith et al. \(2011\)](#) for details regarding the simulation of the data set. In short, a block-design data set was simulated using the DCM forward model ([Friston et al., 2003](#)) which is based upon the nonlinear balloon model ([Buxton et al., 1998](#)). For ease in comparison, we selected simulation set number two (out of a possible 28 simulations) since it was deemed to be the most representative of fMRI data by [Smith et al. \(2011\)](#) and was given the most attention in the original paper. This data set used a structure containing 10 nodes (ROIs) as seen in [Fig. 3](#), with 10 min of data at a TR of 3 s for a total of 200 observations (scans) for each of the 50 simulated participants.

None of the 28 effective connectivity methods (out of 38 methods total) tested by [Smith et al. \(2011\)](#) did well in terms of recovering both the presence of a connection and the directionality of a connection. Those that did well on one construct fared poorly on the other. For instance, the lagged-based effective connectivity methods only identified the existence of a connection about 20% of the time. The Bayes net approaches did much better at identifying the presence of a connection (~90%) but were unable to correctly identify the directionality of the connection above what would be expected by chance. The method which performed best in terms of identifying the correct direction of paths, Patel's  $\tau$ , did so only 65% of the time but detected only the presence of 20% of the true connections.

When the same data were fit to uSEM using GIMME, the parameters were recovered excellently by all criteria, demonstrating that utilizing information across individuals helps arrive at reliable estimates. Since the final models for each individual are arrived at by obtaining individual-level parameter estimates at each step of the iterative search procedure, comparison of weights across individuals, much like that in [Smith et al. \(2011\)](#), is warranted. The distribution of t-scores for the true paths for each individual was substantially higher than seen for false paths ([Fig. 4](#)). The distribution of the estimates for false positives was around zero, which is what one would expect. This pattern is different, and better, than that seen for the majority of effective connectivity methods evaluated by [Smith et al. \(p. 882;](#)



**Fig. 2.** Schematic diagram of GIMME approach for estimating unified SEMs and extended unified SEMs (uSEMs and euSEMs, respectively).

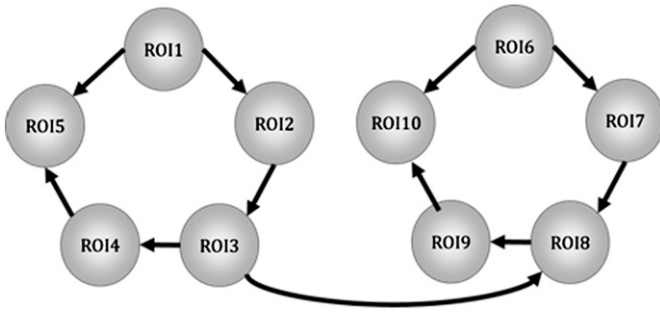


Fig. 3. Connectivity map structure for the Smith et al. (2011) simulations.

2011). It must be stressed that while a group structure is obtained for arriving at a map which is common to the majority of individuals, parameter estimates are conducted on an individual level for both the connections freed which are common to the group and those that are unique for the individual. GIMME correctly identified 100% of the true connections (at above the 95% of the distribution of false positive t-values, called c-sensitivity by Smith et al., 2011) which existed in the model across all individual replications. This is the same rate reported in Ramsey et al. (2011) for the same data (but using a recall metric which does not consider false positives) and is higher than the best approaches out of the 38 methods tested by Smith et al. (2011). For the methods presented in Ramsey et al. (2011) and the present paper, the added accuracy can in part be attributed to increasing the sample size from one individual.

In addition to identifying the presence of a connection, uSEM implemented via GIMME correctly identified the direction of the contemporaneous connections 90% of the time according to Smith et al.'s (2011) criterion for direction sensitivity (d-accuracy). The d-accuracy index indicates the mean percent of true connections for which the true direction estimation was higher than that of the false direction. Identifying directionality 90% of the time offers an improvement upon the individual-level methods tested by Smith et al. (2011). Using the criteria of direction recall (i.e., the number of correct directions recovered out of the total number in the simulated data and used by Ramsey et al. (2011)) the average rate increased to 93%, which is comparable to the ability of the best methods in Ramsey et al. (2011) (91–96%). These findings again highlight the added benefit of utilizing shared information across individuals when arriving at reliable map structures.

When uSEM was estimated for each individual without using the GIMME algorithm, results were similar to the best effective connectivity methods seen in Smith et al.'s (2011) and Ramsey et al.'s (2011) approaches conducted on the individual level. Across individuals, 90% of presence of relations were detected when uSEM was used on the individual level (using the c-sensitivity criteria that true positives must be above 95% of the distribution of false positives); this

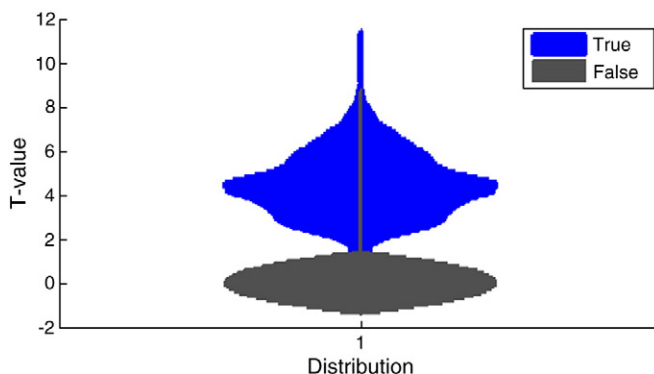


Fig. 4. Violin plots of distributions of true and false paths.

near the best approaches identified in Smith et al. (2011) which were above 90% but do not appear to reach above 95%. On the individual level, the Ramsey et al. (2011) approach identified 85–92% of the relations in the true model when examined in terms of presence recall, or the percentage of true effects recovered out of the total effects used to simulate the data (using this criterion, the individual-level uSEMs had 91% presence recall). In terms of directionality, uSEM again performed similarly to its competitors with identification of directionality being no better than chance (50%) using the d-accuracy criteria used by Smith et al. (2011). The methods in the Smith et al. (2011) paper which were able to identify the existence of a relation did equally poorly with an average identification of true direction only 50% of the time. The individual-level methods in the Ramsey et al. (2011) paper recovered the true direction only 40–42% of the time. While recovery of the presence of connections seems to be reliable at the individual level using uSEM, increasing the sample size appears to assist in identifying the true direction in homogeneous data sets (see Appendix A for the results of solely individual-level analysis and GIMME conducted with groups of 10, 25, and 50 with replacement).

GIMME and other recent methods highlight the added utility of using shared information across individuals. However, it does not work for all group approaches. For this data set, the model obtained on the concatenated series recovered only six of the eleven paths (55%) used to create the data (with no false paths). Five of the paths were in the correct direction (83% according to the criteria used in Smith et al. (2011)). While this model obtained an excellent fit for the concatenated series, when run on individuals the model was not an excellent fit for any individual by any of the five fit indices used discussed in the methods (see Table 1). Hence the group-derived model obtained by concatenating the series failed to describe the individuals comprising the group despite being an excellent fit at the group level. This highlights the need for alternative algorithms for arriving at paths which are common to the group.

The results obtained by GIMME for the majority of data sets simulated by Smith et al. (2011) are supplied in Appendix B. Overall, GIMME excellently recovered the parameters used to create the data. For 92% of the connections across all individuals across the simulations, GIMME estimated the true connection weight as being above the 95%-ile of false connections (i.e., c-sensitivity; Smith et al., 2011). This percentage was brought down from closer to 100% because of GIMME's inability to reliably recover connections when the TR was reduced to an (at this time) unrealistic .25s. In the presence of such short temporal lags with biological data, the autoregressive components estimated in the course of model selection explained such a high degree of variance that few other paths were opened for these simulations. When not considering these two simulations, GIMME's performance (in terms of recovering the existence of true paths) rises to 97%. Of note, GIMME identified the existence of 100% of connections when the source of noise was outside of the researchers' control, such as the case in which there are shared input influencing ROI BOLD activity (simulations 8 and 9; versus 70% in Smith et al. (2011) and 96–98% in Ramsey et al. (2011)) and when there is only one neuronal input driving activity across ROIs (simulation 24; versus 50% in Smith et al. (2011) and 84% in Ramsey et al. (2011)). Hence it appears that group approaches such as GIMME aid in picking up signal from noise across multiple types of scenarios likely to occur in functional MRI research. It should be noted that the Ramsey et al. (2011) metric reported here (adjacency recall) differs from the c-sensitivity index used by Smith et al. (2011) and in Appendix B for the present paper and may yield higher recovery rates because it does not require that true relations recovered be higher than false ones (false connections are considered in the Ramsey et al. (2011) precision metric, which does not assess the proportion of paths in the true model that were recovered and is thus not appropriate to compare with c-sensitivity). Correctly identifying directionality varied more across the simulations. In cases with very short sessions (minutes=

**Table 1**

Average fit indices across varied analytic approaches. “Individual” indicates that analysis was solely conducted on the individual-level. “GIMME (##)” indicates the entire GIMME algorithm was used to arrive at individual-level estimates with the ## denoting the number of individuals used to arrive at a common structure. “Concat.” = concatenated; the line for “1” contains the fit indices for when the structure found from the concatenated series was applied to each individual. The line for 50 displays the fit indices for the concatenated series.

	Analysis	Chi	DF	CFI	NNFI	SRMR	RMSEA
Homogeneity: Smith et al. (2011) sim. 2	Individual	128.3	120	0.99	0.99	0.05	0.01
	GIMME (10)	128.6	121	0.99	0.99	0.05	0.01
	GIMME (25)	127.2	121	0.99	0.99	0.05	0.01
	GIMME (50)	141.2	124	0.98	0.97	0.05	0.05
	Concat. 1	250.8	129	0.88	0.82	0.07	0.06
	50	91.0	129	1.00	1.00	0.05	0.00
Heterogeneity: strength	Individual	42.1	26	0.99	0.98	0.04	0.04
	GIMME (10)	27.9	27	1.00	1.00	0.04	0.01
	GIMME (25)	33.1	26	0.99	0.99	0.04	0.02
	GIMME (50)	25.9	27	1.00	1.00	0.04	0.01
	Concat. 1	33.1	26	0.99	0.99	0.04	0.02
	50	0.7	27	1.00	1.00	0.01	0.00
Heterogeneity: 4 subgroups	Individual	59.0	27	0.99	0.98	0.04	0.05
	GIMME (10)	36.0	27	0.99	0.99	0.03	0.02
	GIMME (25)	40.0	28	0.99	0.99	0.03	0.03
	GIMME (50)	37.4	28	0.99	0.99	0.03	0.02
	GIMME (100)	43.1	28	0.99	0.99	0.03	0.03
	Concat. 1	295.4	28	0.85	0.78	0.08	0.19
	50	66.9	29	0.97	0.96	0.07	0.08
Heterogeneity: random	Individual	128.0	60	0.98	0.96	0.03	0.06
	GIMME (10)	99.0	60	1.00	1.00	0.03	0.04
	GIMME (25)	102.7	59	0.99	0.97	0.03	0.04
	GIMME (50)	111.6	60	0.98	0.95	0.02	0.04
	Concat. 1	254.5	61	0.09	0.04	0.92	0.87
	50	16.0	61	1.00	1.00	0.00	0.01

2.5 min, TRs = 3 s), GIMME had a more difficult time recovering the direction of the effect. Finally, at the present time 50 ROIs (as in simulation 4) is too computationally taxing for Lisrel since it results in a 100 × 100 covariance matrix of lagged and contemporaneous relations; analysis of this simulated data thus could not be completed with GIMME, which calls upon Lisrel for model estimation.

Smith et al. (2011) found lagged-based methods to be particularly poor in terms of recovering both the presence and direction of paths. That the uSEM, which contains in it lagged effects, performed so well here may at first glance seem to contradict these findings. On the contrary, the work presented here further supports Smith et al.'s (2011) findings that contemporaneous models best capture the relations. Prior work has demonstrated that when there are effects which occur at a finer temporal time scale than one TR (i.e., a lag of one in most fMRI analysis), lagged-based methods fail to reliably recover these effects if they do not also include contemporaneous effects (Gates et al., 2010; Kim et al., 2007). Similarly, and most relevant to the present work, if contemporaneous effects are estimated and lagged effects which exist are not included, phantom paths will surface (Gates et al., 2010). Since fMRI data are biological and thus contain, like most biological phenomena, autoregressive components, it improves the reliability of model estimation to include the lagged and contemporaneous effects. Much like the Smith et al. (2011) findings, the relations among ROIs (after controlling for autoregressive effects which surfaced in the search

for all ROIs) were contemporaneous. The lagged effects are used to ensure reliable estimation of the contemporaneous effects of interest.

#### Heterogeneous data sets

Three data sets were generated to demonstrate the ability of GIMME to recover connections from differing types of heterogeneity. The first attends to the condition of varying parameter weights across individuals that have identical underlying connectivity models. The latter two data sets demonstrate GIMME's ability to recover connections that may be missed when using the standard approach of concatenating across individuals for group model discovery. Specifically, the second dataset contains connectivity maps that differ systematically across four equally sized subgroups but share some common characteristics. This approach is informed by one study which identified four subgroups from a sample (n = 79) of individual statistical parametric maps (Kherif et al., 2009). This may be a best-case scenario of heterogeneity; what might be equally likely is that some elements of a connectivity map exist for most individuals in the group while other combinations of elements are entirely unique to the individuals. Evidence exists for processes that exhibit this high degree of seemingly random heterogeneity exists (e.g., Hillary et al., 2011; Miller and Van Horn, 2007).

Since GIMME estimates connectivity maps from an SEM framework, data sets were simulated using generative approaches similar to those seen elsewhere in fMRI literature examining SEM reliability (e.g., Penny et al., 2004). From simple algebraic substitution, Eq. (1) above may be rewritten in the following manner to generate data for one individual with a model containing a lag of one scan:

$$\eta(t) = (I - A)^{-1} \left( \Phi \eta(t-1) + \sum_{j=0}^1 \gamma_j u(t-1) + \tau_1 \eta(t-1) u(t-1) + \zeta(t) \right), \quad (4)$$

where  $\zeta(t)$  are zero mean Gaussian innovations with unit variance and are used to start the sequence and removed prior to analysis. Time series of length 200 observations (e.g., scans) were created for each individual in the sets to follow. The weights used to generate the data differ by individual and by simulation set and will be provided below. The cutoff for  $D_j$ , or proportion of individuals whose model a candidate path must improve, was set at .77 across all sets in accordance with the probability of detection findings in the Smith et al. (2011) paper.

#### Heterogeneous data: varying connection strength

Variation in connection strengths constitutes the most basic level of heterogeneity that may exist across individual connectivity maps. The first set of simulated heterogeneous data presented here contains a set of 50 replications, which may be thought as multiple individuals or multiple sessions for one individual, that contain identical structures of connectivity (see Fig. 5), and were generated to emulate data from resting state conditions with no experimentally manipulated stimuli. The weights of connections in the contemporaneous paths (i.e., parameters in the A matrix) were drawn randomly for each parameter for each person from a normal distribution with mean .6 and standard deviation .3. The autoregressive components (i.e., the diagonal of the  $\Phi$  matrix) were held constant at a weight of .5 across all connections for all individuals.

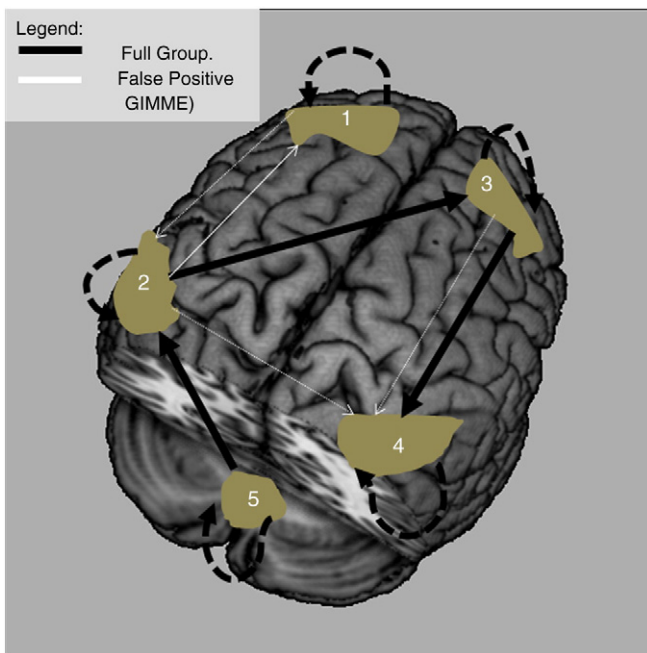
Using uSEM, GIMME recovered all of the connections used to simulate the data across all individuals. Importantly, the estimates related to the true weights used to create the series. The correlation between the simulated parameters and those recovered using GIMME was  $r = .759$  ( $p < .001$ ). The autoregressive components were significant across all individuals. For elements in the A matrix, the estimates for ROI 3 regressed on ROI 2, ROI 4 regressed on ROI 3, and ROI 2 regressed on ROI 5 were significant for 88%, 92%, and 92% of

the individuals, respectively. For this set of data, using shared information across individuals to arrive at some paths that are common to the group improved recovery rates from those obtained when only individual-level analysis is conducted. Still, with as few as 10 individuals the recovery of the presence (97%) and direction (95%) was in acceptable ranges (see Appendix A). As one might expect, the concatenated data series recovered true structure of the model.

#### Heterogeneous data: four subgroups

The second set of data contained four subgroups comprised of 25 individuals (for a total of 100 individuals) who shared some connections in common but differed by a pair of two nonshared elements. All of the beta weights were simulated to equal .70. Fig. 6 displays the results from GIMME as well as the concatenated series. Although this set of data contained clear subgroups, analysis was conducted in an entirely data-driven manner and without prior information of subgroup identification. As above, GIMME first identified connections that existed for the group. Next, the search for individual-level paths identified paths which exist on the individual level. Finally, confirmatory models were run to estimate parameter weights for the group and individual paths simultaneously. GIMME recovered 99% of the group and individual connections with no false positives, and again recovery of estimates was improved by using group analysis (Appendix A). By allowing for individual-level paths to surface, meaningful differences in connectivity structures may emerge that will aid in subgroup identification.

When uSEM was applied to the concatenated series, the resulting model described none of the individuals that comprised the group. Indeed, the connection between ROI 2 and 3 was recovered; however, the directionality was wrong in the contemporaneous effect and the temporal relation was misidentified. Additionally, none of the differences which exist across individuals, which may be meaningful, can be captured by this method. When this model structure was applied to each of the individuals, the model did not fit any of the data excellently by any fit index.



**Fig. 5.** Simulated data network and results for the “varying connection strength” simulated data. Both the GIMME and concatenated approaches recovered all of the parameters. Solid lines denote contemporaneous connections and dashed lines indicate lagged connections of 1 time point (i.e., scan).

#### Heterogeneous data: random maps

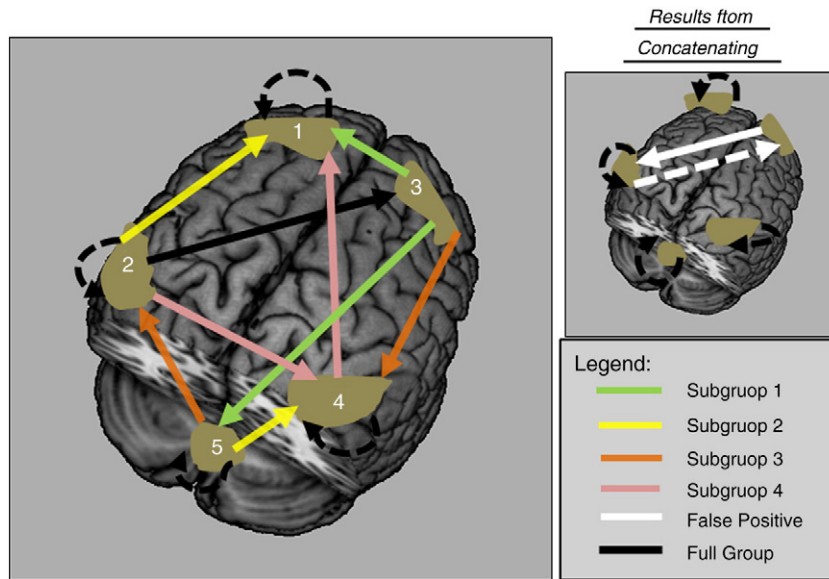
GIMME was next applied to data sets simulated to emulate data obtained during a task implemented with an event-related design when heterogeneity exists in brain connections across individuals across all on and off conditions as well as specifically related to that task. A vector of onset times, which may reflect experimental manipulation, was created such that at each time point there was a .3 chance of the experimental manipulation occurring independent of other time points. This onset vector was convolved with the gamma hemodynamic response function described in Sarty (2007). Series of length 200 TRs ( $T=200$ ) of length 2 s were simulated for a pattern of relations among 5 nodes (ROIs) for 50 replications (or participants). A group structure, or set of paths common to the majority, consisting of five lagged effects (of weight = .70) and two contemporaneous effects among ROIs (60) as well as a direct (.60) and bilinear experimental manipulation effect on the relation between 2 ROIs (−.60) was used as a base network (see Fig. 7).

To emulate the heterogeneity found in empirical network maps (which may indicate unique processing patterns), additional paths were randomly opened for each replication. Each path aside from those in the common structure were opened with a probability of .02 (1/50) independent of other paths which may be opened for the individual. Twenty-nine of the 50 replications had paths in addition to those found in the common structure. The inclusion of a vector of onset times and bilinear effects made the euSEM an optimal modeling method. The time-varying (i.e., bilinear) effect of ROI 4 on ROI 5 as a function of the onset vector representing experimental manipulation is displayed in Fig. 1. Additional bilinear effects that are freed at the individual level could be very important for researchers who are interesting in identifying relations among ROIs that change in the presence of an experimental manipulation.

Results paralleled those seen in the previous examples. First, all of the paths simulated at the group level were recovered in step one of the GIMME procedure. That is, all 9 of the paths used to create all of the individuals' data were recovered in terms of the presence and direction. Furthermore, no false paths surfaced in the group structure. Thus in the face of heterogeneous data, GIMME appropriately picked out signal from noise to obtain a group structure of paths which truly describes the individuals. Starting with the group structure as a base, individual paths were opened that improved model fit for those individuals. GIMME correctly identified the presence and direction of 98% of the paths across individuals. When looking at the directionality of the 98% of connections detected, 100% of the time the correct direction was identified. Similar to result from the “varying parameter strength” condition, conducting euSEM on the concatenated data recovered the true structure. However, when this structure was applied to each of the individuals very poor fits resulted (see Table 1). This underscores that the heterogeneity in the individuals was great enough such that the group-level connections failed to describe the individuals comprising the group.

## Conclusions

GIMME performed excellently in terms of identifying the presence of connections in groups when using the data simulated by Smith et al. (2011), which are homogeneous across replications, as well as arriving at reliable group and individual estimates using in-house data representing heterogeneous processes across individuals. Two aspects of GIMME set it apart from most connectivity methods to date. One, GIMME allows for inclusion of event-related data by implementing euSEM. Simulations contained herein demonstrate the efficacy in detecting direct experimental and bilinear effects at the group and individual levels. These effects help researchers understand how a relationship between two given ROIs changes, or appears, in the presence of experimental manipulation. Few alternatives for



**Fig. 6.** Simulated data network and results for the “four subgroup” simulated data. GIMME obtained no false positives. Solid lines denote contemporaneous connections and dashed lines indicate lagged connections of 1 time point (i.e., scan).

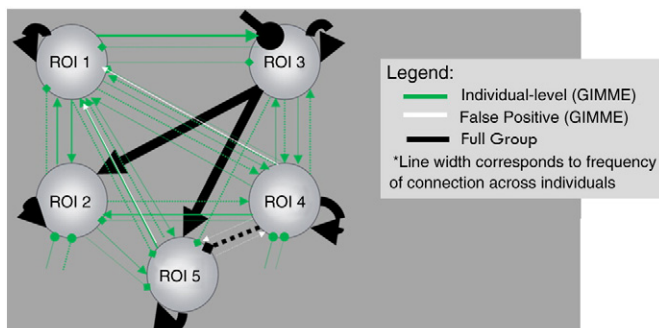
arriving at data-driven models which include the experimental manipulations of event-related designs exist.

Two, GIMME accommodates heterogeneity in samples by arriving at reliable estimates at the individual-level. By first acquiring a group model, GIMME capitalizes on shared information to pick signal out from noise. Data presented herein demonstrate that the standard technique of concatenating for SEM analysis yields spurious results when heterogeneity is present or arrives at models which fail to describe the individuals comprising the group. Having arrived at paths which are common across individuals, individual-level connections may open as needed to obtain excellent model fits. The second step guarantees that connections unique and important to individuals' processing will surface. The degree to which the group-derived map differs from the individual-level maps depends on the degree of true heterogeneity in the sample. In the case of the Smith et al. (2011) data and the data sets with homogenous structures used in *Heterogeneous data: varying connection strength*, the group level model contained all of the paths that were ultimately estimated on the individual level. If there is much heterogeneity, as in the four subgroup and random heterogeneity examples used in *Heterogeneous data: four subgroups* and *Heterogeneous data: random maps*, the

group model may be sparse by comparison to the individual-level models. Since this method obtains reliable results, researchers can be confident that any variation in structure occurring on the individual level is true. Obtaining reliable individual-level connectivity structures and estimates is a prerequisite for analysis into heterogeneity of brain processes across individuals.

The ability to accommodate individual-level differences in model structure while arriving at reliable group inferences thus allows researchers to ask new questions. For instance, one might ask if deviation from the group model relates meaningfully to behavioral, psychological, or genetic indices. Another line of questioning might be, in a sample of individuals with similar behaviors or diagnostic categories (e.g., healthy controls or patients with Alzheimers), how much deviation in brain processing might one expect for specific tasks? Within this line of questioning, researchers interested in the etiology and treatment of illnesses may wish to identify differences in brain connectivity *within* a diagnostic group to arrive at better, person-specific treatment strategies (Seghier et al., 2010). Increasingly, researchers are acknowledging that multiple networks may be necessary to describe even normative functioning, and group-level results for identifying abnormal processing may miss important differences in connectivity patterns that relate meaningfully to a deficit for the individual. GIMME enables investigation into these questions. Finding the answers may lead to an increase in the application to fMRI analysis to real-world problems, such as identifying subgroups of ADHD patients which may respond to specific medication or indications of recovery in the case of traumatic brain injury. It is thus a scientific imperative that statistical methods which yield reliable results are used.

The need for a program which allows for individual nuances, which may be meaningful, is large. Future extensions of algorithms which allow for individual nuances may be used to identify subgroups. One drawback of GIMME's implementation, which it shares with most other approaches, is that it assumes a similar hemodynamic response function across regions. While currently implemented via quasi-maximum likelihood, raw data maximum likelihood implementations will open up the possibility for parameters of the hemodynamic response to be estimated for each region separately. Another improvement would be to allow for correlations between ROIs. Currently, bidirectional relations must be uncovered by having two directed paths between ROIs. Having demonstrated excellent recovery of true connections and their directions, GIMME is well-primed to fill needs of fMRI



**Fig. 7.** Simulated data network and results for the “random heterogeneity” simulated data. GIMME recovered most of the parameters (see Table 1) with few false positives. Concatenating the data recovered all of the paths that are common across all individuals with no false positives. Solid lines denote contemporaneous connections, dashed lines indicate lagged connections of 1 time point (i.e., scan), arrow-heads indicate effects among ROIs, round heads indicate direct influence of experimental stimuli, and square heads indicate bilinear effects of the relationship between ROIs in the presence of experimental stimuli. Multiple types of effects (e.g., bilinear and direct) may occur between two ROIs. Line width corresponds to frequency.



researchers wishing to arrive at valid effective connectivity maps for groups and individuals.

## Acknowledgments

This research was supported in part by the NSF grant nos. 0852147 and 1157220. We thank G. A. James for sharing programming code.

## Appendix A

Direction precision (# of true directions in output/# directions in output); presence precision (# of true connections in output/# connections in output); direction recall (# of true directions in output/# of true directions in simulation); and presence recall (# of true connections in output/# connections in simulation) are presented below for varying sample sizes. The table presents averages from subsamples drawn with replacement in sizes of 10, 25, 50, and 100 (where applicable) and analyzed with the GIMME algorithm as denoted by “GIMME (##)”. “Individual” indicates that the analyses were conducted solely on the individual-level.

		Precision		Recall	
		Direction	Presence	Direction	Presence
Homogeneity:	Individual	43%	85%	46%	91%
Smith et al. (2011) sim. 2	GIMME (10)	83%	89%	94%	99%
	GIMME (25)	88%	91%	94%	97%
	GIMME (50)	85%	91%	93%	100%
Heterogeneity: strength	Individual	65%	78%	68%	85%
	GIMME (10)	92%	95%	94%	97%
	GIMME (25)	85%	91%	91%	97%
Heterogeneity: 4 subgroups	GIMME (50)	99%	99%	100%	100%
	Individual	77%	87%	83%	95%
	GIMME (10)	92%	92%	99%	99%
	GIMME (25)	96%	96%	99%	99%
Heterogeneity: random	GIMME (50)	96%	96%	99%	99%
	Individual	91%	92%	92%	93%
	GIMME (10)	90%	90%	97%	97%
	GIMME (25)	91%	91%	97%	97%
	GIMME (50)	98%	98%	98%	98%

## Appendix B

Results from full Smith et al. (2011) data sets. There reader is referred to Smith et al. (2011) for a full description of the data simulation process or Homogenous data: Smith et al. (2011) simulation 2 above for a brief overview. Simulations numbered 4 and 16 were removed because the former was too computationally taxing (with nodes=50) and the latter simulation did not have information regarding the additional connections public. In the table below, the “simulation” column refers to the simulation number assigned to the datasets in Smith et al. (2011). The “# Nodes” refers to the number of ROIs; “Minutes” the length of the session in minutes; “Noise (%)” the percent of noise added; “HRF Std. Dev.” indicates the degree to which the hemodynamic response function delay parameters varied across simulations; “Other” lists additional conditions modeled in the dataset; “C-sensitivity” reports the percentage of individual replications from the dataset that met Smith et al.’s (2011) criteria for recovering the presence of a connection, defined as a connection whose weight is above the 95%-ile of false connections; and “D-accuracy” indicates the percentage of individual replications for which the direction recovered during the GIMME analysis was greater in value than the weight of the opposite direction (for those true positive connections identified).

Simulation #	Nodes	Minutes	TRs	Noise (%)	HRF std. dev.	Other	C-sensitivity	D-accuracy
1	5	10	3	1	0.5		100	64
2	10	10	3	1	0.5		100	90
3	15	10	3	1	0.5		100	85
5	5	60	3	1	0.5		100	70
6	10	60	3	1	0.5		100	73
7	5	250	3	1	0.5		100	81
8	5	10	3	1	0.5	Shared inputs	100	77
9	5	250	3	1	0.5	Shared inputs	100	80
10	5	10	3	1	0.5	Global mean confound	100	92
11	10	10	3	1	0.5	Bad ROIs	85	80
12	10	10	3	1	0.5	Bad ROIs	100	72
13	5	10	3	1	0.5	Backwards connections	61	70
14	5	10	3	0.1	0.5	Cyclic connections	100	98
15	5	10	3	1	0.5	Stronger connections	100	94
16	5	10	3	1	0.5	More connections	94	93
17	10	10	3	1	0.5		100	77
18	5	10	3	0.1	0		100	82
19	5	10	0.25	0.1	0.5	Neural lag=100ms	35	83
20	5	10	0.25	1	0	Neural lag=100ms	27	88
21	5	10	3	0.1	0.5	Varying connection strengths	96	74
22	5	10	3	1	0.5	Nonstationary connection strengths	89	82
23	5	10	3	1	0.5	Stationary connection strengths	100	96
24	5	10	3	0.1	0.5	Only one strong external input	100	98
25	5	5	3	1	0.5		100	46
26	5	2.5	3	1	0.5		92	51
27	5	2.5	3	0.1	0.5		98	33
28	5	5	3	0.1	0.5		100	74

## References

- Brown, T.A., 2006. *Confirmatory Factor Analysis for Applied Research*. Gilford Press, New York.
- Buxton, R., Wong, E., Frank, L., 1998. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn. Reson. Med.* 39, 855–864.
- Fair, D.A., Posner, J., Nagel, B.J., Bathula, D., Costa Dias, T.G., Mills, K.L., Blythe, M.S., Giwa, A., Schmitt, C.F., Nigg, J.T., 2010. Atypical default network connectivity in youth with attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 68, 1084–1091.
- Friston, K.J., 2007. Dynamic causal modeling. In: Friston, K.J., Ashburner, J.T., Kiebel, S.J., Nichols, T.E., Penny, W.D. (Eds.), *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, Amsterdam, pp. 541–560.
- Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. *Neuroimage* 19 (3), 1273–1302.
- Gates, K.M., Molenaar, P.C.M., Hillary, F., Ram, N., Rovine, M., 2010. Automatic search in fMRI connectivity mapping: an alternative to Granger causality using formal equivalences between SEM path modeling, VAR, and unified SEM. *Neuroimage* 53, 1118–1125.
- Gates, K.M., Molenaar, P.C.M., Hillary, F., Slobounov, S., 2011. Extended unified SEM approach for modeling event-related fMRI data. *Neuroimage* 54, 1151–1158.
- Guttman, M., Mies, C., Dudyycz-Sulicz, K., Diskin, S.J., Baldwin, D.A., Stoekert, C.J., Grant, G.R., 2007. Assessing the significance of conserved genomic aberrations using high resolution genomic microarrays. *PLoS Genet.* 3 (8), e143.
- Hillary, F., Medaglia, J., Gates, K., Molenaar, P., Slobounov, J., Peechatka, A., Good, D., 2011. Examining working memory task acquisition in a disrupted neural network. *Brain J. Neurol.* 134, 1555–1570.

- James, G.A., Kelley, M.E., Craddock, R.C., Hotzheimer, P.E., Dunlop, B.W., Nemeroff, C.B., Mayberg, H.S., Hu, X.P., 2009. Exploratory structural equation modeling of resting-state fMRI: applicability of group models to individual subjects. *Neuroimage* 45, 778–787.
- Jöreskog, K.G., Sörbom, D., 1992. LISREL. Scientific Software International, Inc.
- Kherif, F., Poline, J.B., Mériaux, S., Benali, H., Guillaume, F., Brett, M., 2003. Group analysis in functional neuro-imaging: selecting subjects using global spatial and temporal similarity between subjects. *Neuroimage* 20 (4), 2197–2208.
- Kherif, F., Josse, G., Seghier, M.L., Price, C., 2009. The main sources of intersubject variability in neuronal activation for reading aloud. *J. Cogn. Neurosci.* 21, 654–668.
- Kim, J., Zhu, W., Chang, L., Bentler, P.M., Ernst, T., 2007. Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Hum. Brain Mapp.* 28, 85–93.
- Meek, C., 1997. Graphical Models: Selecting causal and statistical models. PhD thesis, Carnegie Mellon University.
- Miller, M.B., Van Horn, J.D., 2007. Individual variability in brain activations associated with episodic retrieval: a role for large-scale databases. *Int. J. Psychophysiol.* 63, 205–213.
- Miller, M.B., Van Horn, J., Wolford, G.L., Handy, T.C., Valsangkar-Smyth, M., Inati, S., Grafton, S., Gazzaniga, M.S., 2002. Extensive individual differences in brain activations during episodic retrieval are reliable over time. *J. Cogn. Neurosci.* 14, 1200–1214.
- Molenaar, P.C.M., 1985. A dynamic factor model for the analysis of multivariate time series. *Psychometrika* 50, 181–202.
- Molenaar, P.C.M., 2004. A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Measurement* 2 (4), 201–218.
- Molenaar, P.C.M., Campbell, C.G., 2009. The new person-specific paradigm in psychology. *Curr. Dir. Psychol.* 18 (2), 112–117.
- Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Modelling functional integration: a comparison of structural equation and dynamic causal models. *Neuroimage* S264–S274.
- Ramsey, J.D., Hanson, S.J., Hanson, C., Halchenko, Y.O., Poldrack, R.A., Glymour, C., 2010. *Neuroimage* 49, 1545–1558.
- Ramsey, J.D., Hanson, S.J., 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, 838–848.
- Sarty, G.E., 2007. Computational brain activity maps from fMRI time-series images. Cambridge University Press, New York.
- Seghier, M., Lazeyras, F., Pegna, A., Annoni, J.M., Khateb, A., 2008. Group analysis and the subject factor in functional magnetic resonance imaging: analysis of fifty right-handed subjects in a semantic language task. *Hum. Brain Mapp.* 29 (4), 461–477.
- Seghier, M.L., Zeidman, P., Neufeld, N.H., Leff, A.P., Price, C.J., 2010. Identifying abnormal connectivity in patients using dynamic causal modeling of fMRI responses. *Front. Syst. Neurosci.* 4–142.
- Smith, S.M., 2012. The future of fMRI connectivity. *NeuroImage* 62 (2), 1257–1266. <http://dx.doi.org/10.1016/j.neuroimage.2012.01.022>.
- Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., Woolrich, M.W., 2011. Network modeling methods for FMRI. *Neuroimage* 54, 875–891.
- Sörbom, D., 1989. Model modification. *Psychometrika* 54, 371–384.
- Van den Heuvel, M., Mandl, R., Hulshoff Pol, H., 2008. Normalized cut group clustering of resting-state fMRI data. *PLoS One* 3 (4), e2001.