Identifying abnormal connectivity in patients using dynamic causal modeling of fMRI responses

Mohamed L. Seghier1*, Peter Zeidman1, Nicholas H. Neufeld1, Alex P. Leff2 and Cathy J. Price1

1 Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London, UK
2 Institute of Cognitive Neuroscience, Institute of Neurology, University College London, London, UK

*Correspondence:
Mohamed L. Seghier, Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3BG, UK.
E-mail: m.seghier@fli.ion.ucl.ac.uk

INTRODUCTION

Accurate characterization of abnormalities in neural processing is important for understanding pathological conditions and recovery mechanisms. It may also contribute to the tailoring of efficient therapy and intervention procedures. A large fMRI literature has already investigated how pathological conditions change neural processing, usually in terms of activity or signal changes in a set of spatially segregated regions. However, several studies have shown that characterizing such abnormal processes in terms of activation differences in a set of distinct (isolated) brain regions is not enough to provide a comprehensive picture of the abnormal brain (Meyer-Lindenberg et al., 2001; Ween, 2008). This is due to the fact that the function of any brain region cannot be understood in isolation but only in conjunction with the regions it interacts with during active behavior (e.g., McIntosh, 2000; He et al., 2007; Bassett and Bullmore, 2009; Bressler and Menon, 2010; Guye et al., 2010), a principle known as functional integration (Friston, 1994, 2002b, 2007).

Functional integration refers to task-dependent processing that emerges from changes in interactions among brain regions. In this context, deficits in functional integration or connectivity are implied when the influence of one brain region on another is stronger or weaker in patients relative to control subjects (Price et al., 2006; Ween, 2008). There are several approaches that have been proposed to assess functional integration or connectivity (for a review see Ramnani et al., 2004; Harrison et al., 2007; Rogers et al., 2007; Li et al., 2009; Bressler and Menon, 2010), including structural equation modeling (McIntosh and Gonzalez-Lima, 1994), dynamic causal modeling (DCM) (Friston et al., 2003), Granger causality (Roebroeck et al., 2005), psycho-physiological interactions (Friston et al., 1997), dynamic Bayesian networks (Rajapakse and Zhou, 2007), multivariate autoregressive modeling (Harrison et al., 2003), partial correlation analysis (Marrelec et al., 2009), non-linear system identification (Li et al., 2010), and switching linear dynamic systems (Smith et al., 2010). Each method has its advantages and weaknesses (e.g., see Penny et al., 2004; Ramnani et al., 2004; Witt and Meyerand, 2009) and its use should be motivated by the question of interest, level of inference, paradigm design, data acquisition and analysis.

Here, we consider the use of DCM as a flexible and robust tool for assessing effective connectivity (Friston et al., 2003). In contrast to functional connectivity, effective connectivity provides a mechanistic account of the cause of the inter-regional interactions that would explain the emergence of a particular functional pattern (see for more details Friston, 1994, 2002b, 2009b). DCM has been widely used across different imaging modalities, populations and tasks (for a review see Stephan et al., 2007, 2010; Friston, 2009a). For instance, many studies have used DCM of fMRI data to explore inter-regional interactions during different cognitive tasks in normal healthy subjects. This literature has provided interesting insights about how brain regions talk to each other in healthy populations during cognitive and motor processing (Grol et al., 2007; Kasess et al., 2010).
connectivity parameters that present parallels with neuro-physiological models, including psycho-physiological or physio-physiological interactions, (vii) the use of DCM with Bayesian model selection (BMS) and averaging tools enables inferences at different levels, from a set of plausible models (a family of models) to a specific connectivity parameter, and (viii) DCM inferences can be made both at the single subject and group level.

Fundamentally, DCM is only appropriate for describing responses (fMRI, MEG, ERP) that result from controlled external stimuli (Friston, 2009a; Stephan et al., 2010). This requires datasets that also include periods when external stimuli are both present and absent. DCM is therefore not currently suitable for uncontrolled continuous “natural” stimulations such as resting-states. It treats the human brain as a dynamic system that is subject to multiple inputs and produces multiple outputs. Thus, DCM characterizes task-dependent inter-regional interactions.

The starting point for DCM is the selection of a fixed set of regions and their possible connections. Each combination of experimentally modulated connections corresponds to a model, which can then be compared to other alternative models in order to identify which model(s) best predict(s) the data (see illustration in Figure 1 for details of the practical steps). For a given model, DCM estimates three different sets of parameters: (i) input or extrinsic parameters that quantify how brain regions respond to external stimuli (i.e., the external inputs that perturb the model), (ii) endogenous or latent parameters that characterize context-independent (or average) inter-regional interactions, and (iii) modulatory parameters that measure changes in effective connectivity induced by the

WHAT IS DCM?

Dynamic causal modeling aims to explain, quantitatively and mechanistically, how observed fMRI responses are generated (Friston et al., 2003). The key features that make DCM the method of choice for estimating effective connectivity can be summarized as follows (see detailed description in page 3100 of Stephan et al., 2010): (i) DCM is dynamic, in the sense that it uses differential equations to model inter-regional interactions, (ii) DCM is causal as it aims to infer the directionality of the inter-regional interactions and their context-dependent modulations, (iii) DCM is a hypothesis-driven approach that can incorporate any known effect (e.g., stimuli and tasks) to test specific hypotheses that motivated the experimental design, (iv) DCM is Bayesian in all its aspects and uses a generative model to constrain effects by prior knowledge, (v) DCM explicitly uses an empirical forward model that links observed hemodynamic responses to the (hidden) neuronal dynamics, allowing inferences to be made at the neuronal level, (vi) DCM estimates a range of connectivity parameters that present parallels with neuro-physiological models, including psycho-physiological or physio-physiological interactions, (vii) the use of DCM with Bayesian model selection (BMS) and averaging tools enables inferences at different levels, from a set of plausible models (a family of models) to a specific connectivity parameter, and (viii) DCM inferences can be made both at the single subject and group level.

DEFINITIONS AND PRINCIPLES

We need first to define some concepts used throughout this review. Our review of previous work indicates that different terms have been used across studies to refer to the same concept. First, the term “network” has been used to refer to a set of temporally and spatially segregated regions that interact and engage in multiple complex behaviors. A network is by definition dynamic, where nodes and interactions between these nodes change continuously in time and space across different cognitive processes. A network can also be referred to as a “system” or “circuit.” A part of a network can be referred to as a “sub-network” or “subsystem.” Second, the complex dynamics within the network have been described with different terms, including: inter-regional interactions, connectivity, coupling, interaction, interdependency, and co-operative action. Here we refer to these complex dynamics as “inter-regional interactions.” Third, when characterizing how inter-regional interactions differ in patients and controls, previous studies used the terms: altered, abnormal, disturbed, atypical, impaired, or dysfunctional. Throughout the review we will use the term of “abnormal” when referring to inter-regional interactions in patients that are outside the range of those observed in healthy controls.

Using these definitions, we tentatively summarize some of the implicit assumptions that motivate studies of functional integration in patients:

(i) local damage to a part of a network can propagate throughout the whole network (e.g., Alstott et al., 2009; Kim and Horwitz, 2009);

(ii) an abnormal network can be viewed as a “new” network and not simply the normal network minus the damaged parts (e.g., He et al., 2007);

(iii) an abnormal network can comprise sub-networks that correspond to those seen in healthy subjects and novel sub-networks that are not typically seen in healthy subjects;

(iv) abnormal behavior can be an indicator of abnormal inter-regional interactions, however an abnormal network may not necessarily produce abnormal behavior;

(v) an abnormal network is not fixed; it can evolve and change during the course of therapy and recovery; and

(vi) for a given pathology, an abnormal network can vary from patient to patient even within a relatively homogenous population.

...
The bilinear vs. non-linear

The default implementation in DCM codes the rate change in neuronal activity according to the following bilinear evolution or state equation (Friston et al., 2003):

\[
\frac{dz}{dt} = \left(A + \sum_{i=1}^{w} u_i B^{(i)}\right) z + Cu
\]

where \(z\) is the activity of the neuronal population, \(A\) is the first-order (endogenous, latent, or average) inter-regional interactions in the absence of inputs, \(B\) is the second-order interaction between activity and input (the modulatory effect), and \(C\) is the extrinsic effects of inputs "u" on activity. The bilinear term \(B\) is equivalent to the psycho-physiological interactions in the network that reflects how the inter-regional interactions are modulated by a given input/context. This fundamental equation in DCM has been extended experimental conditions. These connectivity parameters are each expressed in Hz within the DCM framework. They are not necessarily constrained by mono-synaptic (i.e., direct) anatomical connections and can be either positive or negative. A positive parameter means that an increase in activity in one region results in an increased rate of change in the activity of another region. Conversely, a negative parameter means that an increase in activity in one region results in a decreased rate of change in the activity of another region. Note that, due to the limited temporal resolution in fMRI, conduction delays in inputs and inter-regional interactions are ignored in DCM of fMRI responses, but not in DCM for EEG or MEG data (Friston et al., 2003). All DCM parameters and their posterior probabilities are assessed with Bayesian inversion by means of the expectation-maximization algorithm (Dempster et al., 1977; Friston, 2002a).

Below, we succinctly present the available implementations of DCM, how to compare alternative models, and the different levels of inference that can be made with DCM.

MULTIPLE IMPLEMENTATIONS

Since the seminal paper reported by Friston et al. (2003), there have been several extensions of the DCM algorithm to increase its flexibility. It is therefore necessary to consider which algorithms are most suitable for the question of interest. This section presents the extensions implemented for DCM of fMRI responses. By default, the most widely used DCM version corresponds to the bilinear and deterministic implementation with one-state equation per region (see below).

**Bilinear vs. non-linear**

The default implementation in DCM codes the rate change in neuronal activity according to the following bilinear evolution or state equation (Friston et al., 2003):

\[
\frac{dz}{dt} = \left(A + \sum_{i=1}^{w} u_i B^{(i)}\right) z + Cu
\]
for the assessment of physio-physiological interactions in a given network. The extension, known as non-linear DCM (Stephan et al., 2008), is suitable for explicitly testing whether the activity in a given region gates or enables interactions between other regions. The state equation in non-linear DCM is given by:

\[
\frac{dz}{dt} = \left( A + \sum_{i=1}^{m} u_{i} B^{(i)} + \sum_{j=1}^{n} \eta_{j} D^{(j)} \right) z + C + \eta
\]  

\[(2)\]

where the new term “D” is a quadratic quantity that represents the physio-physiological interactions in the model or the strength of the activity-dependent gating of connections.

This non-linear DCM equation can be helpful when modeling “neural gain control” mechanisms in short-term synaptic plasticity (Stephan et al., 2008) which are likely to be relevant for characterizing a given pathology. Such mechanisms are critical for various cognitive processes, including top-down modulation of attention and learning (see example in den Ouden et al., 2010).

One-state vs. two-state neuronal equations

Because each region is modeled by one neuronal state equation only (Eq. 1), it is not possible to assess selective changes in excitatory (e.g., glutamatergic) and inhibitory (e.g., GABAergic) subpopulations in each region of the DCM model. This is a very important issue that needs to be taken into account when, for instance, making inferences about potential inhibitory or excitatory mechanisms in DCM. To address this, an extended version of DCM exists that uses two-state neuronal equations (Marreiros et al., 2008) to model excitatory and inhibitory subpopulations in each region. This allows for an explicit description of intrinsic (between subpopulations) connectivity within a region. In other words, each region is modeled by two neuronal subpopulations and DCM can thus estimate the interactions between these two subpopulations. Interestingly, this parameterization confers dynamical stability on the system, enforces positivity constraints on the extrinsic connections, and enables context-dependent changes in the interactions to be modeled as a proportional increase or decrease in connection strength (Marreiros et al., 2008). The two-state DCM can be used, for instance, to specifically test whether patients and controls differ in the interactions between excitatory and inhibitory subpopulations.

Anatomical-based priors vs. default shrinkage priors

Different priors are embodied within the Bayesian framework of DCM. These include priors on the connectivity parameters (see for more details Friston et al., 2003) that are referred to as “shrinkage” priors because they tend to “shrink” posterior means to their prior expectation of zero, particularly when the data are noisy. These shrinkage priors make the detection of significant inter-regional interactions somewhat conservative. Thus only effects that have a significant likelihood with high precision are allowed to deviate from zero (i.e., when new data strongly support an existing effect). However, it has been shown that such priors can be modified (relaxed) to take into account an a priori knowledge about anatomical connectivity (Stephan et al., 2009b). The idea is to assign a stronger belief (i.e., by relaxing the shrinkage priors) to a particular interaction between two regions if one has strong evidence for an existing anatomical connection between the regions. These anatomically-based priors have been shown to provide stronger evidence for anatomically motivated models (Stephan et al., 2009b) and are particularly useful if tractography data (DTI) are available. They can also take advantage of the white matter tracts that are specified in recent atlases (e.g., Mori et al., 2005, 2008; Catani and Thiebaut de Schotten, 2008) and incorporate this information into the DCM when defining the appropriate model structure. In the context of the current paper, it is interesting to note that information about damaged tracts in a group of patients can be explicitly modeled. However, because of the deterministic nature of the current DCM implementation (see below), indirect influences on regions cannot be ruled out even if direct white matter tracts are missing.

Deterministic vs. stochastic

The default implementation of DCM in SPM is deterministic as only the explicitly modeled effects are allowed to influence the inter-regional interactions. Thus, DCM estimates the interactions between the spatially segregated regions that are temporally perturbed by the external inputs included in the model. However, it is obvious that this framework cannot rule out the influence of indirect effects, including interactions with regions not included in the DCM or temporal innovations not modeled in the inputs (see discussion in Smith et al., 2010). For instance, a connection can excite a group of neurons that inhibit another region thereby resulting in an overall effect of inhibition. An extended version, known as stochastic DCM (Daunizeau et al., 2009), allows such subtle indirect effects to be expressed. It extends the previous state equation to:

\[
\frac{dz}{dt} = \left( A + \sum_{i=1}^{m} u_{i} B^{(i)} + \sum_{j=1}^{n} \eta_{j} D^{(j)} \right) z + C + \eta
\]  

\[(3)\]

where \(\eta\) is the state noise that models stochastic innovations in the system. The new term \(\eta\) is defined by Gaussian variables with a mean of zero and a covariance structure that can express any form of innovations. In other words, in this framework, all indirect effects are modeled as stochastic phenomena, thereby allowing the hidden-states causing the data and any non-controlled exogenous inputs to the system to be inferred (Daunizeau et al., 2009). This framework can be very useful because it provides more flexibility in modeling unknown or indirect pathologic effects in patients.

LEVELS OF INFERENCE WITH BAYESIAN MODEL SELECTION

One exciting tool in DCM is its BMS procedure. This procedure compares the evidence for different competing hypotheses. Because the exact mechanisms behind any fMRI responses are unknown, it might not be possible to have an a priori prediction about the “exact” model. Therefore, it is important to specify a range of alternative models and search for the best (most useful) model in the model space, e.g., Leff et al. (2008), Seghier and Price (2010), and Seghier et al. (submitted). This procedure increases the certainty of the best model by testing many other potential explanations of the data.

During the Bayesian inversion of the model, the probability of the data given the model, known as the model evidence, is approximated by the negative variational Free-energy (Friston et al., 2007;
Stephan et al., 2009a). This approximation, as a lower bound on the model evidence, points to the optimal compromise between the accuracy and complexity of a given model. It provides a better estimation for the complexity term, by taking into account the interdependency between the estimated parameters. Thus, model evidence is not an “absolute” measure of how good a model is; instead it is used in BMS to compare between alternative models. Moreover, by using the negative variational Free-energy as the optimal compromise between accuracy and complexity, the current implementation of DCM ensures that (i) model complexity will not increase if additional parameters are “redundant” to existing parameters and (ii) the parameter estimates of a good model are as precise and uncorrelated as possible. In short, BMS estimates the likelihood that a given model has generated the observed data.

**Sampling the model space**

The most critical step when using the BMS tool is the definition of the DCM model space (i.e., the set of specified models) because any inference at the model level depends on the way the model space has been defined (sampled). There are three principles that need to be considered: compatibility, size, and plausibility. First, compatibility between models is attained by ensuring that all models of a given subject have the same regions, that is, the models all contain the same data. Specifically, BMS cannot be applied to models that are fitted to different fMRI data. For group studies, a good practice is to ensure consistent regions between subjects since group BMS implicitly assumes functional compatibility of the modeled effects across subjects. Second, there are an infinite number of possibilities that can explain the data; it is impossible to sample all these possibilities and thus a practical limitation on the number of models (i.e., the size of the model space) is necessary. Third, plausibility reflects a systematic way of defining realistic and interpretable models according to a priori knowledge or some predefined criteria. This will automatically limit the size of the model space and crucially avoid including bad or unrealistic models.

**Random-effects vs. fixed-effects**

Once the model space has been defined and estimated in all subjects, BMS can be used to identify the best (most useful) models at the group level. There are two classic ways of achieving group BMS analyses: fixed-effects and random-effects (see Figure 1 in Stephan et al., 2010). The former assumes that the best models are comparable across subjects because subjects would reasonably perform the task in a similar way (e.g., identical cognitive strategies). The latter may be preferable when studying heterogeneous populations or using complex cognitive tasks where optimal models are likely to vary across subjects. Random-effects BMS is also preferable when outlier effects are suspected because the current implementation (in SPM8) of random-effects BMS uses a robust hierarchical Bayesian approach. It quantifies the likelihood that a specific model generated the data of a subject chosen at random, measured via two quantities: (i) the Dirichlet parameter estimates (alpha) represents a measure of the effective number of subjects in which a given model generated the observed data, and (ii) the “exceedance” probability (xp) describes the belief that a particular model is more likely than any other model given the group data (Stephan et al., 2009a). The exceedance probability “xp” is particularly intuitive as all exceedance probabilities sum to one over all tested models. Note also that there are two sampling schemes in random-effects BMS, the first uses a Variational Bayes method that is optimal and fast when the number of tested models is smaller than the number of subjects (Stephan et al., 2009a), and the second uses a Gibbs sampling method that is optimal and accurate when the number of models is larger than the number of subjects (Penny et al., 2010).

**Inference at the family and the model level**

When comparing models with BMS, the best model would be the winning model that has the most evidence (e.g., significant exceedance probability) compared to other models. However, it is commonly found that one single model does not dominate all other models, particularly when the number of models is very large with many connections shared between models. In this situation, one can look for any similarity in model structures that can better explain the data. To do this, a recent extension of BMS has introduced the ability to make inferences on a “family” of models (Penny et al., 2010). A family is a subset of models that share similar characteristics such as the same driving region or the same modulated connections (see Penny et al., 2010; Seghier and Price, 2010; Seghier et al., submitted). The formation of families should be motivated by the question of interest to ensure all models in the model space are partitioned into different families with no overlap. It is permissible to have different numbers of models in each family. BMS can then be used to compare these competing families and inferences can be made at the family level.

**Inference at the connection level**

The last level of inference assesses the significance of a particular connectivity parameter. Analysis of the connectivity parameters is conducted after comparing models or families because the posterior densities of such parameters are conditional on the particular model or family selected (Stephan et al., 2010). If a winning model has been identified, one can look at the consistency of effects across subjects using random-effects analysis (e.g., t-tests). Alternatively, fixed-effect approaches can be applied using Bayesian parameter averaging where the connectivity parameters are weighted by their precision during the computation of the mean across subjects or sessions. Significant effects at the group level are commonly reported at a corrected p-value (e.g., using a Bonferroni correction based on the number of tested connections, Sonty et al., 2007; Leff et al., 2008; Seghier and Price, 2010). If the inference has been made at the family level, it is possible to use the new Bayesian model averaging (BMA) method where the contribution of each model to the mean effect is weighted by its evidence (Penny et al., 2010).

**Previous DCM Studies of Patients**

In this section we review previous DCM studies of patients. We are particularly interested in the way DCM has been carried out in patients and how the differences between controls and patients have been statistically characterized both at the model and connection level. A PubMed search was conducted with the following inclusion criteria: (i) fMRI studies, (ii) published in English that (ii) used DCM in (iv) patients with any disorder. A total of 28 studies were identified and are listed in Tables 1 and 2 (Bird et al., 2006; Mechelli et al., 2007; Rocca et al., 2007a,b; Sonty et al., 2007).
Mintzopoulos et al., 2009; Shannon et al., 2009; Vaudano et al., 2009; Agosta et al., 2010; Allen et al., 2010; Goulden et al., 2010; Miyake et al., 2010; Rowe et al., 2010. These studies have provided valuable insights into the abnormal connectivity in patients with

### Table 1 | List of previous DCM studies of patients that used one single model (i.e., DCM model space = 1 model). Studies are listed in alphabetical order.

| Study               | Disorder                      | Np/Nc | Task                                                | R  | M | Findings                                                                 |
|---------------------|-------------------------------|-------|-----------------------------------------------------|----|---|--------------------------------------------------------------------------|
| Abutalebi et al. (2009) | Bilingual Aphasia            | 1 pat., 0 con. | Picture naming (in L1 and L2); Longitudinal recovery study (3 sessions). | 5  | 1 | Increased connectivity after therapy between regions associated with “language control.” |
| Agosta et al. (2010)   | Alzheimer’s disease           | 25 pat., 11 con. (with two groups of patients) | A simple motor task with the right hand. | 6  | 1 | Altered endogenous connectivity between patients and controls on the primary sensorimotor cortex. |
| Almeida et al. (2009b) | Major and bipolar depression | 31 pat., 16 con. | Emotion labeling in happy vs. sad faces. | 4  | 1 | Abnormal connectivity between orbitofrontal and amygdala differentiated major from bipolar depressed patients. |
| Bird et al. (2006)     | Autism spectrum disorder     | 16 pat., 16 con. | Attention modulation in faces and houses. | 3  | 1 | Reduced attentional modulation in patients compared to controls. |
| Cao et al. (2008)      | Dyslexia                     | 12 pat., 12 con. (children) | Rhyme judgment task. | 4  | 1 | Reduced connectivity modulation in dyslexics compared to controls between fusiform and parietal regions. Correlations between reading skills and connectivity in left parietal. |
| Crossley et al. (2009) | Schizophrenia                | 26 pat., 13 con. (with two groups of patients) | Working memory task. | 5  | 1 | Connectivity between superior temporal and middle frontal gyri was negative in controls and positive in patients. |
| Eickhoff et al. (2008) | heterotopic hand replantation | 2 pat., 14 con. | Motor (hand movement). | 8  | 1 | Abnormal inhibition from ipsilateral to contralateral M1. |
| Goulden et al. (2010)  | Major depression             | 30 pat., 29 con. | Emotional face processing task. | 4  | 1 | Improved group differences on connectivity parameters when using permutation testing. |
| Grefkes et al. (2008)  | Stroke patients (subcortical lesions) | 12 pat., 12 con. | Motor (hand movement). | 8  | 1 | Reduced coupling between bilateral M1 during stroke-affected hand movements. Correlation between reduced connectivity and degree of impairment. |
| Mechelli et al. (2007) | Schizophrenia                | 21 pat., 10 con. (with two groups of patients) | Voice detection from spoken words task. | 5  | 1 | Abnormal connectivity between anterior cingulate and superior temporal gyri, in particular in patients with verbal hallucinations. |
| Mintzopoulos et al. (2009) | Stroke patients              | 5 pat., 12 con. | Motor (squeezing a robotic device). | 3  | 1 | Reduced endogenous connectivity between M1 and cerebellum and increased connectivity between SMA and M1 in patients relative to controls. |
| Miyake et al. (2010)   | Patients with eating disorders | 36 pat., 12 con. (with 3 groups of patients) | Detection of negative vs. neutral words. | 2  | 1 | Significant group differences in the endogenous connectivity from medial frontal to the amygdala. |
| Rocca et al. (2007a)   | Patients with multiple sclerosis | 12 pat., 14 con. | A simple motor task with the right hand. | 4  | 1 | Stronger endogenous connectivity in patients than controls between right primary sensorimotor cortex and cerebellum. |
| Shannon et al. (2009)  | Externalizing behavior disorder | 21 pat., 11 con. | A reward and non-reward task. | 2  | 1 | Significant differences between controls and patients on both endogenous and non-reward modulatory effects, mainly on the caudate. |

Np, number of patients; Nc, number of controls; R, number of regions; M, number of models.
Table 2 | List of previous DCM studies of patients that used and compared more than one model (DCM model space ≥ 2 models).

| Study                  | Disorder                             | Np/Nc            | Task                                      | R  | M   | Findings                                                                                                                                                                                                 |
|------------------------|--------------------------------------|------------------|-------------------------------------------|----|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Allen et al. (2010)    | Schizophrenia                         | 15 pat., 15 con. | Sentence completion task.                 | 3  | 14  | Best model identical in controls and patients. Increased endogenous connectivity between anterior cingulate and middle temporal regions in patients compared to controls.                                             |
| Almeida et al. (2009a) | Bipolar disorder                      | 21 pat., 25 con. | Emotion labeling in happy vs. neutral faces. | 3  | 4   | Best model identical in controls and patients. Abnormal increase in endogenous connectivity between parahippocampal and cingulate cortex in patients compared to controls.                                           |
| Benetti et al. (2009)  | Schizophrenia (with two groups of patients) | 26 pat., 14 con. | Working memory task.                      | 2  | 4   | Models have different regions. Endogenous connectivity between right hippocampus and inferior frontal gyrus was stronger in controls than patients.                                                           |
| Dima et al. (2009)     | Schizophrenia                         | 13 pat., 16 con. | Perception of the hollow-mask illusion.   | 5  | 2   | Best model identical in controls and patients (but not when using RFX analysis). No significant correlations between connectivity and symptoms.                                                                   |
| Fujii et al. (2009)    | Blind patients                        | 15 pat., 24 con. | Tactile Braille discrimination task.       | 5  | 2   | Model comparison done in patients, and the best model was then used in controls. Modulations between parietal and occipital regions were positive in patients and negative in controls. Early blind showed stronger connectivity than the late blind patients.                                          |
| Grefkes et al. (2010)  | Stroke patients (subcortical lesions) | 11 pat., 0 con. | Motor (hand movement); interference with TMS. | 6  | 4   | TMS applied to the contralesional motor cortex. TMS enhanced endogenous connectivity between ipsilesional SMA and M1. The connectivity parameters of both models were compared between patients and controls. Stronger connectivity in controls than patients during fearful compared to neutral context, in particular on the amygdala. |
| Grezes et al. (2009)   | Autism spectrum disorder              | 12 pat., 12 con. | Perception of fearful or neutral actions (videos). | 6  | 2   | The best model of driving inputs in controls was then used in patients. Endogenous and modulatory effects were different in patients vs. controls, and they were correlated to the severity of the structural damage. |
| Hamandi et al. (2008)  | Epileptic patients                    | 1 pat., 0 con.   | Intercital epileptiform discharges.        | 2  | 2   | Onsets defined as spikes (visual monitoring in EEG); Increased connectivity from left parahippocampal to lingual gyrus during epileptic discharges. The best model was used in controls.                 |
| Rocca et al. (2007b)   | Patients with benign multiple sclerosis | 15 pat., 19 con. | Stroop color-word task.                   | 6  | 2   | The best model was used in controls.                                                                  |
| Rowe et al. (2010)     | Parkinson’s disease                   | 16 pat., 17 con. | Action selection in finger-tapping task; Dopaminergic therapy; Stroop color-word task.      | 4  | 48  | Best model identical in controls and patients. Model selection is reproducible. Connectivity parameters are less reliable across sessions.                                                                          |
| Schlosser et al. (2008) | Major depression                     | 16 pat., 16 con. | Stroop color-word task.                   | 5  | 4   | Best model identical in controls and patients. Higher endogenous connectivity between anterior cingulate regions in patients compared to controls.                                                                    |
| Schlosser et al. (2010) | Obsessive Compulsive disorder         | 21 pat., 21 con. | Stroop color-word task.                   | 6  | 5   | Best model identical in controls and patients. Increased modulation between frontal and cingulate cortex in patients during incongruent trials.                                                          |
| Sonty et al. (2007)    | Primary Progressive Aphasia           | 8 pat., 8 con.   | Semantic word matching.                   | 6  | 5   | Best model identical in controls and patients. Reduced connectivity between Broca and Wernicke’s area in patients compared to controls. Reduced connectivity was correlated with accuracy. Discharges used as driving inputs and enter the system at different regions. The best model showed spike wave discharges input on precuneus. |
| Vaudano et al. (2009)  | Epileptic patients                    | 7 pat., 0 con.   | Generalized spike wave discharges.         | 3  | 3   |                                                                                                                                             |

\(Np, \text{ number of patients}; Nc, \text{ number of controls}; R, \text{ number of regions}; M, \text{ number of models.}\)
a range of pathologies. However, the aim of our review is not to discuss the relevance of their findings. This would necessitate a separate review that considered the models tested, tasks and stimuli used, region selection, driving and modulatory inputs, and the relevance of the findings in light of previous neuropsychological and computational models. Instead, the current review considers the methodological approaches that have been used (see Tables 1 and 2) and we refer to these studies for illustrations of the methods.

Consistent with the most widely used implementation of DCM, all 28 studies were deterministic and used the bilinear one-state neuronal equation without anatomical-based priors. Matched groups of controls were included in all but four studies (Hamandi et al., 2008; Abutalebi et al., 2009; Vaudano et al., 2009; Grefkes et al., 2010). The driving inputs were specified in standard block or event-related designs in all but two studies (Hamandi et al., 2008; Vaudano et al., 2009) where internally generated epileptogenic wave discharges from the epileptic patients being studied were used as the driving inputs.

The models used varied in their complexity, from simple models with two regions (e.g., Benetti et al., 2009) to complex models with eight regions (e.g., Grefkes et al., 2008). All studies involved adult subjects except one study of dyslexic children (Cao et al., 2008). The tasks and stimuli varied extensively between the studies because they were purposely designed to maximize differences between patients and controls. For instance, working memory tasks were used in patients with schizophrenia (Benetti et al., 2009; Crossley et al., 2009), Stroop word-color tasks were used in patients with depression (Schlosser et al., 2008), and semantic tasks were used in aphasic patients (Sonty et al., 2007; Abutalebi et al., 2009). One critical difference between these studies concerns the level of inference made in both patients and controls. We thus divided the 28 studies into two sets: a set of 14 studies that tested one model only (Table 1) and therefore did not require the BMS procedure and a second set of 14 studies that used BMS to compare at least 2 models in patients and controls (Table 2).

STUDIES WITH ONE FIXED DCM MODEL

The 14 studies with only one DCM model in both patients and controls were exclusively interested in how connectivity parameters differed between groups. The same model (i.e., same regions, driving inputs, and modulatory inputs) was defined in patients and controls and the connectivity parameters of that model were compared between patients and controls on a connection basis. The main motivation of such strong reliance on a single model is to ensure the same model is tested on both populations so that connectivity parameters can be compared. This avoids the case when the winning model is not the same in patients and controls. Because the parameters are conditional on the selected model the definition of one unique model can be problematic. Some studies have used a fully connected model because it allows all possible combinations of inter-regional connections to be estimated and tested (e.g., Bird et al., 2006; Mechelli et al., 2007; Rocca et al., 2007b; Cao et al., 2008; Agosta et al., 2010; Miyake et al., 2010). However, we argue that a fully connected model is by definition the most complex and may lead to over-fitting of the data. Moreover, it cannot guarantee that connectivity parameters are estimated within the best model structure given the data. Other studies have used a model that is less complex than the fully connected version when a priori hypotheses concerning the architecture of the optimal model are possible on the basis of previous structural and functional connectivity findings (e.g., Grefkes et al., 2008; Abutalebi et al., 2009; Crossley et al., 2009; Goulden et al., 2010). Irrespective of which model is selected, single model studies were more concerned about the differences in connectivity parameters and thus their findings do not address differences between patients and controls at the network level.

STUDIES WITH MULTIPLE COMPETING DCM MODELS

To avoid heavy reliance on a predefined model other studies have used BMS to identify the best model from competing alternatives (varying from 2 to 48 models, see Table 2). The comparison between connectivity parameters is then performed on the parameters of the best model. This procedure has the advantage of providing the opportunity to make inferences both at the model/system level and the parameter/connection level within the optimal structure. It is perfectly valid (and sufficient) to limit the inference to the system level, for instance by showing whether or not patients are using the same network as the controls (e.g., is the winning model identical in patients and controls?). That said, all previous studies were mainly interested in inferences at the connection level and the BMS was usually presented as an intermediate analysis step. To ensure the possibility of such inference, these studies used different methods to guarantee that the final selected model is the same between controls and patients. This conceptual limitation will be unnecessary when the new BMA tool for making comparisons across multiple models or families is implemented in SPM (see Penny et al., 2010).

Practically, three methodologies have been used to compare the best model in patients and controls. First, the majority of studies used BMS on the same set of alternative models and implemented this independently in patients and controls (e.g., Sonty et al., 2007; Schlosser et al., 2008, 2010; Almeida et al., 2009a; Allen et al., 2010; Rove et al., 2010). All these studies showed the winning model to be identical in patients and controls. Note however that this similarity may depend on whether RFX or FFX methods are used; for instance, patient and control models were found to be identical with FFX analysis but different with RFX analysis (Dima et al., 2009). Second, an alternative approach used by Rocca et al. (2007b) in a study of patients with multiple sclerosis performing a Stroop task involved the use of BMS within controls only to find the best model in controls, followed by a comparison of the parameters of that model in patients and controls (Rocca et al., 2007b). For instance, when faced with the choice of where inputs enter the system, Rocca and colleagues defined two alternative models with two different driving regions. They then identified the best model and used the driving region of the winning model as a driving region for patients as well (Rocca et al., 2007b). The procedure was used by Fujii et al. (2009) to study blind patients during tactile Braille discrimination tasks. It involved identification of the best model in patients and then used this model in comparison to controls (Fujii et al., 2009). Note that the last two procedures rely on the hypothesis that the best model of one group is identical to the best model in the second group.
DIFFERENCES IN CONNECTIVITY PARAMETERS

After defining the best model in both patients and controls, previous studies have tested whether the parameters of that model differ between the two groups. Some studies have limited this comparison to the endogenous connectivity (e.g., Rocca et al., 2007a; Benetti et al., 2009; Miyake et al., 2010), whereas other studies have looked at how patient connectivity differed in the context of stimulus or task changes (e.g., Bird et al., 2006). Abnormal connectivity in patients has been found to correspond to either a change in the strength of connectivity (stronger or weaker parameters than controls) or a change in polarity (positive vs. negative). For instance, Schlösser et al. (2008) found patients with major depression have stronger connectivity between anterior cingulate regions when compared to controls during a Stroop color-word task (Schlösser et al., 2008). Conversely, Sonty et al. (2007) illustrated reduced connectivity by showing that, relative to controls, patients with primary progressive aphasia have weaker connectivity between Broca and Wernicke’s areas during semantic word matching tasks. An example of reversed polarity has also been shown with negative modulatory effects between parietal and occipital regions in controls compared to positive modulations in blind patients (Fujii et al., 2009).

Interestingly, a few studies have also investigated the correlations between abnormal connectivity and other behavioral or diagnostic measures in patients. For instance, Rocca et al. (2007b) showed a significant correlation between abnormal connectivity and the severity of structural damage in patients with multiple sclerosis. Cao et al. (2008) found that reduced modulations between the fusiform and parietal regions correlated with reading skills in dyslexic children (Cao et al., 2008). Another example is provided by Greffkes et al. (2008) who found a significant correlation between reduced connectivity in primary motor regions and the degree of motor impairment in stroke patients with subcortical lesions (Greffkes et al., 2008). Moreover, other studies have used connectivity parameters to distinguish between different pathological groups. For instance, Almeida et al. (2009b) found that the abnormal connectivity between the orbitofrontal cortex and the amygdala differentiated patients with major depression from patients with bipolar depression during a task of emotional labeling of happy vs. sad faces (Almeida et al., 2009b). Miyake et al. (2010) found that endogenous connectivity between the medial frontal gyrus and the amygdala during the detection of negative vs. neutral words showed variable patterns between three subgroups with different types of eating disorders (Miyake et al., 2010). Furthermore, two other studies illustrated how external interventions can be monitored with DCM in diseased populations. The first was a longitudinal study of a bilingual aphasic patient by Abutalebi et al. (2009) and assessed connectivity between five regions of interest during a picture naming task in two languages. Increased connectivity was found after therapy in regions associated with “language control,” consistent with the patients’ behavioral recovery (Abutalebi et al., 2009). The second study combined DCM and transcranial magnetic stimulation (TMS) (Greffkes et al., 2010) and focused on stroke patients with subcortical lesions during hand movement tasks. In their study, Greffkes et al. (2010) found that applying TMS to the contrlesional motor cortex increased connectivity between ipsilesional motor regions (Greffkes et al., 2010).

A SYSTEMATIC PROCEDURE FOR DCM IN PATIENTS

The studies reviewed above have used a variety of procedures to assess effective connectivity in patients. This makes any meaningful comparison between their findings difficult. This diversity also illustrates the different procedures that have been used to address the methodological challenges that arise when DCM is used with patient data. In this section, we propose a practical procedure that can be used to compare effective connectivity in patients and controls. This involves the standard DCM procedures (see Definitions and Principles above) but with some specific modifications that are particularly relevant for patient studies. They relate to region selection, the definition of the driving regions, the definition of plausible models, identifying the best model or family of models, and the comparison of connectivity parameters in patients and controls. Note however we are not claiming this approach is the only correct way of using DCM in patients; the optimal approach should always be tailored to the specific questions, the selected effects of interest, and the definition of the model space.

REGION SELECTION

Models should, by definition, be comparable between patients and controls. This means models must have identical nodes. Accordingly, only commonly activated regions in patients and controls can be included in DCM. Thus, DCM in patients focuses on the characterization of abnormal connectivity in a common network of regions (see schematic illustration in Figure 2). A fuller characterization of patient data would therefore benefit from a report of abnormalities at the level of regions (areas that are more or less activated in patients than controls) in addition to a report of abnormal connectivity within the set of areas commonly activated in patients and controls. Second, there should be minimal intersubject variability in the anatomical location of a given region. This is because large variability in region locations may result in the comparison of functionally different regions across patients and controls (for a similar rationale, see Seghier and Price, 2010; Seghier et al., submitted). This is particularly critical in regions where functional specialization may vary at a high spatial scale. As a general rule, the distance between corresponding regions in different individuals should correspond to the size of the spatial smoothing kernel. Third, the distance between different regions in the same individual needs to ensure that the data included in one region is not also entered into another region. Fourth, we also advocate the definition of proper F-contrasts of interest that can be used to adjust the extracted time-series and thus minimize the contribution of other confounds (e.g., session effect, head motion, incorrect trials).

DRIVING REGIONS

After extracting the ROIs, an important step in the construction of plausible models is the choice of driving regions (i.e., where the inputs enter the system). It is wise to specify primary sensory regions as driving regions if they are included in a model. However, for other regions, the motivation for selecting driving regions should be carefully based on prior anatomical or functional knowledge as all connectivity parameters depend on how the information flow is assumed to propagate from driving regions.
First, it should be stressed that it is better to use the most recently available negative variational Free-energy model evidence measure instead of the suboptimal AIC or BIC measures (for more details see Stephan et al., 2009a). The problem with AIC and BIC (still used in some recent work) is that they may bias model selection toward more simplistic models and do not take into account the interdependency between connectivity parameters. Using the negative variational Free-energy measures, BMS can be run separately on patients and controls to identify the best family of models in patients and controls (assuming the same model space is defined in both groups). Family inference enables the investigator to test and report the evidence for an optimal class of models in patients or controls. If patients and controls show a different winning family or model then this is an interesting result to report, particularly in the context of a commonly activated set of regions (see previous section). This would mean that data from patients and controls cannot be adequately and sufficiently fitted by the same model structure (e.g., Horwitz et al., 1995), suggesting that the patterns of fMRI activation observed in patients have emerged from other alternative models that can be tested and identified with the BMS. Moreover, differences can be investigated at the level of inter-subject variability within and between patients and controls. For example, some patients may have similar winning models to controls and other patients may have different winning models. Inter-subject variability within groups can then be related to behavioral measures acquired in or out of the scanner. This is particularly interesting when the aim of the study is to predict performance at the individual subject level.

DIFFERENCES AT THE CONNECTION LEVEL

If inference is sought at the connection level it is important to only compare models that are tested within the same model space in both groups, including the same winning model or family of models. For the same winning model, two sample t-tests or ANOVAs can

**FIGURE 2** | Illustration of the problem of the “missing nodes” in DCM when comparing patients to controls. Regions A and B are activated in both groups, region D is damaged in patients but present in controls, and region C is a compensatory region that is only activated in patients. The deterministic DCM can assess the interactions between A and B [noted int(AB) on a solid black line], but would ignore the indirect effects of regions D and C (shown with gray lines). The interactions between A and B are thus a complex mixture of these effects (e.g., in patients = int(AB) in the context of C without D; in controls = int(AB) in the context of D without C).

**PLAUSIBLE MODELS**

This step should be systematic in the sense that it needs to explore a set of alternative plausible models for a given set of regions, particularly when prior knowledge of a network is uncertain. Even if the inference of interest is at the connection level, it is difficult to interpret and generalize differences in connectivity parameters in patients that have only been compared to controls in the context of one fixed model. Our concern is based on the fact that the connectivity parameters are conditional on the predefined model. Thus, evidence is stronger when it is known to be based on several competing models. It is also helpful (if possible) to classify models in a set of meaningful families that can then be compared with BMS. For instance, family categorization could be based on the existence of a commonality at the level of an endogenous connection, modulated connection, or driving input. Additional constraints help to limit the size of the DCM model space. This is necessary, since, for example, there are over one billion possible models in the model space of a fully connected DCM with six regions if it is searched in an unconstrained manner for the best modulated connections. Limiting the model space to the most plausible models limits the possibility of diluting evidence with the inclusion of bad or implausible models. It would also be helpful if authors detailed how they constrained the model space as this may aid in the design of future studies.

**BMS FOR THE BEST MODEL OR FAMILY OF MODELS**

First, it should be stressed that it is better to use the most recently available negative variational Free-energy model evidence measure instead of the suboptimal AIC or BIC measures (for more details see Stephan et al., 2009a). The problem with AIC and BIC (still used in some recent work) is that they may bias model selection toward more simplistic models and do not take into account the interdependency between connectivity parameters. Using the negative variational Free-energy measures, BMS can be run separately on patients and controls to identify the best family of models in patients and controls (assuming the same model space is defined in both groups). Family inference enables the investigator to test and report the evidence for an optimal class of models in patients or controls. If patients and controls show a different winning family or model then this is an interesting result to report, particularly in the context of a commonly activated set of regions (see previous section). This would mean that data from patients and controls cannot be adequately and sufficiently fitted by the same model structure (e.g., Horwitz et al., 1995), suggesting that the patterns of fMRI activation observed in patients have emerged from other alternative models that can be tested and identified with the BMS. Moreover, differences can be investigated at the level of inter-subject variability within and between patients and controls. For example, some patients may have similar winning models to controls and other patients may have different winning models. Inter-subject variability within groups can then be related to behavioral measures acquired in or out of the scanner. This is particularly interesting when the aim of the study is to predict performance at the individual subject level.
be used to compare connectivity parameters between patients and controls. Alternative permutation testing (Goulden et al., 2010) can be used in studies with small samples. For the same family of models, the latest BMA procedure is more suitable (see illustration in Penny et al., 2010; Seghier et al., submitted) when the two groups differ in the distribution of model evidence across the model space (e.g., the posterior probabilities of plausible models are different in the two groups). Moreover, it is sometimes interesting to look for correlations between connectivity parameters in patients and their phenotype or genotype. This helps to interpret differences between patients and controls and to determine whether any connectivity pattern can serve as a biomarker for a particular deficit or signature of a particular reorganization mechanism.

CONCEPTUAL AND METHODOLOGICAL ISSUES

Other conceptual and methodological issues should be acknowledged when using DCM in patients (see for instance Daunizeau et al., 2010). We focus here on five issues that we believe are crucial for interpreting or generalizing DCM findings. All five warrant further systematic investigations, on both simulated and real data, in order to characterize their influences.

1. Effective connectivity, like functional responses, varies across subjects. Although group effects are easy to report and important in showing the most consistent effects in a given population, variable connectivity patterns across patients can indicate different ways a given deficit can be expressed and the strategies that patients may be using to compensate for those deficits. It is also possible that variability may change with the level of inference, as shown recently in a group of patients with Parkinson’s disease where inferences at the model level were more reliable and reproducible than inferences at the connection level (e.g., Rowe et al., 2010).

2. The network of areas included in a DCM are typically only part of the complete and complex neural system that supports the task. Therefore, the inter-regional interactions within the selected sub-network are only a parsimonious model of the “true” system.

3. The problem of missing nodes is crucial for DCM (see Smith et al., 2010). As highlighted above, the deterministic DCM will only test the differences in connectivity within the commonly activated network, ignoring the regional effects that are absent or novel in patients compared to controls. For this reason we have suggested that the abnormalities should be reported at the level of regions in addition to the level of connectivity.

4. The remote effect of abnormal connectivity in another network. This problem concerns the influence of abnormal parts in a subsystem that may propagate to other networks causing “indirect” abnormal connectivity in the subsystem of interest. For instance, it is possible that damaged parts within one functional system manifest as abnormal connectivity that is not necessarily related to the main task/process of interest. This effect is linked to the more global problem of the missing nodes that can sometimes be sufficiently strong to invalidate DCM results of a given subsystem (see discussion in Daunizeau et al., 2010).

5. How generalizable are conclusions concerning abnormal connectivity? This issue is particularly critical for studies that aim to define a biomarker for a given pathology. First, the results identified in a particular group of subjects can only be translated to other subjects that show significant effects in all the regions of interest. What then happens when patients have one or two missing nodes? Second, generalizability also depends on abnormalities that manifest outside the network of interest, because the effects could have indirect influences on the sub-network of interest. This problem concerns all levels of inference in the implementation of deterministic DCM and should be carefully considered when generalizations are made about abnormal connectivity. It is even more critical if one is interested in making inferences at the individual level (e.g., if abnormal connectivity is being used as a biomarker for distinguishing between patients and controls or for classifying new patients). Nevertheless, it should be noted that patient and control connectivity parameters represent the same thing and are directly comparable when the analysis of patient and control data is based on the identical model space with identical priors for the model parameters (including those for the forward hemodynamic model and those for neuronal coupling).

FUTURE DEVELOPMENTS

In addition to the methodological issues highlighted above that warrant further investigations, other developments can potentially add more flexibility to the use of DCM in patient studies. We focus here on four future developments: (1) quantify the vulnerability of a given network after damage, (2) sample the model space in an unconstrained manner, (3) include new computational models in DCM, and (4) combine DCM with other approaches.

1. Quantifying the vulnerability of the network to damage. Parallels can be made with the increased interest in graph theory for complex brain network analysis (for review see Bassett and Bullmore, 2009; Bullmore and Sporns, 2009; Bressler and Menon, 2010; Guye et al., 2010). Using graph theory, previous studies have investigated connectivity changes that result from an insult to a given part of the network (e.g., Kaiser et al., 2007; Honey and Sporns, 2008; Alstott et al., 2009), for instance by using structural network properties of the lesion site to predict the functional (i.e., dynamic) consequences of the focal damage. This literature has yielded the development of different measures to quantify the “resilience” or the “vulnerability” of a given network after insult (see different coefficients in Rubinov and Sporns, 2010) and also proposed new statistical approaches for identifying differences in networks between patients and controls (see Zalesky et al., 2010). The translation of these sorts of measures to the DCM framework would be a valuable contribution for future studies of patients.

2. Sampling model space in an unconstrained manner. Defining the model space in a comprehensive way is a challenging step (e.g., see critique in Tauchmanova and Hromcik, 2008), in particular when the number of nodes or inputs is relatively high. One exciting possibility is the ability to sparsely sample
the model space in order to produce an approximation of the best model structure. For instance, using Gaussian Processes, it has been shown that it is possible to identify the optimal model structure based on the selected subsample of models (for more details see Anderson and Lane, 2009). We believe this kind of future development will increase the robustness of DCM when exploring large model spaces and will allow the structure of complex models to be identified.

3. Including new computational models. There is also an exciting opportunity to combine previous computational models with DCM (for detailed discussion see Friston and Dolan, 2010). For instance, previous models from computational neuroscience and theoretical neurobiology can be linked to a particular behavior or pathology (e.g., in patients with schizophrenia, Friston, 1996, or Alzheimer’s disease, Hasselmo, 1997) that can then be tested with DCM. Previous theoretical models constitute an interesting source of hypotheses that can inform the formation of the model space in DCM.

4. Combining DCM with other approaches. One crucial step in DCM is the identification of a subset of regions of interest that is usually based on an a priori knowledge. This hypothesis-driven step may leave several other activated regions unexplored. In this context, data-driven approaches can be used to explore the whole pattern and reduce it to a set of representative nodes that can then be incorporated into DCM (see de Marco et al., 2009). For instance, during a go/no-go task, a previous study has used independent component analysis to identify three response inhibition components and then DCM was used to assess the interactions between these components (Stevens et al., 2007). Recently, a new method called the switching linear dynamic systems was developed to identify important regions that might be left out during the specification of models (Smith et al., 2010). This method explicitly models the task factor as a Markov random variable and it aims to objectively measure the overall quality and sufficiency of an identified model (see illustration with a motor task in Smith et al., 2010).

CONCLUSION

We have succinctly reviewed the available DCM frameworks for fMRI responses that can be applied to characterize brain disorders at the level of networks and their dynamics. The previous literature has demonstrated the feasibility of DCM in diverse populations of patients and has provided valuable insights that characterize the dynamic signatures of focal and non-focal pathologies. We believe the growing use of DCM will shape the field of brain network dynamics and will have an important impact on the future development of theoretical neurobiology and clinical biomarkers.

ACKNOWLEDGMENTS

This work was funded by the Wellcome Trust and the James S. MacDonnell Foundation (conducted as part of the Brain Network Recovery Group initiative).

REFERENCES

Abutalebi, J., Rosa, P. A., Tettamanti, M., Green, D. W., and Cappa, S. F. (2009). Bilingual aphasia and language control: a follow-up fMRI and intrinsic connectivity study. Brain Lang. 109, 141–156.

Acs, F., and Greenlee, M. W. (2008). Connectivity modulation of early visual processing areas during covert and overt tracking tasks. Neuroimage 41, 380–388.

Agosta, F., Rocca, M. A., Pagani, E., Absinta, M., Magnani, G., Marconcini, A., Falautano, M., Comi, G., Gorno-Tempini, M. L., and Filippi, M. (2010). Sensorimotor network rewiring in mild cognitive impairment and Alzheimer’s disease. Hum. Brain Mapp. 31, 515–525.

Alexander, W. H., and Brown, J. W. (2010). Competition between learned reward and error outcome predictions in anterior cingulate cortex. Neuroimage 49, 3210–3218.

Allen, P., Stephan, K. E., Mechelli, A., Day, F., Ward, N., Dalton, J., Williams, S. C., and McGuire, P. (2010). Cingulate activity and fronto-temporal connectivity in people with prodromal signs of psychosis. Neuroimage 49, 947–955.

Almeida, J. R., Mechelli, A., Hassel, S., Versace, A., Kuper, D. J., and Phillips, M. L. (2009a). Abnormally increased effective connectivity between hippocampal gyrus and ventromedial prefrontal regions during emotion labeling in bipolar disorder. Psychiatry Res. 174, 195–201.

Almeida, J. R., Versace, A., Mechelli, A., Hassel, S., Quevedo, K., Kuper, D. J., and Phillips, M. L. (2009b). Abnormal amygdala-prefrontal effective connectivity to happy faces differentiates bipolar from major depression. Biol. Psychiatry 66, 451–459.

Alliot, J., Breakspear, M., Hagmann, P., Cammoun, L., and Sporns, O. (2009). Modeling the impact of lesions in the human brain. PLoS Comput. Biol. 5, e1000408. doi: 10.1371/journal.pcbi.1000408.

Anderson, R., and Lane, T., 2009. Fast Bayesian network structure search using Gaussian processes. Available at http://www.cs.unm.edu/~rreport/tt/09-06/graphic.pdf, p. 1–15.

Bassett, D. S., and Bullmore, E. T. (2009). Human brain networks in health and disease. Curr. Opin. Neurobiol. 22, 340–347.

Benetti, S., Mechelli, A., Picchioni, M., Broome, M., Williams, S., and McGuire, P. (2009). Functional integration between the posterior hippocampus and prefrontal cortex is impaired in both first episode schizophrenia and the at risk mental state. Brain 132, 2426–2436.

Bird, G., Catmur, C., Silani, G., Frith, C., and Frith, U. (2006). Attention does not modulate neural responses to social stimuli in autism spectrum disorders. Neuroimage 31, 1614–1624.

Bressler, S. L., and Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. Trends Cogn. Sci. (Regul. Ed.) 14, 277–290.

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198.

Cao, F., Bitan, T., and Booth, J. R. (2008). Effective brain connectivity in children with reading difficulties during phonological processing. Brain Lang. 107, 91–101.

Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., and Price, C. J. (2009). An anatomici
cal signature for literacy. Nature 461, 983–986.

Catan, M., and Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex 44, 1105–1132.

Chow, H. M., Kaug, P., Raabe, M., and Greenlee, M. W. (2008). Evidence of fronto-temporal interactions for strategic inference processes during language comprehension. Neuroimage 40, 940–954.

Crossley, N. A., Mechelli, A., Fusi-Poli, P., Broome, M. R., Matthiasson, P., Johns, L. C., Bramon, E., Velmagga, L., Williams, S. C., and McGuire, P. K. (2009). Superior temporal lobe dysfunction and frontotemporal disconnectivity in subjects at risk of psychosis and in first-episode psychosis. Hum. Brain Mapp. 30, 4129–4137.

Daunizeau, J., David, O., and Stephan, K. E. (2010). Dynamic causal modelling: a critical review of the biophysical and statistical foundations. Neuroimage doi: 10.1016/j.neuroimage.2009.11.062. (in press).

Daunizeau, J., Friston, K. J., and Kiebel, S. J. (2009). Variational Bayesian identiﬁcation and prediction of stochastic nonlinear dynamic causal models. Physica D 238, 2089–2118.

de Marco, G., Devauchelle, B., and Berquin, P. (2009). Brain functional modeling, what do we measure with fMRI data? Neurosci. Res. 64, 12–19.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. J. R. Stat. Soc. Series B 39, 1–38.

den Ouden, H. E., Daunizeau, J., Roiser, J., Friston, K. J., and Stephan, K. E. (2010). Striatal prediction error modules...
Friston, K. J. (2009a). Causal modelling and brain connectivity in functional magnetic resonance imaging. PLoS Biol. 7, e1000033. doi: 10.1371/journal.pbio.1000033.

Friston, K. J. (2009b). Modalities, modes, and models in functional neuroimaging. Science 326, 399–403.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6, 218–229.

Friston, K. J., and Dolan, R. J. (2010). Computational and dynamic models in neuroimaging. Neuroimage 52, 752–765.

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. Neuroimage 19, 1273–1302.

Friston, K. J., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W.D. (2002). Variational Bayes under the Laplace approximation, in Statistical Parametric Mapping: The Analysis of Functional Brain Images, eds K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nichols, and W. D. Penny. (London: Elsevier), 606–618.

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. Neuroimage 19, 1273–1302.

Friston, K. J., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W.D. (2002). Variational Bayes under the Laplace approximation, in Statistical Parametric Mapping: The Analysis of Functional Brain Images, eds K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nichols, and W. D. Penny. (London: Elsevier), 606–618.

Friston, K. J. (2007). “Functional integration of the anterior cingulate cortex during conflict processing.” Cereb. Cortex 17, 2400–2406.

Fan, J., Hof, P. R., Guise, K. G., Fossella, J. A., and Posner, M. I. (2008). The functional integration of the anterior cingulate cortex during conflict processing. Cereb. Cortex 18, 2268–2279.

Fan, L. Y., Lee, S. H., and Chou, T. L. (2010). Interaction between brain regions during semantic processing in Chinese adults. Lang. Linguist. 11, 159–182.

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. 2, 36–78.

Friston, K. J. (1996). Theoretical neurobiology and schizophrenia. Br. Med. Bull. 52, 644–655.

Friston, K. J. (2002a). Bayesian estimation of dynamical systems: an application to fMRI. Neuroimage 16, 513–530.

Friston, K. J. (2002b). Functional integration and inference in the brain. Prog. Neurobiol. 68, 113–143.

Friston, K. J. (2007). “Functional integration,” in Statistical Parametric Mapping (SPM), The Analysis of Functional Brain Images, eds K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nichols, and W. D. Penny (London: Elsevier), 471–491.
