Title: Principles of Dynamic Causal Modelling and its Application in Psychiatry

Course: MRes Brain Sciences

Candidate Number: **JLPF9**

Word count: 5805

Principles of Dynamic Causal Modelling and its Application in Psychiatry

1. Introduction

1.1 Purpose of the Review

This review is designed to systematically review from existing literatures with a focus on the principles, historical development and the application of the Dynamic Causal Modeling (DCM) in analyzing different mental disorders. This review aims to show how psychiatry research has benefit in using DCM and the with their significant findings as examples. This paper will also discuss about the limitations as well as solutions that could be improve DCM's capability.

1.2 Background Information

In neuroscience the researchers have been working hard with exploring the functional and structural organization of the brain and how the brain regions work together. In traditional studies, neuroscientists were primarily able to employ functional connectivity methods such as correlation analysis and coherence measures to study the brain networks. These include techniques such as Statistical Parametric Mapping (SPM) and voxel-based morphometry (VBM). These methods are able to find out which regions of the brain were active or had different structural densities during certain activities or in given states [1]. Such techniques indeed provided valuable knowledge about the structure and functioning of the human brain; however, they had a disadvantage of being purely descriptive and inability to trace the direction of connections between different regions of the brain[2]. This limitation identified the need of a method that can identify the cause effect relationships in the brain networks and thus gave rise to DCM.

DCM was initially proposed in the field of neuroscience as a significant breakthrough that opens a possibility to discover the causal relations and pathways in the brain[3]. Unlike other functional connectivity analysis methods that provide only correlation patterns, DCM goes a step forward by determining the flow of information, capturing the direction and causality of the connections in the brain. DCM is different than other methods of functional connectivity analysis that only provide correlation patterns as it seeks to determine the flow of information which constitutes the connectivity of the brain. This approach is particularly relevant for mental disorders where the affected connectivity and the disrupted neural communications are frequently found in the pathology of schizophrenia, depression and anxiety disorders[4].

1.3 Relevance to Mental Disorders

Mental disorders have been found to be usually accompanied by alterations in connectivity and activity of the neuronal networks of the brain. Although previous studies have already provided evidence, it is still somewhat limited in identifying how various brain regions are interconnected[5]. To overcome this, DCM could a potential solution because it offers a way of modelling and capturing such causal relationships, and thus discover the underlying neural dynamic in a better manner[6]. This is useful in identifying possible biomarkers for diagnosis, revealing the causes of mental disorders and in coming up with ways of addressing them[7].

Since the development of DCM, it has played a significant role in mental health research, with several studies showing its benefits in uncovering key insights into various mental disorders. For schizophrenia, DCM has demonstrated the dysconnectivity of the important brain networks, especially in the prefrontal and parietal cortices that are associated with cognitive dysfunction and psychosis symptoms. As for depression, DCM has identified the changes in connections within the fronto-limbic network from prefrontal cortex to amygdala. DCM also contributed to the research of bipolar disorder by discovering specific impairments in top-down emotion regulation, supporting the idea that neural dysconnectivity is a core feature of mood dysregulation.

2. Historical Development and Theoretical Background

2.1 Theoretical Foundations

The history of DCM can be traced back to the early 2000s, when Karl Friston and his colleagues proposed the method to overcome challenges of the conventional methods of studying connectivity. These previous methods were essentially of a correlational type and did not shed light on the connectivity of the brain areas[3].

Given the limitations of the conventional methods, DCM was introduced with a more sophisticated theoretical framework to fill the gaps. At the beginning, DCM was designed mainly for fMRI data analysis with focus on the hemodynamic responses related to neuronal activity. Through development over time, its usage has been expanded to include EEG and MEG data [8], which allows the study of the electrical and magnetic signal inside brain. This enhancement further leads to causal interactions in both the spatial and temporal domain thus improving the clarity of the underlying neural mechanisms. When it comes to the development of DCM models, they have been upgraded to be able to include multiple brain regions and to take into account reciprocal connections between them. techniques such as Bayesian model reduction and empirical Bayes were gradually included to improve the computational power and the reliability of DCM analyses[9].

3. Key Concepts and Methodological Framework

DCM integrates multiple components which plays distinct roles in understanding brain connectivity. In essence, DCM uses a biophysical approach in order to describe how the activity of neurons generates the measured data such as fMRI signals [10]. This model includes our current knowledge about neurovascular coupling, synaptic transmission, and neural dynamics allowing for the whole-brain functional representation [3].

Bayesian approach is another important component of DCM which allows for the estimation of model parameters as well as for comparison between different models and thus identify the model that best explains the data [11]. The Bayesian framework ensures the validation of DCM's inferences, by integrating knowledge from previous experimental evidence [6].

3.1 Causal Interactions and Connectivity

Causal interactions are directional influences that one area of the brain has over another area. This is important because the brain is made of so many systems and each system are interdependent with the other and one cannot work in isolation without influencing the rest in a certain way. Understanding of these interactions is useful in the explanation of the mental disorders such as schizophrenia, depression and anxiety disorders where symptoms may be due to aberrant causal interactions [3, 12]. By fitting the model to the data, DCM allows for the identification of the effective connectivity in the brain networks, which can help to understand the ways in which different brain areas are interconnected under various conditions, such as resting states, cognitive tasks, or in the presence of psychiatric symptoms[3, 8, 13].

3.2 Biophysical Models

DCM uses known biophysical processes to simulate the connection between these relationships. For example, in functional Magnetic Resonance Imaging (fMRI), the observed data are blood oxygen level-dependent (BOLD) signals, which are an indirect measure of neural activity because of neurovascular coupling. The modelling of such signals involves the hemodynamic response that describes how changes in the activity of neurons lead to changes in blood flow and oxygenation that culminate to the BOLD signal[3]. In the case of electroencephalography (EEG) and magnetoencephalography (MEG), the measured data is electrical and magnetic fields produced by neural activity. In this case, DCM uses models of how electrical activity spreads in the brain and how the signals are produced [10, 14]

The usage of the neuronal models in biophysical models is to calculate the neural dynamics and connectivity within and between brain regions. The process starts with obtaining the neural state, which describes the activity of a neuron population:

$$\dot{x} = f(x, u, \theta_1) \tag{1}$$

In which $\dot{x}$ is the rate of change of the neural states, x, u is the input to the system, and θ_1 are the parameters of the model.

After obtaining the neural state of a single neuron population, the model then extends to represent the neural state of a brain region, which consists of multiple neuronal populations. Considering two connected brain regions V1 and V2. The neural state equations for these regions, representing the neural activity within each region, can be written as:

$$\frac{\mathrm{d}\chi_1(t)}{\mathrm{d}t} = f_1(\chi_1(t), \chi_2(t), u(t), \theta) \tag{2}$$

$$\frac{d\chi_2(t)}{dt} = f_2(\chi_2(t), \chi_1(t), u(t), \theta) \tag{3}$$

Where $\chi_1(t)$ and $\chi_2(t)$ are the neural states, for example, average membrane potentials, of the two regions respectively. u(t) can be considered as an external stimulus (e.g., a flash of light in a visual related task). θ includes parameters such as coupling strengths between A and B. In this context, f_A and f_B represent how the activity in V1 and V2 changes over time, considering their interactions and external input[10, 12].

Once the neural state of a single region is established, connectivity between regions can be represented by three matrices: matrix ***A*** indicates both intrinsic connections within the regions (which focus on how the interactions of neurons/groups of neurons within that region evolve over time, represented by diagonal elements) and connectivity between regions (which indicates the influence of one region on another, represented by the off-diagonal elements) with no experimental perturbations. Matrix ***B*** represent input-dependent modulations of these connections, which reflects how external stimuli modulates the connections. Matrix ***C*** as the input matrix, representing the direct impact of external stimuli on each brain region. In this case, the state equation can be further expressed as:

$$\frac{d}{dt}\begin{pmatrix}x_1(t)\\x_2(t)\end{pmatrix} = \boldsymbol{A}\begin{pmatrix}x_1(t)\\x_2(t)\end{pmatrix} + \sum_i \boldsymbol{B}_i \begin{pmatrix}x_1(t)\\x_2(t)\end{pmatrix} u_i(t) + \boldsymbol{C}u(t) \quad (4)$$

For example, suppose the ***A***, ***B*** and ***C*** are

$$A = \begin{pmatrix}-1 & 0.5\\0.5 & -1\end{pmatrix}, B = \begin{pmatrix}0 & 0.1\\0.1 & 0\end{pmatrix}, C = \begin{pmatrix}1\\0\end{pmatrix}$$

The negative diagonal elements of ***A***(-1) represent self-inhibition within each region, whereas the positive off-diagonal elements of ***A***(0.5) represent excitatory effect between V1 and V2. **B** represents the effect of an external stimulus $u(t)$ on the connectivity between V1 and V2. Finally, **C** represents direct input to V1[3].

With the biophysical model, researchers can set up the target systum under different conditions, etc.

3.3 Observation Models

The observation model in DCM maps the neural states that are not directly accessible to measurements to the actual data. There is a general form which can be expressed as

$$\mathrm{y(t)} = g(x(t), \theta_0) + \epsilon(t) \quad (5)$$

where $\mathrm{y(t)}$ represents the observed data, and $x(t)$ indicates the hidden neural states, which means the internal brain activity occurring at neuronal level but is not directly observed. $g(\cdot)$ is the observation function, which maps the hidden neural states $x(t)$ to the observed data $\mathrm{y(t)}$, translating the hidden neural states into data type that is measurable (such as transforming neural activity to measurable blood flow changes in fMRI). θ_0 represent the observation model's characteristics, which defines how well the hidden neural states are converted into measurable data. $\epsilon(t)$ is the observation noise, representing random fluctuations in measurements. [8, 11]

fMRI

Because the core of fMRI is the BOLD signal and the hemodynamic response function, the corresponding observation model was designed to describe the relation between non-measurable neural activity and the measurable BOLD, which can be expressed as follow:

$$\mathrm{y(t)} = h(x(t), \theta_h) + \epsilon(t) \tag{6}$$

In this case, $\mathrm{y(t)}$ is the BOLD signal; $h(\cdot)$ is the hemodynamic response function; $x(t)$ is the hidden neural activity; θ_h are parameters of the hemodynamic model; and finally, $\epsilon(t)$ is the observation noise.

The Balloon model is one of the most popular hemodynamic models, which considers multiple parameters such as signal decay, feedback regulation, transit time and vascular conductivity [3, 15]. In detail of the Balloon model, the hemodynamic response function can be described by a set of differential equations, that include the modelling of the dynamics of the vasodilatory signal, blood flow, blood volume, and deoxyhemoglobin content as follows:

1. Vasodilatory Signal Dynamics:

$$\frac{ds(t)}{dt} = x(t) - \frac{s(t)}{\tau_s} - \gamma(f(t) - 1) \tag{7}$$

- $s(t)$: Vasodilatory signal, change in the diameter of the blood vessels in response to neural activity. When neurons are active, they flank a series of chemical reactions that make blood vessels to dilate and hence increase the blood flow supply to the area.
- $x(t)$: Neuronal activity, the main cause for the vasodilatory signal.
- τ_s: Time constant for the decay of the vasodilatory signal.
- $f(t)$: Blood inflow which depends on vasodilatory signal.
- γ: flow-dependent elimination rate.

2. Blood Flow Dynamics

$$\frac{df(t)}{dt} = \frac{s(t)}{\tau_f} \tag{8}$$

- This equation shows the relationship between the vasodilatory signal $s(t)$ and blood flow $f(t)$. The blood flow increases in proportion to the vasodilatory signal but with a delay governed by τ_f.

The above two equations indicate that blood flow initially increases in response to neural activity and then is modulated by the ongoing vasodilation and autoregulatory feedback.

3. Dynamics of Blood Volume and Deoxyhaemoglobin Content

$$\tau_0 \frac{dv(t)}{dt} = f(t) - v(t)^{1/\alpha} \tag{9}$$

$$\tau_0 \frac{dq(t)}{dt} = \frac{f(t)\left(1 - (1 - E_0)^{1/f(t)}\right)}{E_0} - \frac{q(t)v(t)^{1/\alpha}}{v(t)} \tag{10}$$

- $v(t)$: Blood volume, the amount of blood that exists in a certain area of the brain at a certain time.
- $q(t)$: Deoxyhemoglobin content, the amount of oxygen-depleted blood.
- τ_0: Time constant, how fast blood volume and deoxyhaemoglobin concentration change upon changes in blood flow.
- α: Grubb's exponent describes the nonlinear relationship between blood flow and blood volume.
- E_0: The resting oxygen extraction fraction, the proportion of oxygen consumed from the blood due to the neuron activity.

4. The BOLD Signal

$$y(t) = V_0\left[k_1\big(1 - q(t)\big) + k_2\left(1 - \frac{q(t)}{v(t)}\right) + k_3\big(1 - v(t)\big)\right] \qquad (11)$$

- $y(t)$ represents the BOLD signal
- V_0 is a scaling factor representing the resting blood volume fraction.
- k_1, k_2, k_3 are coefficients that weight the contribution of changes in deoxyhaemoglobin and blood volume to the BOLD signal.

Together, these equations tell the story of the entire process from neural activaty to the observed BOLD signal. They are based on the Balloon model which is widely used for the description of the hemodynamic response in fMRI. Initial work by Buxton et al. (1998) and further advances by Friston and colleagues (2000) have acted as the basis for creating these models, which are still widely used in the present fMRI and DCM research[15, 16].

EEG/MEG Observation Model

In the case of EEG and MEG, the observation model consists of relating the neural activity to the measured signals via lead field matrices. The observed signal $\mathrm{y(t)}$ can be modelled as:

$$\mathrm{y(t)} = Lx(t) + \epsilon(t) \qquad (12)$$

where L is the lead field matrix, while $x(t)$ and $\epsilon(t)$ remain to be the hidden neural states and the observation noise respectively[8, 17].

The lead field matrix is the mapping between the neural sources in the brain such as the current dipoles that can be generated by activity in the brain and the signals that can be measured on the scalp in EEG or around the head in MEG. Some of the factors which may influence the lead field matrix include the head geometry, properties of different tissues and position of the sensors. Since the shape and size of the head and skull and brain also differ, it has an impact on the field propagation from neural source to the sensors [18, 19]. In addition, due to the fact that tissues like brain, cerebrospinal fluid, skull and scalp possess different conductivity, such differences may affect the strength and distribution of the signals recorded by EEG and MEG sensors[20, 21]. The spatial arrangement of the electrodes or sensors of EEG

and MEG also significantly influences the signal, as sensors that are closer to the source of activity generate higher amplitude signals [17, 22].

In an EEG setting where there are several electrodes placed on the scalp, one can think of a single active neural source within the brain in the form of an electrical dipole. The lead field matrix L is used to transform the activity of the source to the signals detected by each electrode. L contains the entries that represent the amount of the signal at a given electrode which may be contributed by the source while taking into consideration the position, the direction of the source, and the conductivity of the tissues in between. Given the properties of electromagnetic field, a source has a perpendicular orientation and is close to the scalp will produce a higher amplitude at the nearest electrode than a source which is deeper in the cortex or has a different orientation [23, 24]. Therefore, the lead field matrix can be viewed as the bridge that transforms the physiological electrical activity in the brain into the signals that can be picked up by the sensors, thus allowing the identification of the areas of brain activity. Its estimation and use are crucial in analysing the EEG/MEG data in order to reconstruct the spatial and temporal dynamics of neural activity from surface-level recordings [25, 26].

3.4 Bayesian Framework

The most significant advantage of the Bayesian inference is that it has the ability to construct the probability of a hypothesis based on both prior knowledge and new data. This process is based on Bayes' theorem the tells how the probability of a hypothesis could be updated given new evidence[27]. Such an approach is especially suitable within the context of DCM because it allows for the fitting of rich neural models that depict the causal connections between the various areas of the human brain [3, 11].

The parameter estimation process involves determining the values of the parameters in the models that best explain the observed data. The probability of different parameter values (i.e. posterior distribution) can be calculated by combining the new observed data with the parameters with the prior distribution (assumptions about the parameters based on previous knowledge), updating the prior knowledge with the new evidence. Mathematically, this calculation process can be expressed as:

$$p(\theta \mid y) \propto p(y \mid \theta) \cdot p(\theta) \tag{13}$$

Where θ represents the model parameters, y is the observed data, $p(y|\theta)$ is the prior distribution, which assess how probable the observed data can be fitted given a set of parameters, $p(\theta)$ is the prior distribution of the parameters.

Model evidence or sometimes called marginal likelihood is defined as the probability of the data given the model, which measures how well a model explains the observed data. It answers the question: "Given the model, how likely is it that the data observed would occur?" The model evidence is obtained by integrating the likelihood function with all possible values of the parameters. Each choice of parameter is weighted by its prior probability, which represents how likely the parameter is based on previous knowledge[28].

$$p(y \mid m) = \int p(y \mid \theta, m) \cdot p(\theta \mid m)\, d\theta \tag{14}$$

where m represents the model, $p(y \mid m)$ is the model evidence. $p(y \mid \theta, m)$ represents the probability of observing the data y given a particular set of parameters θ and the model m. $p(\theta \mid m)$ is the prior distribution of the parameters θ given the model m. $\int \ldots d\theta$ indicates the summing over all possible values of the parameters θ.

The model evidence is evaluated by the free energy which balances how well the model can fit the data and how complex the model could be[11]. In order to find the best model, the one with the highest model evidence or the lowest free energy is preferred since that means the best data fit and the least complexity.

3.5 Model Specification

The first step of applying DCM is model specification which concerns the identification of the regions of interest (ROIs) and their connections, this is usually done based on previous knowledge of the areas of the brain that are associated with the cognitive or sensory task of interestt[12]. For instance, when conducting studies related to visual processing, areas like the primary visual cortex or V1 are often chosen together with the secondary visual cortex or V2 because of the known involvement of these areas in visual information processing[29, 30]. While in research on motor control, the motor cortex or cerebellum could be selected as ROIs[31, 32]. As for the connections, there are two types of connections in DCM: one demonstrates feedback loops within an individual ROI, namely intrinsic connection, while the other shows the relationship between the ROIs, namely extrinsic connection[11]. Another important aspect of model specification is the identification of external inputs which influence the system. They are the environmental factors or task conditions that influence the neural activity of the brain[33]. In experimental paradigms, inputs may be in the form of sensory stimuli, for instance, a flash of light or a motor task[34]. These inputs are also modelled with different form to model the brain's response to certain events[8]. In some cases, the modelling for stimuli can be not necessary, for example when dealing with resting state observed data.

The second step will be the parameter estimation. Initial parameters are usually set with a priori knowledge of brain connectivity[35]. For example, the connection between regions in hierarchical sensory systems may be assigned higher values because of their known function in processing sensory information[36]. Additionally, within-region parameters, which include the relative strength of excitation and inhibition, are derived as part of the model fitting process [37]. Although there is parameter optimization process, to what extent the parameters can be adjust must be defined based on prior knowledge. For example, anatomical information about the strength of connections between the regions should be used as prior information, which restricts the estimated connectivity values to be within a biologically reasonable range[11, 35].

The final step is comparing the models and evaluating alternative hypotheses about brain connectivity[11]. To find the best model, DCM not only can select between models, but also enables Bayesian model reduction, which simplifies a given model by removing unnecessary parameters or connections[9]. This helps researchers find the model that best describes the data while having the least number of parameters or connections[38], and speed up computations in

large-scale studies involving multiple brain regions[9]. Besides modelling individual subject, DCM can also be used for Group analysis. The investigation of common connectivity patterns shared by a group of individuals or differing between two populations, such as patients and healthy controls, can be conducted[9]. [39]

4. Clinical Applications in Mental Disorders

4.1 Schizophrenia

Schizophrenia has been a major focus of DCM research due to its strong association with abnormal brain connectivity. One of the earliest applications of DCM in schizophrenia examined perceptual differences, particularly the reduced susceptibility to the hollow mask illusion, where individuals with schizophrenia are less likely to perceive a concave face as normal. Two studies conducted by Dima et al. [40, 41] used fMRI and EEG to show that while healthy individuals rely more on top-down processing (higher brain areas controlling lower sensory regions), schizophrenia patients exhibit a shift toward bottom-up processing (sensory-driven perception). This disruption suggests that schizophrenia impairs the brain's ability to integrate prior knowledge to interpret sensory input.

Further research has explored predictive coding, a theory that the brain constantly generates predictions about sensory input and adjusts based on actual stimuli. Using the mismatch negativity (MMN) paradigm to measure how the brain responds to unexpected sounds, Dima et al. [42] and Ranlund et al. [43] discovered that forward connections (from sensory areas to higher-order regions) were altered in schizophrenia patients, potentially explaining difficulties in sensory prediction, a hallmark of the disorder.

DCM has also been instrumental in understanding working memory deficits in schizophrenia, a cognitive function that is often impaired. Studies by Crossley et al. [44], Deserno et al. [45], and Schmidt et al. [46] revealed reduced effective connectivity between higher cognitive regions like the dorsolateral prefrontal cortex (DLPFC) and sensory processing areas. This suggests that schizophrenia patients struggle with organizing and controlling information within working memory, impairing cognitive tasks.

Besides task-based studies, DCM has been applied to study resting-state networks, particularly the default mode network (DMN), which is active when the brain is at rest. In their study of first-episode psychosis patients, Bastos-Leite et al. found weakened connectivity within the DMN, suggesting that disruptions in this network may play a role in the early stages of schizophrenia [47].

At the synaptic level, DCM has revealed abnormalities in the major receptors involved in schizophrenia, including NMDA receptors. In their study, Moran et al. [48] showed that the drug L-Dopa affected the NMDA and AMPA receptor functions in the schizophrenia patients and therefore proposed that the receptor dysregulation could be a factor in the pathogenesis of

schizophrenia. Gilbert et al. [49] also investigated ion channel dysfunction where they classified schizophrenia patients with particular gene variants from healthy controls via MEG data and proved DCM's efficiency in defining synaptic pathologies.

Lastly, DCM has been applied for grouping schizophrenia based on neural connectivity architecture. Brodersen et al. [50] have used DCM parameters to assign the schizophrenia patients into subgroups and reported that these subgroups are associated with different symptomatology. This approach has the potential of improving the diagnosis and treatment of the disorder by distinguishing unique patterns of brain dysfunction that are present in the schizophrenia spectrum.

4.2 Depression

DCM have proved to be invaluable tools in the study of neural basis of affective dysregulation and cognitive impairment. A large number of studies using DCM has shown that Major Depressive Disorder (MDD) is characterized by dysfunctions in critical networks of the brain which include the networks involved in the regulation of emotions, attention, and self-referential processing.

One of the main focuses has been emotional regulation, with particular emphasis on the disconnection in prefrontal-limbic networks, involving key regions such as amygdala, dorsolateral prefrontal cortex (DLPFC), and anterior cingulate cortex (ACC). Schlösser and colleagues used fMRI and DCM to investigate the FC in patients with major depressive disorder[51]. The ACC activation was found to be significantly stronger in the patient group than the control group with intrinsic connectivity from dorsal to rostral ACC. In addition, the study found a significant correlation between the interference scores and intrinsic connections from the dorsal ACC to the DLPFC in both groups with no statistic difference between the two groups. In addition to intrinsic connectivity, a 2012 study by Lu et al. [52] used MEG with DCM to assess the between-region connectivity in the prefrontal-limbic networks during emotional processing in patients with MDD. They found reduced top–down effective connectivity from DLPFC to the amygdala in patients with depression, suggesting that the prefrontal cortex has a reduced capacity to modulate emotional networks. However, the bottom-up connectivity from amygdala to the ACC was enhanced, indicating an enhanced emotional reactivity in MDD patients, which contribute to the difficulties in managing bad mood. Demonstrating disruption on both intrinsic and between-region connections, these results are in line with current theories of the differential involvement of the fronto-cingulate system in major depression.

Attention and cognitive control have been another major focus in the DCM research. In Desseilles et al. [53] the authors investigated the effects of depression on the attentional networks using fMRI and DCM. They determined that MDD patients had an altered visual attention regulation, with focuses on the top-down control from intraparietal sulcus to the visual areas than the healthy subjects. This impaired attentional filtering could be a reason why depression patients have problems in focusing on task relevant stimuli while at the same time inhibiting other stimuli, which results in inefficiency in cognitive processes in daily activities. Moreover, a recent study extended the focus to larger scale brain networks including DMN, salience network (SN), and the dorsal attention network (DAN). Wang et al. [54] employed spectral DCM

and resting-state fMRI to investigate these networks and revealed enhanced positive connections within the DMN in the MDD patients, which may be attributed to the overactive self-referential processing and rumination. In addition, patients with recurrent depression displayed had even greater changes in effective connectivity than patients with first-episode patients, indicating that disruption of networks is more severe in recurrent depression. In addition to adult patients, efforts had also been made to understand the dysconnectivity patterns in adolescents MDD. Musgrove et al. [55] found decreased bottom-up effective connectivity from amygdala to subgenual anterior cingulate cortex (sgACC) in unmedicated adolescents with MDD. This finding indicates that the processes of emotional regulation in adolescent depression are uniquely impaired, which may lead to lifelong problems with emotional dysregulation.

Besides emotional and cognitive dysregulation, self-referential processing in MDD has been found to have abnormal connectivity patterns. A study conducted by Davey et al. [56] using DCM revealed that the medial prefrontal cortex had an enhanced control over the posterior cingulate cortex in depressed patients. This hyperregulation was linked to negative self-evaluation, a feature of depressive thinking. This result confirms the hypothesis that abnormal self-referential processing networks contribute to the continuation of depressive symptoms.

4.3 Bipolar disorder

Bipolar disorder (BD) has also been identified to be associated with emotional dysregulation and cognitive impairments. Similar to emotional regulation study on depression, DCM studies on BD also focused primarily on fronto-limbic dysfunction, and both studies on BD and MDD shows similar results. Radaelli et al. [57] employed DCM to examine the effective connectivity between these regions in BD patients and observed decreased connectivity from the DLPFC to the amygdala, indicating dysfunctional top-down regulation of emotion in BD, similar to MDD[52]. This is consistent with Breakspear et al.'s study which investigated participants with genetic propensity to BD and noted that the dysfunctional ACC to DLPFC to inferior frontal gyrus hierarchy in the first network level indicated that early network alterations play a part in emotional regulation[58]. In a similar study, Dima et al. [59] took a step forward to include both the BD patients and their unaffected family members. Resilient relatives also had higher fronto-limbic connectivity during the emotional task than the healthy controls, and this was further elevated in the ventral visual stream, which may be a compensatory mechanism that would have protected them from BD. DCM has also been employed to examine treatment responses based on fronto-limbic networks. Vai et al. [60] observed that effective chronotherapy consisting of sleep deprivation and light therapy in BD patients affected the connectivity between the DLPFC and ACC and the modulation of the Amygdala-DLPFC pathway, which means that the enhancement of fronto-limbic connectivity may be useful in predicting the effectiveness of treatment.

Also results for fronto-limbic connectivity are relatively consistent between BD and MDD, studies involving orbito-medial prefrontal cortex (OMPFC) shows opposite results. Almeida et al. [61] compared BD with MDD and observed abnormal bottom-up connectivity from the amygdala to the OMPFC in BD patients in happy face recognition whereas MDD is

characterized by top-down OMPFC to amygdala dysconnectivity. This finding identifies unique neural networks for each of the disorders, which is crucial in differential diagnosis.

In addition to the emotional regulation, Diwadkar et al. [62] investigated the cortical-striatal connectivity in the adolescents with high genetic risk of BD while performing the sustained attention tasks. The researchers identified decreased connectivity of the frontal-striatal network that points to the early signs of attention network impairment associated with BD neurobiology, which can be used as a biomarker for the disorder.

DCM has been proved to be a powerful technique for understanding the neural basis of mental disorders. DCM through effective connectivity analysis between the key brain regions has helped in understanding how dysconnectivity of the networks leads to emotional, cognitive and behavioural manifestation. However, most of the studies have been conducted on these disorders; DCM has a possibility of being used in other psychiatric conditions for instance, anxiety disorders[63, 64],autism spectrum disorders. Although the number studies regarding DCM analysis in the field of psychiatry has been increasing, the existing studies are still not enough to produce robust findings due to the lack of validation.

5. Challenges and Future Directions

Despite the advancements made through the use of DCM to reveal the connectivity of the brain, there are still some drawbacks and issues encountered. The main drawback of the present approach is that DCM is a hypothesis-driven method, which implies that pre-defined regions of interest (ROIs) are used when designing the model. This is not ideal for data exploration. While approaches that include Bayesian Model Reduction (BMR) have been proposed to search for models across reduced model space, and methods for analysing large scale brain networks have been suggested, these still require model specification. Consequently, psychophysiological interaction analysis is usually used for explorative analyses, in particular when some areas of interest are to be suggested, which could be used as ROIs in DCM[13].

Another major issue relates to the computational aspects of DCM with a focus on the variational Bayesian algorithms applied for parameter estimation. These methods based on the Laplace approximation, in which the posterior distribution of the parameters is approximated by a Gaussian distribution. This assumption, however, complicates the process further where highly non-linear models are involved since local minima may result in the free energy not being a good representation of the model evidence[65]. It may however happen that the Laplace approximation is less accurate and the resulting parameter estimates are less reliable in such cases. Although there are more accurate methods of estimating the distribution such as Metropolis-Hastings algorithm, they are costly in terms of time and computational power. Due to this, sampling is most of the time utilized to confirm the validity of the variational approximations in DCM[66].

Despite these difficulties, DCM still offers a valuable way of investigating effective connectivity in the brain. Further developments of the methodological approach will lead to the extension of the range of its potentials and increase in accuracy. Thus, as DCM develops, one of the prospective directions of improvement is the combination with other methods, including machine learning and graph theory, which hold the promise of improving the model specification as well as reduce the computational burden in the models. An additional future

research direction is to use DCM in real-time as a tool for providing real-time feedback in clinical practice. In addition, individual level models can be useful for precision psychiatry to better pinpoint neural patterns and work on

6. Conclusion

Dynamic Causal Modelling is a powerful technique which has revolutionized the field of brain connectivity and has offered potential pathways to the understanding of mental illnesses. In this review, we have discussed the theoretical background, the methodological approaches and the potential of DCM for investigation of the structural and functional connectivity of human brain networks. As compared to more conventional correlation analyses, DCM allows the researchers to investigate the causal connections and mutual connections between different brain areas, which can be considered as a significant improvement when considering the neural dynamics.

It is in the area of psychiatry that one of the most important uses of DCM can be seen especially in schizophrenia, depression, and bipolar disorder. In this review, it has been shown that research that has employed DCM has found that there are always abnormal connectionist signatures in patients with these conditions. For instance, in schizophrenia, the studies have established that the cognitive control network is disrupted, while in depression it has been evidenced that the prefrontal-amygdala interactions are abnormally functional and lead to the dysregulation of emotions. These findings not only have helped advance the knowledge of the neurobiological mechanisms of mental disorders but also have provided new ways of thinking about diagnosis and treatment.

DCM could play a pivotal role in development of personalised medicine. With the development of DCM, it is possible to generate personalized models of brain connectivity that could potentially transform the landscape of psychiatric treatment by employing specific neural patterns for patients. In addition, integrating DCM with other neuroimaging approaches such as functional connectivity and graph theory might improve its value in finding biomarkers of psychiatric disorders with aim of early detection and treatment.

In conclusion, DCM has brought significant changes in the understanding of the human brain especially the connection between the brain and mental disorders. Due to providing a valuable tool for modelling effective connectivity, DCM has played an important role in advancing psychiatry and understanding mental disorders. Although there are still some problems in DCM, the further developments in the methodology of DCM can provide more possibilities for extending the use of DCM and enhance the efficiency of model specification in future mental health studies.

Reference

1. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: A general linear approach. Human brain mapping. 1994;2(4):189-210.
2. Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. Magnetic resonance in medicine. 1995;34(4):537-41.
3. Friston KJ, Harrison L, Penny W. Dynamic causal modelling. NeuroImage (Orlando, Fla). 2003;19(4):1273-302.
4. Fornito A, Zalesky A, Breakspear M. The connectomics of brain disorders. Nature reviews Neuroscience. 2015;16(3):159-72.
5. Fox MD, Greicius M. Clinical applications of resting state functional connectivity. Frontiers in systems neuroscience. 2010;4:19-.
6. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. NeuroImage (Orlando, Fla). 2009;46(4):1004-17.
7. García-Gutiérrez MS, Navarrete F, Sala F, Gasparyan A, Austrich-Olivares A, Manzanares J. Biomarkers in Psychiatry: Concept, Definition, Types and Relevance to the Clinical Reality. Frontiers in psychiatry. 2020;11:432-.
8. Daunizeau J, David O, Stephan KE. Dynamic causal modelling: A critical review of the biophysical and statistical foundations. NeuroImage (Orlando, Fla). 2011;58(2):312-22.
9. Friston KJ, Litvak V, Oswal A, Razi A, Stephan KE, van Wijk BCM, et al. Bayesian model reduction and empirical Bayes for group (DCM) studies. NeuroImage (Orlando, Fla). 2016;128:413-31.
10. Stephan KE, Harrison LM, Penny WD, Friston KJ. Biophysical models of fMRI responses. Current opinion in neurobiology. 2004;14(5):629-35.
11. Penny WD, Stephan KE, Mechelli A, Friston KJ. Comparing dynamic causal models. NeuroImage (Orlando, Fla). 2004;22(3):1157-72.
12. Friston KJ. Functional and effective connectivity: a review. Brain connectivity. 2011;1(1):13-36.
13. Stephan KE, Penny WD, Moran RJ, den Ouden HE, Daunizeau J, Friston KJ. Ten simple rules for dynamic causal modeling. Neuroimage. 2010;49(4):3099-109.
14. Friston K. The free-energy principle: a unified brain theory? Nature reviews Neuroscience. 2010;11(2):127-38.
15. Buxton RB, Wong EC, Frank LR. Dynamics of blood flow and oxygenation changes during brain activation: The balloon model. Magnetic resonance in medicine. 1998;39(6):855-64.
16. Friston KJ, Mechelli A, Turner R, Price CJ. Nonlinear Responses in fMRI: The Balloon Model, Volterra Kernels, and Other Hemodynamics. NeuroImage (Orlando, Fla). 2000;12(4):466-77.
17. HÄMÄLÄInen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. Reviews of modern physics. 1993;65(2):413-97.

18. Sarvas J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Physics in medicine & biology. 1987;32(1):11-22.
19. Mosher JC, Leahy RM, Lewis PS. EEG and MEG: forward solutions for inverse methods. IEEE transactions on biomedical engineering. 1999;46(3):245-59.
20. Nunez PL, Srinivasan R. Electric Fields of the Brain: The neurophysics of EEG. New York: Oxford University Press; 2006.
21. Haueisen J, Ramon C, Eiselt M, Brauer H, Nowak H. Influence of tissue resistivities on neuromagnetic fields and electric potentials studied with a finite element model of the head. IEEE transactions on biomedical engineering. 1997;44(8):727-35.
22. Baillet S, Mosher JC, Leahy RM. Electromagnetic brain mapping. IEEE signal processing magazine. 2001;18(6):14-30.
23. HÄMÄLÄInen MS, Ilmoniemi RJ. Interpreting magnetic fields on the brain : minimum norm estimates. Medical & biological engineering & computing. 1994;32(1):35-42.
24. Fuchs M, Wagner M, Köhler T, Wischmann H-A. Linear and Nonlinear Current Density Reconstructions. Journal of clinical neurophysiology. 1999;16(3):267-95.
25. Gramfort A, Papadopoulo T, Olivi E, Clerc M. Forward Field Computation with OpenMEEG. Computational Intelligence and Neuroscience. 2011;2011:923703-13.
26. Grech R, Cassar T, Muscat J, Camilleri KP, Fabri SG, Zervakis M, et al. Review on solving the inverse problem in EEG source analysis. Journal of neuroengineering and rehabilitation. 2008;5(1):25-.
27. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. Third edition. ed. Milton: Chapman and Hall/CRC; 2013.
28. Information Theory, Inference, and Learning Algorithms. IEEE transactions on information theory. 2004;50(10):2544-5.
29. Lumer ED, Friston KJ, Rees G. Neural Correlates of Perceptual Rivalry in the Human Brain. Science (American Association for the Advancement of Science). 1998;280(5371):1930-4.
30. Mechelli A, Price CJ, Friston KJ, Ishai A. Where Bottom-up Meets Top-down: Neuronal Interactions during Perception and Imagery. Cerebral cortex (New York, NY 1991). 2004;14(11):1256-65.
31. Grefkes C, Eickhoff SB, Nowak DA, Dafotakis M, Fink GR. Dynamic intra- and interhemispheric interactions during unilateral and bilateral hand movements assessed with fMRI and DCM. NeuroImage (Orlando, Fla). 2008;41(4):1382-94.
32. Shipp S, Adams RA, Friston KJ. Reflections on agranular architecture: predictive coding in the motor cortex. Trends in neurosciences (Regular ed). 2013;36(12):706-16.
33. David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ. Dynamic causal modeling of evoked responses in EEG and MEG. NeuroImage (Orlando, Fla). 2006;30(4):1255-72.
34. Friston K, Mattout J, Trujillo-Barreto N, Ashburner J, Penny W. Variational free energy and the Laplace approximation. NeuroImage (Orlando, Fla). 2007;34(1):220-34.

35. Stephan KE, Penny WD, Moran RJ, den Ouden HEM, Daunizeau J, Friston KJ. Ten simple rules for dynamic causal modeling. NeuroImage (Orlando, Fla). 2010;49(4):3099-109.
36. Friston KJ, Bastos AM, Oswal A, van Wijk B, Richter C, Litvak V. Granger causality revisited. NeuroImage (Orlando, Fla). 2014;101:796-808.
37. Pinotsis DA, Moran RJ, Friston KJ. Dynamic causal modeling with neural fields. NeuroImage (Orlando, Fla). 2012;59(2):1261-74.
38. Zeidman P, Jafarian A, Seghier ML, Litvak V, Cagnan H, Price CJ, et al. A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. NeuroImage (Orlando, Fla). 2019;200:12-25.
39. Rosa MJ, Bestmann S, Harrison L, Penny W. Bayesian model selection maps for group studies. NeuroImage (Orlando, Fla). 2010;49(1):217-24.
40. D. Dima, J. P. Roiser, D. E. Dietrich, C. Bonnemann, H. Lanfermann, H. M. Emrich, et al. Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modelling. Neuroimage. 2009;46(4):1180-6.
41. D. Dima, D. E. Dietrich, W. Dillo, H. M. Emrich. Impaired top-down processes in schizophrenia: a DCM study of ERPs. NeuroImage. NeuroImage. 2010;52(3):824-32.
42. Dima D, Frangou S, Burge L, Braeutigam S, James AC. Abnormal intrinsic and extrinsic connectivity within the magnetic mismatch negativity brain network in schizophrenia: A preliminary study. Schizophrenia Research. 2012;135(1-3):23-7.
43. S. Ranlund, R.A. Adams, A. Diez, M. Constante, A. Dutt, M.H. Hall, et al. Impaired prefrontal synaptic gain in people with psychosis and their relatives during the mismatch negativity. Human brain mapping. 2016;37(1):351-65.
44. Crossley NA, Mechelli A, Fusar-Poli P, Broome MR, Matthiasson P, Johns LC, et al. Superior temporal lobe dysfunction and frontotemporal dysconnectivity in subjects at risk of psychosis and in first-episode psychosis. Hum Brain Mapp. 2009;30(12):4129-37.
45. Deserno L, Sterzer P, Wustenberg T, Heinz A, Schlagenhauf F. Reduced prefrontal-parietal effective connectivity and working memory deficits in schizophrenia. J Neurosci. 2012;32(1):12-20.
46. Schmidt A, Smieskova R, Aston J, Simon A, Allen P, Fusar-Poli P, et al. Brain connectivity abnormalities predating the onset of psychosis: correlation with the effect of medication. JAMA Psychiatry. 2013;70(9):903-12.
47. Bastos-Leite AJ, Ridgway GR, Silveira C, Norton A, Reis S, Friston KJ. Dysconnectivity Within the Default Mode in First-Episode Schizophrenia: A Stochastic Dynamic Causal Modeling Study With Functional Magnetic Resonance Imaging. Schizophrenia Bulletin. 2015;41(1):144-53.
48. Moran RJ, Symmonds M, Stephan KE, Friston KJ, Dolan RJ. An in vivo assay of synaptic function mediating human cognition. Curr Biol. 2011;21(15):1320-5.
49. Gilbert JR, Symmonds M, Hanna MG, Dolan RJ, Friston KJ, Moran RJ. Profiling neuronal ion channelopathies with non-invasive brain imaging and dynamic causal models: Case studies of single gene mutations. Neuroimage. 2016;124(Pt A):43-53.

50. Brodersen KH, Deserno L, Schlagenhauf F, Lin Z, Penny WD, Buhmann JM, et al. Dissecting psychiatric spectrum disorders by generative embedding. Neuroimage Clin. 2014;4:98-111.
51. Schlosser RG, Wagner G, Koch K, Dahnke R, Reichenbach JR, Sauer H. Fronto-cingulate effective connectivity in major depression: a study with fMRI and dynamic causal modeling. Neuroimage. 2008;43(3):645-55.
52. Lu Q, Li H, Luo G, Wang Y, Tang H, Han L, et al. Impaired prefrontal–amygdala effective connectivity is responsible for the dysfunction of emotion process in major depressive disorder: a dynamic causal modeling study on MEG. Neuroscience letters. 2012;523(2):125-30.
53. Desseilles M, Schwartz S, Dang-Vu TT, Sterpenich V, Ansseau M, Maquet P, et al. Depression alters "top-down" visual attention: a dynamic causal modeling comparison between depressed and healthy subjects. Neuroimage. 2011;54(2):1662-8.
54. Wang Y, Chen X, Liu R, Zhang Z, Zhou J, Feng Y, et al. Disrupted effective connectivity of the default, salience and dorsal attention networks in major depressive disorder: a study using spectral dynamic causal modelling of resting-state fMRI. Journal of Psychiatry and Neuroscience. 2022;47(6):E421-E34.
55. D.R. Musgrove, L.E. Eberly, B. Klimes-Dougan, Z. Basgoze, K.M. Thomas, B. A. Mueller, et al. Impaired Bottom-Up Effective Connectivity Between Amygdala and Subgenual Anterior Cingulate Cortex in Unmedicated Adolescents with Major Depression: Results from a Dynamic Causal Modeling Analysis. Brain connectivity. 2015;5(10):608-19.
56. Davey CG, Breakspear M, Pujol J, Harrison BJ. A Brain Model of Disturbed Self-Appraisal in Depression. Am J Psychiatry. 2017;174(9):895-903.
57. Radaelli D, Sferrazza Papa G, Vai B, Poletti S, Smeraldi E, Colombo C, et al. Fronto-limbic disconnection in bipolar disorder. Eur Psychiatry. 2015;30(1):82-8.
58. Breakspear M, Roberts G, Green MJ, Nguyen VT, Frankland A, Levy F, et al. Network dysfunction of emotional and cognitive processes in those at genetic risk of bipolar disorder. Brain. 2015;138(11):3427-39.
59. Dima D, Roberts RE, Frangou S. Connectomic markers of disease expression, genetic risk and resilience in bipolar disorder. Transl Psychiatry. 2016;6(1):e706.
60. Vai B, Poletti S, Radaelli D, Dallaspezia S, Bulgarelli C, Locatelli C, et al. Successful antidepressant chronotherapeutics enhance fronto-limbic neural responses and connectivity in bipolar depression. Psychiatry Res. 2015;233(2):243-53.
61. Almeida JR, Versace A, Mechelli A, Hassel S, Quevedo K, Kupfer DJ, et al. Abnormal amygdala-prefrontal effective connectivity to happy faces differentiates bipolar from major depression. Biol Psychiatry. 2009;66(5):451-9.
62. Diwadkar VA, Bakshi N, Gupta G, Pruitt P, White R, Eickhoff SB. Dysfunction and Dysconnection in Cortical-Striatal Networks during Sustained Attention: Genetic Risk for Schizophrenia or Bipolar Disorder and its Impact on Brain Network Function. Front Psychiatry. 2014;5:50.
63. Wei L, Zhang Y, Zhai W, Wang H, Zhang J, Jin H, et al. Attenuated effective connectivity of large-scale brain networks in children with autism spectrum disorders. Front Neurosci. 2022;16:987248.

64. Wang L, Qin Y, Yang S, Jin D, Zhu Y, Li X, et al. Posterior default mode network is associated with the social performance in male children with autism spectrum disorder: A dynamic causal modeling analysis based on triple-network model. Human Brain Mapping. 2024;45(8).
65. Razi A, Seghier ML, Zhou Y, McColgan P, Zeidman P, Park HJ, et al. Large-scale DCMs for resting-state fMRI. Netw Neurosci. 2017;1(3):222-41.
66. Chumbley JR, Friston KJ, Fearn T, Kiebel SJ. A Metropolis-Hastings algorithm for dynamic causal models. Neuroimage. 2007;38(3):478-87.