Fronto-limbic dysconnectivity leads to impaired brain network controllability in young people with bipolar disorder and those at high genetic risk

Jayson Jeganathan[a*], Alistair Perry[a,b*], Danielle S. Bassett[c,d], Gloria Roberts[b,e], Philip B. Mitchell[b,e], Michael Breakspear[a,f]

[a]Program of Mental Health Research, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

[b]School of Psychiatry, University of New South Wales, Randwick, NSW, Australia

[c]Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104 USA

[d]Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104 USA

[e]Black Dog Institute, Prince of Wales Hospital, Randwick, NSW, Australia

[f]Metro North Mental Health Service, Brisbane, QLD, Australia

* The authors contributed equally

Running title: Impaired brain network controllability

Corresponding author

Jayson Jeganathan, QIMR Berghofer Medical Research Institute, 300 Herston Rd, Herston, Queensland 4029, Australia.

Email: jayson.jeganathan@gmail.com

Abstract

Recent investigations have used diffusion-weighted imaging to reveal disturbances in the neurocircuitry that underlies cognitive-emotional control in bipolar disorder (BD) and those at high genetic risk (HR). It has been difficult to quantify how structural changes disrupt the superimposed brain dynamics, leading to the emotional lability that is characteristic of BD. Average controllability is a concept from network control theory that estimates the manner in which local neuronal fluctuations spread from a node or subnetwork to alter the state of the rest of the brain. We used this theory to ask whether structural connectivity deficits previously observed in HR (n=84) individuals, patients with BD (n=38), and age- and-gender-matched controls (n=96) translate to differences in the ability of brain systems to be manipulated between states. Localized impairments in network controllability were seen in the left parahippocampal, middle occipital, superior frontal, and right inferior frontal and precentral gyri in BD and HR groups. Subjects with BD had distributed deficits in a subnetwork containing the left superior and inferior frontal gyri, postcentral gyrus, and insula (p=0.004). HR participants had controllability deficits in a right-lateralized subnetwork involving connections between the dorsomedial and ventrolateral prefrontal cortex, the superior temporal pole, putamen, and caudate nucleus (p=0.008). Between-group controllability differences were attenuated after removal of topological factors by network randomization. Some previously reported differences in network connectivity were not associated with controllability-differences, likely reflecting the contribution of more complex brain network properties. These analyses highlight the potential functional consequences of altered brain networks in BD, and may guide future clinical interventions.

Highlights

- Control theory estimates how neuronal fluctuations spread from local networks.
- We compare brain controllability in bipolar disorder and their high-risk relatives.
- These groups have impaired controllability in networks supporting cognitive and emotional control.
- Weaker connectivity as well as topological alterations contribute to these changes.

Keywords

Bipolar disorder, controllability, brain network, cognition, emotion, genetic risk

Abbreviations

CN = control; HR = high-risk; BD = bipolar disorder; MRI = magnetic resonance imaging; fMRI= functional MRI; dMRI = diffusion MRI.

1. Introduction

For much of the 20th century, neuroscience was predicated on the notion that individual cognitive functions could be attributed to segregated regions of the brain. In recent times, this paradigm has shifted toward a connectivity-based approach which emphasizes the crucial role of network-mediated functional integration (Sporns, 2013). In parallel, the clinical neurosciences have shifted from a predominantly lesion-based approach towards a connectomic framework (Fornito *et al.*, 2015). However, disrupted network connections in the brain impact not only upon directly connected regions, but can also influence distant cortical regions through complex dynamics (Crofts *et al.*, 2011). To characterize the spread of dysfunction in disease states, network connectivity and the superimposed dynamics of brain activity must both be considered (Stam, 2014). In human brains, the former is now readily accessible through advances in diffusion-weighted imaging and tractography (Jbabdi *et al.*, 2015; Farquharson and Tournier, 2016). However, the principles of neuronal dynamics, well-known at the neuron level, remain incompletely understood at the macroscopic scale (Breakspear, 2017).

Studies using diffusion-weighted imaging have identified 'disconnection syndromes' in schizophrenia (Stephan *et al.*, 2009; Zalesky *et al.*, 2011), depression (Bai *et al.*, 2012), attention-deficit hyperactivity disorder (Cao *et al.*, 2013), and epilepsy (Widjaja *et al.*, 2015). More recently, attention has turned to bipolar disorder (BD), a psychiatric condition characterized by episodic disturbances in mood and cognition. While variations in mood around the set point of euthymia are intrinsic to human experience, excursions of mood in BD reach a magnitude and duration that produces substantial distress, dysfunction, and disability. Recent structural imaging studies of BD have consistently reported white matter alterations in circuits involving prefrontal, striatal, and limbic regions (Xekardaki *et al.*, 2011; Nortje *et al.*, 2013; Phillips and Swartz, 2014; Roberts *et al.*, 2016). The disease also has a strong genetic component. Unaffected first-degree relatives often have attenuated affective disturbances and are at significantly increased risk of developing the disorder; their odds-ratio is estimated to lie between 7 and 14 (McGuffin *et al.*, 2003; Mortensen *et al.*, 2003; Purcell *et al.*, 2009; Perich *et al.*, 2015). Identification of structural and functional connectivity differences between BD patients, their unaffected relatives, and healthy controls, have permitted a dissociation of genetic, illness-expression, and adaptive influences (Frangou, 2009; Kempton *et al.*, 2009; Pompei *et al.*, 2011; Sprooten *et al.*, 2011; Meda *et al.*, 2012; Doucet *et al.*,

2017; Ganzola *et al.*, 2017) .The neurodevelopmental risk of their first-degree relatives is supported by effective and functional connectivity studies that reveal alterations in fronto-limbic networks supporting emotion processing and regulation (Pompei *et al.*, 2011; Frangou, 2012; Breakspear *et al.*, 2015; Dima *et al.*, 2016). We recently investigated structural network disturbances in both BD patients and youth at high genetic risk for the disorder (Roberts *et al.*, 2016). The high-risk (HR) group demonstrated unique disturbances in subnetworks of connections centering upon the inferior frontal gyri and insular cortex. These findings speak to an endogenous risk, present as these subjects transition through a critical developmental period.

Connectivity deficits on their own, without an informed understanding of network topology or the overlying dynamics comprise 'correlation' rather than 'causation'. Relating network disturbances in psychiatric conditions to the illness phenotype is an important, yet unfulfilled ambition. In this paper, we employ network control theory to study the link between the structural network changes previously observed (Roberts *et al.*, 2016) and the affective dysregulation seen in BD and, in muted form, their unaffected first-degree relatives. The integration of network control theory with structural connectomic data has recently been used to predict brain dynamics and stability (Gu *et al.*, 2015; Betzel *et al.*, 2016). As neuronal activity evolves in time, functional brain states undergo transitions, traversing a path through a dynamic state-space landscape. Perturbations applied to a set of control nodes, either from an extrinsic source or from internal dynamics, can modulate these trajectories (Bassett and Khambhati, 2017). This input energy depends on the choice of control nodes and the strength of structural connections. Recent work has found intriguing links between the average controllability of different brain systems and their contribution to cognitive function (Betzel *et al.*, 2016; Muldoon *et al.*, 2016; Tang *et al.*, 2017).

In general, high brain network controllability appears to reside in those cortical regions typically associated with executive function, self-monitoring and emotional control including the default, limbic and cognitive control systems (Gu *et al.*, 2015). The notion that dysregulation of affect in BD could arise from the structural disturbances impacting upon the controllability of these systems has intuitive appeal. We sought to test that hypothesis here. We first characterize the network underpinnings of average controllability, and then apply controllability analyses to brain regions and subnetworks where changes in bipolar and high-risk cohorts were previously reported (Roberts *et al.*, 2016). We hypothesized that the

structural dysconnectivity of these subnetworks and regions would be associated with impairments in wider network controllability, particularly in the neurocircuitry that underlies cognitive-emotional control.

2. Materials and methods

2.1. Participants

218 participants comprise three richly-phenotyped groups: (i) 84 participants at high risk (HR) of bipolar disorder (mean age 22.4), (ii) 96 controls (CNs) without a family history of mental illness (mean age 22.6), and (iii) 38 bipolar disorder (BD) patients (mean age 23.9) (Roberts *et al.*, 2016). HR and BD participants were recruited from families who had previously participated in a BD pedigree molecular genetics study or a specialised BD research clinic, as well as from clinicians, mental health consumer organisations, and other forms of publicity. CN subjects were recruited via advertisements in print, electronically, and noticeboards in universities and local communities. Demographic and clinical data are provided in Supplementary Table 1. Further description of the sample ascertainment and clinical assessments for the current population are provided in (Perich et al., 2015).

The study was conducted with approval from the University of New South Wales Human Research Ethics Committee (HREC Protocol 09/097) and the South Eastern Sydney Illawarra Health Service HREC (Protocol 09/104). Written informed consent was obtained from all participants.

2.2. Construction of structural networks

Diffusion-weighted images were obtained using a 3T Philips Achieva X MRI scanner and analyzed as previously described (Roberts *et al.*, 2016). In short, pre-processed data underwent spherical deconvolution followed by probabilistic tractography to generate 5 million streamlines. The AAL template (Tzourio-Mazoyer *et al.*, 2002) was subdivided into 512 regions of uniform volume (Zalesky *et al.*, 2010b) (see https://github.com/AlistairPerry/CCA). Weighted structural networks were produced by combining each subject's tractography with their parcellation template (Roberts *et al.*, 2016). Edge weights represent the number of streamlines connecting two parcels. Structural networks were thresholded with a connection density of 10%, as investigations typically

analyze brain networks with densities centering upon this value (Sporns, 2013; Perry *et al.*, 2015). Analyses of 5%, 15% and maximum densities are provided in Supplementary Table 4, 5, and 6.

2.3. Network control theory

Linear control theory quantifies transient responses to brief perturbations of a linearly stable system. In this study, networks were overlaid with linear discrete-time and time-invariant dynamics evolving according to,

$$x(t+1) = Ax(t) + Bu(t), \quad (1)$$

where $x(t)$ is a vector denoting the simulated state of all nodes i at time t. The matrix A is the structural connectivity matrix, with element A_{ij} representing the number of white matter streamlines connecting regions (or *nodes*) i and j. The input matrix B specifies the control nodes in the brain (those receiving the perturbation),

$$B = [e_1\ e_2\ \dots e_M],$$

where e_i is a vector with '1' in the i^{th} position, and represents the i^{th} control node. The variable $u(t)$ is the energy applied to the set of input nodes B at time t.

The eigenvector states $v(t)$ of each network are derived from this equation by linear algebra. These are patterns of activation satisfying the relation $x(t+1) = Av(t) = \mu v(t)$, where the eigenvalue μ is a constant. Once an eigenvector state is initiated, the relative activation of different nodes within the network will remain unchanged. However, the magnitude of $v(t)$ will be amplified ($\mu > 1$) or damped ($\mu < 1$) with time.

Controllability analyses are predicated upon linear stability of all individual subject networks, which was ensured by dividing all edge weights by a number 10% greater than the largest eigenvalue among all subjects. We chose to rescale all subjects by a common number rather than normalize each subject separately to preserve individual and group differences in mean weighted connectivity. We chose 10% to prevent unstable dynamics in the most strongly-

connected (least stable) subject. Results were consistent across greater degrees of network stabilization (Supplementary Table 7).

This linear system is a simplification of the full nonlinear neural dynamics that would more realistically embody the behavior of large-scale brain dynamics (Breakspear, 2017). However, linear controllability of a structural network is usually sufficient to imply controllability of nonlinear dynamics overlaid on the same structure, because linear dynamics are accurate around the linearization point (local, stable equilibria) (Deco *et al.*, 2008). Our choice of discrete-time dynamics is based on prior work showing that controllability is similar in discrete and continuous-time settings (Gu *et al.*, 2015).

2.4. Network controllability

The energy required to reach a target state is obtained by summing energy input over time. Let the energy of the control input be $\mathrm{E(u)} = \sum_{t=0}^{N-1} \|u(t)\|^2$, the sum of the energy applied to the input nodes. The minimum energy control input that takes the network from the zero state $x(0) = 0$ to the target state $x(N) = x_f$ is,

$$u^*(t) = B^T(A^T)^{N-t-1} W_K^{-1} x_f, \quad (2)$$

where $W = \sum_{t=0}^{\infty} A^t B B^T A^t$ refers to the controllability Gramian.

One can show that $E(u^*) \leq 1/\mu_{min}$, where μ_{min} is the smallest eigenvalue of W. Therefore, μ_{min} sets the energy requirement of the most difficult-to-reach state.

Average controllability for a set of control nodes is the average energy needed to steer the system to any target state in finite time (Fig. 1). This is directly proportional to Trace(W^{-1}). We use Trace(W), the mean impulse response energy, as the measure of average controllability because Trace(W) and Trace(W^{-1}) are highly correlated due to a relation of inverse proportionality, and Trace(W^{-1}) is ill-defined. When the control set is a single node i, B is just the input vector e_i to the i^{th} control node, and Trace(W), the average controllability of node i, simplifies to,

$$\mathrm{Trace}(W) = \sum_j \frac{v_{ij}}{1-\mu_j^2}, \quad (3)$$

where i is a node, v_{ij} is the i^{th} component of the j^{th} eigenvector, and μ_j is the j^{th} eigenvalue. The variable v_{ij} can be considered as the influence of the i^{th} node on the j^{th} eigenmode. In the rest of this paper, we refer to average controllability simply as 'controllability' for brevity, noting that there are other independent controllability metrics which have been studied (Wu-Yan *et al.*, 2017).

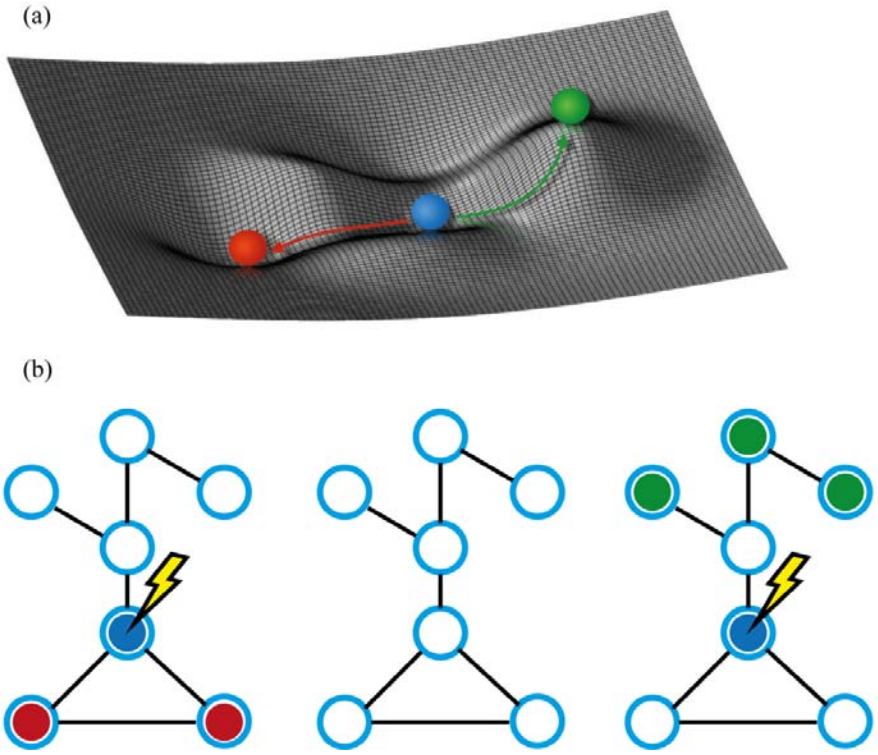

Figure 1. Average controllability. (a) Schematic representation of the mean impulse response energy required to steer a dynamic network from its base "resting" state (blue ball) to either an easy-to-reach state (red) or a hard-to-reach state (green). (b) Schematic representation of corresponding changes from the baseline state (centre) to an easy-to-reach new network state (left, red) or a hard-to-reach state (right, green). Arrow depicts the node receiving the perturbation.

2.5. Contributors to average controllability

Nodal strength is the weighted sum of each node's connection to all other nodes to which it is connected, and has an important influence on controllability (Gu *et al.*, 2015). However, other network features also influence controllability, such as the distribution of paths of different lengths as measured by the communicability matrix (Betzel *et al.*, 2016; Gu *et al.*, 2017).

Some topological factors, such as scale-freeness, are captured in the distribution of node strengths, while others, such as clustering, can occur independently of the strength distribution.

To examine the impact of strength distribution on variability in controllability for each node, that node's controllability was calculated before and after strength-preserving randomization (Maslov and Sneppen, 2002). If these values are perfectly correlated, then strength accounts for all of that node's controllability: Lower correlations imply that other network features influence that node's controllability. Topological network features can contribute to between-subjects and between-nodes variability in controllability. We assessed the presence and strength of such relations separately by correlating controllability values across subjects or across nodes (see Supplementary material for more detail). The coefficient of determination was calculated as the square of the Spearman rank correlation coefficient.

We also examined how other widely adopted nodal measures of network topology contribute to average controllability. To identify which of these measures predict controllability, stepwise linear regression was implemented using entry and removal probabilities of 0.05 and 0.10, respectively. The nodal measures that were entered in the regression model were node strength, local clustering coefficient, subgraph centrality, local efficiency, betweenness centrality, within-module degree *z*-score, and participation coefficient. The latter two measures were calculated based upon the modular decomposition of the group average network from the healthy control participants, detected with the Louvain algorithm for maximization of the modularity quality function (Rubinov and Sporns, 2010). The nodal values for each node were averaged across all CN participants before the stepwise regression was performed.

2.6. Controllability at different levels of network granularity

We calculated group differences in average controllability metrics at three levels of network granularity: single nodes, subnetworks of nodes, and communities derived from intrinsic functional connectivity patterns.

2.6.1. Node level

Our previous analyses of structural connectivity identified seven brain regions with differences in nodal strength between CN, HR and BD groups (Roberts *et al.*, 2016). Here,

we calculated the average controllability of these seven nodes; that is, the ability to control whole brain dynamics through a stimulus to each of these individual nodes.

Group differences in average controllability were elicited with a one-way ANOVA for all seven nodes with a false discovery rate (FDR) correction. We then performed a one-tailed *t*-test of surviving nodes using previously identified group contrasts, and a final FDR correction step. The Benjamini-Hochberg procedure ($\alpha = 0.05$) was used for FDR correction.

2.6.2. Subnetwork level

We previously analyzed group effects in these data using network-based statistic (NBS), a permutation-based method which employs topological inference to control for family-wise error when identifying subnetworks with group-wise differences in connection strength (Zalesky *et al.*, 2010a). Connectivity differences between CN, HR and BD were observed in four subnetworks (Roberts *et al.*, 2016) (Fig. 2). In this study, we calculated the controllability of all nodes within these subnetworks. Subnetwork average controllability is the sum of average controllabilities for all nodes in the subnetwork. Intuitively, this measures the impulse response to a stimulus applied across the subnetwork. This impulse response depends on (i) the subnetwork's external connectivity to surrounding regions, and (ii) the internal connections within the subnetwork, which amplify the signal by internal excitatory feedback (Fig. 3). Because these networks were discovered by the NBS, they have, by definition, weaker internal connectivity but relatively preserved external connectivity. Hence, subnetwork average controllability here is a surrogate for abnormal internal signal amplification.

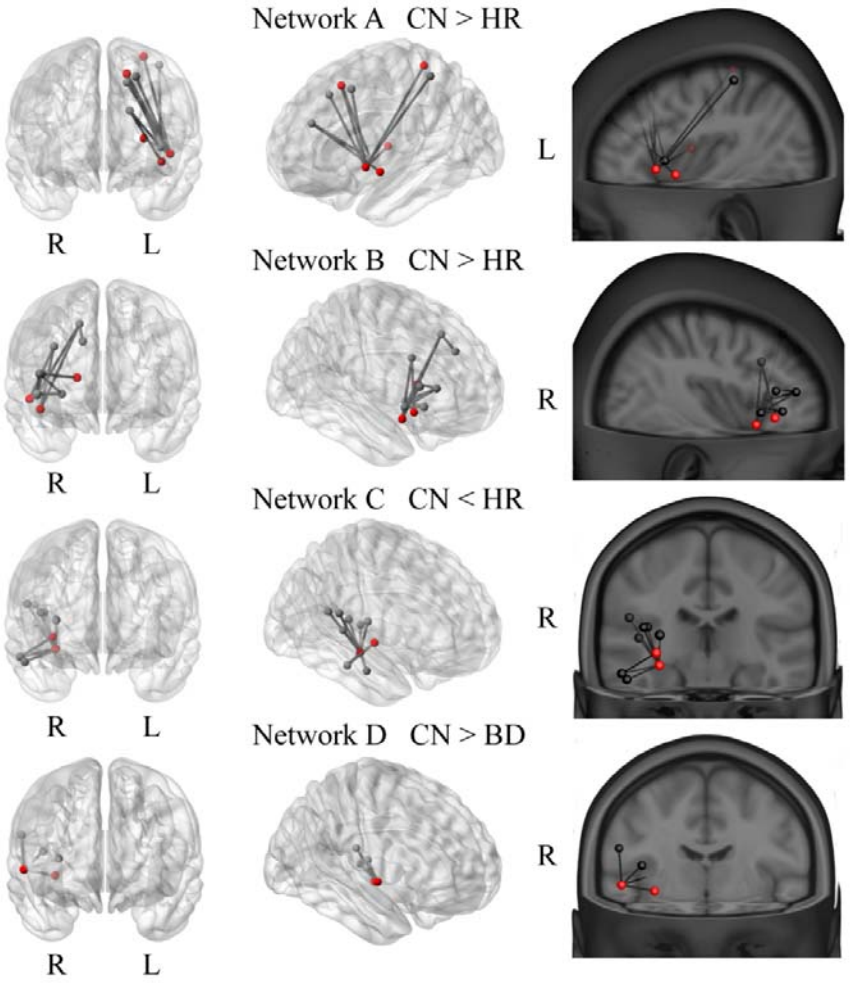

Figure 2. Subnetworks showing significant between-group differences as identified by the network-based statistic in the previous investigation (Roberts *et al.*, 2016). Connections (lines) between nodes (circles) with significant group differences in streamline count, with nodes coloured according to their previously identified node degree. Red circles indicate high-degree hub regions (top 15%), while grey circles represent non-hubs. CN, controls; HR, high risk; BD, bipolar disorder; L, left; R, right; α, azimuth. Figure adapted with permission from (Roberts *et al.*, 2016)

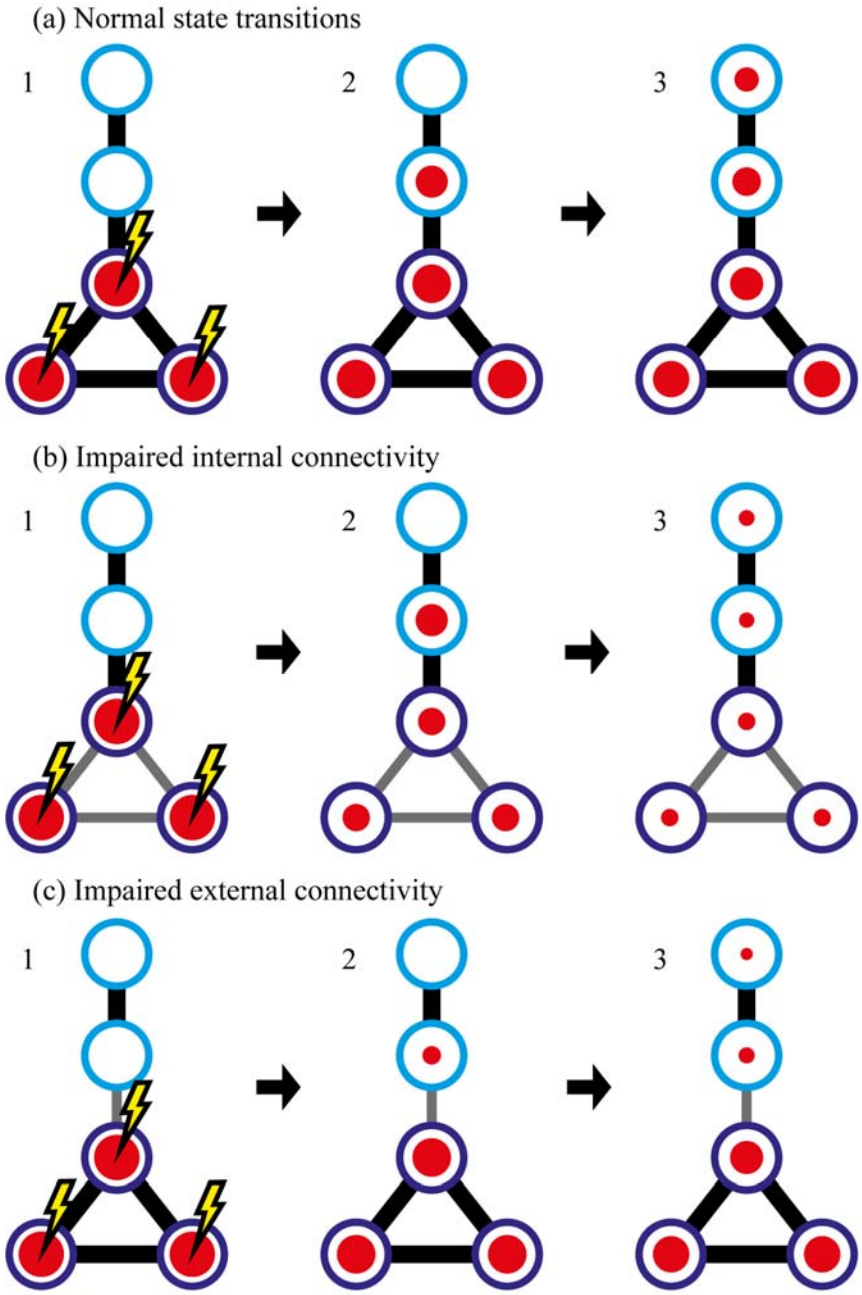

Figure 3. Schematic of signal spread from a closed loop of three nodes (purple circles in left panels). The size of the circle indicates the corresponding magnitude of activation across the subnetwork. Attenuated intranetwork (b) or outgoing connections (c) (grey lines) can diminish the impulse response.

To ascertain the influence of network topology on group differences in subnetwork controllability, the subnetwork analysis was repeated after strength-preserving randomization. If the group differences were preserved, then these differences could be attributed primarily

to differences in the strength of the nodes within each subnetwork. Changes in group differences following strength-preserving randomization reflect contributions from higher order topological features.

The following statistical protocol was used: A one-way ANOVA of group differences in each subnetwork's average controllability, then FDR correction of the four resulting *p*-values. Two-tailed *t*-tests were then performed for surviving binary contrasts followed by a final FDR correction step for the individual possible group effects.

2.6.3. Community level

Nodal-regions were assigned to widely adopted functional network affiliations (Yeo *et al.*, 2011). The community affiliation of a brain-region was identified by its best spatial fit within one of the seven functional networks (Perry *et al.*, 2017) (see https://github.com/AlistairPerry/CCA). As with the subnetwork controllability, the average controllability of a community is the sum of controllabilities for all nodes within that community. A one-way ANOVA was first used to test for group-wise differences in the controllability of each community, followed by FDR-correction.

3. Results

3.1. Contributors to controllability

The nodal strength and average controllability of each node were positively correlated (R=0.89; p<0.0001, see Supplementary material for a formal derivation). This association is consistent with previous analyses of other structural connectomic data (Gu *et al.*, 2015). Regions of high average controllability, which shift the brain easily between functional regimes, are topologically well-connected regions.

The brain network factors underlying average controllability were probed with a randomization scheme that preserved node-strength distribution within the network but destroyed higher-order topological properties. Strength distribution accounted for 87% of between-node variability in average controllability within a subject and 68% of between-subject variability. The remainder can hence be attributed to other topological factors. Figure

4 demonstrates shrinkage in the mean and variance of subject-mean controllabilities after network randomization.

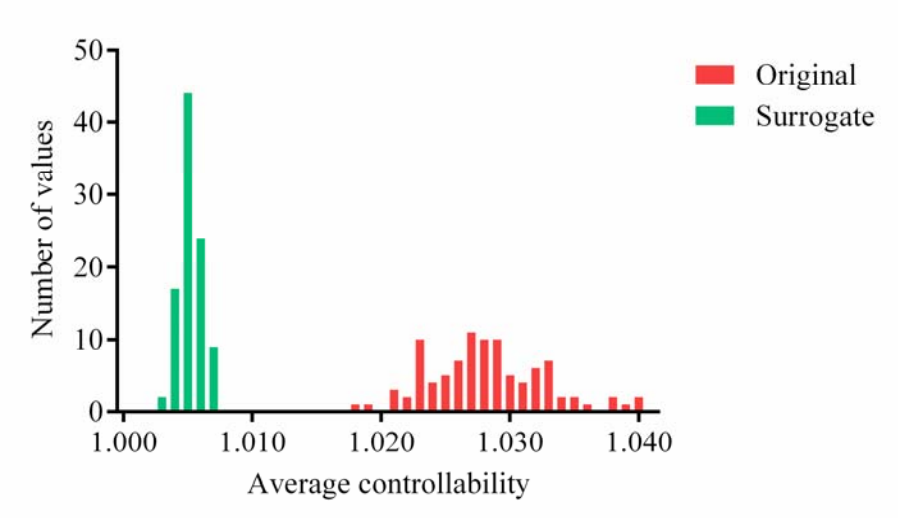

Figure 4. Distribution of average controllabilities averaged across nodes in the original data (green) and in the random surrogate networks (red).

To quantify the contribution of topological factors to variability in controllability, stepwise linear regression was employed to model the putative contributions of various nodal measures (Table 1). As expected, nodal strength was the strongest predictor of average nodal controllability. However, other network features also make substantial contributions; the next measure most strongly predictive of controllability was subgraph centrality, followed by the clustering coefficient. Subgraph centrality is defined by the weighted sum of closed walks of different path lengths starting and ending at the nodal region (Estrada and Rodriguez-Velazquez, 2005). Clustering reflects the number of closed cycles with three edges. Average controllability, the impulse response to a stimulus, is hence magnified by positive feedback via closed walks and loops (Supplementary Figure 1).

Table 1. Stepwise linear regression of average controllability

Variable	Standardized beta	Full correlation	Partial correlation
Node strength	0.648**	0.886	0.587
Subgraph centrality	0.297**	0.608	0.600
Clustering coefficient	0.152**	0.388	0.387
Within-module z-score	0.098*	0.793	0.114

Adjusted $R^2 = 0.887$.

Excluded variables: local efficiency, betweenness centrality, participation coefficient

* $p<0.01$, ** $p<0.001$

3.2. Controllability in high-risk and bipolar disorder subjects

Having established that node strength, subgraph centrality, clustering coefficient, and within-module *z*-score contribute to network controllability in our data, we next studied between-group differences in controllability at the three levels of granularity.

3.2.1. Node level

Six out of seven regions previously identified as having reduced node strength in HR and BD groups also show significant between-group controllability differences. Reductions in average controllability were observed in the left insula, left parahippocampal gyrus, left middle occipital gyrus, left superior frontal gyrus, right inferior frontal gyrus / pars triangularis, and right precentral gyrus (Supplementary Table 2).

3.2.2. Subnetwork level

We next studied average controllability differences in four lateralized subnetworks (Fig. 2, Supplementary Table 3). Significant average controllability differences ($p<0.05$, corrected) were seen in subnetworks A and B, but not C or D (Fig. 5). Controllability of network A, including the left superior/inferior frontal gyri, postcentral gyrus, insula, and pars triangularis (Fig. 1A), was reduced in BD compared to CN ($p<0.004$, $\eta^2 = 0.043$). Average controllability of network B, involving the right superior/middle/inferior frontal gyri, superior temporal pole, putamen, caudate, pars triangularis and pars orbitalis (Fig. 1B) was reduced in HR subjects compared to CNs. ($p<0.008$, $\eta^2 = 0.037$). Network C showed no significant group differences ($p=0.544$). Average controllability differences approached significance (p=0.059) between CN and the BD group in subnetwork D.

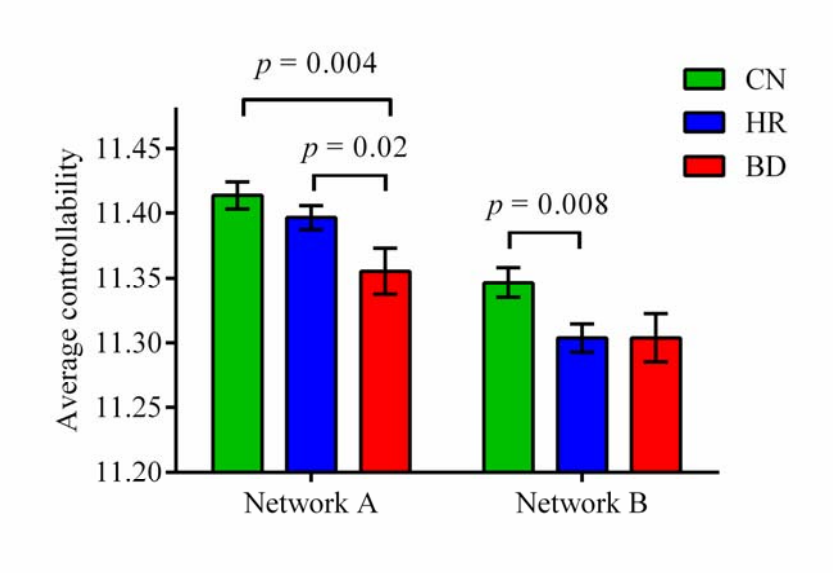

Figure 5. Mean average controllability for subnetworks with identified significant group differences. Error bars show standard error of the mean. CN, controls; HR, high-risk; BD, bipolar disorder.

Group differences in subnetworks A and B were reduced after strength-preserving randomization, highlighting the role of network topology (Table 2).

Table 2. Differences between cohort-mean controllabilities for each group before and after strength-preserving randomization

	Original	**Surrogate**
Network A: CN-BD	0.059	0.012
Network A: HR-BD	0.041	0.008
Network B: CN-HR	0.043	0.008

3.2.3. Community level

We finally addressed average controllability at the scale of functional brain communities whose specialized functional roles have been previously documented (Yeo *et al.*, 2011). There were substantial differences in controllability across these communities, with the limbic and frontoparietal modules showing the highest per-node average controllability (Fig. 6). These differences reflected, but were not strictly enforced by, differences in average node strength. For example, the limbic and frontoparietal communities switch relative positions of strength *versus* controllability. However, no group-wise differences in average controllability of these communities survived FDR correction.

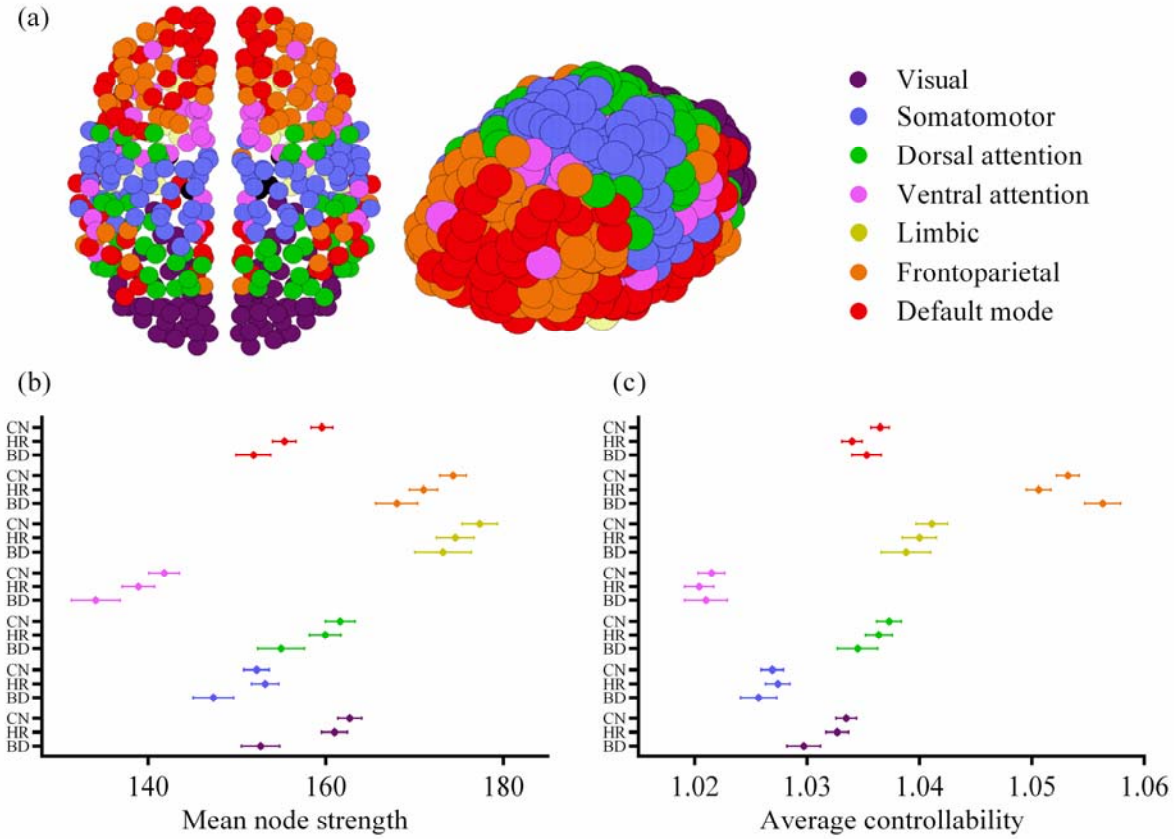

Figure 6. Average controllability of functional communities across the three groups. (a) Corresponding affiliation of each node to the functional communities derived from (Yeo *et al.*, 2011). (b) Mean node strength, and (c) average controllability (normalized by community size) in each group and each functional module. Errors bars show standard error of the mean.

CN, controls; HR, high-risk; BD, bipolar disorder.

3.2.4. Auxiliary analyses

We observed a significant relationship (after FDR-correction) between age and controllability in the CN group (r=0.291, p=0.004). However, for the group contrasts of interest, there were no significant age x group interactions for average controllability (Network A CN vs BD p=0.07; Network A HR vs BD p=0.09; Network B CN vs HR p=0.93; Supplementary Figure 2). The occurrence of a major depressive episode in HR individuals may represent an important early stage in the development of BD (Perich *et al.*, 2015). In this vein we checked for a potential association. Notably, the average controllability for subnetworks A and B in

HR individuals was not significantly modulated by a history of anxiety or major depression (Supplementary Table 10).

Measures of illness severity in the BD group (illness duration and total number of mood episodes) were not significantly associated with the average controllability of the two subnetworks (Supplementary Table 11).

4. Discussion

Quantifying brain structural network disturbances in psychiatric conditions provides novel insights into their neurobiological correlates, yet lacks a direct explanatory link to the illness phenotype. We previously observed structural network changes in those with, or at high genetic risk of BD (Roberts *et al.*, 2016). Here we used network control theory to address the theoretical gap between these structural network disturbances and the dysregulation of cognitive-emotional function that characterizes BD and those at HR. Specifically, we asked whether focal and distributed network disturbances observed in HR and BD subjects translated to impaired brain controllability. Controllability deficits were seen in six key brain regions and also in two subnetworks: a left-sided network involving the insula and postcentral gyrus in BD, and in a right-lateralized PFC-striatal subnetwork in HR. There were no group differences in controllability in seven widely studied functional communities. Nodal strength thus provides an important, but not complete account of controllability. Likewise, group differences in network strength do not fully account for group differences in controllability. We now consider the conceptual and clinical implications of these findings.

The average controllability quantifies the capacity of input to a subnetwork to manipulate the transition of broader brain states (Gu *et al.*, 2015). It reflects the network-wide impulse response to a stimulus applied to a node or subnetwork of interest (Muldoon *et al.*, 2016). This interpretation is instructive, as signal spread depends on (i) amplification by positive feedback within the subnetwork, and (ii) links between the subnetwork and its surrounds. We found that the top three predictors of average controllability in our data were node strength, subgraph centrality, and local clustering coefficient, highlighting the role of closed motifs such as triangles and larger loops in amplifying positive feedback, in addition to the role of strong connections in propagating that energy outwards. Topological factors such as these amplifying motifs were responsible for 32% of the between-subject variability in

controllability. Removal of these topological factors by network randomization reduced between-subject and between-group differences. Hence the present work builds upon our previous observation of between-group subnetwork strength differences (Roberts *et al.*, 2016) by showing how these, as well as other network features, translate into group differences in subnetwork average controllability.

Impaired controllability was seen in six brain regions and two lateralized networks. Although Network A (involving the left superior and inferior frontal gyri, insula, and postcentral gyrus) had reduced connectivity in HR compared to CN, it showed impaired controllability exclusively in the BD group. NBS finds subnetworks by comparing mean edge weights across groups, but is insensitive to topological group differences which would also influence controllability. Such effects are also consistent with the loss of significant group differences with strength-preserving randomization. The regions and connections of Network A overlap with previous demonstrations of changes in fractional anisotropy for tracts involving the left superior frontal gyrus (Adler *et al.*, 2006), decreased effective connectivity between the left dorsolateral prefrontal cortex (PFC) and the left inferior frontal gyrus (Breakspear *et al.*, 2015), hypo-activation in the ventrolateral PFC during emotionally salient tasks (Chen *et al.*, 2011; Phillips and Swartz, 2014), and altered reward processing-related activity in the left striatum and left ventrolateral PFC (Bermpohl *et al.*, 2010; Nusslock *et al.*, 2012). Altered connectivity between these frontal areas and the insula, a region involved in interoceptive perceptions (Critchley *et al.*, 2004), may mediate interplay between dysfunctional cognition and emotional homeostasis in BD. Impaired signal amplification from these circuits could hence produce the affective dysregulation that characterizes BD.

We also observed reduced controllability in Network B (which encompassed the right frontal gyri, superior temporal pole, putamen, and caudate nucleus), in this case, for HR compared to CN. Existing research has also noted structural and functional changes in the right frontal white matter (Bruno *et al.*, 2008), right ventrolateral PFC (Hajek *et al.*, 2013), putamen and caudate (Haller *et al.*, 2011), and right amygdala. (Torrisi *et al.*, 2013). PFC dysregulation has been a central theme in the psychiatric neuroimaging literature. This region is involved in executive function, cognition, planning, and reward. Attenuated controllability of this right-lateralized frontostriatal reward loop, seen in HR subjects alone, may reflect these young HR participants' vulnerability to the manic phenotype - impulsivity, reduced attention span, and increased goal-oriented activity (Lau *et al.*, 2017). Longitudinal follow-up will allow us to

test the validity of these inferences by seeing if individual differences in controllability confer increased risk of later conversion to BD.

Although we also selected Networks C and D for further analyses, based on our prior observation of group differences in intra-network connectivity (Roberts *et al.*, 2016), we did not observe group differences in average controllability. Therefore, network strength differences alone are not sufficient to confer differences in controllability.

The different network controllability effects in the HR group compared to BD are intriguing. Controllability deficits that are specific to BD may reflect a purer form of illness risk and expression than those in HR, which rather reflect a mix of risk and resilience but not (yet) BD expression. Effects in BD alone may also be state markers of medication exposure or the consequences of recurrent mood episodes. The left-hemisphere disturbances seen in BD alone may reflect the left-lateralization of autonomic processes in fronto-temporal areas (Guo *et al.*, 2016). Moreover, the edges in network A are more distributed across the cortex – potentially allowing propagation of fluctuations to cognitive, emotional, interoceptive, and somatosensory areas.

Several caveats of this study need to be considered. Interpreting subnetwork controllability as a surrogate for signal amplification assumes that external edges are preserved. Edges that are outside a subnetwork discovered through statistical testing may be weakened but with insufficient magnitude to be included in the final supra-threshold subnetwork. Therefore, average controllability should be conservatively interpreted as a combination of subnetwork outflows and internal signal amplification. Selective analysis of controllability relies substantially, but not solely, on strength differences in the subnetworks identified previously (Roberts *et al.*, 2016). However, we observed that there was a strong influence of topological features, which contribute to 32% of inter-subject variability. Subgraph centrality and the clustering coefficient were included as predictors in the regression model independent of their collinearity with node strength, demonstrating the influence of complex network motifs such as closed walks and loops. Therefore, the present analyses provide additional insights not provided through interrogation of edge strength alone.

Psychiatric comorbidity in the HR group is a potential confounder, with 8% of this group suffering recurrent major depressive episodes. However, controllability in the HR group did

not co-vary with comorbid anxiety or depressive episodes, illness duration or total number of mood episodes. It is also important to note that prior depression or anxiety was not an exclusion factor for our control cohort, ensuring that a proportion of these participants also had some psychiatric comorbidity (and hence avoiding a "super-healthy" cohort). Controllability hence appears to speak to trait rather than state differences between our three cohorts. Again, longitudinal follow-up will be useful in identifying the controllability precedents to the development of manic symptoms in the HR individuals whom do transition to BD.

Finally, controllability is predicated upon the reduction of a complex nonlinear system to the simple linear model given in Equation (1). Several points pertain to this simplification. First, although fine-grained, fast dynamics bear the fingerprints of highly nonlinear dynamics such as multistability (Freyer *et al.*, 2011), linear models have been found to predict much of the slow fluctuations and time averaged variance in empirical recordings (Honey *et al.*, 2009; Deco *et al.*, 2013). Second, complex nonlinear dynamics organize around simple underlying attractors, including fixed points (Cocchi *et al.*, 2017). Linear models provide a sufficient approximation for the behavior of such dynamics in the neighborhood of such fixed points (Hirsch *et al.*, 2013). Linear theory, therefore, arguably provides a unique, albeit incomplete, account of brain network controllability. Future work could harness the potential of nonlinear control theory, although the challenges in doing so remain substantial (Slotine and Li, 1991; Cornelius *et al.*, 2013).

Quantifying structural network differences in high risk and psychiatric cohorts is an important step in developing novel diagnostic markers for risk and resilience. However, computational frameworks are required to understand the link between these illness correlates and the phenotype (Fornito *et al.*, 2015). In this vein, control theory can lend early insights into distinguishing between state markers of disease, trait markers, and even adaptive changes that may be harnessed in future clinical interventions. Network B, a putative frontostriatal reward loop, had attenuated controllability in the HR group alone. Longitudinal follow-up of the HR group is currently underway to assess whether network changes are protective adaptations or reflect trait risk to the disorder. The absence of controllability deficits in Networks C and D, despite edge strength deficits, suggests broader compensatory topological changes: This conjecture also suggests developmental processes that could be tested through longitudinal follow-up. Alternatively, targeted non-invasive stimulation of these brain regions

in healthy subjects represents another means of testing such predictions. Non-invasive stimulation of weakened networks holds theoretical promise to mitigate disease transition in high-risk individuals, or perhaps even restore cognitive-emotional control in BD patients.

Acknowledgements

None

Funding

This study was funded by the Australian National Medical and Health Research Council [grant numbers APP1037196, APP1118153), the Brother's Reid and the Lansdowne Foundation. DSB also acknowledges support from the Alfred P. Sloan Foundation and the John D. and Catherine T. MacArthur Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

Conflict of Interest Disclosures

All authors report no financial interests or other potential conflicts of interest.

References

Adler CM, Adams J, DelBello MP, Holland SK, Schmithorst V, Levine A, *et al.* Evidence of white matter pathology in bipolar disorder adolescents experiencing their first episode of mania: A diffusion tensor imaging study. American Journal of Psychiatry 2006; 163: 322-4.

Bai F, Shu N, Yuan Y, Shi Y, Yu H, Wu D, *et al.* Topologically convergent and divergent structural connectivity patterns between patients with remitted geriatric depression and amnestic mild cognitive impairment. J Neurosci 2012; 32(12): 4307-18.

Bassett DS, Khambhati AN. A network engineering perspective on probing and perturbing cognition with neurofeedback. Annals of the New York Academy of Sciences 2017; 1396(1): 126-43.

Bermpohl F, Kahnt T, Dalanay U, Hagele C, Sajonz B, Wegner T, *et al.* Altered representation of expected value in the orbitofrontal cortex in mania. Human Brain Mapping 2010; 31: 958-69.

Betzel RF, Gu S, Medaglia JD, Pasqualetti F, Bassett DS. Optimally controlling the human connectome: the role of network topology. Sci Rep 2016; 6: 30770.

Breakspear M. Dynamic models of large-scale brain activity. Nature Neuroscience 2017; 20(3): 340-52.

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.

Bruno S, Cercignani M, Ron MA. White matter abnormalities in bipolar disorder: A voxel-based diffusion tensor imaging study. Bipolar Disorder 2008; 10: 460-8.

Cao Q, Shu N, An L, Wang P, Sun L, Xia MR, *et al.* Probabilistic diffusion tractography and graph theory analysis reveal abnormal white matter structural connectivity networks in drug-naive boys with attention deficit/hyperactivity disorder. J Neurosci 2013; 33(26): 10676-87.

Chen CH, Suckling J, Lennox BR, Ooi C, Bullmore ET. A quantitative meta-analysis of fMRI studies in bipolar disorder. Bipolar Disord 2011; 13(1): 1-15.

Cocchi L, Gollo LL, Zalesky A, Breakspear M. Criticality in the brain: a synthesis of neurobiology, models and cognition. Progress in Neurobiology 2017; in press.

Cornelius SP, Kath WL, Motter AE. Realistic control of network dynamics. Nature Communications 2013; 4: 1942.

Critchley HD, Wiens S, Rotshtein P, Öhman A, Dolan RJ. Neural systems supporting interoceptive awareness. Nature Neuroscience 2004; 7(2): 189-95.

Crofts JJ, Higham DJ, Bosnell R, Jbabdi S, Matthews PM, Behrens TEJ, *et al.* Network analysis detects changes in the contralesional hemisphere following stroke. Neuroimage 2011; 54(1): 161-9.

Deco G, Jirsa VK, Robinson PA, Breakspear M, Friston K. The dynamic brain: from spiking neurons to neural masses and cortical fields. PLoS Comput Biol 2008; 4(8): e1000092.

Deco G, Ponce-Alvarez A, Mantini D, Romani GL, Hagmann P, Corbetta M. Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. The Journal of Neuroscience 2013; 33(27): 11239-52.

Dima D, Roberts RE, Frangou S. Connectomic markers of disease expression, genetic risk and resilience in bipolar disorder. Transl Psychiatry 2016; 6: e706.

Doucet GE, Bassett DS, Yao N, Glahn DC, Frangou S. The Role of Intrinsic Brain Functional Connectivity in Vulnerability and Resilience to Bipolar Disorder. The American journal of psychiatry 2017: in press.

Estrada E, Rodriguez-Velazquez JA. Subgraph centrality in complex networks. Phys Rev E Stat Nonlin Soft Matter Phys 2005; 71(5 Pt 2): 056103.

Farquharson S, Tournier J-D. High Angular Resolution Diffusion Imaging. In: Van Hecke W, Emsell L, Sunaert S, editors. Diffusion Tensor Imaging: A Practical Handbook. New York, NY: Springer New York; 2016. p. 383-406.

Fornito A, Zalesky A, Breakspear M. The connectomics of brain disorders. Nature Reviews Neuroscience 2015; 16(3): 159-72.

Frangou S. Risk and resilience in bipolar disorder: rationale and design of the Vulnerability to Bipolar Disorders Study (VIBES). Biochemical Society Transactions 2009; 37(5): 1085-9.

Frangou S. Brain structural and functional correlates of resilience to Bipolar Disorder. Frontiers in Human Neuroscience 2012; 5(184).

Freyer F, Roberts JA, Becker R, Robinson PA, Ritter P, Breakspear M. Biophysical mechanisms of multistability in resting-state cortical rhythms. The Journal of Neuroscience 2011; 31(17): 6353-61.

Ganzola R, Nickson T, Bastin ME, Giles S, Macdonald A, Sussmann J, *et al.* Longitudinal differences in white matter integrity in youth at high familial risk for bipolar disorder. Bipolar Disorders 2017; 19(3): 158-67.

Gu S, Betzel RF, Mattar MG, Cieslak M, Delio PR, Grafton ST, *et al.* Optimal trajectories of brain state transitions. NeuroImage 2017; 148: 305-17.

Gu S, Pasqualetti F, Cieslak M, Telesford QK, Yu AB, Kahn AE, *et al.* Controllability of structural brain networks. Nat Commun 2015; 6: 8414.

Guo CC, Sturm VE, Zhou J, Gennatas ED, Trujillo AJ, Hua AY, *et al.* Dominant hemisphere lateralization of cortical parasympathetic control as revealed by frontotemporal dementia. Proceedings of the National Academy of Sciences 2016; 113(17): E2430-E9.

Hajek T, Cullis J, Novak T, Kopecek M, Blagdon R, Propper L, *et al.* Brain structural signature of familial predisposition for bipolar disorder: replicable evidence for involvement of the right inferior frontal gyrus. Biological Psychiatry 2013; 73: 144-52.

Haller S, Xekardaki A, Delaloye C, Canuto A, Lovblad KO, Gold G, *et al.* Combined analysis of grey matter voxelbased morphometry and white matter tract-based spatial statistics in late-life bipolar disorder. Journal of Psychiatry & Neuroscience 2011; 36: 391-401.

Hirsch MW, Smale S, Devaney RL. Differential equations, dynamical systems, and an introduction to chaos. 3rd ed. Waltham, MA: Academic Press; 2013.

Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, *et al.* Predicting human resting-state functional connectivity from structural connectivity. Proceedings of the National Academy of Sciences 2009; 106(6): 2035-40.

Jbabdi S, Sotiropoulos SN, Haber SN, Van Essen DC, Behrens TE. Measuring macroscopic brain connections in vivo. Nat Neurosci 2015; 18(11): 1546-55.

Kempton MJ, Haldane M, Jogia J, Grasby PM, Collier D, Frangou S. Dissociable Brain Structural Changes Associated with Predisposition, Resilience, and Disease Expression in Bipolar Disorder. Journal of Neuroscience 2009; 29(35): 10863.

Lau P, Hawes DJ, Hunt C, Frankland A, Roberts G, Mitchell PB. Prevalence of psychopathology in bipolar high-risk offspring and siblings: a meta-analysis. European child & adolescent psychiatry 2017.

Maslov S, Sneppen K. Specificity and Stability in Topology of Protein Networks. Science 2002; 296(5569): 910-3.

McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R, Cardno A. THe heritability of bipolar affective disorder and the genetic relationship to unipolar depression. Archives of General Psychiatry 2003; 60(5): 497-502.

Meda SA, Gill A, Stevens MC, Lorenzoni RP, Glahn DC, Calhoun VD, *et al.* Differences in resting-state functional magnetic resonance imaging functional network connectivity between schizophrenia and psychotic bipolar probands and their unaffected first-degree relatives. Biological Psychiatry 2012; 71(10): 881-9.

Mortensen PB, Pedersen CB, Melbye M, Mors O, Ewald H. Individual and familial risk factors for bipolar affective disorders in Denmark. Archives of general psychiatry 2003; 60(12): 1209-15.

Muldoon SF, Pasqualetti F, Gu S, Cieslak M, Grafton ST, Vettel JM, *et al.* Stimulation-Based Control of Dynamic Brain Networks. PLoS Comput Biol 2016; 12(9): e1005076.

Nortje G, Stein DJ, Radua J, Mataix-Cols D, Horn N. Systematic review and voxel-based meta-analysis of diffusion tensor imaging studies in bipolar disorder. Journal of affective disorders 2013; 150(2): 192-200.

Nusslock R, Almeida JR, Forbes EE, Versace A, Frank E, Labarbara EJ, *et al.* Waiting to win: elevated striatal and orbitofrontal cortical activity during reward anticipation in euthymic bipolar disorder adults. Bipolar Disorder 2012; 14: 249-60.

Perich T, Lau P, Hadzi-Pavlovic D, Roberts G, Frankland A, Wright A, *et al.* What clinical features precede the onset of bipolar disorder? Journal of psychiatric research 2015; 62: 71-7.

Perry A, Wen W, Kochan NA, Thalamuthu A, Sachdev PS, Breakspear M. The independent influences of age and education on functional brain networks and cognition in healthy older adults. Human Brain Mapping 2017; 38(10): 5094-114.

Perry A, Wen W, Lord A, Thalamuthu A, Roberts G, Mitchell PB, *et al.* The organisation of the elderly connectome. NeuroImage 2015; 114: 414-26.

Phillips ML, Swartz HA. A critical appraisal of neuroimaging studies of bipolar disorder: toward a new conceptualization of underlying neural circuitry and a road map for future research. The American journal of psychiatry 2014; 171(8): 829-43.

Pompei F, Dima D, Rubia K, Kumari V, Frangou S. Dissociable functional connectivity changes during the Stroop task relating to risk, resilience and disease expression in bipolar disorder. Neuroimage 2011; 57(2): 576-82.

Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 2009; 460(7256): 748-52.

Roberts G, Perry A, Lord A, Frankland A, Leung V, Holmes-Preston E, *et al.* Structural dysconnectivity of key cognitive and emotional hubs in young people at high genetic risk for bipolar disorder. Molecular Psychiatry 2016; 00: 1-9.

Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and interpretations. NeuroImage 2010; 52(3): 1059-69.

Slotine JJE, Li W. Applied nonlinear control. Englewood Cliffs, N.J: Prentice Hall; 1991.
Sporns O. Structure and function of complex brain networks. Dialogues in Clinical Neuroscience 2013; 15(3): 247-62.
Sprooten E, Sussmann JE, Clugston A, Peel A, McKirdy J, Moorhead TWJ, *et al.* White Matter Integrity in Individuals at High Genetic Risk of Bipolar Disorder. Biological Psychiatry 2011; 70(4): 350-6.
Stam CJ. Modern network science of neurological disorders. Nat Rev Neurosci 2014; 15(10): 683-95.
Stephan KE, Friston KJ, Frith CD. Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-monitoring. Schizophrenia Bulletin 2009; 35(3): 509-27.
Tang E, Giusti C, Baum GL, Gu S, Pollock E, Kahn AE, *et al.* Developmental increases in white matter network controllability support a growing diversity of brain dynamics. Nature Communications 2017; 8: 1-16.
Torrisi S, Moody TD, Vizueta N, Thomason ME, Monti MM, Townsend JD, *et al.* Differences in resting corticolimbic functional connectivity in bipolar I euthymia. Bipolar Disord 2013; 15(2): 156-66.
Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, *et al.* Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. Neuroimage 2002; 15(1): 273-89.
Widjaja E, Zamyadi M, Raybaud C, Snead OC, Doesburg SM, Smith ML. Disrupted Global and Regional Structural Networks and Subnetworks in Children with Localization-Related Epilepsy. American Journal of Neuroradiology 2015; 36(7): 1362-8.
Wu-Yan E, Betzel RF, Tang E, Gu S, Pasqualetti F, Bassett DS. Benchmarking measures of network controllability on canonical graph models. arXiv 2017; 1706.05117 [physics.soc-ph].
Xekardaki A, Giannakopoulos P, Haller S. White Matter Changes in Bipolar Disorder, Alzheimer Disease, and Mild Cognitive Impairment: New Insights from DTI. J Aging Res 2011; 2011: 286564.
Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 2011; 106(3): 1125-65.
Zalesky A, Fornito A, Bullmore ET. Network-based statistic: identifying differences in brain networks. NeuroImage 2010a; 53(4): 1197-207.
Zalesky A, Fornito A, Harding IH, Cocchi L, Yücel M, Pantelis C, *et al.* Whole-brain anatomical networks: Does the choice of nodes matter? Neuroimage 2010b; 50(3): 970-83.
Zalesky A, Fornito A, Seal ML, Cocchi L, Westin CF, Bullmore ET, *et al.* Disrupted axonal fiber connectivity in schizophrenia. Biological Psychiatry 2011; 69(1): 80-9.