Neurosci. Bull.
DOI 10.1007/s12264-017-0190-6

www.neurosci.cn
www.springer.com/12264

ORIGINAL ARTICLE

Neurocognitive Graphs of First-Episode Schizophrenia and Major Depression Based on Cognitive Features

Sugai Liang[1,2] · Roberto Vega[3] · Xiangzhen Kong[4] · Wei Deng[1,2] · Qiang Wang[1] · Xiaohong Ma[1] · Mingli Li[1] · Xun Hu[5] · Andrew J. Greenshaw[6] · Russell Greiner[3] · Tao Li[1,2]

Received: 4 May 2017 / Accepted: 18 August 2017

Abstract Neurocognitive deficits are frequently observed in patients with schizophrenia and major depressive disorder (MDD). The relations between cognitive features may be represented by neurocognitive graphs based on cognitive features, modeled as Gaussian Markov random fields. However, it is unclear whether it is possible to differentiate between phenotypic patterns associated with the differential diagnosis of schizophrenia and depression using this neurocognitive graph approach. In this study, we enrolled 215 first-episode patients with schizophrenia (FES), 125 with MDD, and 237 demographically-matched healthy controls (HCs). The cognitive performance of all participants was evaluated using a battery of neurocognitive tests. The graphical LASSO model was trained with a one-*vs*-one scenario to learn the conditional independent structure of neurocognitive features of each group. Participants in the holdout dataset were classified into different groups with the highest likelihood. A partial correlation matrix was transformed from the graphical model to further explore the neurocognitive graph for each group. The classification approach identified the diagnostic class for individuals with an average accuracy of 73.41% for FES *vs* HC, 67.07% for MDD *vs* HC, and 59.48% for FES *vs* MDD. Both of the neurocognitive graphs for FES and MDD had more connections and higher node centrality than those for HC. The neurocognitive graph for FES was less sparse and had more connections than that for MDD. Thus, neurocognitive graphs based on cognitive features are promising for describing endophenotypes that may discriminate schizophrenia from depression.

Keywords Schizophrenia · Major depressive disorder · Neurocognition · Neurocognitive graph · Graphical LASSO

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12264-017-0190-6) contains supplementary material, which is available to authorized users.

Sugai Liang and Roberto Vega have contributed equally to this work.

✉ Tao Li
litaohx@scu.edu.cn

[1] Mental Health Centre, West China Hospital, Sichuan University, Chengdu 610041, China

[2] Huaxi Brain Research Centre, West China Hospital, Sichuan University, Chengdu 610041, China

[3] Department of Computing Science, University of Alberta, Edmonton, AB T6G 2R7, Canada

[4] Language and Genetics Department, Max Planck Institute for Psycholinguistics, 6525 XD Nijmegen, The Netherlands

[5] Huaxi Biobank, West China Hospital, Sichuan University, Chengdu 610041, China

[6] Department of Psychiatry, University of Alberta, Edmonton, AB T6G 2R7, Canada

Introduction

Neurocognitive deficits are frequently observed in individuals diagnosed with schizophrenia or major depressive disorder (MDD). Estimates indicate that approximately half of patients with schizophrenia have comorbid depression [1]. Although the interpretation of such neurocognitive endophenotypes has provided some evidence for the identification of transdiagnostic processes between schizophrenia and MDD [2], it remains unclear whether intrinsic associations between these neurocognitive endophenotypes (referred to as neurocognitive graphs) can be used to differentiate between the two disorders.

Published online: 02 November 2017

Analysis of this distinction, related to associations between neurocognitive phenotypes, is highly relevant in the context of the Research Domain Criteria project of NIMH which focuses on "new ways of classifying psychiatric diseases on multiple dimensions of biology and behavior" [3]. Using this approach to advance the appropriate differential diagnosis for schizophrenia and MDD may represent a significant step towards increasing the understanding and efficacy of treatment of both disorders [4].

Neurocognitive graphs, measured as Gaussian Markov random fields by conditional independence structure, can be built using the links between multiple cognitive features corresponding to different cognitive functions (and their associated brain structures). Characterizing the interactions between multiple neurocognitive variables in this way can provide new insights into understanding brain functions in the context of these psychiatric disorders. From our current perspective, neurocognitive functions arise from the interactions within and between distributed brain areas operating in large-scale structural and functional networks [5], and both schizophrenia and MDD are considered to be disorders of brain network disorganization [6, 7]. From these and other findings from neuroimaging studies, it is plausible to assume that the neurocognitive graphs of schizophrenia and MDD have specific characteristics which may be discriminable from those of normal controls.

To the best of our knowledge, no studies have investigated neurocognitive graphs based on cognitive features for individuals with schizophrenia and MDD in comparison with healthy controls (HCs). In the present study, to construct a neurocognitive graph for each group, a large cohort of 577 participants completed a battery of neurocognitive tests. Using the performance of individuals in these tests as variables, we adopted the graphical Least Absolute Shrinkage and Selection Operator (LASSO) model to learn the conditional independent structure of neurocognitive features of individuals with first-episode schizophrenia (FES) or MDD. The conditional independent structure of neurocognitive features was used to explore the neurocognitive graph for each group.

Materials and Methods

Participants

Five hundred and seventy-seven participants were recruited for this study: 215 with FES, 125 with MDD, and 237 HCs. All groups were matched for age, gender, and educational level (demographic characteristics in Table S1). All participants were right-handed Han Chinese aged 16 to 50 years. Written informed consent was given by all participants. The current study was approved by the Ethics Committee of West China Hospital, Sichuan University, in accordance with the Declaration of Helsinki.

Neuropsychological Assessments

General intelligence was estimated on initial assessment of all participants using the short version of the Wechsler Adult Intelligence Scale – Revised in China (WAIS-RC) [8]: the seven subtests of the WAIS-RC comprised information, arithmetic, digital symbol, digital span test, block design, picture completion, and similarities.

Immediate and delayed logical memory were evaluated with the Wechsler Memory Scale – Revised in China [9]. Lower raw scores represent poorer performance in logical memory.

The computerized Cambridge Neurocognitive Test Automated Battery (CANTAB- http://www.cambridgecognition.com), comprising a series of visuo-spatial tasks, is accepted as predictive for psychosocial functioning in individuals with schizophrenia and other mental disorders [10, 11]. Seven CANTAB tests are thought to be sensitive to frontal (including frontostriatal, frontotemporal, and frontoparietal), cingulate, and temporal brain functions [12]. CANTAB included the following: Big Circle/Little Circle (BLC), Rapid Visual Information Processing (RVP), Delayed Matching to Sample (DMS), Pattern Recognition Memory (PRM), Spatial Working Memory (SWM), Stockings of Cambridge (SOC), and Intra/extra Dimensional Set Shift (IED). Variables of interest across tasks included latency (reaction time), percent correct, errors, trials completed, and strategy [13–15].

Processing speed during attention and task switching were measured with the Trail Making Test, parts A (TMA) and B-Modified. Scores were recorded as the total time required to complete the task. Higher scores indicate poorer performance [16].

Fifty-four features were specified for each participant. Detailed measures and features are listed in Supplementary Tables S2 and S3.

Gaussian Markov Random Fields and Graphical LASSO

Here, we represented the association among different features for each class of instance (HC, FES, and MDD) using an undirected graphical model, where each feature $X_i \in \boldsymbol{X} = \{X_1, \ldots, X_{54}\}$ corresponds to a node, and edges collectively determine the conditional independence structure. Assuming these 54 selected features follow a multivariate Gaussian distribution, $p(X_1, \ldots, X_{54}) \sim N(\mu; \Sigma)$, a pair of features X_i and X_j are conditionally independent given the other features $\left(X_i \perp X_j | \boldsymbol{X} - \{X_i, X_j\}\right)$ only if the (i, j)

entry in the inverse covariance matrix is zero ($\Sigma_{ij}^{-1} = 0$) [17]. Therefore, learning the structure of the graph, known as the Gaussian Markov random field, reduces the problem of inputting zero entries in the inverse covariance matrix Σ^{-1} (Fig. 1).

The graphical LASSO learner learns this sparse graphical model by estimating the inverse covariance matrix Σ^{-1} using an L_1-norm penalty [18], from a dataset $\mathbf{S} = \{S^1, \ldots, S^n\}$, where each $S^k = \{x_1^k, \ldots, x_{54}^k\}$, $k \in \{1, \ldots, n\}$. In general, the log-likelihood of a dataset **S** from a Gaussian $N(\mu, \Sigma)$ is:

$$L(S) = \frac{n}{2}\left[\log\left(\left|\Sigma^{-1}\right|\right) - \mathrm{tr}\left(\frac{1}{n}(S-\mu)^T(S-\mu)\Sigma^{-1}\right) + c\right] \quad (1)$$

where $\det|X|$ represents the determinant of the matrix X, $\mathrm{tr}(X)$ represents the trace of the matrix X, and c is a constant term independent of the data. (Here "$S - \mu$" involves subtracting the mean from each participant.) Graphical LASSO seeks the inverse covariance matrix that maximizes the L_1-penalized likelihood function,

$$\max_{\Sigma^{-1}} \log\left(\left|\Sigma^{-1}\right|\right) - tr\left((S-\mu)^T(S-\mu)\Sigma^{-1}\right) - \lambda \parallel \Sigma^{-1} \parallel_1 \quad (2)$$

which uses the L_1-norm $\parallel M \parallel_1 = \sum_{i,j} |M_{i,j}|$ which is the sum of the absolute values of the matrix entries, and λ is a regularization parameter that controls the degree of sparsity in the Σ^{-1}; a higher λ value signifies increased regularization and a sparser inverse covariance. (Here, the value of λ is determined in the internal cross-validation.)

The learning process uses the FES instances to learn the parameters of a Gaussian distribution $(\mu_{FES}, \Sigma_{FES})$, where μ_{FES} is the mean of the instances, and Σ_{FES} (actually, Σ_{FES}^{-1}) is based on the graphical LASSO learner; it similarly learns the Gaussian parameters for MDD and HC.

The one-*vs*-one scenarios consist in fitting one classifier per class pair and require to fit n_classes * (n_classes − 1)/2 classifiers. At performance time, given a novel instance S_{new}, its likelihood is computed based on each of the two Gaussians in the one-*vs*-one scenarios, then assign S_{new} to the class y_{new} that is most likely:

$$y_{new} = argmax_k \ln\left(\left|\Sigma_k^{-1}\right|\right) - (S_{new} - \mu_k)^T \Sigma_k^{-1} (S_{new} - \mu_k) \quad (3)$$

where μ_k is the vector which contains the mean of each feature in the dataset *X*.

$$\Sigma^{-1} = \begin{bmatrix} 0.3125 & -0.125 & 0 \\ -0.125 & 0.5823 & 0.3333 \\ 0 & 0.3333 & 0.3333 \end{bmatrix}$$

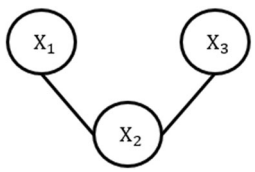

Fig. 1 Conditional independence in the graphical model. The conditional independence of two variables is represented as zero values in the inverse covariance matrix. In the graphical model, conditional independence is represented by the lack of an edge between two variables.

Machine Learning Analysis

The overall approach of machine learning analysis involved the following steps: (1) The data were cleaned and normalized to the Z-scores. (2) To minimize the potential for over-fitting, the dataset was randomly divided into 70% for training the graphical model (training dataset), and the remaining 30% subset was reserved for testing (holdout dataset), balanced by class distribution. (3) The graphical LASSO was applied to the training dataset (using R package *glasso*) to learn the graphical structure with the one-*vs*-one scenario for each group (FES *vs* HC, MDD *vs* HC, and FES *vs* MDD). We considered $\lambda \in \{0.01, 0.03, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, and selected the values with the highest five-fold cross-validation accuracy over the training set. (4) Given the graphic model, the likelihood of each instance was computed in the holdout dataset belonging to each learned model for FES, MDD, or HC (see Equation 3). (5) We ran this process-steps (2), (3), and (4) – 30 times, based on different splits of the data. See Fig. 2 for the data analysis flowchart. The average of performance metrics (accuracy and F_1 score) over these 30 runs are reported. Supplementary material provides details of the performance metrics.

Partial Correlation Model

The partial correlation model was designed to quantify how closely two nodes are connected with each other, considering the other nodes. This model represents the linear correlation between two variables by removing the linear effects of other nodes (see Equation (S3) in Supplementary material). In this study, the partial correlation model was computed on the whole dataset from each group of the one-*vs*-one scenarios with the best λ selected from the five-fold cross-validation. Graphical models of FES *vs* HC were based on $\lambda = 0.01$, graphical models of MDD *vs* HC were based on $\lambda = 0.03$, and graphical models of FES *vs* MDD were based on $\lambda = 0.01$. All non-zero weight connections in the neurocognitive graph were reported and used for visual representation.

Three theoretical properties were also calculated for the graphs based on the absolute values of the entries in the partial correlation matrix: node strength centrality,

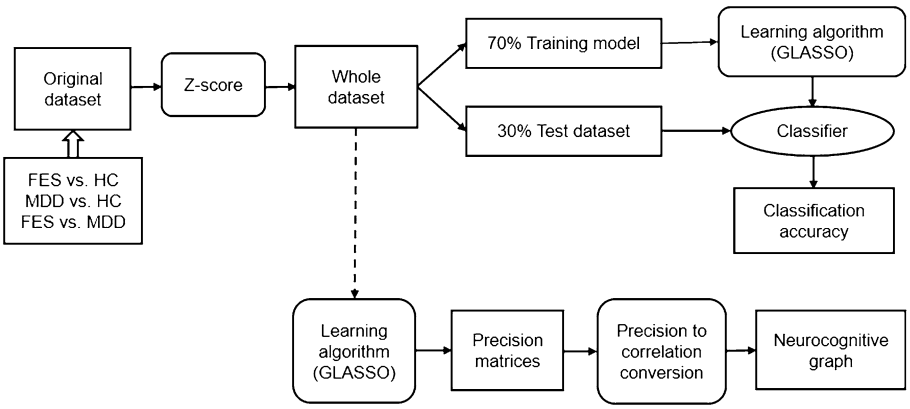

Fig. 2 Flowchart of data processing.

closeness centrality, and betweenness centrality. The centrality of a node represents its relative importance within a network [19]. According to the study of Opsahl, Agneessens, Skvoretz [20], node strength centrality is defined as the sum of all associations of a given feature with all other nodes, closeness centrality represents how close a feature is to all other nodes, and betweenness centrality represents the shortest path length connecting any two features. The higher the measure of node centrality, the more "important" a given node is in the graph [21]. In this study, node strength centrality was revealed in the main report as closeness, and betweenness centrality was substantially correlated with node strength centrality. The results of closeness and betweenness are presented in Supplementary Figs. S1 and S2. We used the R packages *igraph* and *tnet*.

Permutation tests were performed to compare each node strength centrality value across different groups. We created a distribution by assigning group labels randomly to each of the groups 1000 times and then estimated the difference between groups each time. If the observed difference between two groups was within 2.5% on either end of the distribution, we considered the difference to be significant at the 5% level.

Results

Neurocognitive Graphs of FES *vs* HC

The neurocognitive graph of FES had 890 edges with non-zero weights and featured no unconnected nodes (Fig. 3B). The neurocognitive graph of HC had 600 edges, and the *SOC_MM* was conditionally independent of the other nodes (Fig. 3A). The results showed that the graph of FES had more connections than that of HC. The top 3 positive and negative connections in neurocognitive graphs of FES *vs* HC are reported in the Supplementary material.

The node strength centrality for FES *vs* HC is illustrated in Fig. 5A. The permutation tests to compare the node strength centrality of FES and HC showed significant differences in 24 nodes (*VIQ, PIQ; BLC_CRL; DMS_AP, DMS_Ld, DMS_Ls, DMS_PCs, PC0D*, *PC4D*, *PC12D; PRM_Ld, PreED_E; IED_CSE, IED_ATT; SOC_MIT, SOC_MST; RVP_ML; BetwErr, WithErr, TotalErr, SWM _MFR, SWM _MLR, SWM _MTP,* and *Strategy*). The neurocognitive graph of FES had a higher level of strength centrality on these nodes compared to those of HC. The cognitive functions related to these 24 nodes included general intelligence, motor speeding, perceptual sensitivity, visual and pattern recognition memory, shifting, planning, sustained attention and inhibition, working memory, and strategy. The closeness and betweenness centralities for FES and HC are illustrated in Supplementary Figs. S1A and S2A.

In the classification of FES and HC, the accuracy was 73.41% and the F_1 score was 0.64.

Neurocognitive Graphs of MDD *vs* HC

The number of edges with non-zero weights in the neurocognitive graphs was 501 for MDD and 390 for HC (Fig. 4). Both graphs had conditionally independent nodes, i.e. *SOC_MM* and *strategy in SWM*. But the *BLC_PC* node was also conditionally independent in the graph of HC. The results indicated that the neurocognitive graph of MDD had more connections than that of HC. The top 3 positive and negative connections in neurocognitive graphs of MDD *vs* HC are presented in the Supplementary material.

The node strength centrality for MDD and HC is illustrated in Fig. 5B. The permutation tests comparing the node strength estimates of MDD and HC differed significantly in 15 nodes (*ILM; BLC_CRL; DMS_AP, DMS_BDP, DMS_PCs, PC4D; PRM_Li, PRM_PCi, PRM_Ld; PreED_E, EDS_E, IED_CSE, IED_ATT; SOC_PSM,* and *SOC_MIT*). The neurocognitive graph of MDD was more highly connected on these nodes than that

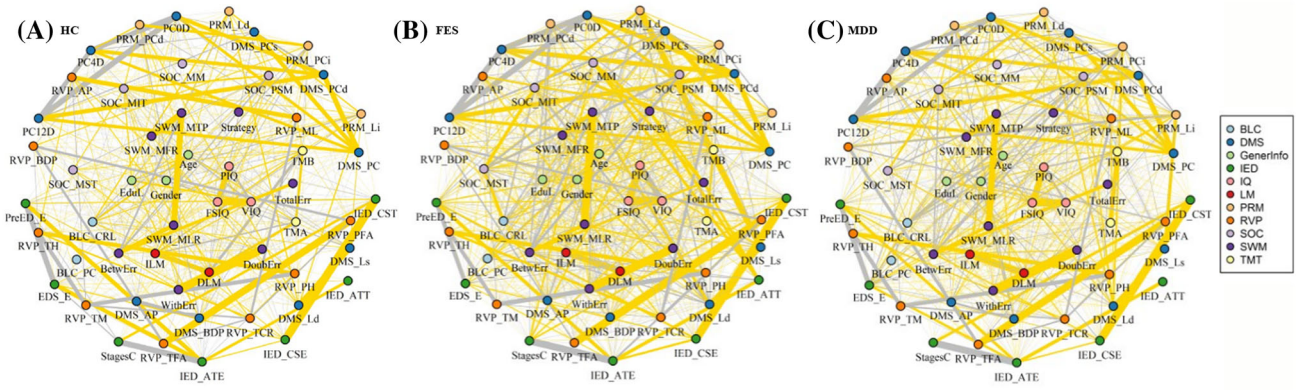

Fig. 3 Neurocognitive graphs for healthy controls, and patients with schizophrenia or depression. **A** Neurocognitive graph of HC, **B** neurocognitive graph of FES, **C** neurocognitive graph of MDD. Each graph contains 54 nodes. Nodes with the same color are from the same neurocognitive test. Yellow lines, positive associations; gray lines, negative relations; edge thickness, association strength; $\lambda = 0.01$ for the three neurocognitive graphs.

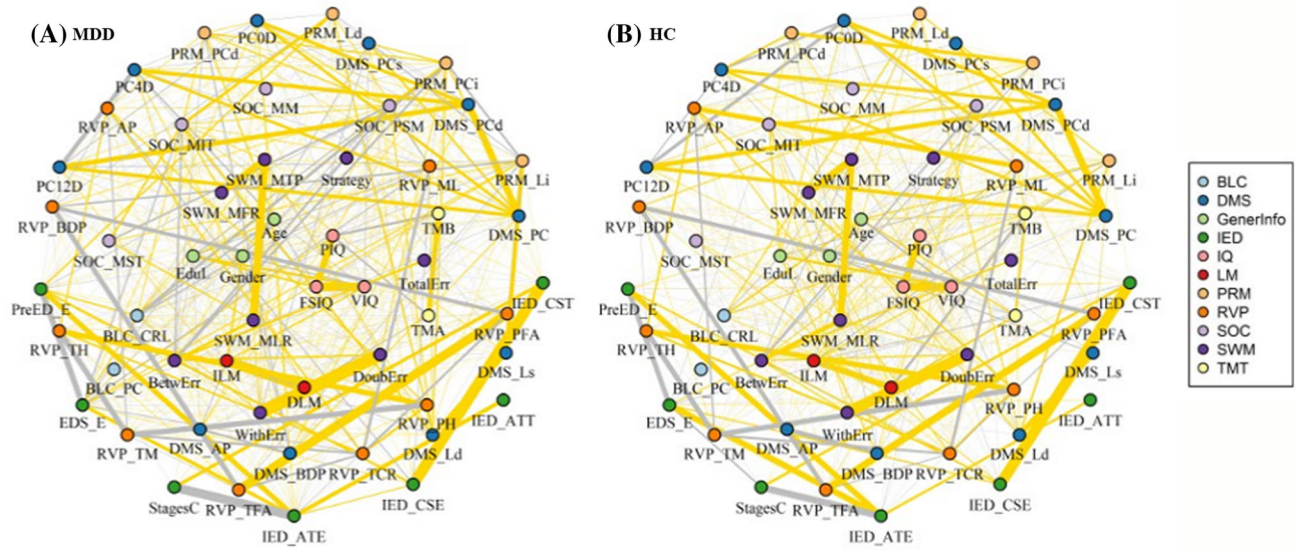

Fig. 4 Neurocognitive graphs for depression and healthy controls. **A** Neurocognitive graph of MDD, **B** neurocognitive graph of HC. Each graph contains 54 nodes. Nodes with the same color come from the same neurocognitive test. Yellow lines represent positive associations, gray lines represent negative relations. Thickness of an edge indicates the association strength. $\lambda = 0.03$ for those two neurocognitive graphs.

of HC. The cognitive functions related to these 15 nodes included logical memory, motor speeding, perceptual sensitivity, visual and pattern cognition memory, shifting, and planning. The visualization of closeness and betweenness centralities for MDD and HC are illustrated in Supplementary Figs. S1B and S2B.

In the classification of MDD and HC, the accuracy was 67.07% and the F_1 score was 0.76.

Neurocognitive Graphs of FES *vs* MDD

The neurocognitive graph of FES had 890 edges with non-zero weights and featured no unconnected nodes (Fig. 3B). The neurocognitive graph of MDD had 724 edges, and *SOC_MM* was conditionally unconnected with the others (Fig. 3C). The results showed that the graph of FES had more connections than that of MDD. The top 3 positive and negative connections in the neurocognitive graphs of FES *vs* MDD are reported in the Supplementary material.

The node strength centrality of FES and MDD are shown in Fig. 5C. The permutation tests to compare the node strength estimates of MDD and FES were significantly different in nine nodes (*ILM, TMA*, *PRM_Li, PRM_PCi, EDS_E*, *SOC_PSM, SOC_MST, SOC_MM,* and *SWM_MTP*). Compared to that of MDD, the neurocognitive graph of FES had higher strength centrality on

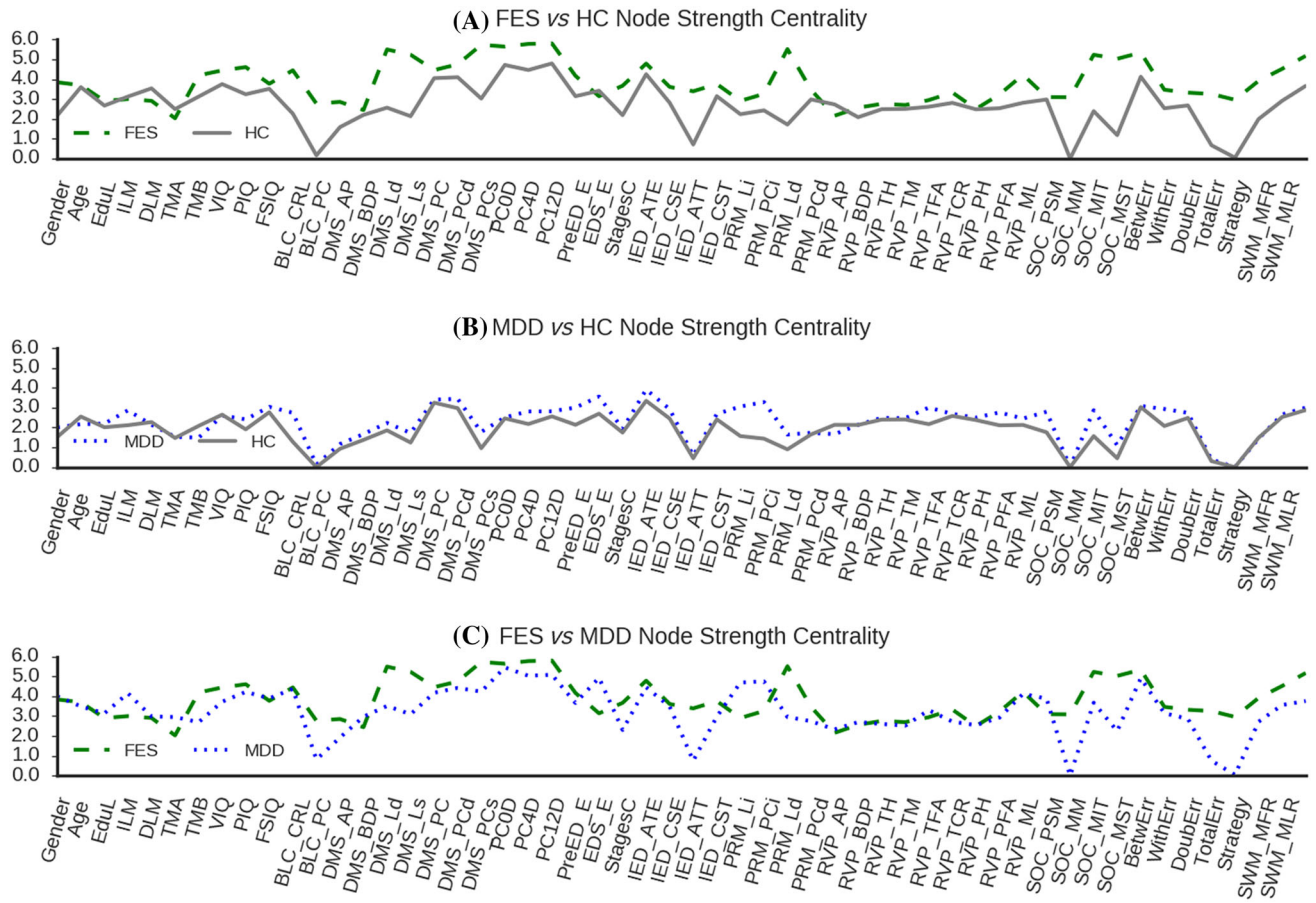

Fig. 5 Node strength centrality in one-*vs*-one scenarios. **A** Node strength centrality for schizophrenia and HC; **B** node strength centrality for MDD and HC; **C** node strength centrality for schizophrenia and MDD. Green dashed line, node strength centrality of FES; gray solid line, node strength centrality of HC; blue dotted line, node strength centrality of MDD.

the nodes *ILM, TMA*, *PRM_Li, PRM_PCi, EDS_E,* and *SOC_PSM* which related to cognitive functions including immediate logical memory, processing speed, pattern recognition memory, shifting, and planning. The neurocognitive graph of FES had lower strength centrality on the nodes *SOC_MST, SOC_MM, SWM_MTP* which related to cognitive functions including planning and working memory. The visualization of closeness and betweenness centralities for FES and MDD are illustrated in Supplementary Figs S1C and S2C.

In the classification of FES and MDD, accuracy was 59.48% and the F_1 score was 0.60.

Discussion

To the best of our knowledge, this is the first report on neurocognitive graphs of FES, MDD, and HC on the basis of cognitive assessments by combining both theory-driven and data-driven approaches. The purposes of this investigation were: (1) to estimate the accuracy of neurocognitive graphs in classifying individuals with FES or with MDD in the one-*vs*-one classification scenario; (2) to examine neurocognitive features to develop further applications of neurocognitive graphing of the heterogeneous neuropsychological profiles of the illnesses; and (3) to provide additional evidence for possible transdiagnostic mechanisms in FES and MDD.

The graphical LASSO model was trained and used to classify participants with the highest likelihood using 54 features such as general intelligence, working memory, visual memory, planning, and shifting. The results showed that patients with FES and patients with MDD had different neurocognitive graphs, that of the former being less sparse.

In this study, the neurocognitive graph of FES had more connections, featured no conditionally-independent nodes and a higher level of node centrality, indicating that FES has hyperconnections in the neurocognitive graph. Differences of node strength centrality between FES-related and MDD-related cognitive functions included general intelligence, motor speeding, perceptual sensitivity, visual and pattern recognition memory, shifting, planning, sustained

attention and inhibition, working memory, and strategy. Consistent with previous studies, patients with FES had more impairments in premorbid IQ, memory, attention, processing speed, and learning than patients with MDD [22–24]. Neuroimaging studies have demonstrated that patients with schizophrenia have more severe deficits in prefrontal and superior temporal activation during the performance of working memory tasks [25, 26]. In this study, the interactions of these cognitive functions in the neurocognitive graph of FES differed greatly from those of HC, indicating a significant neurocognitive endophenotype associated with schizophrenia. This study also suggested that patients with schizophrenia exhibited a higher demand for cognitive reserve to complete cognitive tasks than HCs [27], consistent with the characteristics of the neurocognitive graph of FES. From the aspect of neuroimaging studies, there is evidence that schizophrenia is associated with disrupted functional networks which may transfer parallel information less efficiently at a relatively high cost [28]. In this context, the hyperconnectivity and hyperactivity of the default mode network could contribute to disturbances of thought and risk for the illness [29, 30]. The current study provides evidence for increased connections reported in previous brain network studies using the neurocognitive graph approach.

The neurocognitive graph of MDD also had more connections, fewer conditionally-independent nodes, and higher node centrality than HC. The differences of node strength centrality between MDD-related and HC-related cognitive functions included logical memory, motor speeding, perceptual sensitivity, visual and pattern cognition memory, shifting, and planning. The interactions of these cognitive functions differed between the MDD and HC neurocognitive graphs. Previous studies have also shown that patients with MDD have deficits in these cognitive domains [31, 32]. As patients with MDD have psychomotor retardation and an increased demand for cognitive reserve in motor speeding [33, 34], it may be the case that the neurocognitive graph of HCs had an independent *BLC_PC* node when compared to that of MDD. Based on neuroimaging studies, MDD is associated with hyperconnectivity not only within the default mode network but also within the frontoparietal network [35], and increased functional connectivity of the anterior medial cortex may be positively associated with rumination in depression [36, 37].

In this study, the neurocognitive graph of FES was less sparse with more connections and fewer conditionally-independent nodes than that of MDD. Node centrality differences between the neurocognitive graphs of MDD and FES indicate that immediate logical memory, processing speed, pattern recognition memory, shifting, planning, and working memory differ between these disorders. This distinction may be important for guiding future research to elucidate transdiagnostic mechanisms. When undergoing cognitive assessments, patients with schizophrenia exhibit more demand for cognitive reserve than those with MDD. Based on related neurocognitive studies, MDD with psychotic symptoms is associated with higher levels of cognitive deficit [38]. Schizophrenia is associated with more serious cognitive impairments in selective attention and working memory than MDD [39, 40].

Both schizophrenia and depression are associated with symptoms such as psychomotor slowing, difficulty concentrating and remembering details, and a decline in working memory and learning. Neuroimaging studies have supported the involvement of brain regions closely associated with emotion-processing and for models of psychotic symptoms in both schizophrenia and depression. These regions include the hippocampus, insula, prefrontal cortex, and inferior parietal cortex [38, 41]. Genetic research has also provided some evidence related to the pathogenesis of both schizophrenia and depression, including retinoic acid-inducible or induced gene 1, the α-1C subunit of the L-type voltage-gated calcium channel, and immune genes [42–45]. These convergent findings may explain the relatively lower accuracy in the classification of FES and MDD in this study. In addition to these overlapping features, lower accuracy may also be related to the unbalanced sample size for these two groups.

Although the results of the present study are encouraging, possible limitations should be considered. First, it is important to note that although the battery of neurocognitive tests used provided reliable evaluations, neurocognitive graphs could be constructed with different cognitive features. Given that the neurocognitive tests of this study mostly concerned general intelligence and executive function, it will be important to include social cognition and emotion in future. Second, it remains unclear how neurocognitive graphs vary over the course of longitudinal studies, particularly across exposure to a range of therapies using antipsychotics, antidepressants, or physical treatments. Thus, further longitudinal studies are required. Third, distinguishing psychotic from affective symptoms remains a dilemma of psychiatric classification. In future studies, it will be fruitful to evaluate both psychotic and affective symptoms in individuals with schizophrenia and depression, respectively, in an attempt to build a differential diagnosis paradigm for comorbid schizophrenia and depression symptoms. Fourth, the sample size of the current study is modest; larger sample sizes are needed and the model needs to be validated with new instances.

In conclusion, we introduce the neurocognitive graph as a novel tool with which to explore the intrinsic associations between cognitive features in psychopathology. Specifically, we adopted the graphical model and machine

learning with the one-*vs*-one scenario to predict whether an individual has depression, schizophrenia, or is healthy, and to generate neurocognitive graphs for each group. The results revealed neurocognitive graphs with less sparsity in both FES and MDD relative to HC. The neurocognitive graphs of schizophrenia and major depression may prove beneficial for longitudinal monitoring of the probability of relapse and therapeutic advances.

Acknowledgements We gratefully acknowledge the participants in this study for their generous participation. This work was partially funded by National Nature Science Foundation of China Key Projects (81130024, 91332205, and 81630030), the National Key Technology R & D Program of the Ministry of Science and Technology of China (2016YFC0904300), the National Natural Science Foundation of China/Research Grants Council of Hong Kong Joint Research Scheme (8141101084), the Natural Science Foundation of China (8157051859), the Sichuan Science & Technology Department (2015JY0173), the Canadian Institutes of Health Research, Alberta Innovates: Centre for Machine Learning, and the Canadian Depression Research & Intervention Network.

References

1. Buckley PF, Miller BJ, Lehrer DS, Castle DJ. Psychiatric comorbidities and schizophrenia. Schizophr Bull 2009, 35: 383–402.
2. Bentall RP, Rowse G, Shryane N, Kinderman P, Howard R, Blackwood N, *et al.* The cognitive and affective structure of paranoid delusions: a transdiagnostic investigation of patients with schizophrenia spectrum disorders and depression. Arch Gen Psychiatry 2009, 66: 236–247.
3. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, *et al.* Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 2010, 167: 748–751.
4. Siris SG. Depression in schizophrenia: perspective in the era of "Atypical" antipsychotic agents. Am J Psychiatry 2000, 157: 1379–1389.
5. Bressler SL, Menon V. Large-scale brain networks in cognition: emerging methods and principles. Trends Cogn Sci 2010, 14: 277–290.
6. Rubinov M, Bullmore E. Schizophrenia and abnormal brain network hubs. Dialogues Clin Neurosci 2013, 15: 339–349.
7. Hamilton JP, Chen G, Thomason ME, Schwartz ME, Gotlib IH. Investigating neural primacy in Major Depressive Disorder: multivariate Granger causality analysis of resting-state fMRI time-series data. Mol Psychiatry 2011, 16: 763–772.
8. Gong Y. Wechsler adult intelligence scale-revised in China Version. Hunan Medical College, Changsha, Hunan/China 1992.
9. Gong Y. Wechsler Memory Scale–Revised in China. Hunan Medical College, Changsha, Hunan/China 1989.
10. Levaux MN, Potvin S, Sepehry AA, Sablier J, Mendrek A, Stip E. Computerized assessment of cognition in schizophrenia: promises and pitfalls of CANTAB. Eur Psychiatry 2007, 22: 104–115.
11. Johnston BA, Coghill D, Matthews K, Steele JD. Predicting methylphenidate response in attention deficit hyperactivity disorder: a preliminary study. J Psychopharmacol 2015, 29: 24–30.
12. Sahakian B, Owen A. Computerized assessment in neuropsychiatry using CANTAB: discussion paper. J R Soc Med 1992, 85: 399.
13. Robbins TW, James M, Owen AM, Sahakian BJ, Lawrence AD, McInnes L, *et al.* A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: implications for theories of executive functioning and cognitive aging. Cambridge Neuropsychological Test Automated Battery. J Int Neuropsychol Soc 1998, 4: 474–490.
14. Wu MJ, Passos IC, Bauer IE, Lavagnino L, Cao B, Zunta-Soares GB, *et al.* Individualized identification of euthymic bipolar disorder using the Cambridge Neuropsychological Test Automated Battery (CANTAB) and machine learning. J Affect Disord 2016, 192: 219–225.
15. Haring L, Muursepp A, Mottus R, Ilves P, Koch K, Uppin K, *et al.* Cortical thickness and surface area correlates with cognitive dysfunction among first-episode psychosis patients. Psychol Med 2016, 46: 2145–2155.
16. Lu L, Bigler ED. Performance on original and a Chinese version of Trail Making Test Part B: a normative bilingual sample. Appl Neuropsychol 2000, 7: 243–246.
17. Koller D, Friedman N. Probabilistic Graphical Models: Principles and Techniques-Adaptive Computation and Machine Learning. The MIT Press, 2009.
18. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. Biostatistics 2008, 9: 432–441.
19. Yang Y, Dong Y, Chawla NV. Predicting node degree centrality with the node prominence profile. Sci Rep 2014, 4: 7236.
20. Opsahl T, Agneessens F, Skvoretz J. Node centrality in weighted networks: Generalizing degree and shortest paths. Soc Networks 2010, 32: 245–251.
21. Benzi M, Klymko C. A matrix analysis of different centrality measures. arXiv preprint arXiv:1312.6722 2014.
22. Terachi S, Yamada T, Pu S, Yokoyama K, Matsumura H, Kaneko K. Comparison of neurocognitive function in major depressive disorder, bipolar disorder, and schizophrenia in later life: A cross-sectional study of euthymic or remitted, non-demented patients using the Japanese version of the Brief Assessment of Cognition in Schizophrenia (BACS-J). Psychiatry Res 2017, 254: 205–210.
23. Reichenberg A, Harvey PD, Bowie CR, Mojtabai R, Rabinowitz J, Heaton RK, *et al.* Neuropsychological function and dysfunction in schizophrenia and psychotic affective disorders. Schizophr Bull 2009, 35: 1022–1029.
24. Jeste DV, Heaton SC, Paulsen JS, Ercoli L, Harris J, Heaton RK. Clinical and neuropsychological comparison of psychotic depression with nonpsychotic depression and schizophrenia. Am J Psychiatry 1996, 153: 490–496.
25. Barch DM, Sheline YI, Csernansky JG, Snyder AZ. Working memory and prefrontal cortex dysfunction: specificity to schizophrenia compared with major depression. Biol Psychiatry 2003, 53: 376–384.
26. Walter H, Vasic N, Hose A, Spitzer M, Wolf RC. Working memory dysfunction in schizophrenia compared to healthy controls and patients with depression: evidence from event-related fMRI. Neuroimage 2007, 35: 1551–1561.
27. Stern Y. The concept of cognitive reserve: a catalyst for research. J Clin Exp Neuropsychol 2003, 25: 589–593.
28. Liu Y, Liang M, Zhou Y, He Y, Hao Y, Song M, *et al.* Disrupted small-world networks in schizophrenia. Brain 2008, 131: 945–961.
29. Whitfield-Gabrieli S, Thermenos HW, Milanovic S, Tsuang MT, Faraone SV, McCarley RW, *et al.* Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-

degree relatives of persons with schizophrenia. Proc Natl Acad Sci U S A 2009, 106: 1279–1284.
30. Hu ML, Zong XF, Mann JJ, Zheng JJ, Liao YH, Li ZC*, et al.* A review of the functional and anatomical default mode network in schizophrenia. Neurosci Bull 2017, 33: 73–84.
31. McIntyre RS, Cha DS, Soczynska JK, Woldeyohannes HO, Gallaugher LA, Kudlow P*, et al.* Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions. Depress Anxiety 2013, 30: 515–527.
32. Snyder HR. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. Psychol Bull 2013, 139: 81–132.
33. White DA, Myerson J, Hale S. How cognitive is psychomotor slowing in depression? Evidence from a meta-analysis. Aging Neuropsychol Cogn 1997, 4: 166–174.
34. Albus M, Hubmann W, Wahlheim C, Sobizack N, Franz U, Mohr F. Contrasts in neuropsychological test profile between patients with first-episode schizophrenia and first-episode affective disorders. Acta Psychiatr Scand 1996, 94: 87–93.
35. Kaiser RH, Andrews-Hanna JR, Wager TD, Pizzagalli DA. Large–scale network dysfunction in major depressive disorder: A meta-analysis of resting-state functional connectivity. JAMA Psychiatry 2015, 72: 603–611.
36. Zhu X, Wang X, Xiao J, Liao J, Zhong M, Wang W*, et al.* Evidence of a dissociation pattern in resting-state default mode network connectivity in first-episode, treatment-naive major depression patients. Biol Psychiatry 2012, 71: 611–617.
37. Zhang K, Zhu Y, Wu S, Liu H, Zhang W, Xu C*, et al.* Molecular, functional, and structural imaging of major depressive disorder. Neurosci Bull 2016, 32: 273–285.
38. Busatto GF. Structural and functional neuroimaging studies in major depressive disorder with psychotic features: a critical review. Schizophr Bull 2013, 39: 776–786.
39. Egeland J, Sundet K, Rund BR, Asbjornsen A, Hugdahl K, Landro NI*, et al.* Sensitivity and specificity of memory dysfunction in schizophrenia: a comparison with major depression. J Clin Exp Neuropsychol 2003, 25: 79–93.
40. Egeland J, Rund BR, Sundet K, Landro NI, Asbjornsen A, Lund A*, et al.* Attention profile in schizophrenia compared with depression: differential effects of processing speed, selective attention and vigilance. Acta Psychiatr Scand 2003, 108: 276–284.
41. Bernstein HG, Ortmann A, Dobrowolny H, Steiner J, Brisch R, Gos T*, et al.* Bilaterally reduced claustral volumes in schizophrenia and major depressive disorder: a morphometric postmortem study. Eur Arch Psychiatry Clin Neurosci 2016, 266: 25–33.
42. Haybaeck J, Postruznik M, Miller CL, Dulay JR, Llenos IC, Weis S. Increased expression of retinoic acid-induced gene 1 in the dorsolateral prefrontal cortex in schizophrenia, bipolar disorder, and major depression. Neuropsychiatr Dis Treat 2015, 11: 279–289.
43. Green EK, Grozeva D, Jones I, Jones L, Kirov G, Caesar S*, et al.* The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. Mol Psychiatry 2010, 15: 1016–1022.
44. Fillman SG, Cloonan N, Catts VS, Miller LC, Wong J, McCrossin T*, et al.* Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol Psychiatry 2013, 18: 206–214.
45. Bufalino C, Hepgul N, Aguglia E, Pariante CM. The role of immune genes in the association between depression and inflammation: a review of recent clinical studies. Brain Behav Immun 2013, 31: 31–47.