Brain Imaging and Behavior
https://doi.org/10.1007/s11682-018-9947-4

ORIGINAL RESEARCH

Machine learning technique reveals intrinsic characteristics of schizophrenia: an alternative method

Junhua Li[1,2,3] · Yu Sun[1] · Yi Huang[4,5] · Anastasios Bezerianos[1] · Rongjun Yu[1,4]

Abstract

Machine learning technique has long been utilized to assist disease diagnosis, increasing clinical physicians' confidence in their decision and expediting the process of diagnosis. In this case, machine learning technique serves as a tool for distinguishing patients from healthy people. Additionally, it can also serve as an exploratory method to reveal intrinsic characteristics of a disease based on discriminative features, which was demonstrated in this study. Resting-state functional magnetic resonance imaging (fMRI) data were obtained from 148 participants (including patients with schizophrenia and healthy controls). Connective strengths were estimated by Pearson correlation for each pair of brain regions partitioned according to automated anatomical labelling atlas. Subsequently, consensus connections with high discriminative power were extracted under the circumstance of the best classification accuracy. Investigating these consensus connections, we found that schizophrenia group predominately exhibited weaker strengths of inter-regional connectivity compared to healthy group. Aberrant connectivities in both intra- and inter-hemispherical connections were observed. Within intra-hemispherical connections, the number of aberrant connections in the right hemisphere was more than that of the left hemisphere. In the exploration of large regions, we revealed that the serious dysconnectivities mainly appeared on temporal and occipital regions for the within-large-region connections; while connectivity disruption was observed on the connections from temporal region to occipital, insula and limbic regions for the between-large-region connections. The findings of this study corroborate previous conclusion of dysconnectivity in schizophrenia and further shed light on distribution patterns of dysconnectivity, which deepens the understanding of pathological mechanism of schizophrenia.

Keywords Resting-state fMRI · Functional connectivity · Schizophrenia · Hemispherical distribution of connections · Large-region connectivity

Introduction

Schizophrenia is a neuropsychiatric illness with a global prevalence of approximately 1% (Marín 2012; Walsh et al. 2008) and is more frequently diagnosed during the period from late adolescence to early adulthood (Os and Kapur 2009). A growing body of research is examining the use of machine learning technique in the diagnosis of schizophrenia (Anderson and Cohen 2013; Arbabshirani et al. 2013; Cheng et al. 2015; Davatzikos et al. 2005; Fan et al. 2011; Kim et al. 2016;

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11682-018-9947-4) contains supplementary material, which is available to authorized users.

✉ Junhua Li
juhalee.bcmi@gmail.com; juhalee@nus.edu.sg

✉ Yu Sun
yusun@zju.edu.cn

✉ Rongjun Yu
psyyr@nus.edu.sg

[1] Singapore Institute for Neurotechnology (SINAPSE), Centre for Life Sciences, National University of Singapore, 28 Medical Drive, Singapore, Singapore

[2] Laboratory for Brainbionic Intelligence and Computational Neuroscience, Wuyi University, Jiangmen, China

[3] Centre for Multidisciplinary Convergence Computing (CMCC), School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an, China

[4] Department of Psychology, National University of Singapore, Singapore, Singapore

[5] NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore, Singapore

Published online: 29 August 2018

Shen et al. 2010; Yoon et al. 2012; Yu et al. 2013; Zarogianni et al. 2013). Clinical physicians can give more confident diagnosis supplemented by the diagnostic information derived by machine learning technique. Alternatively, machine learning technique can be utilized to extract discriminative features that represent differential characteristics between the schizophrenia group and the healthy group. For instance, Shen et al. demonstrated that the abnormal functional connectivity between frontal cortex and cerebellum significantly contributed to this differentiation (Shen et al. 2010), implying therefore that inter-regional dysconnectivity might be an important cause for schizophrenia. Additionally, functional connectivity abnormalities in temporal, default mode (Fan et al. 2011), auditory, and visual networks (Arbabshirani et al. 2013) also contributed greatly to schizophrenia differentiation. All these studies utilized the leave-one-out cross validation (CV) to assess the performance of schizophrenia differentiation. This leads to inflated performance and might subsequently bias the detection of functional connections that significantly contribute to schizophrenia differentiation. Therefore, the findings derived from these biased functional connections might not accurately reveal the intrinsic characteristics of schizophrenia.

It has been repeatedly reported that schizophrenia is associated with functional dysconnectivity between disparate brain regions, rather than anomalies only within isolated regions (Fitzsimmons et al. 2013; Friston et al. 2016; Stephan et al. 2009). This functional dysconnectivity implicates numerous brain areas (Pettersson-Yeo et al. 2011; Zarogianni et al. 2013), particularly frontal and temporal regions (Stephan et al. 2009). For instance, using a seed-based functional connectivity method, it has been revealed that the frontal region exhibited reduced connectivity to other regions (e.g., the temporal region and basal ganglia) (Anticevic et al. 2012; Lawrie et al. 2002; Yoon et al. 2013). In contrast, enhanced functional connectivity in the default mode network was also observed during working memory task (Whitfield-Gabrieli et al. 2009). These findings suggested a potential co-existing pattern of elevation and reduction of functional connectivity pertaining to schizophrenia (Pettersson-Yeo et al. 2011; Stephan et al. 2009; Venkataraman et al. 2012), which is widely distributed throughout the brain (Chen et al. 2013; Liang et al. 2006). In addition, Bleich-Cohen et al. showed that diminished inter-hemispheric connectivity was related to schizophrenia through exploring functional connectivity between homologue regions (Bleich-Cohen et al. 2012). This concept of inter-hemispheric dysconnectivity was also supported by a recent study of voxel-mirrored homotopic connectivity (Chang et al. 2015).

Of note, the existing inter-hemispheric investigations only focused on paired connectivity between homologue regions (Bleich-Cohen et al. 2012; Chang et al. 2015; Guo et al. 2014), leaving functional connectivity from one region to all other regions located in contralateral hemispheres unexplored. Furthermore, the proportions of intra-, inter-, left-, and right-hemispherical connections that significantly contribute to schizophrenia differentiation have not yet been determined, which requires further investigation. To date, almost all brain regions have been implicated in schizophrenia according to prior reports, but little is known about the relative contribution of each brain region to schizophrenia. We therefore explored functional connectivity using machine learning technique to address the questions. As mentioned above, the leave-one-out CV might lead to bias detection of functional connections. We proposed n-fold CV to diminish the bias through seeking consensus functional connections that occur in every fold validation.

Materials and methods

Participants

A total of 148 participants were recruited for imaging scanning. Two of them were excluded due to missing category information (i.e., patient or healthy control). One more participant was excluded because of short length of volume scanning. The remaining number of participants is 145. The dataset can be downloaded from http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html, which was released by the Center for Biomedical Research Excellence (COBRE), University of New Mexico (Mayer et al. 2013). All participants included in the imaging scanning did not meet any of the following criteria: (1) history of neurological disorder; (2) history of mental retardation; (3) history of severe head trauma with loss of consciousness greater than 5 min; or (4) history of substance dependence or abuse within the last 12 months. The diagnosis was implemented according to the DSM-IV diagnostic criteria (American Psychiatric Association 1994). At the time of scanning, all patients regularly received antipsychotic medications.

Data acquisition

All scans were performed in a 3-T Siemens Trio scanner with foam padding and paper tape to restrain head motion. All images were acquired parallel to anterior-commissure-posterior-commissure line using an auto-align technique, which included high-resolution T1-weighted MRI and resting-state fMRI. Specifically, the resting-state fMRI was performed with single-shot full k-space echo-planar imaging (EPI) with ramp sampling correction using the inter-commissural line as a reference (volume number = 150, slice number = 33, repetition time [TR] = 2000 ms, echo time [TE] = 29 ms, field of view [FOV] = 256 × 256, acquisition matrix = 64 × 64, voxel resolution = 3 × 3 × 4 mm^3). A five-echo multiecho magnetization-prepared rapid gradient echo sequence was used to acquire a whole brain high-resolution T1-weighted image (TR = 2530 ms, TE = [1.64, 3.5, 5.36, 7.22, 9.08] ms, flip angle = 7°, slab thickness = 176 mm,

FOV = 256 × 256 mm, acquisition matrix = 256 × 256, voxel resolution = 1 × 1 × 1 mm^3).

Data preprocessing

All of fMRI data were preprocessed using statistical parametric mapping (SPM12, http://www.fil.ion.ucl.ac.uk/spm), resting-state fMRI data analysis toolkit (REST) (Song et al. 2011), and data processing assistant for resting-state fMRI advanced edition (DPARSFA) (Yan 2010), operating in a MATLAB programming environment (Mathworks, Inc., Natick, Massachusetts, USA). Briefly, it included the processing steps of the first ten volumes removal, motion correction, slice timing correction, spatial normalization to the Montreal Neurological Institute (MNI) space, signal regression (including regressors of 24 head-motion parameters, cerebrospinal fluid, and white matter), temporal band-pass filtering (0.01 Hz ~ 0.08 Hz), and spatial smoothing with a 6-mm full-width half-maximum Gaussian kernel. In order to eliminate the interference of large head motion, participants were excluded when the maximal inter-scan motion exceeded 2.5 mm translation or 2.5 degrees rotation in any direction. This resulted in the exclusion of 14 participants. The remaining participants consisted of 60 patients (age: 38.42 ± 14.21, mean ± standard deviation; gender: 47/13, male/female) and 71 healthy controls (age: 36.23 ± 11.63; gender: 48/23). In addition, we inspected whether there was significant difference in the residual effect of head motion between groups. We compared the number of volumes with relatively large framewise displacement (FD > 0.4) representing the displacement of each volume as compared to the previous volume, and found that there was no significant difference between the patient group and the healthy group ($t_{129} = 1.51$, $p = 0.134$). The detailed demographic and clinical characteristics of the remaining participants were shown in Table 1.

To define the network nodes, we employed a previously validated, automatically labelled template (Tzourio-Mazoyer et al. 2002) to parcellate the brain into 116 regions of interest (ROIs) (see Supplementary Table S1). The representative time series of an ROI was obtained by averaging the time series of each voxel within that region. Inter-regional connectivity strengths were then calculated between all possible pairs of ROIs using Pearson correlation. Correlation coefficients were subsequently converted to z values by Fisher's r-to-z transformation to improve the normality of values.

Feature selection and classification methods

Kendall Tau method has been used in feature selection for schizophrenia differentiation (Shen et al. 2010; Tang et al. 2012). Additionally, t-test and Fisher's score can also be utilized for feature selection, which were applied to other differentiation tasks such as mental state classification and stroke identification (Li et al. 2016; Rehme et al. 2014). Therefore, we adopted these three methods in our study to evaluate discriminative power of each functional connection. All connections were then sorted in descending order and the top discriminative connections were fed into a classifier. Four

Table 1 Characteristics of the participants involved in this study*

Variable	Group (patients/controls = 60/71)		Statistic
	Patients (SCZ)	Healthy Controls (HC)	
Age (years)	38.42 ± 14.21 (18–65)	36.23 ± 11.63 (18–65)	$t_{(129)} = 0.97$, $p = 0.33$[a]
Gender (Male/Female)	47/13	48/23	$\chi^2_{(1)} = 1.88$, $p = 0.17$[b]
Handedness (Left/Right/Ambidexterous)	10/48/2	1/68/2	$\chi^2_{(2)} = 9.96$, $p = 0.007$[b]
Education (years)	13.16 ± 1.80 (10–20)[c]	13.90 ± 1.73 (10–18)[d]	$t_{(117)} = -2.30$, $p = 0.023$[a]
Age of illness onset (years)	22.54 ± 9.12 (5–61)[e]	–	–
Duration (years)	15.69 ± 12.35 (0–47)[e]	–	–
PANSS Positive Symptoms	15.10 ± 4.82 (7–28)	–	–
PANSS Negative Symptoms	14.53 ± 4.73 (8–29)	–	–
PANSS General Symptoms	29.18 ± 8.33 (16–56)	–	–

PANSS, Positive and Negative Syndrome Scale; *SCZ*, Schizophrenia; *HC*, Healthy Control

*Data are presented as Mean ± SD (Minimum-Maximum) unless otherwise indicated

[a] Two-tailed two-sample t-test

[b] Pearson Chi-square test

[c] Missing for four participants

[d] Missing for eight participants

[e] Missing for one participant

Table 2 Performances of all combinations of feature extraction methods and classification methods

Classifier	Feature selection	Accuracy	Sensitivity	Specificity	PPV	NPV
k-NN	t-test	70.99%	56.67%	83.10%	73.91%	69.41%
	Kendall Tau	70.23%	60.00%	78.87%	70.59%	70.00%
	Fisher's score	68.70%	68.33%	69.01%	65.08%	72.06%
L-svm	t-test	72.52%	66.67%	77.46%	71.43%	73.33%
	Kendall Tau	70.23%	68.33%	71.83%	67.21%	72.86%
	Fisher's score	71.76%	70.00%	73.24%	68.85%	74.29%
RBF-svm	t-test	68.70%	55.00%	80.28%	70.21%	67.86%
	Kendall Tau	67.18%	53.33%	78.87%	68.09%	66.67%
	Fisher's score	67.18%	53.33%	78.87%	68.09%	66.67%
LDA	t-test	76.34%	70.00%	81.69%	76.36%	76.32%
	Kendall Tau	70.23%	66.67%	73.24%	67.80%	72.22%
	Fisher's score	69.47%	68.33%	70.42%	66.13%	72.46%

The best performances for each evaluation metric are highlighted in boldface

k-NN, k-nearest neighbours; *L-svm*, Linear support vector machine; *RBF-svm*, support vector machine with Radial Basis Function kernel; *LDA*, Linear Discriminant Analysis; *PPV*, Positive Predictive Value; *NPV*, Negative Predictive Value

classifiers (k-nearest neighbours, k-NN, parameter k = 5; linear support vector machine, L-svm, parameter C = 1; support vector machine with radial basis function kernel, RBF-svm, parameters C = 1 and Gamma = 1/feature number; and linear discriminant analysis, LDA) were employed, which were recognized as either good performance or classical methods for classification. All possible combinations of feature selection methods and classifiers with different number of selected features were compared using 10-fold CV in this study. Permutation test was used to assess performances of methods in schizophrenia differentiation. More details of methodological definition and description can be found in the supplementary section of methodology.

Results

Differentiation performance

In general, all combinations of feature selection methods and classification methods accomplished comparable performance in distinguishing patients with schizophrenia from healthy people (see Table 2). The highest classification accuracy (76.34%) was achieved by the combination of t-test and LDA using the 442 most discriminative functional connections (see Fig. 1a, and see Supplementary Fig. S1 for other methods). Permutation test (repetition times: 100,000) demonstrated that the classification accuracy was significantly higher than the random ($p < 10^{-5}$, see Fig. 1b, and see Supplementary Fig. S2 for other methods). In terms of sensitivity, specificity, positive predictive value, and negative predictive value, the combination of t-test and LDA was better than the other combinations except the specificity (see the details in Table 2).

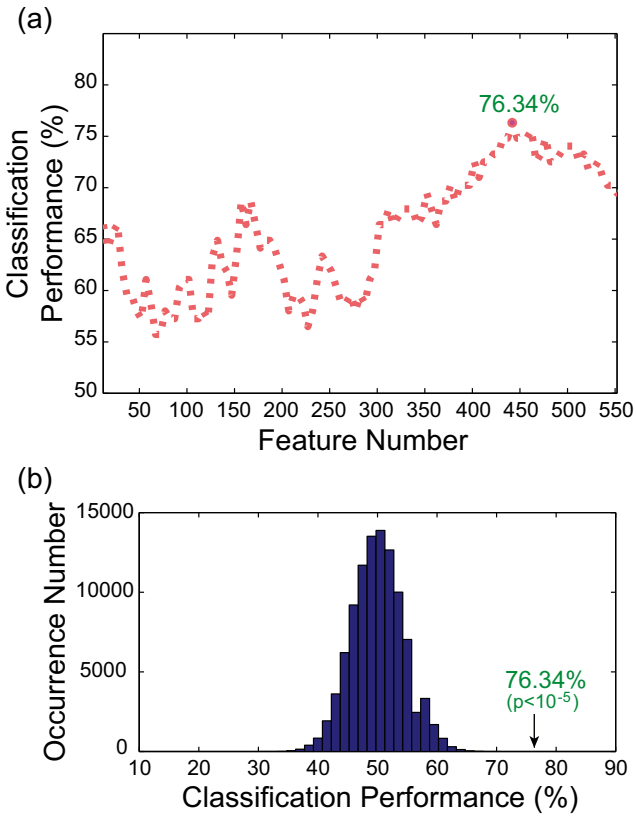


Fig. 1 **a** Classification performance with respect to the number of selected features. All features were first assigned scores that represent discriminative power and then sorted in a descending order. Different numbers of features (from 12 to 552 with incremental step size of 5) were fed to the linear discriminant analysis (LDA) to obtain classification accuracies. **b** The permutation distribution of the estimate using 442 features and LDA. In each permutation, class labels of samples were randomly permuted prior to training and classification. The permutation was repeated 100,000 times, resulting in 100,000 classification accuracies based on random labels. In this case, the classification accuracy (76.34%) based on the true labels exceeded all classification accuracies in permutation, declaring that the classifier can reliably learn the relationship between the features and the labels with a probability higher than 0.99999

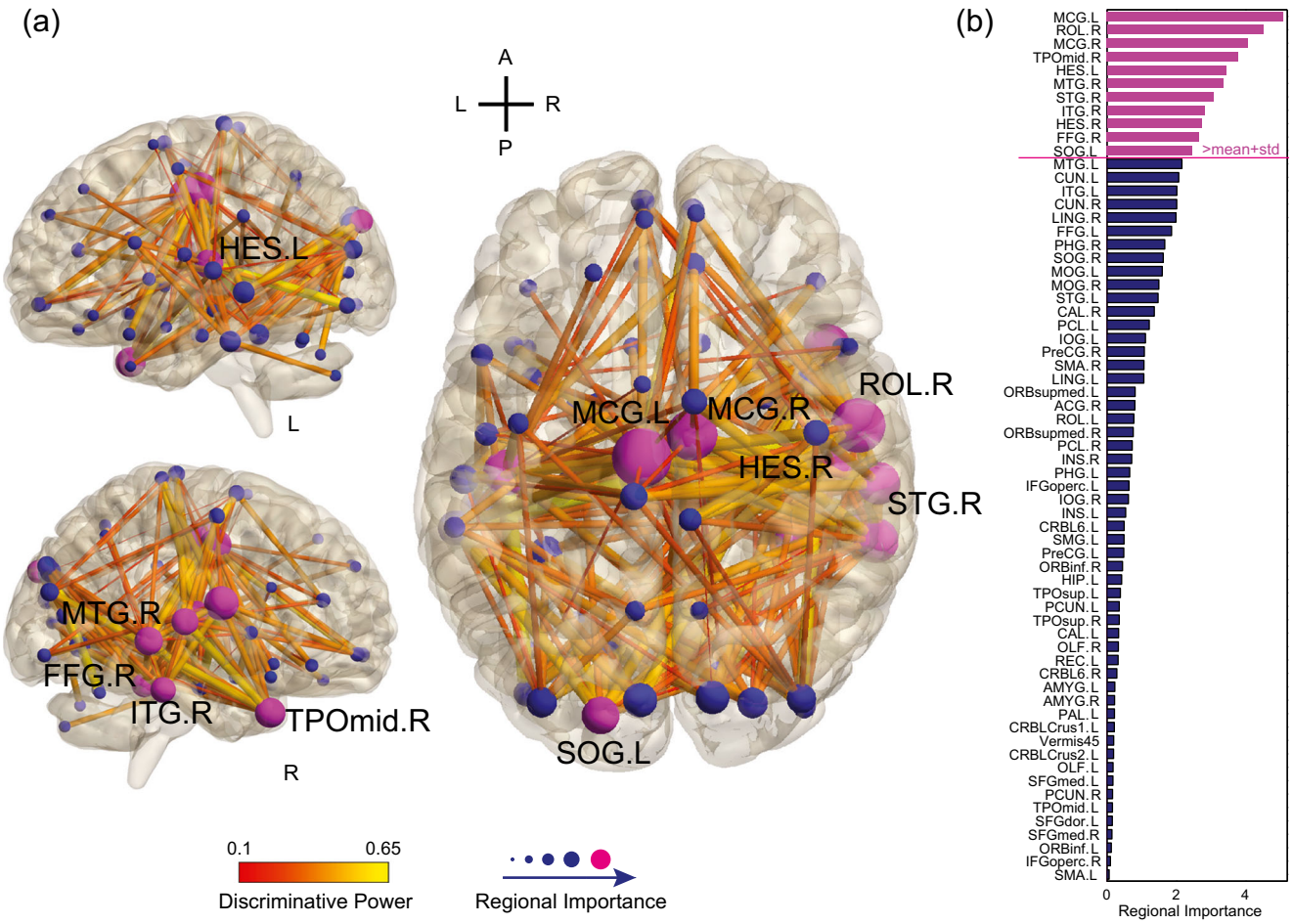


Fig. 2 **a** Left sagittal, right sagittal, and axial views of the consensus connections. Consensus connections were those selected as features at every time in the 10-fold cross-validation. Colour and width of lines encode discriminative power. Sphere size represents regional importance, which is the sum of discriminative powers of all edges connecting to that region. **b** Regional importance. The magenta bars indicates top discriminative regions whose discriminative powers are one standard deviation more than the mean of powers of all discriminative regions (> mean + std). The regions marked by magenta bars on the right panel were also illustrated the same colour in the left connectivity views

Dominant connections and regions

In this study, 442 functional connections were selected by means of the 10-fold CV when the highest performance was achieved. Among them, 235 consensus functional connections were found, which were consistently selected at every time of the 10-fold CV (see consensus connections in Fig. 2a). These consensus connections involved a number of regions spreading across the entire brain. We detected 11 important regions, which are defined by that their discriminative powers are one standard deviation more than the mean of powers of all discriminative regions (see regions highlighted in magenta colour in Fig. 2b). In particular, the temporal lobe contains the majority of important regions. These regions are left heschl gyrus [HES.L], right heschl gyrus [HES.R], right superior temporal gyrus [STG.R], right middle temporal gyrus [MTG.R], right temporal pole, middle temporal gyrus [TPOmid.R], and right inferior temporal gyrus [ITG.R]. Besides, two important regions (i.e., left superior occipital gyrus [SOG.L] and right fusiform gyrus [FFG.R]) reside in the occipital lobe and two other important regions (i.e., left and right median cingulate and paracingulate gyri [MCG.L][MCG.R]) reside in the limbic lobe. Another important region (right rolandic operculum [ROL.R]) resides in the frontal lobe.

Hemispherical distribution

Almost all consensus functional connections (99.57%, 234 out of 235 connections) exhibited fairly weaker strength in the linkages between brain regions for patients with schizophrenia compared to healthy controls (see Fig. 3a). Only one functional connection showed stronger connectivity strength in patients. Furthermore, inter-hemispherical connectivity strength of a region was calculated by averaging connectivity

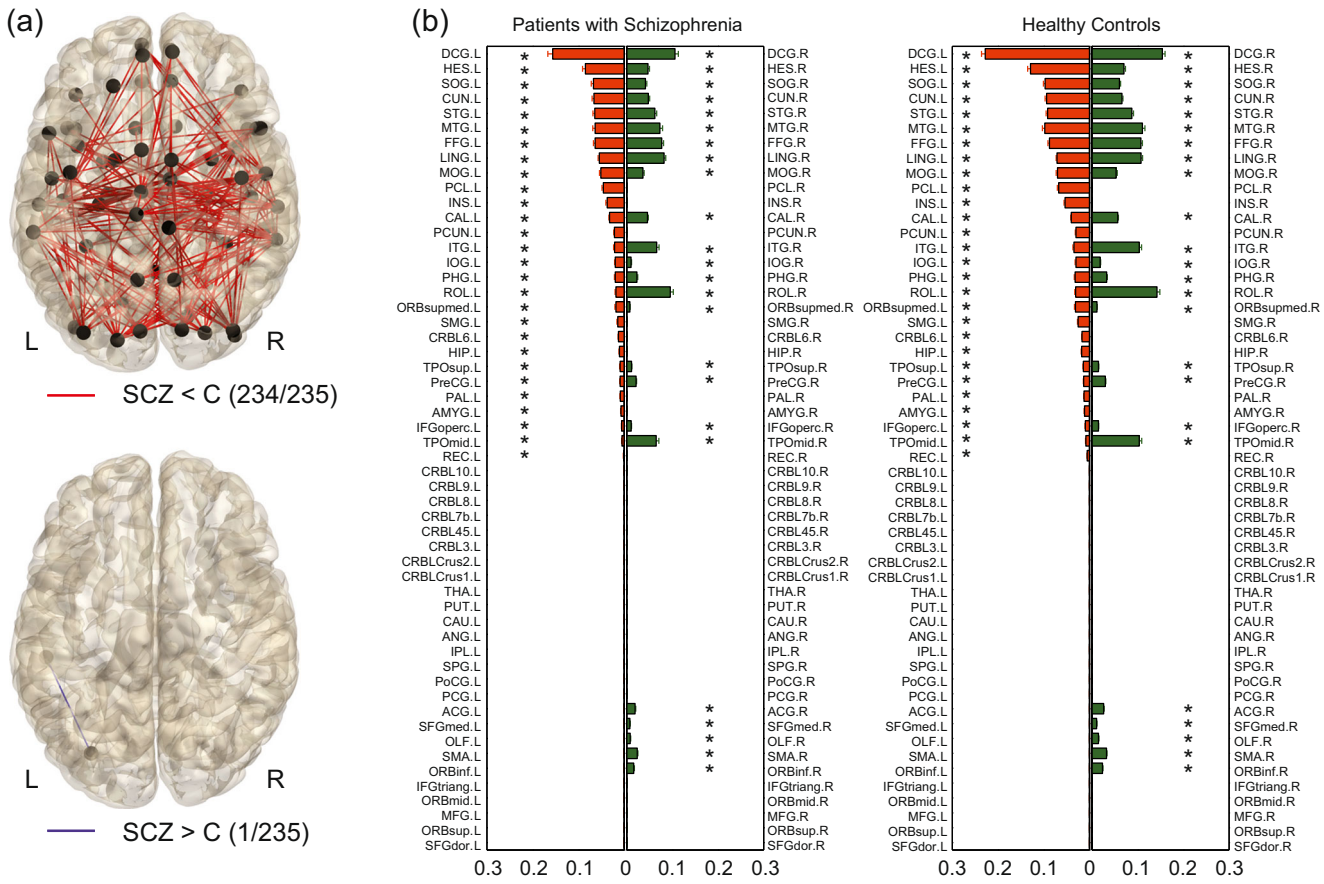


Fig. 3 **a** Comparisons of connectivity strength between patient group and control group for the most discriminative consensus edges, which were selected as features at every time in the 10-fold cross-validation. Red lines in the upper diagram indicate the connections whose strength means averaged across patients were weaker than that averaged across controls. Blue lines in the bottom diagram show the opposite case. **b** Comparisons of inter-hemispheric connectivity strength between patients with schizophrenia and controls. Inter-hemispheric connectivity strength of a region was obtained by averaging across all pairs between this region and all regions located in the opposite hemisphere. Red bars represent inter-hemispheric connectivity strength of each region located in the left hemisphere, while green bars illustrated for regions of the right hemisphere. Error bars show standard errors over subjects. Asterisks indicate regions with significant difference between patients with schizophrenia and controls in inter-hemispheric connectivity ($p < 0.05$, Bonferroni-corrected). All these regions exhibit reduced inter-hemispheric connectivity for patients. For illustrative purposes, regional names shown in the right bar panel (**b**) were sorted in a descending order according to inter-hemispheric connectivity strength of the left hemisphere of patients with schizophrenia

strength across all pairs between this region and all regions located in the opposite hemisphere to investigate inter-hemispherical connectivity. We found a distributed inter-hemispherical connectivity reduction in patients with schizophrenia compared to controls (see Fig. 3b). 28 regions in the left hemisphere and 24 regions in the right hemisphere showed significantly weaker inter-hemispherical connectivity for patients with schizophrenia ($p < 0.05$, Bonferroni-corrected). Within these consensus connections, the number of inter-hemispherical connections is comparable to the number of intra-hemispherical connections (see Fig. 4). However, the number of connections located in the right hemisphere is always more than that of connections located in the left hemisphere regardless of the proportion of consensus connections included for counting.

Connectivity between large regions

Small regions were grouped into 7 large regions (i.e., frontal, insula and limbic, occipital, parietal, subcortical, temporal, and cerebelum) to explore large-region connectivity patterns (see Fig. 5). In the top 25% of the most discriminative consensus connections, within-large-region connections appeared in the temporal and occipital regions (see Fig. 5a). The between-large-region connections were mainly observed from the temporal region to the occipital, insula and limbic regions. This phenomenon was consistent for both intra-hemispherical and inter-hemispherical connections. The same inter-large-region pattern was also preserved when less number of consensus connections (top 10% of consensus connections) was included for the exploration (see Fig. 5b). When all consensus

connections were included for the exploration, this pattern still dominantly appeared, although more connections were observed in other regions (see Fig. 5c).

Discussions

In this study, we identified differential functional connections that greatly contributed to distinguishing patients from healthy people, and further determined consensus functional connections. Almost all consensus functional connections exhibited reduced connectivity strength in patients with schizophrenia compared to healthy people. These disrupted connections involved a number of brain regions, from which 11 important regions were shown to be highly implicated in schizophrenia differentiation. The majority of important regions resided in the temporal lobe. Furthermore, we found that the number of consensus connections between hemispheres (inter-hemispherical connections) was comparable to that of consensus connections within hemispheres (intra-hemispherical connections). However, the number of consensus connections in the right hemisphere was more than that of consensus connections in the left hemisphere. Through the exploration of large-region connectivity, we found that the occipital and temporal regions were predominant in the disruption of within-large-region connectivity while the majority of inter-large-region connections were from the temporal region to the occipital, insula and limbic regions. Our study corroborated the previous finding of dysconnectivity in schizophrenia and revealed new characteristics in the distribution of abnormal connections.

Three feature selection methods and four classification methods were employed for schizophrenia differentiation. The results showed that performance was comparable among different combinations of feature selection and classification methods. Functional connections can be utilized for schizophrenia identification, which was also reported in other studies (Shen et al. 2010; Yu et al. 2013). In these studies, leave-one-out CV was used to evaluate classification performance, which leads to inflated performance (i.e., an optimistic estimate of classification accuracy), especially for the case of small sample size (Anderson and Cohen 2013). To circumvent this problem, our study utilized a scheme of 10-fold CV, which resulted in a more accurate and credible evaluation for the performance. This scheme could also benefit the detection of robust differential functional connections and subsequent analysis because only consensus connections selected at every time of the 10-fold CV are included for further analysis. As an alternative method for detecting intrinsic characteristics in schizophrenia, machine learning technique seems more intuitive and straightforward compared to statistical methods. For statistical methods, a significance level have to be specified so that those connections whose significances exceed this significance threshold can be extracted. The threshold is relatively arbitrarily determined (e.g., setting as 0.01 or 0.05 after multiple comparison correction) and does not follow a definitive manner. This issue is avoided by using machine learning technique. Differential connections can be screened out in the context of maximizing classification performance.

In the exploration of consensus functional connections, the decline of connectivity strength was observed in almost all consensus connections in patient group compared to healthy group and only one functional connection exhibited enhanced connectivity strength. This co-existing pattern of increasing and decreasing in functional connectivity is in accordance with previous findings derived from group analyses (Chen et al. 2013; Öngür et al. 2010; Venkataraman et al. 2012). This finding could also provide a possible explanation why some studies reported decreased functional connectivity for schizophrenia (Bleich-Cohen et al. 2012; Segal et al. 2010; Zhou et al. 2008) while others reported increased functional connectivity (Whitfield-Gabrieli et al. 2009). Besides supporting previous finding of co-existing pattern, our study further addressed that the decrease of functional connectivity is dramatically predominant, reflecting that the disruption of functional connectivity was crucial factor in schizophrenia. It conferred added evidence supporting the notion of dysconnectivity in schizophrenia (Friston et al. 2016; Stephan et al. 2009) from the perspective of machine learning. There is emerging evidence to indicate that schizophrenia was associated with brain structural anomalies Collinson et al. 2009, (2014; Takao et al. 2010), which was also observed across developmental phases of schizophrenia (Wheeler and Voineskos 2014). All these

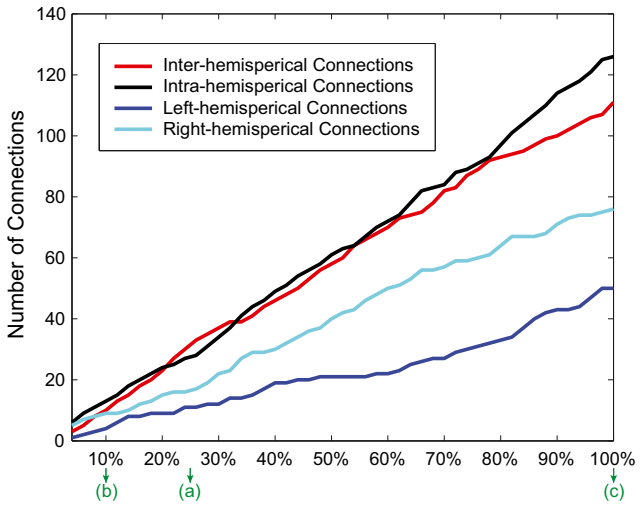
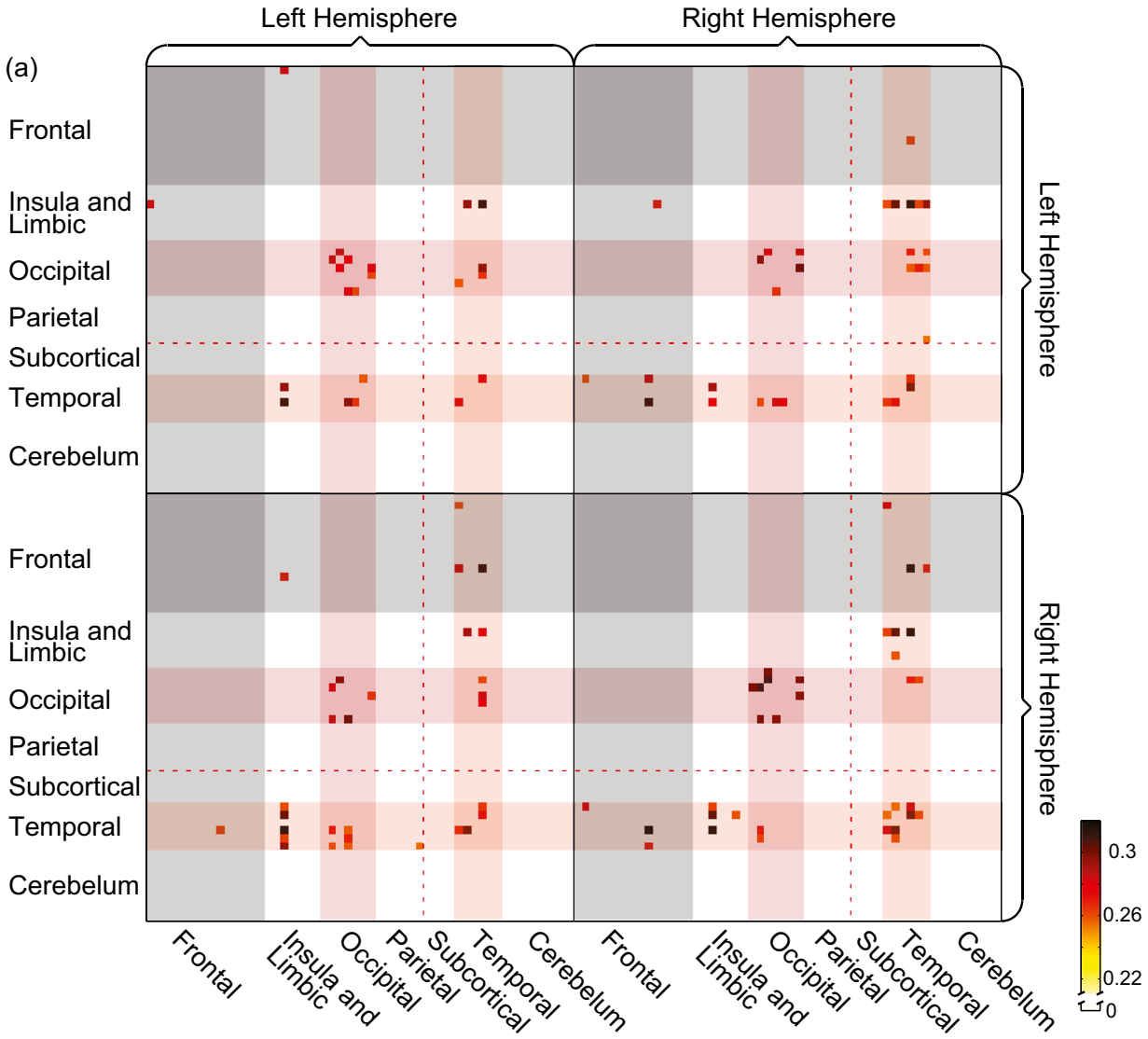
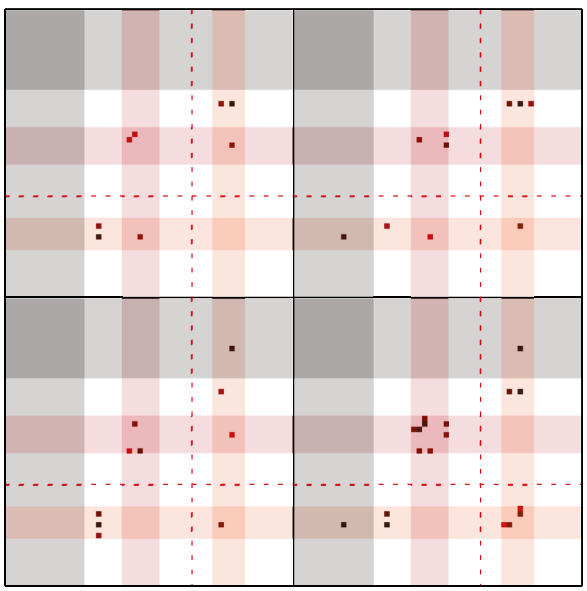


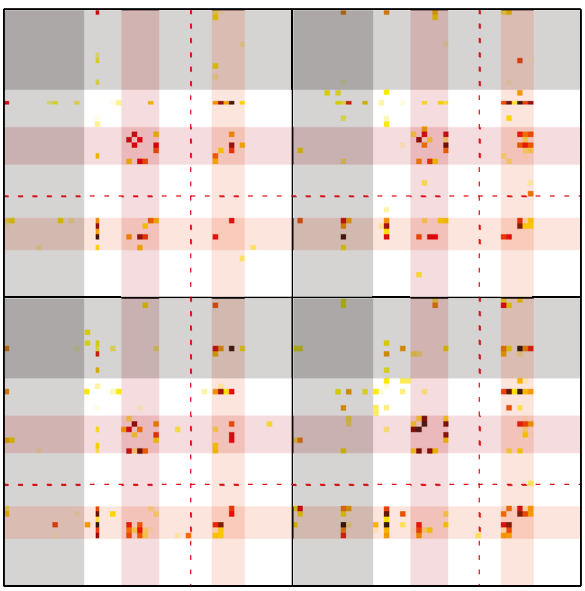
Fig. 4 The numbers of inter-, intra-, left-, and right-hemispherical connections against with the percentages of the consensus functional connections. The number of intra-hemispherical connections is comparable to that of inter-hemispherical connections, but the number of right-hemispherical connections is consistently more than that of left-hemispherical connections. Three typical percentages are selected (indicated by (**a**), (**b**), and (**c**)) and their corresponding connections are illustrated in Fig. 5


Left Hemisphere
Right Hemisphere
(a)
Frontal
Insula and Limbic
Occipital
Parietal
Subcortical
Temporal
Cerebelum
Left Hemisphere
Right Hemisphere
0.3
0.26
0.22
0

(b)



(c)



findings suggest that both structural and functional abnormalities could attribute to schizophrenia (Camchong et al. 2011) and they might be related to each other (Sugranyes et al. 2012). More recent study showed that brain structure impairments might lead to abnormal functional connectivity (Fornito and Bullmore 2015). A broad range of structural anomaly could result in widespread functional dysconnectivity as observed in our study.

The brain regions involving in aberrant connections are widespread according to the observations in our study, which is in agreement with the finding of many previous studies (Andreasen and Pierson 2008; Bleich-Cohen et al. 2012; Chang et al. 2015; Gur et al. 2000; Shen et al. 2010; Zhou et al. 2008; Zipursky et al. 1992). Widespread aberrant connections might explain the diversity of dysfunctions in schizophrenia. The connectivity disruption of a specialized brain area could result in the dysfunction corresponding to that area. As found in our study, abnormal connections largely reside in the temporal and occipital lobes. These lobes are implicated in auditory and visual hallucinations (Ford et al. 2009; Rolland et al. 2015), which are prevalent symptoms in patients with schizophrenia. The regions in the cerebellum were found to be involved in disrupted functional connectivity according to the current study, which might link to abnormal motor behaviour as cerebellum serves to coordinate motor function (Andreasen and Pierson 2008). Abnormalities of functional connectivity in different brain areas could lead to high-level cognitive problem such as disorganized thinking, which generally involves a wide range of brain areas.

The studies exploring homologue connectivity found significant reduction in connective strength (Bleich-Cohen et al. 2012; Chang et al. 2015; Guo et al. 2014; Hoptman et al. 2012), which is in accordance with our finding of strength decline in the inter-hemispherical connections. Our research extended these studies by exploring connectivity from one region to all regions in the contralateral hemisphere, rather than only restricting to homologue connectivity, showing wide connectivity reduction between hemispheres. We speculated that connectivity reduction between hemispheres might be associated with abnormal asymmetry in brain structure (Sun et al. 2015), especially with disruption in the corpus callosum (Collinson et al. 2014). Furthermore, our study explored the hemispherical distribution of consensus functional connections and found that the number of disrupted connections was comparable between inter-hemispherical connectivity and intra-hemispherical connectivity. However, the number of disrupted connections in the right hemisphere was consistently more than that of disrupted connections in the left hemisphere (see Fig. 4), implying severer affection in the right hemisphere. This right laterality is in line with the laterality found in the distribution of important regions (see regions marked in magenta colour in Fig. 2b). The exploration of large-region connectivity revealed that the within-large-region connections were mainly observed in temporal and occipital regions, implying that connectivity disruption in these regions might lead to failures in the processing of auditory and visual information. The inter-large-region connections between temporal region and occipital region were predominately observed, showing the interruption of functional connectivity. Both disrupted within-large-region and inter-large-region connections could result in improper information processing and integration, which might cause hallucinations.

The use of machine learning technique for schizophrenia investigation is shown to be promising. Although the findings in our study are intriguing, there are a few limitations: (1) the sample size used in this study is relatively large compared with the majority of the reported studies (Arbabshirani et al. 2013; Cheng et al. 2015; Davatzikos et al. 2005; Du et al. 2012; Fan et al. 2011; Yoon et al. 2012), but more samples are preferable to draw a more confident conclusion; (2) the effect of treatment medication might confound observations in regional changes and inter-regional connections (Andreasen et al. 2011); However, an opposing viewpoint was also proposed, suggesting that medication is unlikely to be a confounding factor and may even exert a normalizing influence against the randomization (Rubinov et al. 2009). Therefore, the medication effect on functional connectivity is undetermined and requires further investigations; (3) the current study only explored brain connectivity at one time point, which does not allow us to inspect changes over time. To this end, a longitudinal study is required for further exploring brain connectivity with machine learning technique.

Compliance with ethical standards

Conflict of interest RY has received MOE Tier 2 grant (MOE2016-T2–1-015) from the Ministry of Education, Singapore. RY declares that the funder had no role in study design, implementation and data analysis, decision to publish, or preparation for the manuscript, and he has no conflict of interest. The data used in this study are publicly available. The owner of the data declares that all procedures performed in experiments involving human participants were in accordance with the ethical standards of the institutional review board of the University of New Mexico and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All participants provided their informed consent forms.

◀ **Fig. 5** Distribution of connections over large regions for three typical percentages (as indicated in Fig. 4). Squares represent connections while colours show absolute differences of average strengths in connections between healthy controls and patients with schizophrenia. **a** Distribution of connections who are ranked in the top 25% of the consensus connections. **b** Distribution of the top 10% consensus connections. **c** Distribution of all consensus connections (100%)

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, D.C.: APA.

Anderson, A., & Cohen, M. S. (2013). Decreased small-world functional network connectivity and clustering across resting state networks in schizophrenia: An fMRI classification tutorial. *Frontiers in Human Neuroscience, 7*, 520.

Andreasen, N. C., & Pierson, R. (2008). The role of the cerebellum in schizophrenia. *Biological Psychiatry, 64*, 81–88.

Andreasen, N. C., Nopoulos, P., Magnotta, V., Pierson, R., Ziebell, S., & Ho, B. C. (2011). Progressive brain change in schizophrenia: A prospective longitudinal study of first-episode schizophrenia. *Biological Psychiatry, 70*, 672–679.

Anticevic, A., Repovs, G., Krystal, J. H., & Barch, D. M. (2012). A broken filter: Prefrontal functional connectivity abnormalities in schizophrenia during working memory interference. *Schizophrenia Research, 141*, 8–14.

Arbabshirani, M. R., Kiehl, K. A., Pearlson, G. D., & Calhoun, V. D. (2013). Classification of schizophrenia patients based on resting-state functional network connectivity. *Frontiers in Neuroscience, 7*, 1–16.

Bleich-Cohen, M., Sharon, H., Weizman, R., Poyurovsky, M., Faragian, S., & Hendler, T. (2012). Diminished language lateralization in schizophrenia corresponds to impaired inter-hemispheric functional connectivity. *Schizophrenia Research, 134*, 131–136.

Camchong, J., MacDonald, A. W., Bell, C., Mueller, B. A., & Lim, K. O. (2011). Altered functional and anatomical connectivity in schizophrenia. *Schizophrenia Bulletin, 37*, 640–650.

Chang, X., Xi, Y.-B., Cui, L.-B., Wang, H.-N., Sun, J.-B., Zhu, Y.-Q., Huang, P., Collin, G., Liu, K., Xi, M., Qi, S., Tan, Q.-R., Miao, D.-M., & Yin, H. (2015). Distinct inter-hemispheric dysconnectivity in schizophrenia patients with and without auditory verbal hallucinations. *Scientific Reports, 5*, 11218.

Chen, Y. L., Tu, P. C., Lee, Y. C., Chen, Y. S., Li, C. T., & Su, T. P. (2013). Resting-state fMRI mapping of cerebellar functional dysconnections involving multiple large-scale networks in patients with schizophrenia. *Schizophrenia Research, 149*, 26–34.

Cheng, H., Newman, S., Goñi, J., Kent, J. S., Howell, J., Bolbecker, A., Puce, A., O'Donnell, B. F., & Hetrick, W. P. (2015). Nodal centrality of functional network in the differentiation of schizophrenia. *Schizophrenia Research, 168*, 345–352.

Collinson, S. L., Mackay, C. E., OJ, James, A. C. D., & Crow, T. J. (2009). Dichotic listening impairments in early onset schizophrenia are associated with reduced left temporal lobe volume. *Schizophrenia Research, 112*, 24–31.

Collinson, S. L., Gan, S. C., Woon, P. S., Kuswanto, C., Sum, M. Y., Yang, G. L., Lui, J. M., Sitoh, Y. Y., Nowinski, W. L., & Sim, K. (2014). Corpus callosum morphology in first-episode and chronic schizophrenia: Combined magnetic resonance and diffusion tensor imaging study of Chinese Singaporean patients. *The British Journal of Psychiatry, 204*, 55–60.

Davatzikos, C., Shen, D., Gur, R. C., Wu, X., Liu, D., Fan, Y., Hughett, P., Turetsky, B. I., & Gur, R. E. (2005). Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. *Archives of General Psychiatry, 62*, 1218–1227.

Du, W., Calhoun, V. D., Li, H., Ma, S., Eichele, T., Kiehl, K. A., Pearlson, G. D., & Adali, T. (2012). High classification accuracy for schizophrenia with Rest and task fMRI data. *Frontiers in Human Neuroscience, 6*, 1–12. https://doi.org/10.3389/fnhum.2012.00145.

Fan, Y., Liu, Y., Wu, H., Hao, Y., Liu, H., Liu, Z., & Jiang, T. (2011). Discriminant analysis of functional connectivity patterns on Grassmann manifold. *Neuroimage, 56*, 2058–2067.

Fitzsimmons, J., Kubicki, M., & Shenton, M. E. (2013). Review of functional and anatomical brain connectivity findings in schizophrenia. *Current Opinion in Psychiatry, 26*, 172–187.

Ford, J. M., Roach, B. J., Jorgensen, K. W., Turner, J. A., Brown, G. G., Notestine, R., Bischoff-Grethe, A., Greve, D., Wible, C., Lauriello, J., Belger, A., Mueller, B. A., Calhoun, V., Preda, A., Keator, D., O'Leary, D. S., Lim, K. O., Glover, G., Potkin, S. G., & Mathalon, D. H. (2009). Tuning in to the voices: A multisite fMRI study of auditory hallucinations. *Schizophrenia Bulletin, 35*, 58–66.

Fornito, A., & Bullmore, E. T. (2015). Reconciling abnormalities of brain network structure and function in schizophrenia. *Current Opinion in Neurobiology, 30*, 44–50.

Friston, K., Brown, H. R., Siemerkus, J., & Stephan, K. E. (2016). The dysconnection hypothesis (2016). *Schizophrenia Research, 176*, 83–94.

Guo, W., Xiao, C., Liu, G., Wooderson, S. C., Zhang, Z., Zhang, J., Yu, L., & Liu, J. (2014). Decreased resting-state interhemispheric coordination in first-episode, drug-naive paranoid schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 48*, 14–19.

Gur, R. E., Turetsky, B. I., Cowell, P. E., Finkelman, C., Maany, V., Grossman, R. I., Arnold, S. E., Bilker, W. B., & Gur, R. C. (2000). Temporolimbic volume reductions in schizophrenia. *Archives of General Psychiatry, 57*, 769–775.

Hoptman, M. J., Zuo, X. N., D'Angelo, D., Mauro, C. J., Butler, P. D., Milham, M. P., & Javitt, D. C. (2012). Decreased interhemispheric coordination in schizophrenia: A resting state fMRI study. *Schizophrenia Research, 141*, 1–7. https://doi.org/10.1016/j.schres.2012.07.027.

Kim, J., Calhoun, V. D., Shim, E., & Lee, J. H. (2016). Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *Neuroimage, 124*, 127–146.

Lawrie, S. M., Buechel, C., Whalley, H. C., Frith, C. D., Friston, K. J., & Johnstone, E. C. (2002). Reduced frontotemporal functional connectivity in schizophrenia associated with auditory hallucinations. *Biological Psychiatry, 51*, 1008–1011.

Li, J., Wang, Y., Zhang, L., Cichocki, A., & Jung, T.-P. (2016). Decoding EEG in cognitive tasks with time-frequency and connectivity masks. *IEEE Transactions on Cognitive and Developmental Systems, 8*, 298–308. https://doi.org/10.1109/TCDS.2016.2555952.

Liang, M., Zhou, Y., Jiang, T., Liu, Z., Tian, L., Liu, H., & Hao, Y. (2006). Widespread functional disconnectivity in schizophrenia with resting-state functional magnetic resonance imaging. *Neuroreport, 17*, 209–213.

Marín, O. (2012). Interneuron dysfunction in psychiatric disorders. *Nature Reviews. Neuroscience, 13*, 107–120.

Mayer, A. R., Ruhl, D., Merideth, F., Ling, J., Hanlon, F. M., Bustillo, J., & Cañive, J. (2013). Functional imaging of the hemodynamic sensory gating response in schizophrenia. *Human Brain Mapping, 34*, 2302–2312. https://doi.org/10.1002/hbm.22065.

Öngür, D., Lundy, M., Greenhouse, I., Shinn, A. K., Menon, V., Cohen, B. M., & Renshaw, P. F. (2010). Default mode network abnormalities in bipolar disorder and schizophrenia. *Psychiatry Research: Neuroimaging, 183*, 59–68.

Os, J. V., & Kapur, S. (2009). Schizophrenia. *Lancet, 374*, 635–645.

Pettersson-Yeo, W., Allen, P., Benetti, S., McGuire, P., & Mechelli, A. (2011). Dysconnectivity in schizophrenia: Where are we now? *Neuroscience and Biobehavioral Reviews, 35*, 1110–1124.

Rehme, A. K., Volz, L.J., Feis, D.-L., Bomilcar-Focke, I., Liebig, T., Eickhoff, S.B., Fink, G.R., Grefkes, C., (2014). Identifying neuroimaging markers of motor disability in acute stroke by machine learning techniques. *Cerebral Cortex, 25*, 3046–3056.

Rolland, B., Amad, A., Poulet, E., Bordet, R., Vignaud, A., Bation, R., Delmaire, C., Thomas, P., Cottencin, O., & Jardri, R. (2015).

Resting-state functional connectivity of the nucleus accumbens in auditory and visual hallucinations in schizophrenia. *Schizophrenia Bulletin, 41*, 291–299.

Rubinov, M., Knock, S. A., Stam, C. J., Micheloyannis, S., Harris, A. W. F., Williams, L. M., & Breakspear, M. (2009). Small-world properties of nonlinear brain activity in schizophrenia. *Human Brain Mapping, 30*, 403–416.

Segal, D., Mehmet Haznedar, M., Hazlett, E. A., Entis, J. J., Newmark, R. E., Torosjan, Y., Schneiderman, J. S., Friedman, J., Chu, K. W., Tang, C. Y., Buchsbaum, M. S., & Hof, P. R. (2010). Diffusion tensor anisotropy in the cingulate gyrus in schizophrenia. *Neuroimage, 50*, 357–365.

Shen, H., Wang, L., Liu, Y., & Hu, D. (2010). Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. *Neuroimage, 49*, 3110–3121.

Song, X. W., Dong, Z. Y., Long, X. Y., Li, S. F., Zuo, X. N., Zhu, C. Z., He, Y., Yan, C. G., & Zang, Y. F. (2011). REST: A toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS One, 6*. https://doi.org/10.1371/journal.pone.0025031.

Stephan, K. E., Friston, K. J., & Frith, C. D. (2009). Dysconnection in schizophrenia: From abnormal synaptic plasticity to failures of self-monitoring. *Schizophrenia Bulletin, 35*, 509–527.

Sugranyes, G., Kyriakopoulos, M., Dima, D., O'Muircheartaigh, J., Corrigall, R., Pendelbury, G., Hayes, D., Calhoun, V. D., & Frangou, S. (2012). Multimodal analyses identify linked functional and white matter abnormalities within the working memory network in schizophrenia. *Schizophrenia Research, 138*, 136–142.

Sun, Y., Chen, Y., Collinson, S.L., Bezerianos, A., Sim, K., (2015). Reduced hemispheric asymmetry of brain anatomical networks is linked to schizophrenia: A Connectome study. *Cerebral Cortex, 27*, 602–615.

Takao, H., Abe, O., Yamasue, H., Aoki, S., Kasai, K., & Ohtomo, K. (2010). Cerebral asymmetry in patients with schizophrenia: A voxel-based morphometry (VBM) and diffusion tensor imaging (DTI) study. *Journal of Magnetic Resonance Imaging, 31*, 221–226.

Tang, Y., Wang, L., Cao, F., & Tan, L. (2012). Identify schizophrenia using resting-state functional connectivity: An exploratory research and analysis. *Biomedical Engineering Online, 11*, 50.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15*, 273–289.

Venkataraman, A., Whitford, T. J., Westin, C.-F., Golland, P., & Kubicki, M. (2012). Whole brain resting state functional connectivity abnormalities in schizophrenia. *Schizophrenia Research, 139*, 7–12.

Walsh, T., McClellan, J. M., McCarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M., Nord, A. S., Kusenda, M., Malhotra, D., Bhandari, A., Stray, S. M., Rippey, C. F., Roccanova, P., Makarov, V., Lakshmi, B., Findling, R. L., Sikich, L., Stromberg, T., Merriman, B., Gogtay, N., Butler, P., Eckstrand, K., Noory, L., Gochman, P., Long, R., Chen, Z., Davis, S., Baker, C., Eichler, E. E., Meltzer, P. S., Nelson, S. F., Singleton, A. B., Lee, M. K., Rapoport, J. L., King, M.-C., & Sebat, J. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science, 320*, 539–543.

Wheeler, A. L., & Voineskos, A. N. (2014). A review of structural neuroimaging in schizophrenia: From connectivity to connectomics. *Frontiers in Human Neuroscience, 8*, 653.

Whitfield-Gabrieli, S., Thermenos, H.W., Milanovic, S., Tsuang, M.T., Faraone, S. V, McCarley, R.W., Shenton, M.E., Green, A.I., Nieto-Castanon, A., LaViolette, P., Wojcik, J., Gabrieli, J.D.E., Seidman, L.J., (2009). Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. Proceedings of the National Academy of Sciences of the United States of America 106, 1279–1284.

Yan. (2010). DPARSF: A MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Frontiers in Systems Neuroscience, 4*, 1–7. https://doi.org/10.3389/fnsys.2010.00013.

Yoon, J. H., Nguyen, D. V., McVay, L. M., Deramo, P., Minzenberg, M. J., Ragland, J. D., Niendham, T., Solomon, M., & Carter, C. S. (2012). Automated classification of fMRI during cognitive control identifies more severely disorganized subjects with schizophrenia. *Schizophrenia Research, 135*, 28–33.

Yoon, J. H., Minzenberg, M. J., Raouf, S., D'Esposito, M., & Carter, C. S. (2013). Impaired prefrontal-basal ganglia functional connectivity and substantia nigra hyperactivity in schizophrenia. *Biological Psychiatry, 74*, 122–129.

Yu, Y., Shen, H., Zhang, H., Zeng, L.-L., Xue, Z., & Hu, D. (2013). Functional connectivity-based signatures of schizophrenia revealed by multiclass pattern analysis of resting-state fMRI from schizophrenic patients and their healthy siblings. *Biomedical Engineering Online, 12*, 10.

Zarogianni, E., Moorhead, T. W. J., & Lawrie, S. M. (2013). Towards the identification of imaging biomarkers in schizophrenia, using multivariate pattern classification at a single-subject level. *NeuroImage Clin., 3*, 279–289.

Zhou, Y., Shu, N., Liu, Y., Song, M., Hao, Y., Liu, H., Yu, C., Liu, Z., & Jiang, T. (2008). Altered resting-state functional connectivity and anatomical connectivity of hippocampus in schizophrenia. *Schizophrenia Research, 100*, 120–132.

Zipursky, R.B., Lim, K.O., Sullivan, E. V, Brown, B.W., Pfefferbaum, A., (1992). Widespread cerebral gray matter volume deficits in schizophrenia. Archives of General Psychiatry 49, 195–205.