Nodal centrality of the resting state functional network in the differentiation of schizophrenia using a support vector machine

Hu Cheng[1], Sharlene Newman[1], Jerillyn S. Kent[1], Josselyn Howell[1], Amanda Bolbecker[1], Aina Puce[1], Brian F. O'Donnell[1], and William P. Hetrick[1]
[1]*Indiana University, Bloomington, IN, United States*

INTRODUCTION Graph analysis of the resting state fMRI signal can provide a functional connectivity viewpoint on how the 'functional network' as a whole behaves**.** The data are composed of correlations of time courses between different brain regions. Alterations in the functional network have been found in individuals with brain disorders or neurodegenerative diseases[1,2]. Head motion and other data acquisition related sources of variability can affect these correlations[3]. Therefore, we attempted to use selected graph analysis measures that are relatively unaffected by noise and variability to compare the nodal centrality of the resting state functional network of schizophrenia subjects (SZ) and normal controls (NC). By applying a support vector machine on the rank of centrality for a certain number of nodes, we were able to distinguish schizophrenia from normal subjects with a high accuracy rate of 77%.

METHODS Subjects: 19 SZs (12 male, mean age 33.1 ± 10.9 years) and 29 NCs (13 male, mean age 27.8±8.5 years) were completed the study protocol. SZ diagnosis was determined using the Structured Clinical Interview for the DSM-IV (SCID-IV) and medical chart review. MRI data acquisition: Subjects were scanned on a Siemens TIM Trio 3 T scanner using a 32-channel headcoil. The high resolution (1 mm3) anatomical scan was performed with an MP-RAGE sequence. This was followed by the resting state fMRI scan performed with EPI sequences (TR/TE = 2500/30 ms, FOV = 220 mm, 128×128 matrix, iPAT2, 200 volumes), during which the subjects were at rest with eyes closed. Head motion characterization: All functional data were motion corrected in FSL (http://fsl.fmrib.ox.ac.uk/). We computed the translation motion and rotational motion from the output of motion correction[3]. Characterization of the functional network: Aided by the anatomical image, the functional images were parcellated using a new parcellation scheme[4] based on functional homogeneity of resting state functional data of 79 healthy subjects. After regressing out head motion, white matter and the CSF time signal, and band-pass filtering between 0.01-0.10 Hz, time courses were extracted from 281 brain ROIs and averaged. The functional network was obtained from a weighted correlation matrix for the 281 ROIs. For each constructed network, first we computed the mean of the weight of positive edges, and used this as a threshold to remove all edges with weights smaller than the mean. This step can effectively minimize the contribution of variance of the overall correlation strengths of the subjects due to motion or other data acquisition factors. Betweenness centrality: Betweenness centrality is defined as the fraction of all shortest paths in the network that pass through a given node. It characterizes the influence of a single node in the communication between other nodes. It was computed using the Brain Connectivity Toolbox (https://sites.google.com/site/bctnet/). We combined the network of all the subjects to obtain a mean network and computed the betweenness centrality of each node of this mean network. The nodes were then ordered based on their centrality. Ten nodes with highest centrality (identified as hubs) were selected for further analysis. Support Vector Machine (SVM): The SVM analysis was performed on the nodal centrality data using the Matlab toolbox from LIBSVM (http://www.csie.ntu.edu.tw/~cjlin/libsvm/). Because the computed centrality is susceptible to the variability of the constructed functional network and the effectiveness of the SVM may also suffer from large inter-subject variability, four different feature spaces were used: centrality of all nodes, centrality of previously selected hubs, centrality rank of all nodes, centrality rank of previous selected hubs. We adopted the 'leave-one-out scheme' to evaluate the accuracy rate of classifying the subject to the corresponding group based on the centrality information. For different methods, the regularization parameter was adjusted to obtain highest prediction accuracy rate.

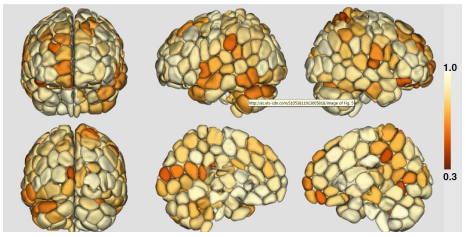
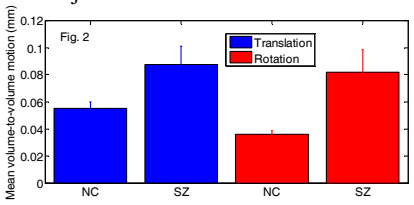


Fig. 1. Parcellation schem with 281 regions. The cross-subject reproducibility is indicated by the shading of each subunit in the parcellation.

RESULTS Figure 2 (right) compares both mean translational (blue) and rotational (red) motions between the SZ and NC groups. Error bars denote standard errors within each group. The SZ group has larger motion than the normal controls. The two sample t-test shows a significant difference for rotation (p = 0.001) and translation (p = 0.01). The variation of the motion is also larger for the schizophrenia subjects.

Prediction accuracies of support vector machine analysis using different methods are shown in Table 1. Although all methods achieve relatively high overall accuracy rates, the specific accuracy on schizophrenic subjects is low when all the nodes were included in the analysis. Also, using rank of centrality rather than actual value gives rise to slightly higher prediction rate.

We also used different sets of nodes in SVM while keeping the same number of features. Fig. 3 shows the prediction rates as we chose 10 nodes with continuous ranks of betweenness centrality starting from lowest to highest. Although many node sets give rise to high overall accuracy rate, only three achieve accuracy rates above 0.6 in predicting SZ.



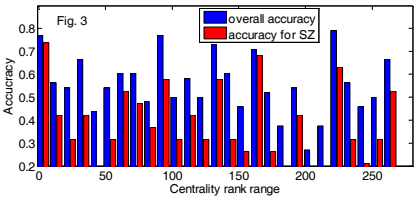


Table 1. Classification accuracy rate for different SVM feature spaces

	Centrality, all nodes	Centrality, top ten hubs	Rank of centrality, all nodes	Rank of centrality, top ten hubs
Overall accuracy	0.73	0.75	0.75	0.77
Accuracy on NC	0.93	0.79	0.90	0.79
Accuracy on SZ	0.42	0.68	0.53	0.74

DISCUSSION We demonstrated that by using a small fraction of nodes with highest betweeness centrality, support vector machine learning can predict if a subject belongs to the schizophrenia or healthy group with reasonably high accuracy. The prediction rate, however, goes down significantly for most other nodes. Our results suggest that schizophrenia can be interpreted as a network disease. Network hubs play an important role, but other nodes may also be involved.

REFERENCES 1. Lynall M. et al., J. Neurosci. 2010; 10:829–843. 2. Pievani M. et al., Lancet Neurol. 2011; 30:9477–9487. 3. Van Dijk K.R., et al., Neuroimage. 2012; 59:431-438. 4. Shen X., et al., Neuroimage. 2013; 82:403-415.