## Supplementary Materials

## 1. Dosage information

Supplementary Table S1 displays the antidepressant administration of SSRI in the three cohorts.

One-way analysis of variance and Pearson chi-square tests were executed in order to determine whether medication types or medication dosages were significantly different across the cohorts. Results show that three cohorts had not significant differences in medication types (*χ2* = 3.683, *p* = 0.159) and escitalopram dosages (*F* = 0.499, *p* = 0.61). Since there were a restricted number of patients who received sertraline, two-sample t test was not performed. However, the average dosages and standard deviations of sertraline were similar between cohort A and B, indicating that the dosages of sertraline did not vary by cohorts. In conclusion, there were not significant differences upon treatment strategies across the cohorts.

## 2. Optimal group threshold selection

The group-averaged weighted *Mgroup* was computed based on all controls across the three datasets. The method was grounded on the rationale that only connections which were detected in a large percentage of the subjects, specified by group threshold, were considered as present. The group threshold was set by finding the minimum difference between *Mgroup* density and the densities of individual networks. The network density was employed as the measure which was defined as the number of existing edges divided by the total number of all possible edges. We explored the optimal group threshold within the range of 30% to 90%, with a step size of 10%. The results are shown in Supplementary Figure S1 which indicates that 50% is an optimal choice.

## 3. Graph kernel Computation

Graph kernel is a type of kernel based on graphs which measure the topological similarities between two graphs. As a result, each sample has a feature map consisting of its similarities with other individuals. In this study, similarity matrix F (Figure S2 C) was constructed and each element *F*(*i,j*) represented the topological similarities between network of sample *i* and network of sample *j*.

The Weisfeiler-Lehman (WL) test was implemented to detect the isomorphism of graphs [1]. Supplementary Figure S2 demonstrates the work flow of graph kernel computation over two networks. The graph kernels were determined by first assigning each node to an original label which was defined as the number of edges that were connected to the node. The original label was then augmented by the sorted set of node labels of its neighboring nodes and these augmented labels were compressed into new labels, called the compressed labels. This process was iterated until the number of repetition reached the predefined maximum value *h* which was set as 10. As a result, there was a label set *L= {l1,l2,**l3,…, ln}* which contained all the original labels and compressed labels of graph *Gi*and *Gj*. Graph kernel was computed by the inner product spaces as follows:

where,

Here, is the number of occurrences in the corresponding label *li* of graph G and is the feature vector which takes into account discrete node labels at different levels. is an inner product between two feature vectors which measures the similarities based on isomorphic structures of graphs *Gi* and *Gj*.

## 4. Reproducibility and validation

#### 4.1. Methods

In order to determine whether the choice of rich club threshold affected the main results, rich club regions were modified. Previous studies indicated that the top 10%-20% nodes with high degrees were selected as rich club regions [2]. With the purpose of testing the reproducibility of our findings on rich club’s definition, rich club regions were selected from the top 10% nodes instead of the original 20% with high degrees and the whole network was re-constructed into five subnetworks. Discriminative performances on these new hierarchical systems were then reappraised.

Discriminative power over individual recognition was explored via information encoding strategy and another pattern characterization based on graph metrics was employed. The pre-defined subnetworks were re-applied so as to compute their topological measures as input features for MVPA, including betweenness centrality [3], clustering coefficient [4], degree [5], and local efficiency [4]. All topological measures were calculated by the Brain Connectivity Toolbox (BCT, http://www.nitrc.org/projects/bct) on Matlab platform. Since these topological measures contained redundant information, we employed the maximum relevant and minimum redundancy (mRMR) strategy to select the most discriminative features [6]. The discriminative pattern based on this new feature design was re-investigated.

#### 4.2. Robustness of subnetworks’ performance in relation to rich-club definition

In order to test the robustness of results regarding rich club level setting, rich club nodes were modified to a 10% threshold and group discrimination were re-assessed over the five re-constructed hierarchical subnetworks in cohort A. The redefined rich club regions are listed in Supplementary Table S4. Similar discriminative performance superiority over these hierarchical subnetworks was achieved. The rich-feeder subnetwork had the best performance when differentiating pre-treatment patients from HCs with an accuracy of 85%, while feeder-local subnetwork had the highest discriminative performance in pre-treatment patients relative to post-treatment patients with accuracy of 74%.

#### 4.3. Consistent subnetworks’ performance with nodal topological features

Four types of nodal topological features were additionally performed for MVPA instead of subnetworks’ graphs in order to assess the robustness of our findings. When comparing pre- to post-treatment patients, the feeder-local network again exhibited the best performance. Notably, there is a long plateau with feature numbers ranging from 36 to 81, indicating that the identification ability was stable despite of feature pools (Figure S4). In distinguishing pre-treatment patients and controls, rich-feeder network displayed the best performance. In contrast to the information encoding strategy with graph structure, nodal topological features still displayed a priority on subnetworks’ discriminative abilities, suggesting consistent network patterns of lesions caused by depression and treatment recovery.

## Reference

[1] Shervashidze N, Schweitzer P, Leeuwen EJv, Mehlhorn K, Borgwardt KM. Weisfeiler-lehman graph kernels. Journal of Machine Learning Research. 2011;12:2539-61.

[2] Daianu M, Jahanshad N, Nir TM, Jack CR, Jr., Weiner MW, Bernstein MA, et al. Rich club analysis in the Alzheimer's disease connectome reveals a relatively undisturbed structural core network. Hum Brain Mapp. 2015;36:3087-103.

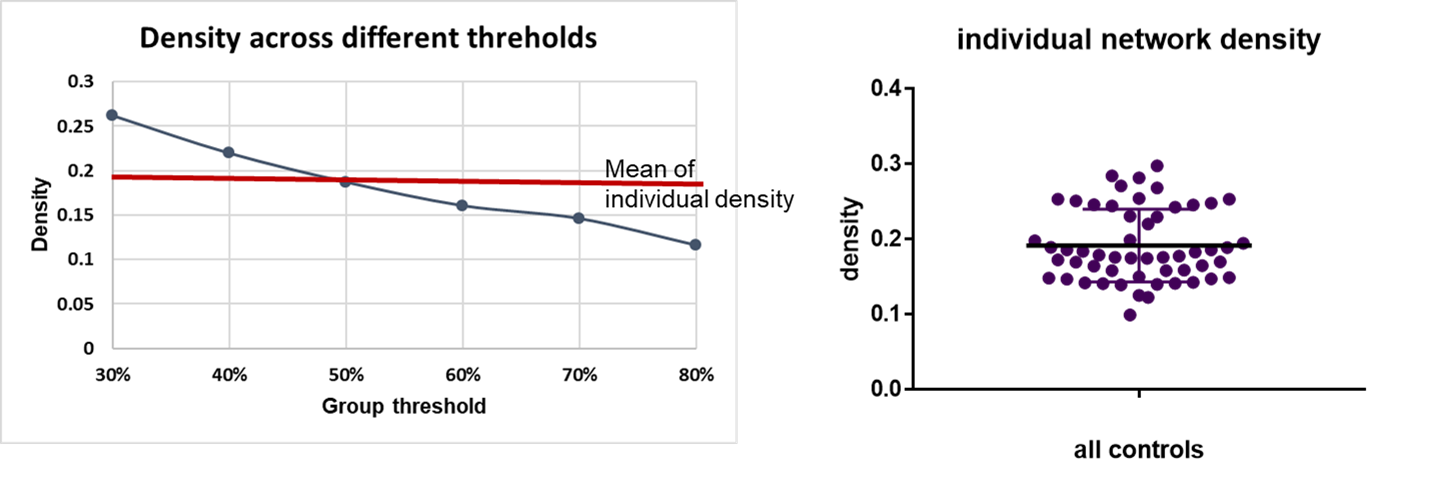
[3] Brandes U. A faster algorithm for betweenness centrality. Journal of mathematical sociology. 2001;25:163-77.

[4] Onnela J-P, Saramäki J, Kertész J, Kaski K. Intensity and coherence of motifs in weighted complex networks. Physical Review E. 2005;71:065103.

[5] Sporns O, Zwi JD. The small world of the cerebral cortex. Neuroinformatics. 2004;2:145-62.

[6] Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence. 2005;27:1226-38.

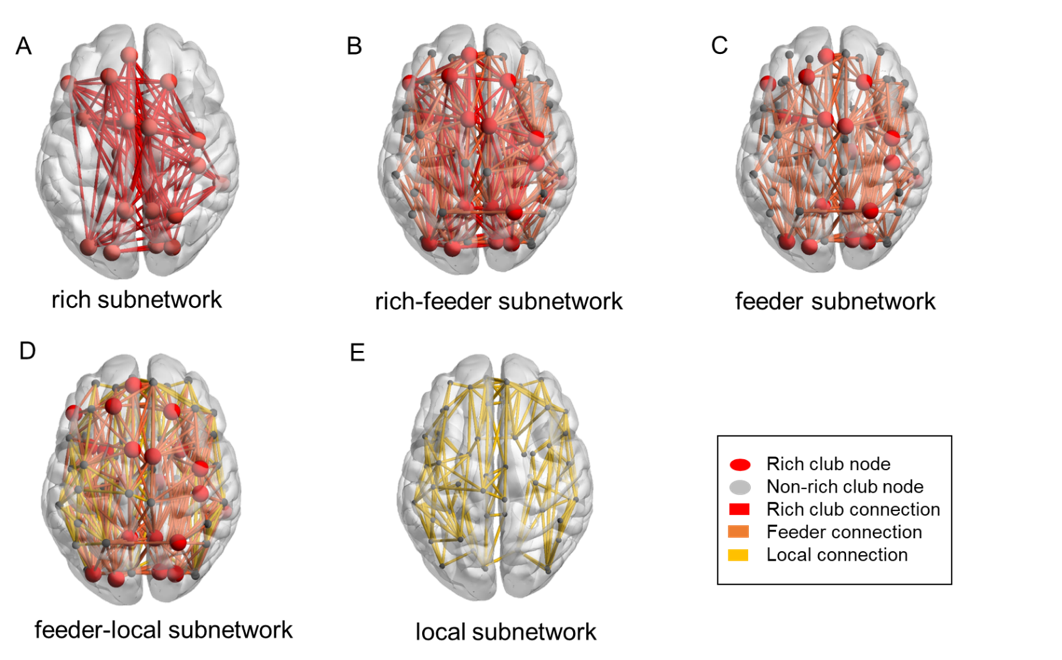
## Supplementary Figures



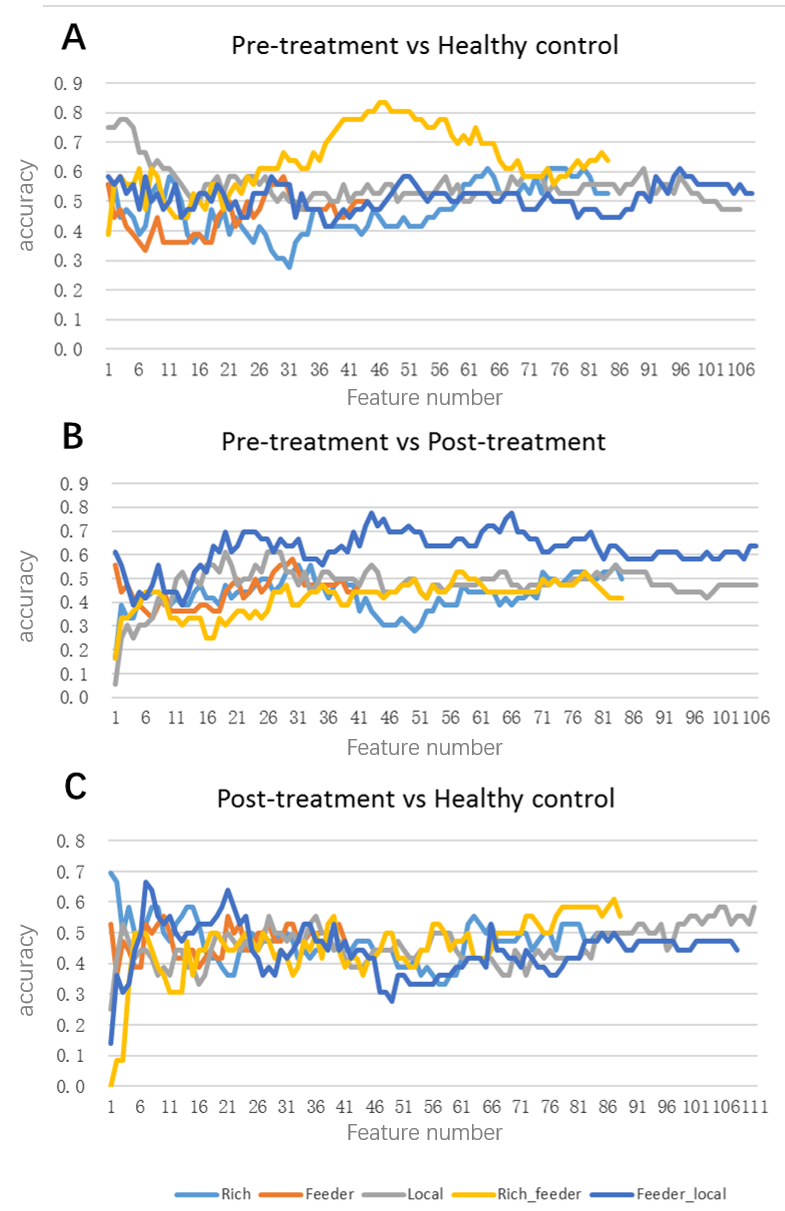
**Figure S1.** Network density over a range of group threshold levels and individual network orignal density. Error bars represent standard deviation.



**Figure S2.** Computation of graph kernel with an iteration of h=10 for two graphs.



**Figure S3.** Graphical representation of five subnetworks. Nodes are colored to indicate node types: rich club nodes in red and non-rich club nodes in grey. Edges were divided into three differing types: rich club connections (red line), feeder connections (orange line) and local connections (yellow line). (A) The rich network constructed via rich club nodes. (B) The rich-feeder subnetwork including feeder connections. (C) The feeder subnetwork comprising of feeder connections and all nodes. (D) The feeder-local subnetwork consisting of feeder connections, local connections and all nodes. (E) The local network consisting of non-rich nodes and local connections.



**Figure S4.** Classification results over the hierarchical subnetworks with nodal topological measures. A: Accuracy curves for comparison of pre-treatment patients versus HCs. Distinct curves represent five types of hierarchical subnetworks. The horizontal axis shows the number of features designated by mRMR feature selection and the vertical axis indicates classification accuracy; B: Accuracy curves for comparison of pre-treatment patients versus post-treatment patients; C: Accuracy curves for comparison of post-treatment patients versus HCs.

## Supplementary Tables

**Table S1.** SSRI administration in MDD patients from three cohorts.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Number of patients (n)** | **Medication** | | | **Dosage** | |
| **Escitalopram (n)** | | **Sertraline (n)** | **Escitalopram (mg/day)** | **Sertraline (mg/day)** |
| **Cohort A** | 18 | 15 | 3 | | 18±4 | 150±0 |
| **Cohort B** | 20 | 16 | 4 | | 17.5±3.95 | 137.5±21.65 |
| **Cohort C** | 17 | 17 | 0 | | 18.82±3.22 | 0 |
| ***p*-value** | - | **0.159a** | - | | **0.61b** | **-c** |

The dosage is presented as mean ± standard deviation; **a**. The *p* value was obtained by Pearson chi-square tests; **b**. The *p* value was obtained by one-way analysis of variance; **c**. The statistic was not performed due to limited sample size.

**Table S2.** Regions in Rich subnetwork.

|  |  |
| --- | --- |
| **Regions Name** | |
| Left dorsolateral superior frontal cortex | Right dorsolateral superior frontal cortex |
| Right postcentral gyrus | Right inferior frontal gyrus, triangular part |
| Left insula | Right superior parietal gyrus |
| Left middle temporal gyrus | Right middle temporal gyrus |
| Left precuneus | Right precuneus |
| Left putamen | Right putamen |
| Left thalamus | Right thalamus |
| Right supplementary motor area | Left superior medial frontal gyrus |
| Left middle occipital gyrus | Left inferior frontal gyrus, opercular part |
| Left superior occipital gyrus | Right superior occipital gyrus |
| Right Calcarine fissure | Right Cuneus |

**Table S3.** Regions in Local subnetwork.

|  |  |
| --- | --- |
| **Regions Name** | |
| Left Precental gyrus | Left Calcarine fissure and surrounding cortex |
| Right Precental gyrus | Left Cuneus |
| Left Superior frontal gyrus, orbital part | Right Superior frontal gyrus, orbital part |
| Left Middle frontal gyrus | Left Lingual gyrus |
| Right Middle frontal gyrus | Right Lingual gyrus |
| Left Middle frontal gyrus, orbital part | Right Middle frontal gyrus, orbital part |
| Right Inferior frontal gyrus, opercular part | Right Middle occipital gyrus |
| Left Inferior frontal gyrus, triangular part | Left Inferior occipital gyrus |
| Left Inferior frontal gyrus, orbital part | Right Inferior occipital gyrus |
| Right Inferior frontal gyrus, orbital part | Left Fusiform gyrus |
| Left Rolandic operculum | Right Fusiform gyrus |
| Right Rolandic operculum | Left Postcentral gyrus |
| Left Supplementary motor area | Left Superior parietal gyrus |
| Left Olfactory cortex | Left Inferior parietal, but supramarginal and angular gyri |
| Right Olfactory cortex | Right Inferior parietal, but supramarginal and angular gyri |
| Right Supramarginal gyrus | Left Supramarginal gyrus |
| Right Superior frontal gyrus, medial | Left Angular gyrus |
| Left Superior frontal gyrus, medial orbital | Right Angular gyrus |
| Right Superior frontal gyrus, medial orbital | Left Paracentral lobule |
| Left Gyrus rectus | Right Paracentral lobule |
| Right Gyrus rectus | Left Caudate nucleus |
| Right Insula | Right Caudate nucleus |
| Left Anterior cingulate and paracingulate gyri | Left Lenticular nucleus, pallidum |
| Right Anterior cingulate and paracingulate gyri | Right Lenticular nucleus, pallidum |
| Left Median cingulate and paracingulate gyri | Left Heschl gyrus |
| Right Median cingulate and paracingulate gyri | Right Heschl gyrus |
| Left Posterior cingulate gyrus | Left Superior temporal gyrus |
| Right Posterior cingulate gyrus | Right Superior temporal gyrus |
| Left Hippocampus | Left Temporal pole: superior temporal gyrus |
| Right Hippocampus | Right Temporal pole: superior temporal gyrus |
| Left Parahippocampal gyrus | Left Temporal pole: middle temporal gyrus |
| Right Parahippocampal gyrus | Right Temporal pole: middle temporal gyrus |
| Left Amygdala | Left Inferior temporal gyrus |
| Right Amygdala | Right Inferior temporal gyrus |

**Table S4.** Re-defined rich club organization.

|  |  |
| --- | --- |
| **Name of Rich Club Regions** | |
| Right supplementary motor area | Left superior medial frontal gyrus |
| Right Cuneus | Left middle occipital gyrus |
| Left superior occipital gyrus | Right superior occipital gyrus |
| Left precuneus | Right precuneus |
| Left putamen | Right putamen |
| Right middle temporal gyrus |  |