Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Hyo Yul Byun 3/25/2014

Support Vector Machine Classification of Resting State fMRI Datasets Using Dynamic Network Clusters

By

Hyo Yul Byun

Helen S. Mayberg, M.D.
Advisor

Department of Mathematics and Computer Science

Helen S. Mayberg, M.D.
Advisor

James Lu, Ph.D.
Committee Member

Patricio Riva Posse, M.D.
Committee Member

Cengiz Gunay, Ph.D

Committee Member

2014

Support Vector Machine Classification of Resting State fMRI Datasets Using Dynamic Network Clusters

By

Hyo Yul Byun

Helen S. Mayberg, M.D.
Advisor

An abstract of

a thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Sciences with Honors

in Computer Science

Department of Mathematics and Computer Science

2014

Abstract

Support Vector Machine Classification of Resting State fMRI Datasets Using Clustered Dynamic Networks

By Hyo Yul Byun

Resting state functional magnetic resonance imaging (rsfMRI) is a powerful tool for investigating intrinsic and spontaneous brain activity. Investigations on rsfMRI data are challenging due to its high dimensionality and the complex nature of brain functioning. The application of univariate and multivariate methods such as multi voxel pattern analysis have been instrumental in localizing neural correlates to various cognitive states and psychiatric diseases. However, many existing methods of rsfMRI analysis are insufficient for investigating the true mechanisms of the brain since they make implicit assumptions that are agnostic to the temporal and spatial dynamics of brain activity.

The proposed method in this thesis aims to create a superior feature space for representing brain activity and to create interpretable generalizations on these features for studying group differences by taking advantage of machine learning algorithms. k-means clustering is used to decompose dynamic resting state functional connectivity networks into discreet and holistic centroid networks. Next, the expression of these ideal centroid networks are computed for each subject and used as a new feature space for a support vector machine classifier. The interpretation of the generalizations computed by

the SVM regarding the classification problem can be revealing since the feature space has been carefully designed to represent discreet dynamic brain states.

The novel method was tested for proof of concept using simulated and randomized data. Experiments show moderate success in classification ability for a real rsfMRI dataset including subjects with major depressive disorder and healthy controls. Although further research and optimization is necessary, the method holds promise as an investigative tool for studying group differences with rsfMRI and may hold future use in clinical applications.

Support Vector Machine Classification of Resting State fMRI Datasets Using Clustered Dynamic Networks

By

Hyo Yul Byun

Helen S. Mayberg, M.D.
Advisor

a thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Sciences with Honors

in Computer Science

Department of Mathematics and Computer Science

2014

Acknowledgements

First, I owe my deepest gratitude to Dr. Helen Mayberg for providing the support, resources, and freedom to peruse this thesis and countless other intellectually stimulating projects. Thank you to lab members: Justin Rajendra, Callie McGrath, Ki Seung Choi, and Dr. Patricio Riva Posse for their assistance, lessons, and laughs.

Thank you to Dr. Cengiz Gunay and Dr. James Lu for lending their time, advice, and expertise for this thesis. This work would not have been possible without their direction and their classroom lessons. Furthermore, thank you to the Department of Mathematics and Computer Science of Emory University.

Also, thank you to my brother Woo for his constant encouragement. Lastly, thank you to my parents Thank Sang Byun and Jin Sook Hong for providing me with life, love, inspiration, and care packages.

Contents

List of Figures

List of Tables

Introduction

Machine learning is a branch of artificial intelligence that deals with algorithms that learn and make generalizations from data. Such algorithms have been applied to a wide range of scientific and commercial pursuits with great success. Not only is their learning and generalizing ability useful for decision making, the mechanisms and models resulting from machine learning systems make them a powerful tool for data mining and extracting usable information from large and complex datasets.

These robust features of machine learning algorithms make them ideal for studying complex neuroimaging datasets. Resting state functional magnetic resonance imaging (rsfMRI) is a neuroimaging technique that is sensitive to spontaneous and natural correlates of brain activity. It is a popular tool used for investigating the mechanisms of the brain and its various disorders. rsfMRI scans produce a 4D image with an extremely large feature space. Such scans can contain a wealth of information, however, extracting useful information from raw scan data remain challenging. Unsupervised machine learning algorithms provide several methods to reduce these high dimensional datasets into lower dimensions that contain less redundant information. However, reducing dimensions while preserving interpretability and information relevant to classification problems can difficult.

In the proposed method, the unsupervised k-means algorithm was used to find discreet and stable rsfMRI network states that appeared across time in subject scans. These clustered network states were then used to compute new feature spaces for subject rsfMRI scans. This was accomplished by calculating the relative expression of each clustered network for each subject scan. Next, supervised classifiers Support Vector

Machines (SVM) were trained to classify between various subject groups within the new feature space for simulated and real rsfMRI datasets containing real subjects. The classification was attempted with a rsfMRI dataset containing subjet groups with major depressive disorder and healthy controls. The performance of SVM on the new feature space was examined while taking the theoretical usefulness and interpretability of the classifier's generalizations into consideration

k-means Clustering

For the following methods, k-means will be used to find stable network states occurring across time in rsfMRI datasets. K-means clustering is an unsupervised learning method that partitions a set of observations into k clusters. In summary, k means is able to find and group observations with similar features together without being provided any labels on the observations. It will find k cluster centroids within the feature space that minimize the sum of squares distance from an observation to the centroid for all points assigned to a cluster. The algorithm accomplishes this by initially setting random centroids and iteratively converging to a local minimum solution. Given a set of clusters S and a training dataset D with observations x_i in p dimensions,

$$D = \{x_i \mid x \in \mathbb{R}^p\}_{i=1}^{n}$$

The k-means algorithm attempts to partition each x_i into a cluster $S_j \in S$ where $k \leq n$ and $|S| = k$. Let μ_j be a vector in p dimensions that is the centroid (mean) of a cluster and d be a distance function, the final optimization problem becomes

$$\min_{S} \sum_{i=1}^{k} \sum_{x_j \in S_i} d(x_j, \mu_j)^2$$

Common distance functions include Euclidian (L2) and Manhattan (L1) distance. Other distance measures based on correlation or cosine similarity can be also used.

In order to solve this optimization problem, an iterative refinement algorithm called Lloyd's algorithm is used. However, before the algorithm can begin, there must be an initial value for the centroid centers. There are two ways of initialing the centroid before applying Loyd's algorithm. The random partition method randomly assigns a cluster in S to all x_i. The Forgy method randomly chooses k observations in x and assigns them as the centroids for a corresponding cluster in S.

Once the centroids have been initialized, the iterative algorithm can begin. The algorithm has an assignment step and an update step. In the assignment step, each observation is assigned to a cluster set S_j that results in the least within-cluster distance.

$$S_j = \left\{ x_i : \ d(x_j, \mu_j)^2 \leq \ d(x_j, \mu_j)^2 \forall \ l, 1 \leq l \leq k \right\}$$

Next, in the update step, the new means of the observations in each cluster is calculated to be the centroid.

$$\mu_j = \frac{\sum_{x_i \in S_j} x_i}{|S_j|}$$

Eventually, the algorithm will converge and there will be no changes to the cluster sets in the assignment step. This algorithm will converge to a local optimum – not neccesarily to the global optimum. The resulting output will include the centroid

coordinates and observation to cluster assignments. The centroids will resemble the ideal features of each cluster or grouping found.

Support Vector Machines

A support vector machine (SVM) is a supervised learning method proposed by Vladimir N. Vapnik and Corinna Cortes in 1995 that can be used to generalize and classify data. It works by finding the optimal hyperplane that discriminates the two classes in the feature space. SVM is a supervised machine learning technique since training datasets are necessary in order to create the generalizations in a model. After being trained, the SVM model can be used to classify new observations not seen in the original training dataset. The traditional algorithm supports classification between two classes for an unbounded number of features. The algorithm was published

The training dataset will be defined as D where

$$D = \{(x_i, y_i) \mid x_i \in \mathbb{R}^p, y_i \in \{-1,1\}\}_{i=1}^{n}$$

Each x_i represents an observation out of n total observations. x_i is a p dimentional vector with a dimension per feature. Each x_i has a corresponding y_i being the class or group. y_i can be either -1 or 1. -1 and 1 can map to any arbitrary set of binary classes depending on the problem. One example may be disease states where -1 can map to patients and 1 can map to healthy control.

The support vector machine must be trained by finding the optimal hyperplane separating the two binary patients within the feature space. A hyperplane can be defined as

$$w \cdot x - b = 0$$

where w is the weight vector and b is the bias. w and b can be scaled to represent any hyperplane in the feature space. Note that in a perfectly separable dataset, there are an infinite number of hyperplanes that perfectly separate between classes but only one hyperplane that maximizes the distance between the margin and x_i. In the case that the data is not perfectly separable, a soft margin must be used. The margins are defined by the equations

$$w \cdot x - b = 1$$

$$w \cdot x - b = -1$$

The distance between the two hyperplanes is $\frac{2}{\|w\|}$. Since the optimal separating hyperplane will have the largest possible margin, $\|w\|$ must be minimized. Note that $\frac{1}{2}\|w\|^2$ will be used without changing the solution for mathematical convenience. The following constraints must be defined in order for the observations to stay outside the margin for all x_i

$$y_i\ (w \cdot x_i - b) \geq 1$$

In many real world implementations, observations are not perfectly separable by a hyperplane. A slack variable ξ_i can be used to create a soft margin

$$y_i\ (w \cdot x_i - b) \geq 1\ - \xi_i$$

Thus the final optimization problem becomes

$$\min_{w, \xi, b} \{\frac{1}{2}\|w\|^2 + \mathrm{C} \sum_{i=1}^{n} \xi_i\ \}$$

subject to for all *i*

$$y_i\,(w \cdot x_i - b) \geq 1 \; - \xi_i$$

and

$$\xi_i \geq 0$$

where C is a tuning parameter that scales the cost of misclassification. Since w can be very high dimensions, and only a subset of x is necessary for classification, the dual form of the optimization problem is solved. Additionally, a kernel trick can be performed by mapping the feature values in x to another high dimensional feature space. The kernel is defined by

$$k(x_i x_j) = x_i \cdot x_j$$

The dual form of the soft margin SVM is the following optimization problem

$$\tilde{\mathrm{L}}(\alpha) = \sum_{i=1}^{n} \alpha_i \; - \sum_{i,j} \alpha_i \alpha_j \, y_i y_j k(x_i x_j)$$

Subject to the constraints for i

$$0 \leq \alpha_i \leq C$$

and

$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

This dual form is also a quadratic optimization problem that can be solved through various methods. Note that there is an optimal and unique solution.

Multi Class Classification

For many real world classification problems, two classes are insufficient. Knerr et. Al. (1990) have come up with a "one-against-one" approach for binary classifiers. If there are k classes, an SVM classifier for every combination of pairs of classes are created. Thus $\binom{k}{2}$ or $k(k-1)/2$ classifiers are constructed. During classification, each classifier "votes" on a class for each observation. The class with the most number of votes will determine the output of the machine.

Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a widely used imaging technique. Its non-invasiveness have facilitated its use various medical and clinical applications. Through modification of the scanning protocols, MRI can be used to investigate a wide range of anatomical and functional properties of the brain and body.

Principles of MRI

Magnetic resonance imaging takes advantage of powerful magnetic fields and the electromagnetic emissions of excited hydrogen nuclei in the body. In order to collect an image, the body is first subject to a strong uniform magnetic field (B_0). The strength of the magnetic field in commercial scanners range from .5 Tesla to 7 Tesla. The strong magnetic field encourages the hydrogen nuclei magic spins to line up parallel to the field. While most will align parallel to the field, some will align in the anti parallel direction. While a very small number of nuclei will align, enough will align in such a way that the net effect will yield a net magnetic vector (M) pointing in the direction of B_0.

Next an electromagnetic pulse at a certain frequency forces M to become transverse to B_0. The electromagnetic radio frequency is applied at the resonance (Lamor) frequency. This frequency is dependent on B_0. When the pulse is over, the magnetic moments of the hydrogen nuclei precess. The precession continues to occur after the excitation is over while slowly returning to alignment with the original magnetic field. Since this precession causes a change in magnetic flux, an electrical signal can be induced and finally read as a voltage by a receiver coil. Since the precession frequency affects the frequency of the change in magnetic flux, a spatial gradient in the magnetic field allows for signal localization. A fourier transform of the spatial frequency domain results in an image.

There are three primary intrinsic factors that affect image contrast. The first is the Hydrogen proton density. There are also two relaxation time constants, T_1 and T_2. These time constants will depend on the chemical and material properties of tissue. By changing parameters in the electromagnetic excitation and signal acquisition, different contrasts in the MRI can be acquired.

Functional Magnetic Resonance Imaging

The development of various functional neuroimaging techniques has allowed researchers to correlate tasks, sensory processes, affective, cognitive, and conative features to discrete areas of the brain. Functional Magnetic Resonance imaging is able to capture signals associated with brain activity.

Functional magnetic resonance imaging is a MRI technique that allows for the measurement of the blood-oxygen-level dependent (BOLD) signal, which is an indirect measure of brain activity. When neurons in the brain are activated, an increase in

metabolism occurs which increases the demand for oxygen. In order the meet the increased demand for oxygen, the supply of oxygenated hemoglobin is increased to the activated area. The amount of oxygenated hemoglobin delivered in response will be greater than the increased demand. Thus, when a neuron becomes active, the ratio of oxygenated hemoglobin increases relative to deoxygenated hemoglobin. This response occurs in around two seconds. The peak ratio of oxygenated hemoglobin occurs around 4-6 seconds before returning to resting levels. Since the degree of magnetization of oxygenated hemoglobin and deoxygenated hemoglobin is different, a contrast occurs in the signal. Despite the low signal to noise ratio of fMRI, it has been used extensively to localize areas of the brain to various tasks and disease features.

Regions of Interests, Functional Connectivity

fMRI can be used to measure brain activity across time across brain areas. Using this information, certain aspects of tasks, cognition, conation, or affect can be correlated to discrete areas in the brain through experiment design. These discrete regions that have been correlated with various functions can be called regions of interests (ROIs).

In order to study the relationship between two or more discrete brain areas, functional connectivity analysis can be used. Functional connectivity is defined as the correlation, covariance, spectral coherence, or phase-locking of the BOLD timeseries between ROIs in fMRI analysis. It is important to note that functional connectivity does not infer any causality or mechanism behind the “connections”. It only gives a measurement of coherence between two regions against the null hypothesis. These connections are hypothesized to correlate to neuronal populations that are firing together with a common purpose due to shared connections (Cole, Smith, & Beckmann,

2010).

Resting State Functional Magnetic Resonance Imaging

Resting state functional magnetic resonance imaging (rsfMRI) is a functional brain imaging technique that attempts to capture brain activity in subjects that are not performing a specific task. rsfMRI datasets have been successfully used in finding features that are correlated to various neuropsychiatric disorders (Greicius, 2008).rsfMRI is a contrast to task based fMRI. In task based fMRI subjects are prompted with a specific task repeatedly so that discrete brain areas may be correlated with the task. In rsfMRI however, subjects are told to simply lie in a scanner and not think about anything in particular, thus measuring the basal activity of the brain. A collection of functional connections, correlation, or cohererence observed across multiple observed in an rsfMRI scan may be defined as a resting state network (RSN). Additionaly, rsfMRI datasets have been successfully used in finding features that are correlated to various neuropsychiatric disorders (Greicius, 2008).

Previous research has found consistent patterns of coherent brain activity across patients during resting states. Studies have found high reliability and replicability of RSNs across different observations and subjects and variabilities of RSNs in diseases. Furthermore rsfmRI is also being utilized to investigate basic mechanisms mediating these basal functional patterns (Biswal et al., 2010). The human connectome project is collecting and making available a large rsfMRI dataset for this purpose (Van Essen et al., 2013).

Dynamic Functional Connectivity

Many studies have examined fMRI datasets by examine the averaging signal or functional connectivity across a defined time period, however, these analysis require the mplicit assumption that brain activity is static across the scanning period. Research has found that temporal dynamics of functional connectivity are lost with an averaged functional connectivity analysis done across an entire scan period (Hutchison et al., 2013). In order to capture the temporal dynamics of functional connectivity, sliding windows are used across the scan period. Sliding windows representing 30-60 seconds in length have been used for scans usually lasting 5 minutes or more. In each sliding window, a functional connectivity analysis is done. Thus, resting state networks are represented across time thought the entire scanning period. Various studies have found reproducible, discreet, brain networks that are represented at differing times during the scanning period (Hutchison et al., 2013).

Limitations of MVPA and Univariate Statistical Analysis of fMRI

Univariate Analysis

fMRI is routinely used in order to localize various processes of the brain to specific areas. Countless investigations have been done to localize neural correlates of various tasks and conditions by comparing against control conditions. By treating each voxel in a functional dataset as an independent variable, they can be tested for increased or decreased activity through univariate statistical analysis. For resting state images, the functional connectivity value of an ROI to other regions can be tested for statistical significane. Univariate methods, such as t-tests or ANOVAs, can be run in parallel for all

voxels to create a statistical map. The general linear model is also commonly used in a massively univariate way by incorporating t-tests or ANOVAs .

While these univariate methods have been excellent at localizing correlates of various brain functions, voxelwise univariate methods are insufficient in capturing the complex interactions involved in the highly connected brain. Voxelwise findings are merely correlates of a network of regions functioning in dynamic cooperation with other sets of regions. In order to study the mechanisms responsible for localized correlates of brain activity, taking advantage a multivariate analysis may better take advantage of the complex interactions between connections and regions.

These univariate methods may also be applied to functional connectivity where the functional connectivity value between two regions may be tested against a hypothesis. In other words, Univarate analysis can be used to test whether the connections between two areas are significantly different between subject groups. While such analyses may get closer to investigating network functioning, again, such approaches are limited since functional connectivity values across two regions are only a single component in a set of connections that dynamically function in combination with others.

Multi Voxel Pattern Analysis (MVPA)

Multi Voxel Pattern Analysis is an increasingly popular multivariate fMRI analysis technique that takes advantage of support vector machines. While MVPA has been applied to various task and disease investigations with great success, the method still suffers from lack of interpretability and high dimensionality. MVPA treats the amplitude of each voxel as a feature. Often, the original feature space is reduced using

various feature selection methods. Unfortunately, the original feature space is immense. For the dataset discussed in this study, the feature space would include a significant fraction of 122,880 dimensions. The predictive power of a classifier decreases with an increased dimension space, training a classifier with a tremendous feature space will often be possible for typical cohort sizes (Hughes, 1968). Furthermore, optimal feature selection is an intractable combinatory problem since there are $2^{\#features}$ possible sets of features to select from. While SVM is robust against overfitting due to regularization, statistical power is reduced when feature selection is performed with limited observations with immense dimensions. This suggests that a theoretically informed method of feature space reduction is needed.

Even if MVPA's feature space was restricted to an intelligently selected set of voxels or ROIs, the generalization of the classifier offers limited interpretability. Similar to univarate methods, MVPA will merely localize combinations areas that most effectively classify between groups. MVPA is difficult to interpret unless one takes the assumption that only one static pattern of brain activity is responsible for certain subject classes. If there is a mixture of relevant brain activity patterns occurring temporally independent of each other, the MVPA method will not be able to resolve the independent components. While MVPA will resolve multivariate relations between areas, it does not allow for the accurate delineation the mechanism behind these relations. The MVPA method is agnostic to the true dynamic and connected nature of brain activity.

Major Depressive Disorder

In order to test the performance and behavior of the method to be discussed, a dataset containing scans from subjects diagnosed with Major Depressive Disorder will be used. Major Depressive Disorder (MDD) is a widespread illness that causes profound human suffering and incurs large economic costs. It is mainly characterized by depressed mood and diminished interest or pleasure in activities. Additional symptoms include changes in weight and sleep, increased agitation, psychomotor retardation, fatigue, feelings of worthlessness or excessive guilt, diminished ability to concentrate, indecisiveness, and thoughts of death. These symptoms can cause significant impairment in social, work, or other important areas of everyday functioning (Center for Substance Abuse Treatment, 2008). At 2010, the prevalence of MDD worldwide was around 4.33% (Vos et al., 2013). Not only does Major Depressive Disorder inflict profound suffering to a significant portion of the global population, the disease comes with major economic costs. Depression's economic burden in the United States alone was estimated to be $83.1 billion in 2000. Of the $83.1 billion, $26.1 billion (31%) were direct medical costs, $51.5 billion (62%) were work related costs, and $5.4 billion were suicide related mortality costs (Greenberg et al., 2003).

A diverse range of treatments exists for MDD: the most common first line treatments include medication and psychotherapy. However, remission to first line treatment rarely exceeds 40% and treatment remission is not uncommon. A subgroup of 24% of patients receiving typical treatment achieve remission in their depressive episode (Blais et al., 2013). There is ongoing research on biomarker that may predict treatment response (McGrat et al. 2013). This study will test the proposed method with rsfMRI scans from depressed and control groups in order to investigate the performance

and utility of the method as a an investigative tool and as a potential clinical classifier for eventual treatment selection.

Methods

fMRI provides a wealth of high dimensional information regarding brain activity. Since brain activity derived from the BOLD signal is a function of a complex and dynamic system, investigating mechanism and function of brain activity is an extremely difficult process. While univariate analysis of fMRI data is informative in revealing various correlates of the BOLD signal, it is insufficient for investigating the deeper layers of systems responsible for the localized correlates found from univarte analysis.

The following method first decomposes many resting state fMRI scans into a finite but interpretable set of brain network states by taking advantage of clustering algorithms. Next, it quantifies the expression of these brain state networks for each subject such that they their variabilities across subject classes can be studied through the use of a classifier algorithm. The method will be utilized on a rsfMRI dataset collected for investigations into major depressive disorder.

Subjects

This study was done with anonymized data collected from multiple studies being conducted by the Emory University School of Medicine's Department of Psychiatry and Behavioral Sciences. These studies were approved by the Emory Institutional review board and informed consent was provided by all subjects. The subject groups were the healthy control group (HC), major depressive disorder group(MDD), and treatment

resistant depression group (TRD). The subjects in the healthy controls group had no current or prior instances of major depressive disorder or major psychiatric illnesses. The major depressive disorder group included patients diagnosed with clinical depression of mild to moderate severity determined by the Hamilton Depression Rating Scale. The treatment resistant group had diagnoses of severe depression and numerous past instances of depressive episodes with failure to respond to multiple treatments.

Image Acquisition

Scans were acquired for all subjects using a 3.0 Tesla Siemens Tim Trio human MRI whole body scanner. T1 weighted anatomical scans were collected using an optimized magnetization-prepared rapid gradient-echo imaging protocol (MP-RAGE). The echo time (TE) was 5 milliseconds with a repetition time (TR) of 35. Each TR represents a frame or volume. The resulting image was a 3D matrix with dimensions 256 x 208 x 196 at 1mm isotropic resolution.

The resting state functional magnetic resonance (rsfMRI) images were T2* weighted echo-planar images. Subjects were ordered to fixate on a crosshair with eyes open. The zSAGA sequence was used (Heberlein & Hu, 2004) in order to minimize sinus-cavity artifacts often seen in fMRI accusations. The parameters used were a repetition time (TR) of 2920 ms, echo time (TE) of 35 ms, and flip angle (FA) of 5 degrees. Each resulting image at each time point was a 64 x 64 x 30 dimension image. All scans had at least 140 TRs. The resulting final image format was a 4D 140 x 64 x 64 x 30 DICOM image.

rsfMRI Image Pre-Processing

The resting state functional magnetic resonance images were preprocessed using the Analysis of Functional NeuroImages (AFNI) toolkit from the NIMH (Cox, 1996) and the FMRIB Software Library (FSL) from the FMRIB Analysis Group from the University of Oxford in the UK (Smith et al., 2004).

The dicom outputs of the T1 anatomical and EPI resting state scans were first converted to NIFTI format. The first TR was removed in order to avoid T1 saturation effects on the EPI images. The images were corrected for slice timing using the Fourier method. The slice timing correction interpolates the signal such that all slices contained in a volume are from the same time point. Next, the EPI volumes were aligned with the base volume. The anatomical images were skull stripped and a dilated whole brain mask was created for all subjects. Transforms for the resting state EPI and T1 anatomical images were computed for alignment, and then transforms for anatomical space into the standard MNI space was computed for all subjects. The resting state EPI images were segmented by white matter and CSF. Local slice by slice white-matter signal regressors were created and hardware related artifacts were regressed out from the EPI images (Jo et al, 2010). The resting state EPI images were de spiked , each voxel time series was scaled to a mean of 100, and then detrended. Six degrees of motion were regressed out including changes in the superior-inferior axis, anterior-posterior axis, left-right axis, roll degree, pitch degree and yaw degree. The EPI images were bandpass filtered in the range of 0.01 to 0.1 Hertz. Next, the EPI images were aligned with the MNI standard brain template using the transforms computed previously. The resulting data was blurred with an 8mm Gaussian kernel. The resulting images were carefully quality

controlled for excessive motion or artifacts. Any images not meeting quality control standards were not used in this study.

Regions of Interest Signal Extraction

A set of 40 regions of interest (20 ROIs unilaterally) was selected due to their relevancy in MDD from previous research. Specific ROI coordinates were defined in standard MNI space by an experienced neuroanatomist. These coordinates would define the center of an aliased sphere of 5mm radius in MNI space with 1 x 1 x 1 mm spatial resolution. The radius was selected for optimal coverage of all brain areas while staying within signal boundaries of the fMRI images. While the expected volume for a 5mm sphere is 523.599 mm^3 , due to aliasing, the volumes of the ROIs were 485 mm^3 / voxels. Binary masks for all ROIs were rendered in AFNI. The set of these 40 ROIs will be referred to as R containing $\{r_1, r_2, r_3, \dots, r_{40}\}$. The 40 selected ROIs locations can be seen in figure 1 and table 1 in the appendix.

Next, timeseries for the ROIs were extracted for all subjects using AFNI. This extraction was done on a Debian system with an Intel Core –i7-2600K at 3.4 GHz CPU with 16 GB of RAM. For each subject, for each ROI, the BOLD signal was first averaged across all voxels for all 40 ROIs.

Collection of Dynamic Functional Connectivity Networks

Dynamic functional connectivity matrices were generated and collected for all subjects. The dynamic functional connectivity of an rsfMRI scan of a subject will be a series of functional connectivity matrices across time. All following analysis were completed on Mathworks' Matlab R2013a on a Windows 7 computer with a AMD Phenom II X6 1045T Processor at 2.7GHz with 8 GB of RAM.

Given a rolling window length w, speed v, and timeseries length l, $\left\lceil \frac{l-w}{v} \right\rceil$ is the number of rolling windows available for sampling in a timeseries of length l. Some time points at the end may be excluded from the rolling window analysis depending on the three variables. The rolling window length w and speed v are adjustable parameters. For this study, the window length w was chosen as 10 TRs and the speed/overlap v as 4 TRs. This gave an overlap percentage of 40% between consecutive windows. These values define the temporal resolution of the set of dynamic function connectivity matrices for a given subject. Generally, choosing lower values of w and v would be preferable for this study with the constraint that w be greater than 2 due to the following Pearson correlation computation. The window length and speed/overlap was chosen to minimize leftover TRs and to optimize computation time while maintaining enough spatial resolution for a single window to capture discreet cognitive states. Unfortunately, the last four TRs were excluded from the analysis given that the rsfMRI datasets were 149 TRs in length.

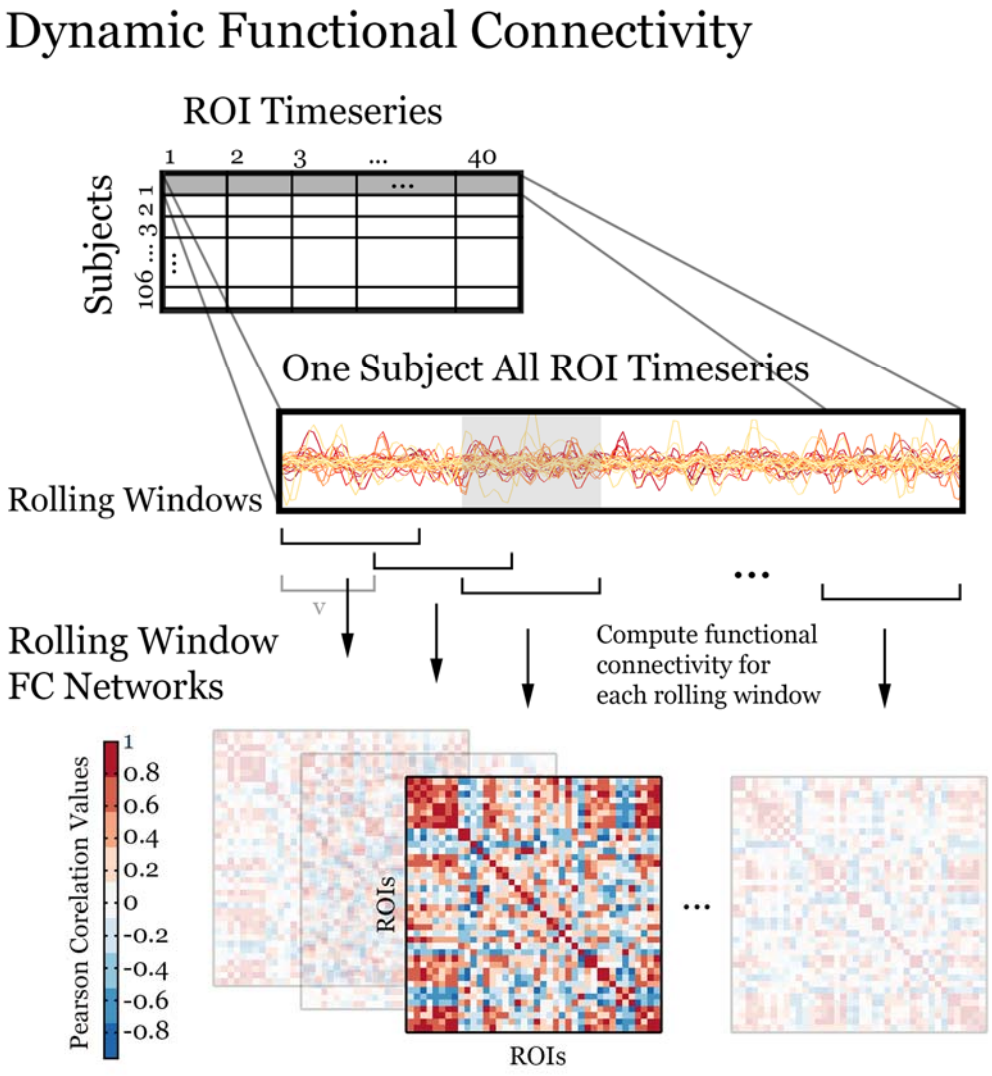

Figure 2. Dynamic Functional Connectivity Computation.

A functional connectivity matrix M was computed for each subject k for the t^{th} rolling window. Each M is a 2 dimensional 40 x 40 matrix. The length of each dimension represents an ROI in our ROI set R. M_{ij} is the Pearson correlation coefficient between the timeseries of ROIs r_i and r_j for one subject within a single rolling window. Thus, M_{ij} is the functional connectivity value between ROIs r_i and r_j. There will be a total of $\left\lceil \frac{l-w}{v} \right\rceil$ Ms computed for each subject.

Next, all functional connectivity matrices (Ms) generated for all subjects were collated into a single 4 dimensional data structure C. C_{ktij} would refer to the functional connectivity between ROIs r_i and r_j for the t^{th} rolling window for the k^{th} subject. In this study, the initial dimensions of C was (106,17,40,40).

The collection of dynamic functional connectivity matrices C was resized in preparation for the following analysis. First all observed functional connectivity matrices across rolling windows and subjects were combined into one dimension by combining the 1st and 2nd dimensions. Next, the 3rd and 4th dimensions representing the functional connectivity between two ROIs were flattened into a single dimension, making the final dimension of C to be (1802,1600). In summary, C is now a matrix containing 1,802 windowed functional connectivity networks. In summary, the dataset C represents a large pool of 1,802 cognitive state network observations with a feature space of 1,600 functional connectivity values between all possible pairs of ROIs.

k-means Clustering of Dynamic FC Windows

Next, the dataset C containing 1802 observations of windowed function connectivity networks were clustered into k clusters using the k-means implementation included in Mathworks' Matlab R2013a. In essence, this clustering step partitions all observed functional connectivity networks from all subjects and rolling windows into k clusters. The optimal value of k depends upon the fundamental structure of the input dataset which is usually not known at the time of clustering. While several algorithms exist for the estimation of k, since the clustering step is a precursor to the SVM evaluation, k is searched for during the parameter search step with the leave-one-out

cross validation success percentage as the optimization criteria. This will be elaborated upon in the following sections.

Assuming that we have chosen a *k*, the k-means clustering algorithm will minimize the sum of the Manhattan distances between all points to the respective centroid for all clusters. Manhattan distance was chosen over Euclidian distance due to research suggesting that Manhattan distance is a superior metric for high dimensional spaces (Aggarwal,Hinneburg, & Keim, 2010). The output of the k-means clustering algorithm will include a vertex labeling all input data points (dynamic FC network windows) to one of k clusters and cluster means (centroids) for the k clusters. For our analysis, the vertex labeling is discarded as we are only interested in taking advantage of the clustering mechanism of k-means.

The other output of the k-means algorithm will be a *k* by 1600 matrix containing *k* centroids. Each centroid can be reconstructed to form a 40 x 40 matrix representing a functional connectivity network. The set of centroids can be thought of as an idealized clustering of network states observed in all subject datasets across time.

Computation of Subject-Centroid Similarities

After clustering, the next step is to compute the similarities between all sets of subjects and dynamic network centroids. This was done by computing the Euclidian distance between the centroid networks and whole scan resting state networks for each subject. This gave relative metric for the level of expression of the clustered networks for each subject.

The whole scan resting state networks can be generated in a method similar to the dynamic functional connectivity matrices. One matrix will be generated for each subject- this is equivalent to a single "rolling" window where window length equals the scan length. The whole scan resting state network is an average of a subject's functional connectivity network over the entire period of a scan. Each resulting subject's average resting state network will be a 40 x 40 matrix with each entry representing the Pearson correlation coefficient between a pair of two ROIs. This matrix will also be flattened to a 1,600 length vector.

Now with the average resting state network for all subjects, the Euclidian distance between each subject's network and the dynamic network centroids will be computed. Since there are 106 subjects and *k* centroids, t his will result in a 106 by k matrix. Each entry in this matrix will represent a relative metric for the level of expression of a clustered network for each subject.

Support Vector Machine Training and Cross Validation

The subject-centroid similarities will function as the input for the SVM Classifier. LIBSVM version 3.17 (Chang & Lin, 2011) was used with a MATLAB interface. The subject-centroid similarities and the group labels for each subject was input. The group sets checked within the 106 dataset include {MDD & TRD, HC} and {MDD, TRD, HC}. Since SVMs are normally formulated as a binary classifier, the voting method was utilized for multi class problems as discussed in the introduction.

There are several adjustable parameters for the SVM. First is the kernel function- for maximum interpretability and generalization of the output, a linear SVM was used. The two main parameters of concern in this stage are the C and ε-insensitive loss

function parameters of SVM. Both C and ε are depend on the fundamental structure and distribution of the training dataset. Also, the incorrect selection of these parameters will increase overfitting or negatively affect the classifier's ability to generalize. The parameter C controls the balance between training errors and hard margins. The parameter ε determines the level of satisfactory accuracy in the SVM training step and has been found to scale with the range of values in the input dataset and noise. Each of these parameters were investigated for our dataset during the parameter search.

Leave One Out Cross Validation

When evaluating a performance of an SVM classifier with an input dataset, the successful evaluation of the same input dataset is of little significance due to the possibility of overfit. Thus, the classification ability of an SVM defined with parameters C and ε for an input dataset was measured using leave one out cross validation (LOOCV).

In order to perform LOOCV, multiple SVMs must be trained and evaluated. For each input observation (subject), it is excluded from the input dataset. Next, the SVM is trained using the given parameters and input dataset excluding a subject. After training, the excluded observation is classified by the trained classifier. Next, the resulting classification label is compared to the actual class of the training observation. This step simulates the practical use of the classifier. Thus an SVM is trained to exclude every subject in the input dataset. The reported LOOCV accuracy is the average success rate of the classification of the excluded subjects for all subjects.

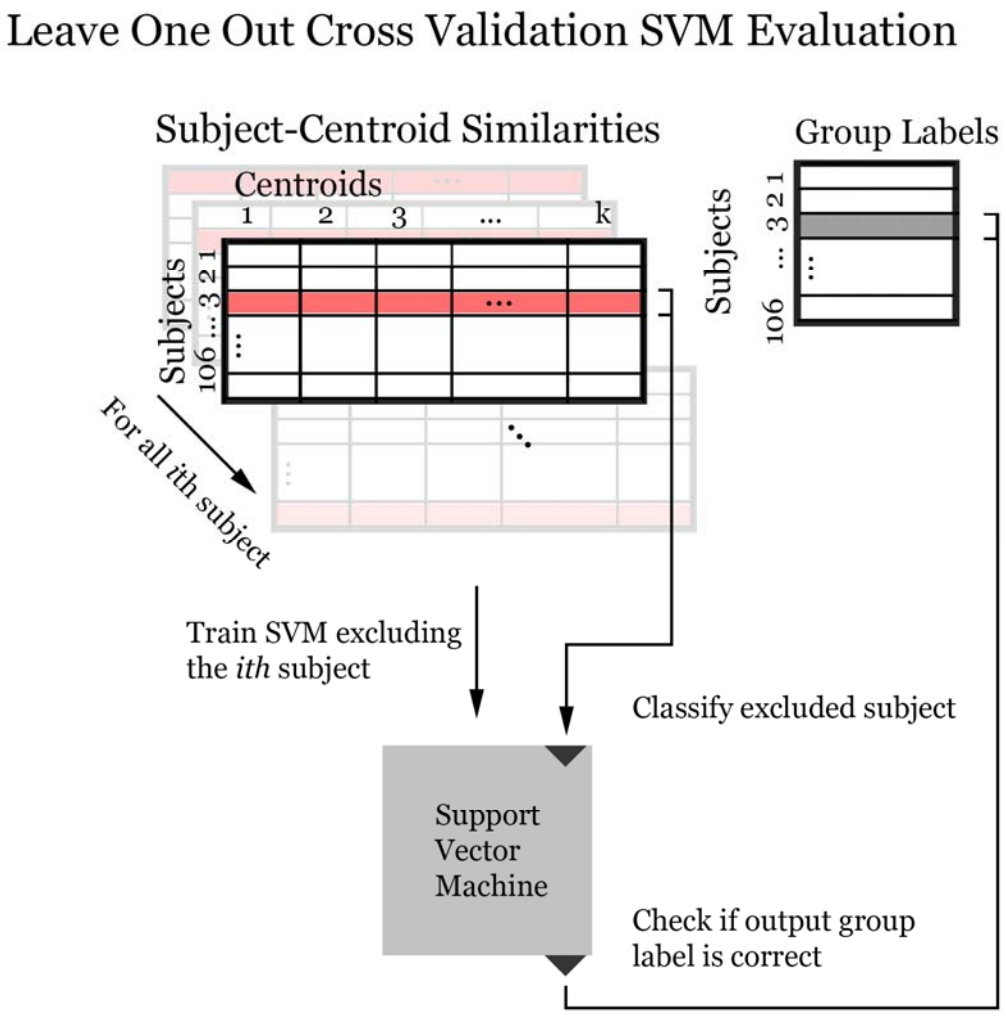

Figure 3. Leave One Out Cross Validation and SVM Evaluation Steps.

Parameter Search

There are 3 primary clustering and SVM parameters that are unknown for this dataset: k, C, and ε. These three parameters were explored using a grid search. Many combinations of k, C, and ε were explored for clustering and SVM training. The classification accuracies of the parameters were evaluated using leave one out cross validation accuracy.

Due to the computational burden of the grid search and leave one out cross validation, the ranges and intervals of the search had to be carefully chosen. A preliminary parameter search was done. Out of the three parameters, *k* was the most crucial to theoretical utility of this method. Thus, it was first investigated for all values in the range of 1 to 50. The parameters C however was found to have negligible effect on the generalization performance after a certain threshold (Cherkassky & Ma, 2004). Thus only 5 initial values of C were investigated, ranging from 0.01 to 10 in multiples of 10. The parameter ε was investigated for the same range of values. After the results of the preliminary parameter search, the parameters and their effects were investigated in further detail.

Results

Subjects

After quality control, a total of 37 healthy controls, 46 non treatment resistant major depressive disorder, and 23 treatment resistant depressive disorder subjects remained. The entire dataset consisted of 106 scans.

Proof of Concept on Simulated Data

An artificial dataset was generated in order to show proof of concept on this method. This artificial dataset included 6 subjects with 3 class types and 9 regions of interest. BOLD signals were simulated in the ROIs by using a sine wave signal and adding noise. Subjects 1 and 2 were in class 1 , subjects 3 and 4 were in class 2, subjects 5 and 6 were in class 3. 3 Networks states (NS) were simulated. The network states were

simulated such that subjects in group 1 expressed more network state 1, group 2 expressed more network state 2, and group 3 expressed more network state 3. Network state 1 had ROIs 1-3 synchronized in activity, network state 2 had ROIs 4-6 synchronized and network state 3 had ROIs 7-9 synchronized.

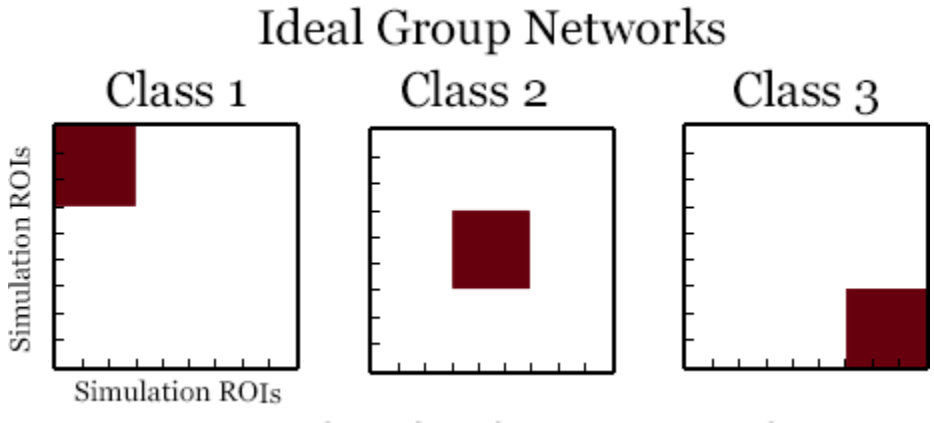

Figure 4. Simulated Group Networks

The simulation was done with noise simulated with an SNR of 2. Expressions in network states were chosen such that the subjects expressed their respective group network state 70% of the time and the other network states 15% of the time. The value for C was chosen as 10 and for ε was chosen as .1 . and k was explored from 1-10. The k-means clustering was completed 100 times in order to account for varability in clustering. Thus, SVM was trained 100 times for each k. The reported LOOCV classification accuracy is the average of these trials. Signals were generated using various phase shifted sine waves and additive sine waves.

Subjects	Group Label	% time in NS 1	% time in NS 2	% time in NS 3
1,2	Group 1	70	15	15
3,4	Group 2	15	70	15
5,6	Group 3	15	15	70

Table 2. Subjects, Group Network Labels, and Network State Mix Percentages

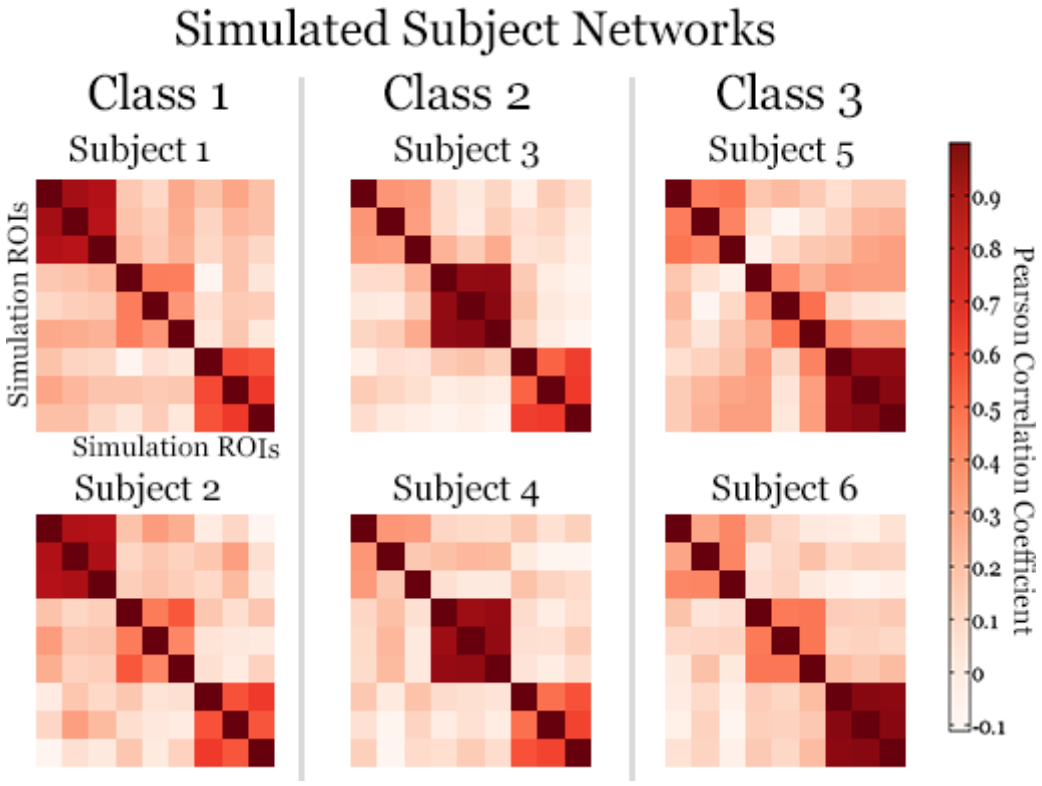

Figure 5. Static Functional Connectivity Networks per Subject

Figure 6 illustrates the k-means clustering centroid output networks for k-means run on the subject's dynamic rolling windows for k=3. Despite network signal mixing and noise addition, the k-means algorithm was able to successfully find the subject's simulated dynamic functional connectivity networks into the original signal sources.

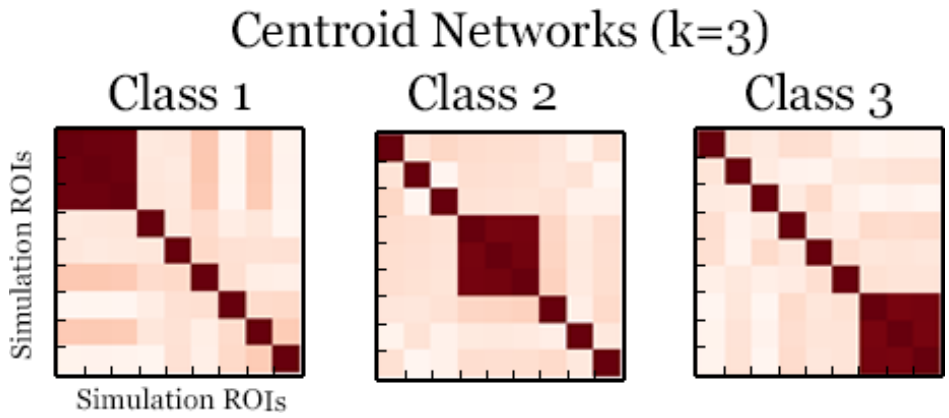

Figure 6. Centroid Networks Output from K-Means.

The SVM was trained using the expression metric of each centroid network for each subject. The classification accuracy quickly reached near 100% accuracy at the expected k of 3.

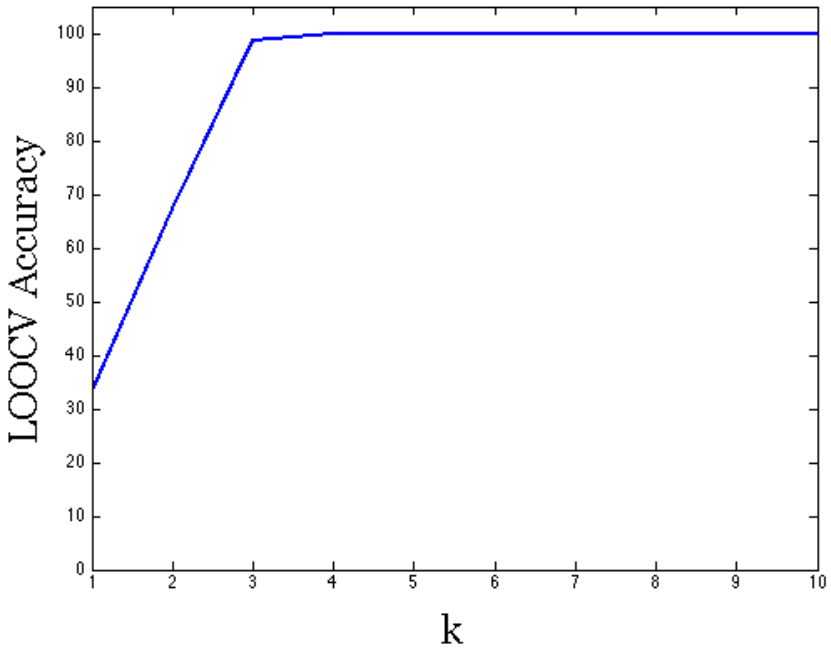

Figure 7. Average SVM LOOCV Accuracy For 100 k-means Solutions

The methods performed as expected on the simulation dataset. Even with noise and mixed expressions of the group networks, the k-means clustering algorithm was successfully able to decompose the signals into the original functions. Furthermore, the SVM classifier successfully generalized the three way classification problem.

Testing Resistance to Overfitting using Randomized Data

Next, another simulated dataset containing randomized dataset was tested with the method in order to ensure that the classifier was resistant to overfitting. All subject ROI timeseries were replaced with normally distributed random values. The SVM parameters C was set at 1 and ε was set at .1. The k-means clustering step was run 10 times in order to account for local minima solutions. The real dataset including subjects with MDD and HC was replaced with randomized values. The randomized test dataset

included a total of 106 subjects, with 69 subjects labeled as MDD, and 37 labeled as HC. The LOOCV accuracy was computed for all 10 iterations form k=1 to 50. Figure 8 displays the results.

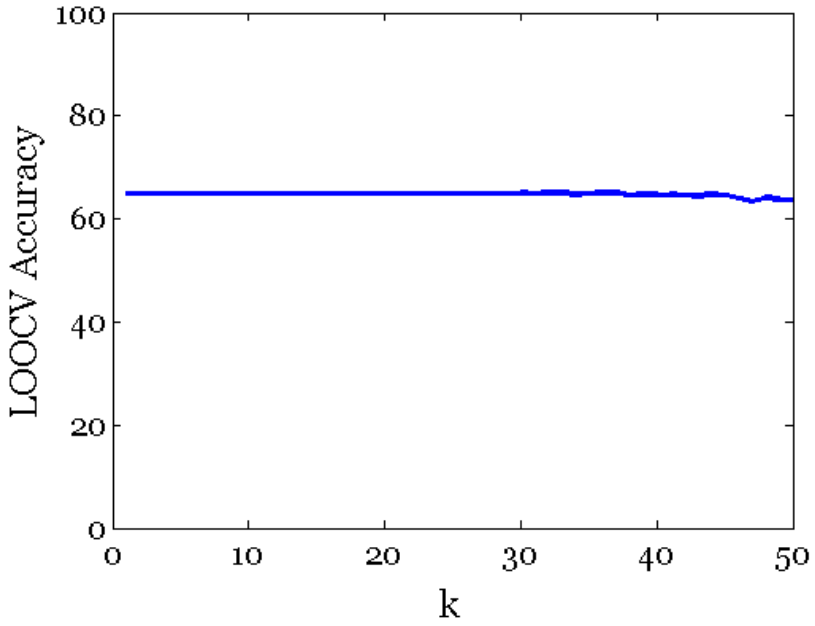

Figure 8. Average SVM LOOCV Accuracy Over 10 Iterations of Clustering and Training on a Randomized Dataset

The LOOCV accuracy of the classifier was around 65.09% for all k. This value is also represents the percentage of the number of the largest subject group. There were 69 subjects labeled as MDD (65.09%) and 37 subjects labeled as HC (34.91%). For all k, the classifier had taken advantage of the naïve choice of always classifying observations as the largest subject group (MDD). This suggests that the benchmark for the classifier accuracy rate can be set as the percentage of the largest subject group.

Parameter Search on Subject Data

An initial parameter search was done for parameter k in k-means and C and ε ranging from 0.01 to 10 in multiples of 10 in order to observe the behavior of the various parameters on the final SVM LOOCV classification percentage. The input dataset included the original subjects split into subject group labels: HC, MDD, and TRD-MDD. K-means was run 10 times in order to account for variability in the non-deterministic, local minimum solutions. The SVM LOOCV thus was measured 10 times per a (k,C,ε) Tuple. The reported LOOCV classification percentages in the figure are averaged across the 10 trials.

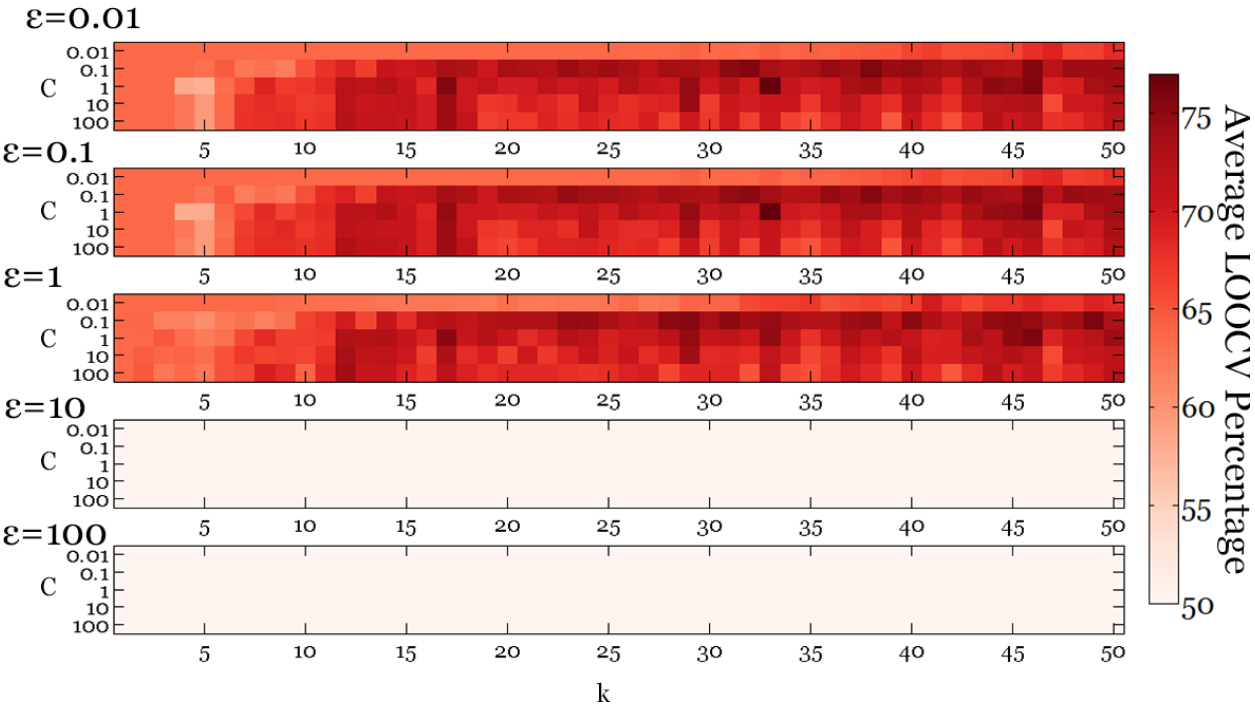

Figure 9. Average SVM LOOCV Accuracy Over 10 Iterations for Ranges of Parameters k, C, and ε .

The brute force parameter search generated a 3D matrix with entries containing the average LOOCV percentage across 10 sets of k-means centroids. Several conclusions can be made from the parameter search, however, the exhaustive parameter search did not give trivial optimum parameter choices.

k-means Parameters

The parameter k determines the number of centroids or clusters from the large set of dynamic network states. Unfortunately, this parameter is perhaps the most difficult to select. There is no expected unique correct solution. For smaller k, each centroid or clustered resting state dynamic networks the number of network components represented in each clusters become larger and more averaged. As the k becomes larger, the number of centroids become larger and the networks described in each may represent smaller parts of component networks in greater detail. Since different functional network components may become "split" at different k, the optimum choice of k is indeterminate without examining the effect of SVM parameters on k.

SVM Parameters

The ε insensitivity loss function made little difference to the LOOCV accuracy after below threshold of 1. This was expected since the ε insensitivity loss function determines the point at which the SVM s optimizing its solution. As the ε become smaller, the complexity of the models increased. The ε parameter was selected as .1 in order to balance model complexity and overfitting. The effect of the cost parameter C on the LOOCV accuracy appeared to be partially dependent on the choice of k or the number of centroids representing dynamic functional connectivity networks. Observe

that the LOOCV accuracy begins to increase at lower values of C when the number of features determined by k is higher. This effect maybe accounted for by the fact that with lower values of C, the ability for the SVM to overfit becomes decreased, yet, with enough features available (increase in k), an informative feature becomes more likely due to the increased feature dimensions. Note that the LOOCV accuracy appears to level off at a certain k. Perhaps as k increases, the number of theoretically meaningful dynamic state network centroids is increased and there is quick improvement upon the classification ability of the centroids. Yet, as the number of centroids increases, the networks that are useful to classification are further fragmented into networks that do not increase in classifiable information. The classifiable information reaches a maximum at a certain number of network centroids. C was chosen as 1 since it was responsive to increasing k at smaller numbers. This suggested that classifiers with C of 1 was successful in responding to the increasing numbers of discriminatory information in increasing numbers of clustered network states without relying on the performance gained by overfitting due to larger numbers of features.

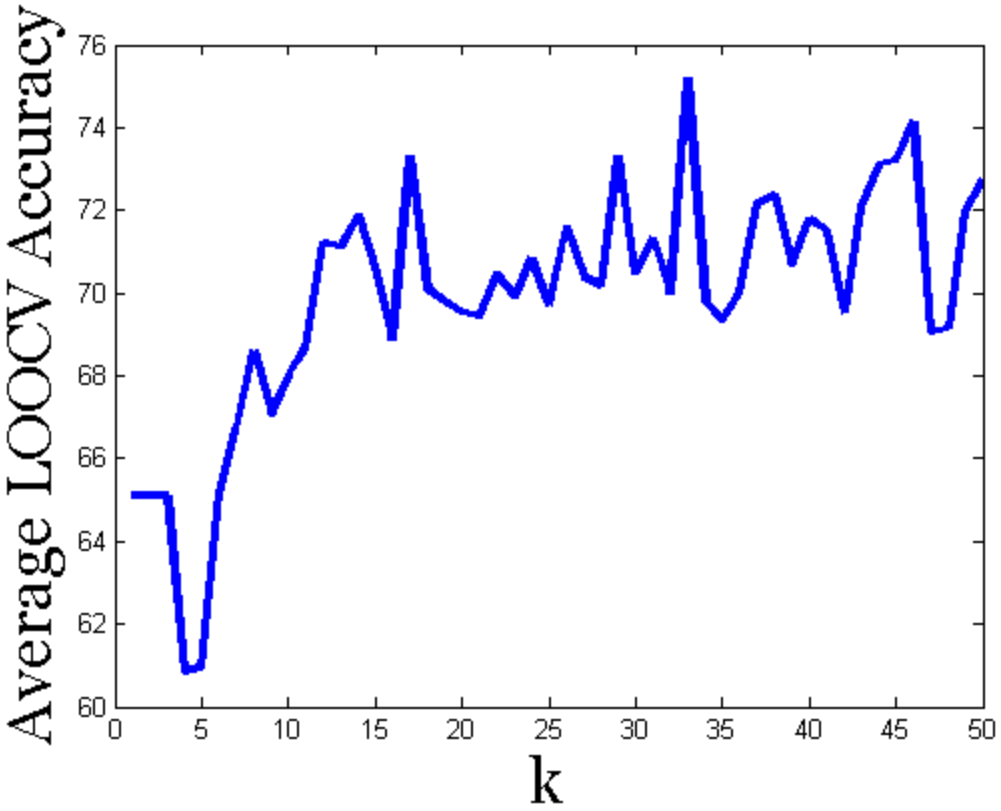

Figure 10. Average SVM LOOCV Accuracy Over 10 Iterations for Ranges of Parameters k

In summary, parameter selection was non-trivial. C was chosen to be 1 since performance gains appeared to be independent of overfitting as the number of centroid networks increased. The ε parameter did not appear to significantly change LOOCV accuracy when being below a certain threshold. It was chosen to be .1. k was chosen dependent on the elbow point of the LOOCV accuracy and k curve.

Data Preconditioning

The subject-centroid distances were also classified after feature scaling. The data was standardized by computing the mean and standard deviation for all features and

subtracting each feature by the respective mean and then dividing by the result by the standard deviation. The resulting accuracies and patterns in accuracies were nearly identical to the results from unconditioned data. This may be due to the fact that all features were homogeneous metrics and had marginal variance in the mean and standard deviation.

Classification Results

Disease State Binary Classification (MDD, HC)

First, the classifier was tested for binary disease state classification. All patients diagnosed with MDD, including the treatment resistant group were labeled as the same group. The centroid network generation step using k-means was completed 10 times with dynamic resting state networks of window length 10 TRs sampled at intervals of 4 TRs. The SVM parameter C was set at 1 and ε was set at .1 as informed by the initial parameter search.

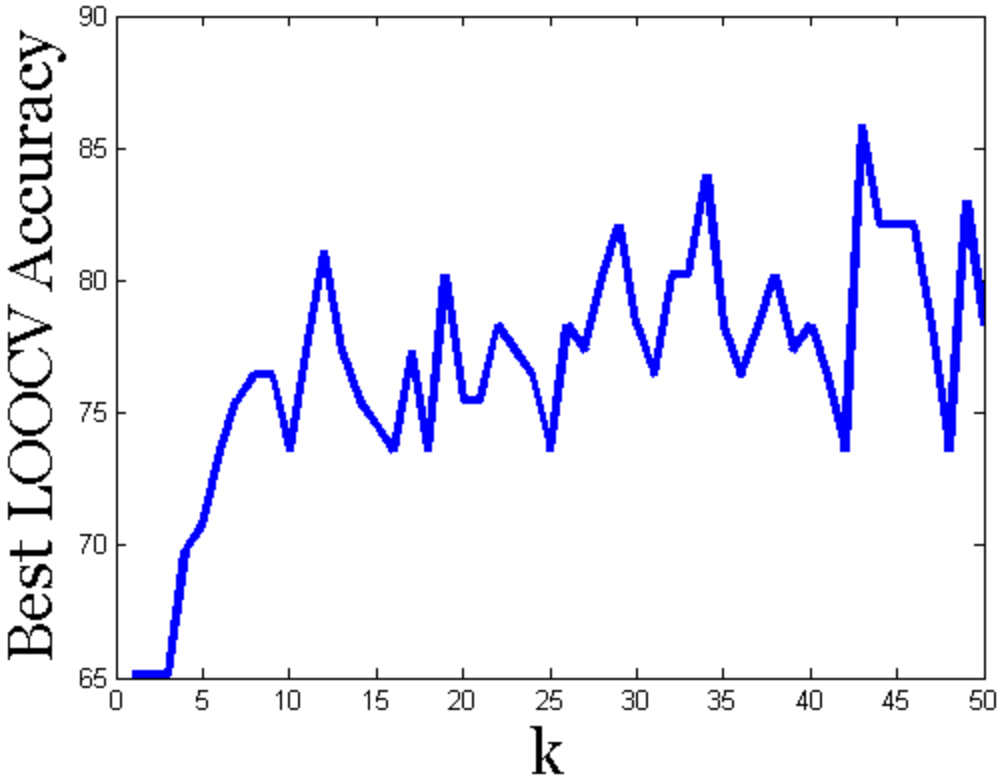

Figure 11. Best SVM LOOCV Accuracy Over 10 Iterations for Ranges of Parameters k For Binary Disease State Classification

The classifier achieved moderate success with a classification accuracy of up to 85.85% for a k of 43. The model with k=12 was chosen for discussion for its simplicity which had achieved a classification accuracy of 81.13%. The following is the confusion matrix for the two-way classifier:

		Actual Class	
		MDD	HC
Predicted Class	MDD	62	13
	HC	7	24

Table 3. Confusion Matrix for Two Way Classification

Although using this classifier on rsfMRI scans to diagnose MDD is impractical, the generalizations created by this classifier can be used to further study the mechanisms behind the disorder. By computing the primal variable w, the networks that contain the most classifying power can be identified. Perhaps by examining these networks states, the functional mechanisms behind MDD can be further studied. Since the simplicity of the model was preferred, the best model near the elbow point was chosen which had a k of 12. The Figure 12 shows all 12 centroid networks ordered by the amplitude of the primal variable w.

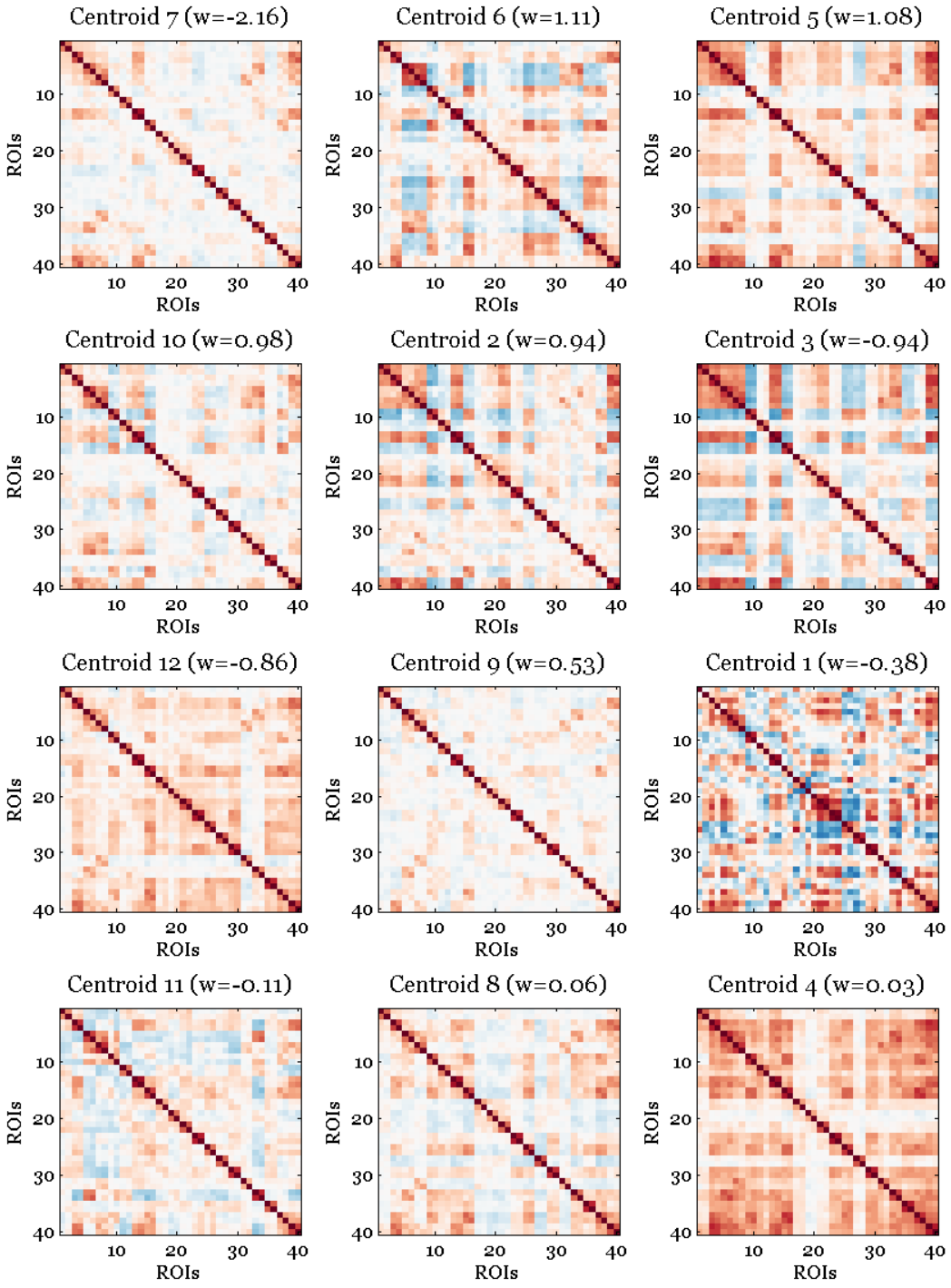

Figure 12. Centroid Clusters Ordered by Importance in Classification

The higher the amplitude of w, the more important the network was for classification. Positive weights imply more expression in the HC group, and negative weights imply more expression in the MDD group.

Three way Classification (MDD, TRD, HC)

Next, the classifier was tested for disease severity classification. This was a three way classification problem with treatment resistant MDD group, MDD group, and HC group. The centroid network generation step using k-means was completed 10 times with dynamic resting state networks of window length 10 TRs sampled at intervals of 4 TRs. The SVM parameter C was set at 1 and ε was set at .1 as informed by the initial parameter search.

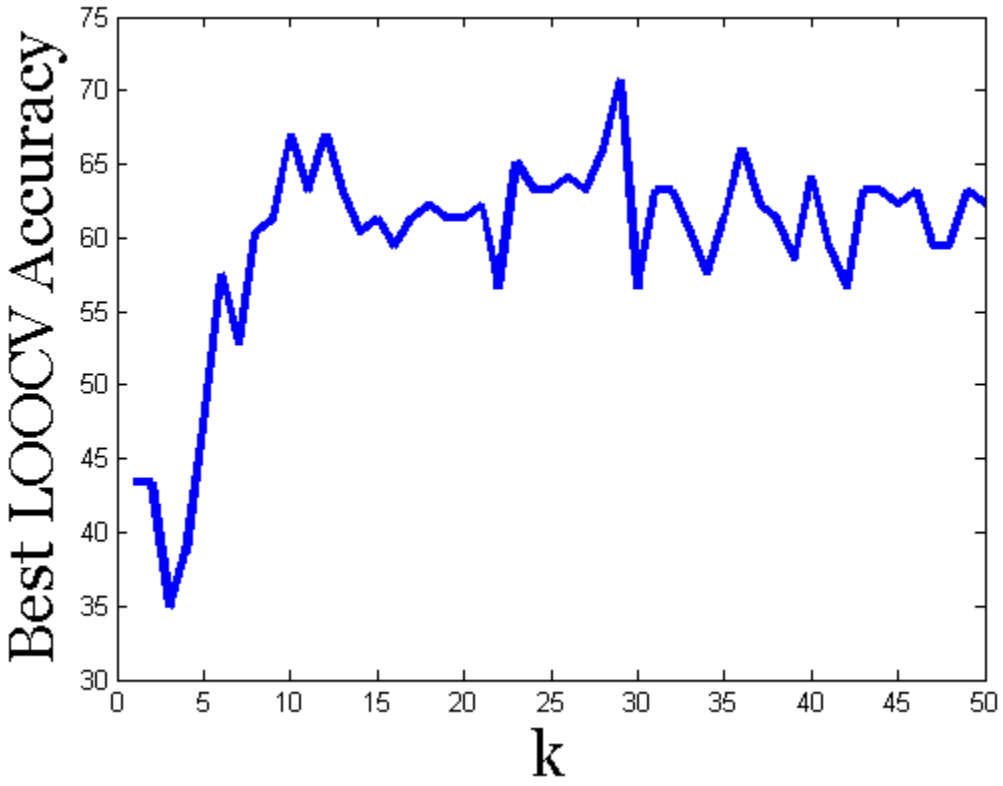

Figure 13. Best SVM LOOCV Accuracy Over 10 Iterations for Ranges of Parameters k For Three Way Disease State Classification

The classifier achieved moderate success with a classification accuracy of up to 70.75% for a k of 29. However, the model created using a k of 10 (66.98% LOOCV accuracy) was chosen for discussion due to the greater simplicity of the model over the small accuracy gain attained at the cost of greater complexity. The following table is the confusion matrix for the three way classifier:

		Actual Class		
		TRD	MDD	HC
Predicted Class	TRD	13	5	3
	MDD	6	33	9
	HC	4	8	25

Table 4. Confusion Matrix for Three Way Disease State Classification

Next, the feature weights for all three classifiers created during the multi-way classifications were computed from the SVM models. The following diagrams show the unordered 10 centroid networks used for the three way classification. Table 5 lists the centroids ordered by their computed feature weights for each binary classifier.

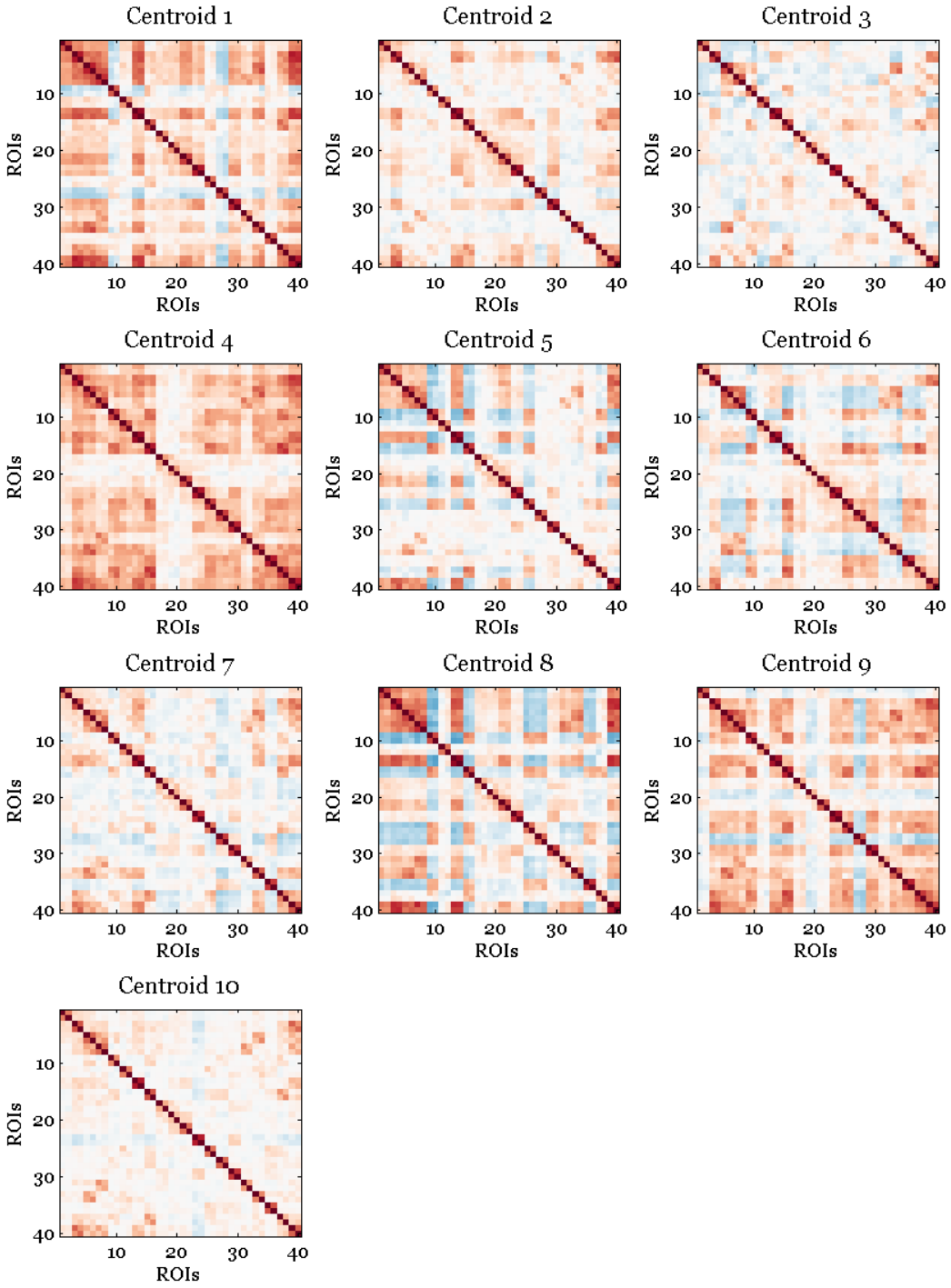

Figure 14. Unordered Centroids for the Three Way Disease State Classification

TRD vs. HC		MDD vs. HC		TRD vs. MDD	
Centroid #	w	Centroid #	w	Centroid #	w
10	-1.25	5	1.06	10	-1.07
2	-0.8	6	0.94	2	-0.85
6	0.75	9	0.88	4	0.78
7	0.54	8	-0.85	3	0.7
9	0.51	4	-0.62	5	-0.58
3	0.39	10	-0.57	6	0.57
1	0.28	7	-0.39	7	0.54
5	0.15	3	-0.39	8	0.51
4	0.15	1	0.38	9	0.28
8	-0.09	2	-0.12	10	0.06

Table 5. Centroids Networks Ordered by Weight Amplitude for Each Binary Classifier from Three Way Disease State Classification

Higher amplitude weights imply greater importance for classification. Negative weights signify greater expression (greater similarity/less distance) of the feature network for first class, and positive weights imply greater expression (greater similarity/less distance) of the feature network for the second class. Centroids are not labeled in any particular order.

Discussion

The proposed classifier for resting state fMRI networks performed as expected with simulation data and classified real world data with moderate success. The k-means clustering algorithm applied to dynamic functional connectivity ROI networks successfully decomposed the subject data into an interpretable and holistic feature space. Using this new feature space, the SVM classifier was able to successfully create generalizable models. By attempting to model random data using the SVM classifier, it was shown that the classifier was resistant to over fitting. Leave one out cross validation

ensured generalizability of the models. Furthermore, the method of feature space generation and classification escaped the limitations of univariate and MVPA analysis.

Two Way Disease State Classification (MDD vs. HC)

The classifier achieved moderate success with a classification accuracy of up to 85.85% for a k of 43 from a naive guess expectation of 65.09%. For analysis and discussion, the model with 12 clusters (k=12) was chosen for discussion for the greater simplicity. This model had achieved a classification accuracy of 81.13%. During practical use of this classifier for exploratory data analysis, a judgment must be made on the selection of k. There are several considerations to make during this selection. First, if the simplicity of the model is preferred, models with smaller k would provide a more manageable feature space to investigate. However, classifiers with greater features are likely to provide better classification accuracies (this may not always be the case). This may be due to superior centroid networks found with greater numbers of k.

Figure 12 displays the centroid networks ordered by the amplitude of the computed weights (primal variable w representing the SVM hyperplane). The relative amplitude of w for each feature or network centroid can be interpreted as the relative utility of the network in the classification problem. A positive weight indicates that the network was expressed more in the depressed group, and a negative weight indicates that the network was expressed more in the healthy control group. Weights closer to 0 imply that they are not informative in distinguishing between the two classes. Note that the interpretability of the weights is only possible with linear support vector machines. The use of other kernels will create weights and variables that do not trivially convey the importance of various features.

By examining the important centroids in this binary classification problem, insights can be made regarding the various brain network states that are associated with MDD. This method could easily be applied to other cohorts including other psychiatric disorders. By examining these important network centroids in combination with expert knowledge and literature findings, the classifier's generalizations may lead to powerful insights into the pathophysiology of brain disorders.

Three Way Disease State Classification (TRD vs. MDD vs. HC)

The three way disease state classification yielded moderate success. The LOOCV accuracy was 70.75% for a k of 29 with the baseline naïve guess rate at 43.39%. However, the model created using a k of 10 (66.98% LOOCV accuracy was preferred for discussion due to its simplicity. During usage, k must be selected by considering tradeoff between model simplicity and classification capability. Again, these conditions may not always be mutually exclusive, however, in this investigation, marginal amounts of accuracy were by gained increasing the number of centroid networks.

Multiclass problems introduce new possible insights and considerations. Table 5 shows the confusion matrix that includes the predicted and actual class of each observation during LOOCV. The confusion matrix can give insights into which classes were relatively more difficult to classify. Observe that the classification between the TRD and HC groups were the least difficult for the classifier. The classification between MDD and HC was the most difficult for the classifier. Unfortunately, making inferences from the relative differences between the classifications is difficult. The intuitive reason for the differences in difficulty may be that some group pairs are more similar to each other than other pairs of groups, thus they would be more difficult to distinguish. However,

there are other confounding factors that may attribute to the differences in the classification accuracy between different pairs of classes. For example, the features available to the classifier may have unequal utility to classification group pairs. The features that were found using the k-means clustering may have found networks that are effective at distinguishing between control and TRD groups, however, it may not have found networks that are effective at distinguishing between MDD versus control Groups. Perhaps the network expressions function at another kernel space between MDD and control groups, thus the linear classification was less effective for these groups. In summary, the confusion matrix shows information on which groups were more difficult to classify between, however, the cause of the variance in difficulties between group pairs cannot be determined easily.

Interesting conclusions can be drawn from examining the various feature weights computed for each binary classifier in the multi class classification. Recall that the voting method was utilized for the multiclass classification: this required the training of a classifier for all pairwise groups. By utilizing the same feature set for all classifiers, each feature's weight, thus each feature's importance, can be compared for various group types. Table 5 lists the centroid networks for all pairwise classifiers. The centroid networks are labeled from 1 to 10 in no particular order. Observe that centroid network 10 was ranked the highest for classification between TRD vs. MDD and TRD vs. HC, but was ranked much lower for the MDD vs. HC classification. This suggests that network 10 was exclusive to classification for TRD. The negative sign of the weight for both classifiers implies that patients with TRD expressed network 10 yet was absent for all other patient groups. Similar conclusions can be made regarding network 2 which

ranked 2nd highest for the classification between TRD vs. MDD and TRD vs. HC. Network 5's absence strongly indicated that the subjects had MDD over HC, however, its absence was only moderately predictive in the classification between MDD and TRD. Network 6's absence was indicative of both MDD and TRD as it ranked highly for both MDD vs. HC and TRD vs. HC classifications.

Examining feature rankings between the binary classifiers for multi class problems in such a way can give insights into the relationships between various networks and group types.

Recommendations for Future Studies and Limitations

Treatment Selection Tool

This method or variations of it may also hold utility as a future clinical tool. Although much more investigation is necessary, this classifier could be trained to generalize functional differences in brain networks that correlate to treatment response for a particular disorder where several evidence based treatments are available. For example, in MDD, psychotherapy and medication are standard first treatments. However, there are no available procedures to select the best treatment for a given patient. As such, the treatment selection often consists of trial and error with less than 40% of patients receiving a first treatment that leads to clinical remission. By utilizing the classification method discussed, the classifier could select an optimal treatment after a rsFMRI scan, given that there is a functional networks or combinations of them that correlate with various response to treatments. While the use of the classifier in such a way would have great clinical usage, further optimizations and research must be done on the reproducibility of the classifiers.

Testing Constructed Networks

The proposed method in this thesis utilizes clustering of dynamic functional connectivity windows in order to generate the feature space. While this unsupervised algorithm is useful in the case that the structure and function of the underlying networks are unknown, a hypothesis regarding the discriminating ability of a manually constructed network can be easily tested.

In order to test the discrimatory ability of a certain functional connectivity network or sub-network, the network may be manually inserted in addition to the centroids. Since the feature space for the classifier is generated by taking a similarity measure for each subject to the centroids, replacing a centroid with a test network will allow the SVM classifier to integrate the hypothesized network into respective classification problems. After parameter selection and SVM training, the primal variable w can be computed for the respective similarity measure. The absolute value of w will give a relative measurement of the importance of the manually constructed network for classification.

Optimizing Feature Space Generation

The limitations in classification accuracy may be due to limitations in the feature space generation method. The data decomposition method investigated in this thesis involves clustering dynamic resting state functional connectivity networks. This method was chosen due to its advantages over traditional voxelwise univarate analysis and MVPA. The selection of the feature space is crucial for good classification performance. However, the selection of the feature space is difficult since the theoretical correctness

and interpretability must be accounted. Modifications and other methods of data decomposition may be used in order to generate feature spaces.

Regions of Interest Selection

The regions of interest used in this investigation were selected for with expert knowledge and literature findings. While this restricted ROI selection yield moderate success during classification, by using a more inclusive ROI set, better classification accuracy may be achieved. The limitations of the 40 ROI sets were mostly derived from univariate methods that may have resulted in the exclusion of regions that may hold useful information when examined in a multivariate analysis that takes advantage of temporal dynamics. Future investigations of this method should be performed using whole brain parcellations so that all brain areas may be accounted for during dynamic connectivity clustering.

Dynamic Functional Connectivity

Dynamic functional connectivity is a relatively new development in functional neuroimaging. More research in dynamic functional connectivity is necessary in order to determine the optimal window lengths and speeds when collecting network samples previous to k-means clustering. Various window length ranges would capture network dynamics at differing temporal resolutions. While the theoretical understanding of the various lengths of dynamic network states is currently limited, a more inclusive feature set may be generated for this method by collecting the centroids of networks for varying window lengths.

Alternative Distance Measures in Clustering

Other distance metrics may be utilized during computing subject-centroid similarities. The method in this study minimizes the Manhattan distance for each dynamic windowed network to the centroids. Other possible distance measures may be utilized during k-means clustering. Some possibilities may include metrics based on cosine similarity or correlation. Depending on the choice of the distance function during clustering, the resulting output and interpretation of the centroid networks can be drastically altered.

Dynamic Functional Connectivity Similarity Metrics

The similarity measurement of each centroid network to a subject is computed using a Euclidian distance. While this gives an approximate average metric of expression for each centroid, this metric does not distinguish between a subject expressing a centroid network in high amplitude for a short period of time from a subject expressing a centroid network in low amplitude for a long period of time. The similarity measure can refined by adding unequal weights to amplitude or periods of centroid network expression.

Independent Component Analysis

As an alternative to the k-means clustering on dynamic functional connectivity networks, independent component analysis (ICA) could be used in to generate component networks. ICA can decompose fMRI data into independent spatially additive components. Similar to k-means, the number of components must be chosen for ICA analysis. Using the ICA component outputs, an expression of each component may be calculated and used as the feature space in the classifier.

The current method of k-means clustering on dynamic functional connectivity windows takes the assumption that FC networks will occur temporally exclusive of each other. ICA would be able to find networks that occur statistically independently and thus will be able to decompose networks that are temporally mixed. Furthermore, network selection using ICA on whole brain datasets would bypass the need for brain parcellation or ROI selection.

Conclusion

A novel method was developed in order to decompose and classify rsfMRI data. The high dimensional rsfMRI data was converted into an alternate feature space with interpretability and theoretical relevance in mind. The dynamic resting state connectivity networks were collected from each subject and clustered using k-means. This decomposition method provided holistic centroid networks based on dynamic networks observed in subject scans. Next, the expression of each centroid network was computed for each subject- this became the new feature space used for the SVM classifier.

The method was tested using simulated data in order to test proof of concept and performed as expected. Testing on randomized data indicated that the method was resistant to overfitting. The application of the method on real data on subjects with MDD showed moderate success with 85.85% LOOCV accuracy when classifying between depressed and healthy controls and 70.75% LOOCV when classifying between patients with treatment resistant depression, non-treatment resistant depression, and healthy controls. The generalizations created by the SVM were investigated by examining feature weights from the linear SVM models.

The method discussed in this thesis holds promise for its data mining ability during exploratory data analysis for a diverse range of fMRI problems and may hold potential as a clinical tool in treatment selection for psychiatric illnesses. Further investigation is necessary in order to optimize the feature space generation step. Improvements in the method of feature space generation may improve classifier accuracy.

Bibliography

Aggarwal, C. C., Hinneburg, A., & Keim, D. A. (2001). On the surprising behavior of distance metrics in high dimensional space (pp. 420-434). Springer Berlin Heidelberg.

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2012). Tracking whole-brain connectivity dynamics in the resting state. Cerebral Cortex, bhs352.

American Psychiatric Association. (2013). DSM 5. American Psychiatric Association.

Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., ... & Windischberger, C. (2010). Toward discovery science of human brain function.Proceedings of the National Academy of Sciences, 107(10), 4734-4739.

Blais, M. A., Malone, J. C., Stein, M. B., Slavin-Mulford, J., O'Keefe, S. M., Renna, M., & Sinclair, S. J. (2013). Treatment as usual (TAU) for depression: A comparison of psychotherapy, pharmacotherapy, and combined treatment at a large academic medical center. Psychotherapy, 50(1), 110.

Center for Substance Abuse Treatment. (2008). Managing Depressive Symptoms in Substance Abuse Clients During Early Recovery . In Treatment Improvement Protocol (TIP) Series. Rockville, MD: Substance Abuse and Mental Health Services Administration (US).

Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST),2(3), 27.

Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. Neural networks, 17(1), 113-126.

Cole, D. M., Smith, S. M., & Beckmann, C. F. (2010). Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. Frontiers in systems neuroscience, 4, 8.

Cortes, C., & Vapnik, V. (1995). Support vector machine. Machine learning,20(3), 273-297.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical research, 29(3), 162-173.

Craddock, R. C., Holtzheimer, P. E., Hu, X. P., & Mayberg, H. S. (2009). Disease state prediction from resting state functional connectivity. Magnetic resonance in Medicine, 62(6), 1619-1628.

Greenberg, P. E., Kessler, R. C., Birnbaum, H. G., Leong, S. A., Lowe, S. W., Berglund, P. A., & Corey-Lisle, P. K. (2003). The economic burden of depression in the United States: how did it change between 1990 and 2000?.Journal of clinical psychiatry, 64(12), 1465-1475.

Greicius, M. (2008). Resting-state functional connectivity in neuropsychiatric disorders. Current opinion in neurology, 21(4), 424-430.

Harris, E. C., & Barraclough, B. (1997). Suicide as an outcome for mental disorders. A meta-analysis. The British Journal of Psychiatry, 170(3), 205-228.

Heberlein, K. A., & Hu, X. (2004). Simultaneous acquisition of gradient-echo and asymmetric spin-echo for single-shot z-shim: Z-SAGA. Magnetic resonance in medicine, 51(1), 212-216.

Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. Information Theory, IEEE Transactions on, 14(1), 55-63.

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., ... & Chang, C. (2013). Dynamic functional connectivity: promise, issues, and interpretations. Neuroimage, 80, 360-378.

Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S., & Menon, R. S. (2013). Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. Human brain mapping, 34(9), 2154-2177.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651-666.

Jo, H. J., Saad, Z. S., Simmons, W. K., Milbury, L. A., & Cox, R. W. (2010). Mapping sources of correlation in resting state FMRI, with artifact detection and removal. Neuroimage, 52(2), 571-582.

McGrath, C. L., Kelley, M. E., Holtzheimer, P. E., Dunlop, B. W., Craighead, W. E., Franco, A. R., ... & Mayberg, H. S. (2013). Toward a Neuroimaging Treatment Selection Biomarker for Major Depressive DisorderTreatment-Specific Biomarker for Major DepressionTreatment-Specific Biomarker for Major Depression. JAMA psychiatry, 70(8), 821-829.

Nash, J., & Nutt, D. (2007). Antidepressants. Psychiatry, 6(7), 289-294.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., ... & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage, 23, S208-S219.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn human connectome project: an overview.NeuroImage, 80, 62-79.

Vos, T., Flaxman, A. D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., ... & Brooker, S. (2013). Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet, 380(9859), 2163-2196

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: a meta-analysis. Psychological bulletin, 132(1), 132.

Appendix

ROI Index	Name	Side	Volume (mm^3)	Actual Volume(mm^3)	X	Y	Z
1	R_cg25	R	523.599	485	6	24	-11
2	L_cg25	L	523.599	485	-6	24	-11
3	R_ACC24	R	523.599	485	7	40	0
4	L_ACC24	L	523.599	485	-7	40	0
5	R_dMF10	R	523.599	485	9	64	14
6	L_dMF10	L	523.599	485	-9	64	14
7	R_vMFp10	R	523.599	485	7	66	1
8	L_vMFp10	L	523.599	485	-7	66	1
9	R_dlPF9	R	523.599	485	35	49	23
10	L_dlPF9	L	523.599	485	-35	49	23
11	R_vlPF47	R	523.599	485	48	33	-6
12	L_vlPF47	L	523.599	485	-48	33	-6
13	R_vPCC	R	523.599	485	7	-47	24
14	L_vPCC	L	523.599	485	-7	-47	24
15	R_MMC24	R	523.599	485	7	24	28
16	L_MCC24	L	523.599	485	-7	24	28
17	R_NucAcc	R	523.599	485	15	7	-12
18	L_NucAcc	L	523.599	485	-15	7	-12
19	R_Amyg	R	523.599	485	20	-4	-24
20	L_Amyg	L	523.599	485	-20	-4	-24
21	R_Hipp	R	523.599	485	30	-24	-13

22	L_Hipp	L	523.599	485	-30	-24	-13
23	R_Thalamus	R	523.599	485	8	-12	7
24	L_Thalamus	L	523.599	485	-8	-12	7
25	R_aINS	R	523.599	485	36	15	-6
26	L_aINS	L	523.599	485	-36	15	-6
27	R_preMotor	R	523.599	485	30	-9	61
28	L_preMotor	L	523.599	485	-30	-9	61
29	R_caudate	R	523.599	485	12	16	4
30	L_caudate	L	523.599	485	-12	16	4
31	R_BA11	R	523.599	485	13	63	-10
32	L_BA11	L	523.599	485	-13	63	-10
33	R_BA10	R	523.599	485	14	59	27
34	L_BA10	L	523.599	485	-14	59	27
35	R_Putamen	R	523.599	485	26	4	2
36	L_Putamen	L	523.599	485	-26	4	2
37	R_dACC	R	523.599	485	9	36	28
38	L_dACC	L	523.599	485	-9	36	28
39	R_mF1032	R	523.599	525	7.4	53.4	1.8
40	L_mF1032	L	523.599	525	-7.4	53.4	1.8

Table 1. Regions of Interest Locations in MNI Space (LPI)

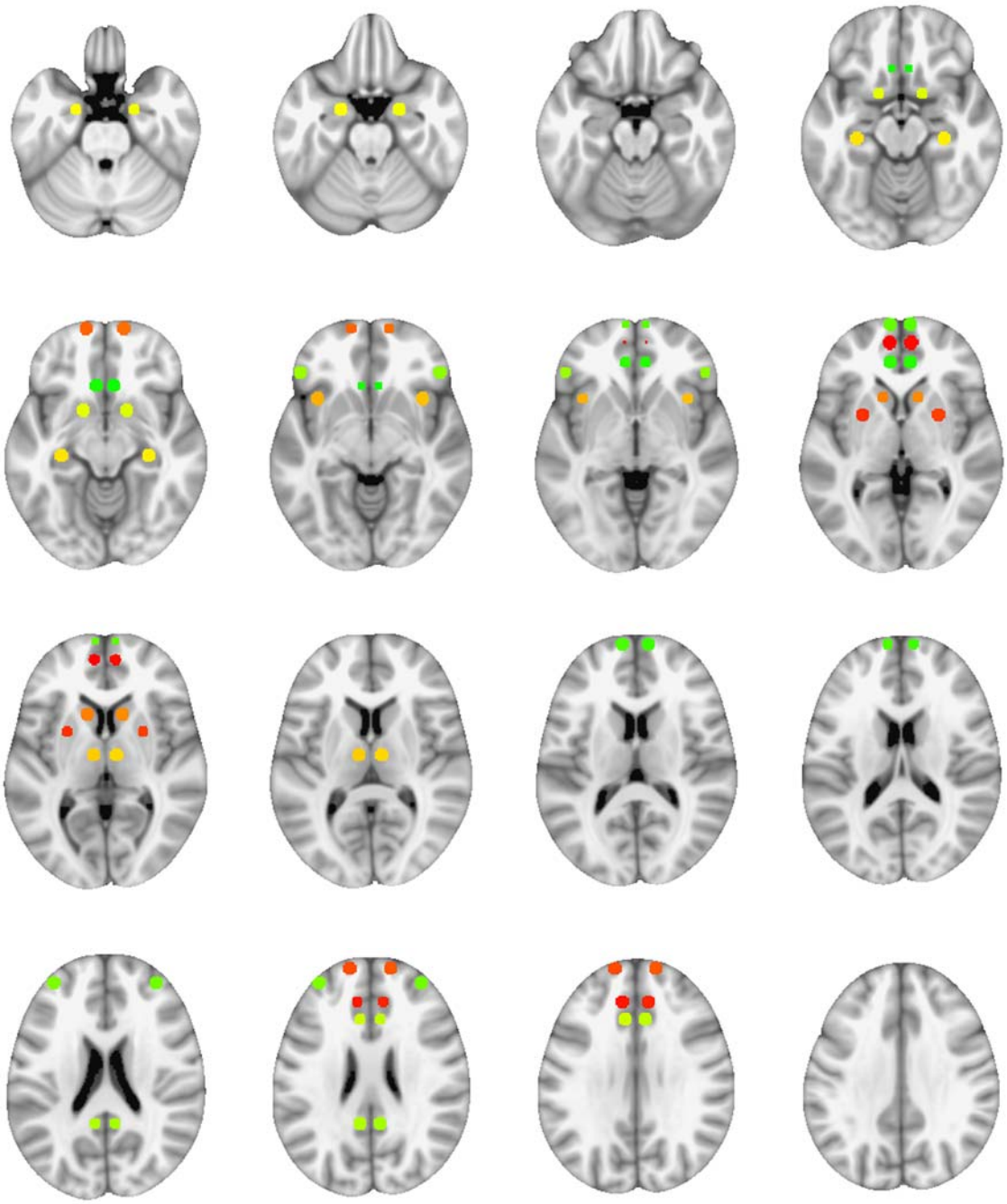

Figure 1. ROI Locations