Predicting Conversion to Psychosis in Clinical High Risk Patients using Resting-State Functional MRI Features

Jolie McDonnell[a], William Hord[a], Jenna Reinen[b], Pablo Polosecki[b], Irina Rish[b*], Guillermo Cecchi[b]

[a]Dobbs Ferry High School, 505 Broadway, Dobbs Ferry, NY 10522, USA

[b]IBM T. J. Watson Research Center, 1101 Kitchawan Rd, Yorktown Heights, NY 10598, USA

ABSTRACT

Recent progress in artificial intelligence provides researchers with a powerful set of machine learning tools for analyzing brain imaging data. In this work, we explore a variety of classification algorithms and functional network features derived from resting-state fMRI data collected from clinical high-risk (prodromal schizophrenia) patients and controls, trying to identify features predictive of conversion to psychosis among a subset of CHR patients. While there are many existing studies suggesting that functional network features can be highly discriminative of schizophrenia when analyzing fMRI of patients suffering from the disease vs controls, few studies attempt to explore a similar approach to actual prediction of future psychosis development ahead of time, in the prodromal stage. Our preliminary results demonstrate the potential of fMRI functional network features to predict the conversion to psychosis in CHR patients. However, given the high variance of our results across different classifiers and subsets of data, a more extensive empirical investigation is required to reach more robust conclusions.

Keywords: schizophrenia, functional magnetic resonance imaging (fMRI), functional networks, multivariate predictive modeling, classification, predictive features

1. INTRODUCTION

One of the central hypothesis in psychiatric research, starting from to the original work by Wernicke[1] and Bleuler[2], is that schizophrenia is associated with somehow disrupted functional connectivity ("dysconnection" hypothesis[3]). In the past decade, exploring functional connectivity disruptions in functional MRI of patients with schizophrenia, as well as building multivariate discriminative models based on functional network features is a fast-growing area of research on the intersection of neuroimaging and machine learning[4-7].

However, a much less investigated question is whether one can predict conversion to schizophrenia (development of psychosis) given the brain imaging data of clinical high-risk (CHR) patients. Developing accurate predictive models based on modern machine-learning techniques could provide a valuable tool for identification of highest-risk prodromal patients, leading to earlier intervention and potentially better treatment outcomes.

In this work, we focus on exploring a variety of machine-learning approaches and several types of features extracted from resting-state functional MRI data collected from the two groups including healthy controls and CHR patients, where a subgroup of these prodromal patients converted to psychosis later in the study. We extracted both region of interest (ROI) and voxel-level functional networks and used their features to evaluate several types of predictive models in order to discriminate between the converters and the rest of the subjects, including non-converters and controls. Note that this problem is much more challenging than classifying controls vs actual schizophrenic patients with several years of disease history and more pronounced connectivity disruptions, as reported in previous studies[4-6]. Nevertheless, we obtain encouraging preliminary results, both indicating the promise of such discriminative approaches, and the need for further exploration of different neuroimaging features, machine-learning techniques, and experimental setups which can help to improve psychosis prediction based on brain imaging.

2. BACKGROUND AND RELATED WORK

Schizophrenia is a debilitating mental disorder that affect 1-2% of the world population, and approximately 2.2 million people in the U.S. alone[11]. On average, individuals with schizophrenia die at an earlier age than healthy people. Suicide

* rish@us.ibm.com; phone 1 914 945 1896.

is the single largest contributor to this high mortality rate, which is 10 to 13 percent higher in schizophrenia sufferers as compared to the general population. Other contributors to the high mortality rate include schizophrenic patients' increased rates of accidents, diseases, and homelessness[12]. Historically, schizophrenia has been characterized by hallucinations, delusions, and abnormal emotion regulation[13].

More specifically, schizophrenia symptoms include positive symptoms and negative symptoms[14]. Positive symptoms include abnormal behaviors/experiences present in schizophrenic patients but not in healthy individuals, such as hallucinations, delusions, and varied motor behavior[11], usually caused by general dissociative thinking. Negative symptoms are defined as normal functions that are diminished or lost due to schizophrenia. These included alogia, or the slowing or lethargic speech, and catatonia, meaning a decline in ability or desire to move. The losses also applied to social and emotional affects, such as an observable withdrawal, known as blunted affect. A more specific acute emotional loss was also observed, known as anhedonia, which is a significant decrease in a person's ability to feel pleasure or motivation[15]. The variety of positive and negative symptoms showed that schizophrenic symptoms fall on a wide spectrum of cognitive abnormalities and each can be experienced at different levels of intensity, depending on the individual[11].

Besides multiple symptoms on the behavioral level, schizophrenics have been shown to have an array of brain structure abnormalities. The vast majority of structural abnormalities were observed in the limbic system and a loss of gray matter was found to be common among schizophrenic patients[11]. Furthermore, abnormal functional connectivity in the brain as measured by functional MRI have been frequently observed in schizophrenic patients[20]. This included both global and local dysconnectivity in specific areas, such as the frontal and parietal cortices, as well as across entire brain regions[16]. A general dysconnectivity would occur while schizophrenic patients were completing tasks in addition to a constant task-independent dysconnectivity[16].

Typically, schizophrenia symptoms first manifest themselves in adolescence or early twenties, although there are also cases when schizophrenia develops later in life, as well as relatively rare cases of childhood-onset schizophrenia. Often, they would progress to a chronic state of cognitive impairment[11]. It was observed that people who clearly experienced sub-psychotic symptoms were at risk for developing schizophrenia or prodromal schizophrenia. There is a 30% risk for transitioning into the full-blown disorder within two years of experiencing sub-psychotic symptoms[17]. Recently, researchers have witnessed an upsurge in research on the potential benefits of early intervention during the prodromal stage. The sooner people start the treatment at their first psychotic episode, the better the symptom relief and subsequent neural functioning[18].

2.1 Functional MRI and Machine Learning Methods

The possibility that both structural and functional abnormalities in schizophrenic patients could be better understood through functional magnetic resonance imaging (fMRI) has been a topic of interest for many years[20]. In Functional Magnetic Resonance Imaging (fMRI), a MR scanner non-invasively records a subject's blood-oxygenation-level dependent (BOLD) signal, known to be correlated with neural activity, as a subject performs a task of interest, or is simply resting without any specific task, as in case of resting-state fMRI. Such scans produce a sequence of 3D images, where each image typically has on the order of 10,000-100,000 subvolumes, or voxels, and the sequence typically contains a few hundreds of time points, or TRs (time repetitions). Standard fMRI analysis approaches, such as the General Linear Model (GLM) examine mass-univariate relationships between each voxel and the stimulus in order to build so-called statistical parametric maps that associate each voxel with some statistics that reflects its relationship to the stimulus.

Identifying neuroimaging-based features that can serve as reliable "statistical biomarkers" of mental disorders such as schizophrenia remains a challenging open problem. Most prior work on biomarker discovery involves mass-univariate hypothesis testing to identify individual features that have significantly different empirical distributions across two populations, for example patients with schizophrenia vs. healthy controls. However, we aim at finding biomarkers that are capable of predicting the disorder on previously unseen subjects. We adopt multivariate machine-learning approaches, as opposed to standard mass-univariate hypothesis testing.

Machine learning has been investigated as a promising field in computer science, specifically with respect to neurology. Classification algorithms or "classifiers" were used as a type of machine learning program for use with fMRI data that would be able to adjust its own programming and produce predicted outcomes. It has been shown that classifiers were able to generate different, testable, predictive models. The extraction of application-specific variables, or "features", was crucial to the development of the most accurate models[20].

3. MATERIALS AND METHODS

3.1 Data

Participants were recruited from the New York State Psychiatric Institute. Resting state fMRI data for n=79 subjects were collected from 44 healthy controls and 35 clinical high risk (CHR) patients, out of which 11 later converted to psychosis.

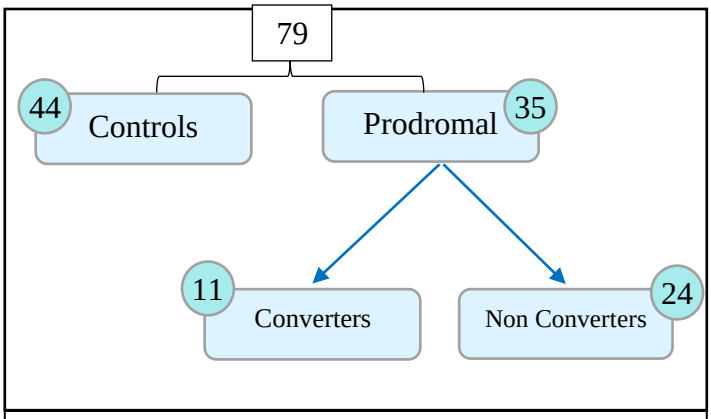


Figure 1: Subject types for n=79 dataset. Each circled number represents the number of subject type.

Resting state scans were acquired during a designated clinical baseline prior to longitudinal follow-up. Participants were matched on age, gender, scanner motion, and minority status. Psychosis risk was assessed using the Structured Interview for Psychosis-Risk Syndromes/Scale for Prodromal Symptoms[8]. Clinical outcome was determined by the last follow-up visit using the Global Functioning: Role scales[9]. For additional details, see Colibazzi et al[10]. Functional and structural MRI scans were acquired using a GE Signa 3T scanner. Participants were in a supine position and asked to keep their eyes open and were allowed to let their minds wander. Total acquisition time for the functional resting state scans per run was 5 minutes and 21 seconds. Resting state scans were comprised of echoplanar images (TR=2200ms, TE=30ms, flip angle = 90 degrees, slice thickness = 3.5mm, 34x24x24cm field of view). Participants needed to have at least one functional and one structural scan that passed assessment for motion and SNR in order to be included in analysis. Additional details are found in the supplemental materials of Colibazzi et al, 2017. Preprocessing was performed using the Nipype software package. Motion correction using MCFLIRT was used to account for motion-related noise. Spatial smoothing (5mm), a high-pass temporal filter (cutoff-100s), low-pass temporal filter (cutoff = 2.5TR) and registration to the MNI template were performed. Images that did not pass thresholds for motion or SNR cutoffs, as well as those with significant dropout, were not included in the analysis.

3.2 Feature Extraction and Feature Subset Selection

In order to learn a classifier, we first had to extract predictive features, or variables, from the raw fMRI data. Herein, we used the voxel degree features extracted from the resting-state functional networks as follows. Pair-wise Pearson correlation coefficients were computed among all pairs of time-series ($v_i(t)$, $v_j(t)$) where $v_i(t)$ corresponds to the BOLD signal of i-th voxel. The degree of each voxel was then computed as the total number of voxel's correlations exceeding $R > 0.6$ in the full-brain correlation matrix. For each subject, and each session, voxel degrees formed a separate feature vector, labeled by the type of the subject (binary label denoted whether a subject converted to psychosis or not). We used both voxel-level and ROI-level features (activity averaged over each ROI), including logarithmic degrees, interquartile range of logarithmic degrees, interquartile range of ROI partial correlations, and interquartile range of ROI regular correlations. We performed feature selection in order to achieve best classification results for each classifier, using a simple filter-based approach, which consists in ranking all features using some function that measures the relevance of the feature to the class label, and thus the discriminative ability of the feature. Next, a subset of the top-ranked features is selected. Commonly used ranking criteria include mutual information between feature and class variables, correlation between the two, or some other measure of feature's relevance. Herein, we used as a ranking function the p-value resulting from the paired t-test, with the null-hypothesis being that the feature's values corresponding to samples of different classes came from distributions with equal mean.

3.3 Classifiers

In order to predict whether a CHR subject will convert to psychosis based on the above features, we trained several different classifiers, including SVM, Naive Bayes, Logistic Regression, Decision Tree, Random Forest, SDCA SVM Elastic, k-Nearest Neighbors (kNN), and AdaBoost. The classifiers' error rates were charted for different numbers of top

features. This methodology was aimed at determining which feature files would produce the lowest error rates for different classifiers. We also used 'chance' classifier as a baseline, where the most common label in the training dataset was always used to label the previously unseen test samples. Namely, the error rates for chance level were always equal to the percentage of the least common label, as those were the ones that would be guessed incorrectly. A brief description of the other classifiers we used is provided below.

Support Vector Machines (SVM) with linear kernel: a linear classifier which constructs a hyperplane separating the examples of two classes, where the hyperplane has the largest margin (distance to the nearest training-data point of any class) among all such hyperplanes.

Naive Bayes: a probabilistic classifier making a simplifying assumption that all features were independent of each other given the class label, and uses Bayes rule to infer the class given the feature values. In geometric terms, Naive Bayes is also a linear classifier, but unlike SVM, its separating hyperplane is not necessarily maximizing the margin. Despite its simplifying assumption, Naive Bayes has been shown to perform surprisingly well in practice and have advantages of being scalable and very easy to learn and apply.

Logistic regression estimates the probability of a binary response given a set of predictor variables (features). By thresholding the output of the logistic regression for a given input sample at 0.5 (i.e., replacing it by 1 if logistic regression predicts above the threshold, and by 0 otherwise), a classification model is obtained. A standard approach to handling the data with the number of dimensions (features) significantly exceeding the number of samples is to add some type of regularization (i.e. constraints) to the model. A common type of regularization, also used herein, was the l2-norm penalty on the vector of the model's parameters.

Decision Tree classifier learns a predictive model in the form of a tree structure, where leaves represent class labels and nodes correspond to conditions on features of a given sample; thus, tree branches represented conjunctions of such conditions that lead to class labels/leaves. The predicted outcome is the class to which the data belongs.

Random Forest is an ensemble method consisting of a group of independent random decision trees. Each tree is grown using a randomly selected subset of features. For each input, outputs of all trees are computed, and the class with majority of votes is selected. The number of estimators for the random forest was varied from 5 to 20.

SDCA SVM Elastic: This was a form of a regularization model, used to prevent overfitting and maximize interpretability.

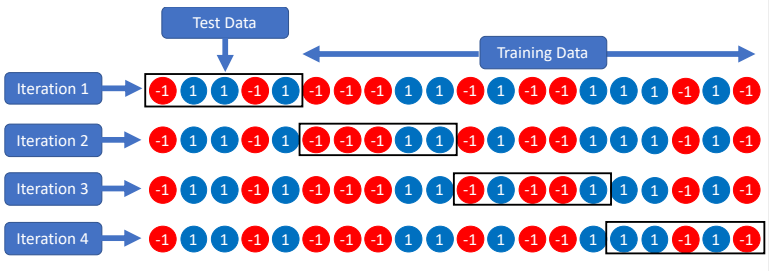
Nearest Neighbors (kNN): the k-Nearest Neighbors classification algorithm (kNN) assigns a class label to a given (unlabeled) test sample via majority vote of its neighbors, i.e. the sample is assigned to a class most common among the sample's k nearest neighbors.

AdaBoost (short for Adaptive boosting) is a classifier that combines multiple different machine learning algorithms by averaging their outputs. Each algorithm used only has to be better than chance level, and then AdaBoost would be expected to have good performance. One drawback of the classifier was that it was sensitive to noisy data.

Many of these classification models were based off of algorithms that are included in Python's machine learning library, Scikit Learn.

The data were comprised of multiple subjects labelled patients or controls (typically 1 for control and -1 for patient). Different variables were measured for each subject, including BOLD signals and symptom severity. The program extracted useful features that could build a discriminative model. By properly selecting relevant features, models were easier to interpret, program run times were shortened, and other benefits were reached[19].

Figure 2: 4-Fold Classification Diagram. For each iteration (run of classification algorithm to build a model), a different subset of the data is used as testing data, and the rest are training data. The final model incorporates all iterations into one.



The feature file produced was then split into "folds," or subsets. The script ran, leaving 1-fold out at a time as testing data. These data were left out of the model to check for accuracy. The rest of the data, called training data, were then fit to a graphical model. When the script was done running, average

error rates were plotted based on how well the model predicted which class the test data belonged to[19].

4. RESULTS

Initially, for n=79 dataset with 11 converters vs. all others (68, both healthy controls and non-converters), each classifier had low error rates. However, chance level is 11/79, or about 14%, so the accuracy of the classifiers was difficult to determine with such a low chance level; best results of the classifiers were only meeting the chance level. However, in the next run, we explored an even split of 11 converters and 11 completely healthy controls. Prodromal non-converters were excluded from this analysis. Here, chance level would be at 50%. Multiple classifiers performed clearly above the chance level. Though this undermined the large sample size provided, the smaller subset offered a more class-balanced view. AdaBoost Classifier had an error rate as low as 12%, which was substantially lower than the (theoretical) 50% chance level for the interquartile range of ROI partial correlation. AdaBoost also demonstrated consistency in low error rates over different amounts of top variables. Other classifiers like SDCA SVM Elastic and Linear SVM had notable accuracies as well, ranging from an 18-20% error rate. We show in Figure 1 the results for two types of network features, at ROI and at the voxel level.

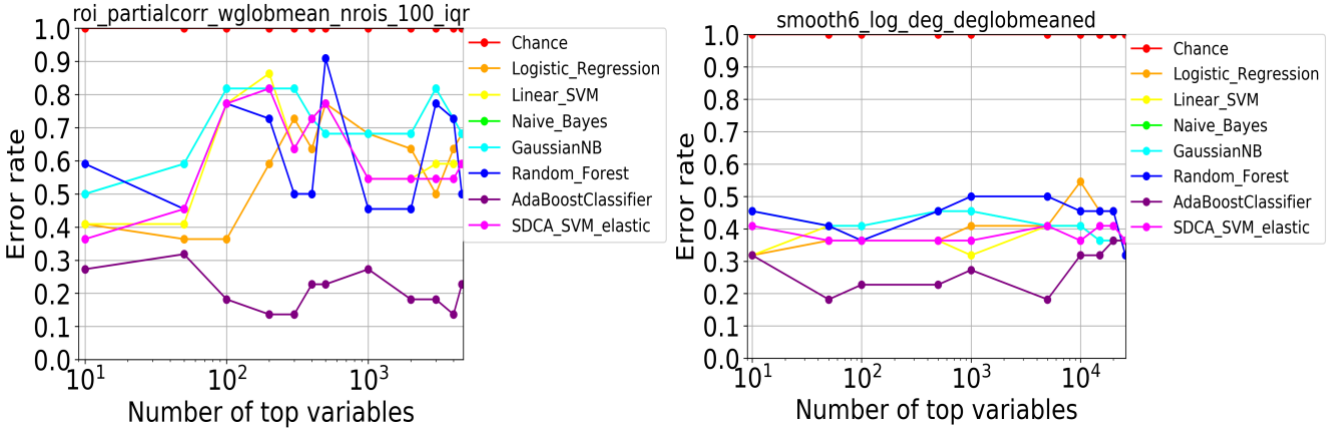


Figure 3: Classification error for balanced subset of data, as a function of different subsets of top-ranked features (x-axis). Extracted features: (a) interquartile range of ROI partial correlations,(b) smooth logarithmic degrees. Chance level: 0.5.

Overall, the results are preliminary but promising as they demonstrate a potential for functional network features to predict the conversion to psychosis in CHR patients, at least for some classifiers and data subsets. However, given the high variance of our results across different classifiers and subsets of data, a more extensive empirical investigation is required to reach more robust conclusions.

REFERENCES

[1] Friston, K. J. & Frith, C. D. Schizophrenia: a disconnection syndrome? *Clin Neurosci* **3**, 89-97 (1995).

[2] Wernicke, C. Grundrisse der Psychiatrie. *Thieme* (1906).

[3] Bleuler, E. *Dementia Praecox or the Group of Schizophrenias*. (International Universities Press, 1911).

[4] Gheiratmand, M., Rish, I., Cecchi, G. A., Brown, M. R., Greiner, R., Polosecki, P. I., . . . Dursun, S. M. Learning stable and predictive network-based patterns of schizophrenia and its clinical symptoms. Npj Schizophrenia, 3(1). doi:10.1038/s41537-017-0022-8 (2017).

[5] Guillermo Cecchi, Irina Rish, Benjamin Thyreau, Bertrand Thirion, Marion Plaze, et al.. Discriminative Network Models of Schizophrenia. Neural Information Processing Systems, Dec 2009, Vancouver, Canada. 2009.

[6] Rish, I. *et al.* Schizophrenia as a network disease: disruption of emergent brain function in patients with auditory hallucinations. *PLoS One* **8**, e50625, doi:10.1371/journal.pone.0050625 (2013).

[7] Wolfers, T., Buitelaar, J. K., Beckmann, C. F., Franke, B. & Marquand, A. F. From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. *Neurosci Biobehav Rev* **57**, 328-349, doi:10.1016/j.neubiorev.2015.08.001 (2015).
[8] Miller, Tandy J., et al. "Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability." Schizophrenia bulletin 29.4 (2003): 703-715.
[9] Cornblatt BA, Auther AM, Niendam T, et al. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. Schizophr Bull. 2007;33(3):688-702.
[10] Colibazzi, Tiziano, et al. "Aberrant Temporal Connectivity in Persons at Clinical High Risk for Psychosis." Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 2.8 (2017).
[11] Breedlove, S. M., Rosenzweig, M. R., & Watson, N. V. (2008). Biological psychology: An introduction to behavioral, cognitive and clinical neuroscience. Sunderland, MA, MA: Sinauer Associates.
[12] Treatment Advocacy Center. (2008, August 13). Schizophrenia – Fact Sheet. Retrieved from http://www.treatmentadvocacycenter.org/evidenceand-research/learn-more-about/25-schizophrenia-fact-sheet
[13] Tamminga, C. A., & Medoff, D. R. (2000). The biology of schizophrenia. Dialogues in clinical neuroscience, 2(4), 339-48.
[14] Andreasen, N. C., Flaum, M., Arndt, S., Alliger, R., & Swayze, V. W. (1991). Positive and Negative Symptoms: Assessment and Validity. Negative Versus Positive Schizophrenia, 28-51. doi:10.1007/978-3-642-76841-5_3
[15] Kring, A. M. (1999). Emotion in Schizophrenia. Current Directions in Psychological Science, 8(5), 160-163. doi:10.1111/1467-8721.00038
[16] Fornito, A., Zalesky, A., Pantelis, C., & Bullmore, E. T. (2012). Schizophrenia, neuroimaging and connectomics. NeuroImage, 62(4), 2296-2314. doi:10.1016/j.neuroimage.2011.12.090
[17] Chung, Y., & Cannon, T. D. (2015). Brain imaging during the transition from psychosis prodrome to schizophrenia. The Journal of nervous and mental disease, 203(5), 336-41.
[18] Cheng, S. C., & Schepp, K. G. (2016). Early Intervention in Schizophrenia: A Literature Review. Archives of Psychiatric Nursing, 30(6), 774-781. doi:10.1016/j.apnu.2016.02.009
[19] Lindquist, M. A. (2008). The Statistical Analysis of fMRI Data. Statistical Science, 23(4), 439-464. doi:10.1214/09-sts282
[20] Rish, I., Thyreau, B., Thirion, B., Plaze, M., Paillere-martinot, M. L., Martelli, C., ... & Cecchi, G. A. (2009). Discriminative network models of schizophrenia. In *Advances in Neural Information Processing Systems* (pp. 252-260).