

Reproducing fMRI Data Analysis on Brain Connectivity

Li, Jie
Jay4869

Li, Zeyu
lizeyuyuz

Yun, Chuan
ay2456

Zhang, Qiangyuan
amandazhang

December 9, 2015

Abstract

In this research paper, we analyse the fMRI data from the study done by Repov and Barch exploring the relationship between schizophrenia and brain connectivity [1, 2]. Instead of reproducing ANOVA analysis of within and between brain network connectivity, we focus on analysing fMRI scans of one subject using generalized linear model. Before data analysis, we smoothed the data spatially and corrected noise in the fMRI data. On the preprocessed data, we perform student t test to identify the voxels responding to designed stimulus. Having localized the activated voxels, we closely examine the behaviour of these voxels and validated our assumptions.

1 Introduction

The paper of which our research is based is titled “Working memory related brain network connectivity in individuals with schizophrenia and their siblings” [1, 2]. Schizophrenia is a chronic, severe, and disabling brain disorder. Previous studies have shown that changes in the function of a single brain region, or even a brain system, cannot explain the functional impairments seen in this illness [1]. However, Repovs and Barch attempts to show that individuals with schizophrenia have reduced connectivity within and between neural networks, which could be a forward step to understanding schizophrenia.

Repov and Barch systematically examine changes in functional connectivity across rest and different task states in order to make an inference on characteristics of schizophrenia. They asseess connectivity using blood oxygen level dependent (BOLD) time series acquired using fMRI. They find four types of participates, individuals with schizophrenia, the siblings of individuals with schizophrenia, healthy controls and the siblings of healthy controls. They designed three working memory loads of an N-back task and designated four regions of interest (ROIs). The four brain networks are: (1) Dorsal fronto-parietal network (FP), (2) Cingulo-opercular network (CO), (3) Cerebellar network (CER) and (4) ”Default mode” network (DMN). The objective of the study is to examine the altered functional connectivity within and between these four brain networks when the participant perform a designated N-back task. With ANOVA analysis, it is found that individuals with schizophrenia and their siblings show consistent reductions in connectivity between both the FP and CO networks with the CER network.

We narrow our focus and analysis on a single schizophrenic subject. The fMRI data in the study shows the changes in blood flow in the brain, and analysing this data can help us understand how schizophrenic brains respond during task performance. We detail in the following the steps we take to analyse the data. We attempted both spatial and temporal smoothing on the voxels in order to reduce noise. We also performed spectral analysis on fMRI data to correct noise in the data. After preprocessing the data, we fit multiple linear regression model using two different sets of experiment conditions. Having obtained coefficient estimates, we performed Student t-tests on each voxel to examine whether the particular voxel has show significant activity subject to appropriate thresholds. Performing t-tests relying on a set of assumptions that we need to validate. We performed the Shaprio-Wilk test to check the normality assumption.

2 Data

The data used in their published paper is available on the website OpenFMRI.org. The are many data files grouped by subjects, but we will be focusing on only one schizophrenic subject—subject 001, who

is male, Caucasian and Schizophrenic. The data has four dimensions (91 x 109 x 91 x 137), consisting of three dimensional voxels that are moving across time. We removed the first four outlying data along the time dimension, which leaves us 133 images. (Need to add how to come to understand the BOLD signal = signal + noise, and the necessity of remove noise)

3 Linear modelling methods

The basic idea of linear models is to find the relationship between a response variable and one or more explanatory variables (regressors). For our purposes, we attempt to define a relationship between the BOLD signals and the working memory tasks. Mathematically, our model is as follows:

$$y = X\beta + \epsilon, X = \{1, x_1, x_2, \dots, x_k\}$$

To perform hypothesis tests using this model, we assume linearity of the relationship, and $E[\epsilon|X] = 0$, $Var[\epsilon|X] = \Omega$. We solve for linear regression using generalized least square: $\hat{\beta} = (X^T\Omega^{-1}X)^{-1}X^T\Omega^{-1}Y$. After finding estimates for the β s, we perform Student t-test for each β_i , where $H_0 : \beta_i = 0, i = \{1, 2, \dots, k\}$.

3.1 Block design

3.1.1 Convolution and design matrix

We need a good design matrix, containing data on the explanatory variables, to be used in our linear models to explain as much variation in the observed fMRI data as possible. We studied the design of the experiment closely. fMRI scans were acquired while participants perform specified memory tasks. They are to respond for each letter shown whether it was the same as a pre-specified letter (0-back), the same as the immediately preceding letter (1-back), or the same as the letter shown two trials previously (2-back). For our subject, there were three BOLD runs, each consisting of two blocks of 0-back, 1-back, or 2-back working memory task.

We examined the seven study conditions closely. The conditions are unusual from what we are familiar with in that they are consisted of fractions of durations and negative amplitudes. The first condition consists of the start-cues, the second condition consists of the task targets (amplitudes are all one), the third condition consists of the task non-target (the positive amplitude indicates correct identification of the target and the negative amplitude indicates the non-response to the non-target), the fourth condition consists of end-cues, and the fifth condition consists of durations of the two task blocks. There is no information on what the sixth file mean, and the seventh condition consists of the erroneous responses.

Since we were only familiar with convolution for block design, we initially used only the condition one, four and five, which give us the start time, the duration, and end time of the two task blocks. Due to the delay of signal after the onset of neural activity, we make use of the hemodynamic response function (HRF). The idea is to model the signal response as the convolution of the stimulus function with the HRF. Convoluting with stimulus timing allows us to get idealized response as an explanatory variable (a regressor included in design matrix). We bear in mind the limitations of this design matrix due to its omission of trials within each of two task blocks.

3.1.2 Linear regression results

We included three convolved response in our design matrix along with a constant term. We obtained three beta estimates for each voxel, and we perform t-test where null hypothesis claims that $\beta_1 + \beta_2 + \beta_3 = 0$.

3.2 Event-related design

3.2.1 Convolution and design matrix

Previously we used block design that has too long of duration, and therefore unable to explain much of the variation in the BOLD signals. We revised convolution of stimulus using a finer and higher resolution on time. With trials within task blocks convolved and included in the design matrix, we expect the result from linear regression have higher explanatory power.

3.2.2 Data Preprocessing

To reduce noise and improve signal in the fMRI data, we attempted smoothing of the fMRI data. We experimented with smoothing temporally using discrete cosine transformations as well as Gaussian kernel smoother. We added a set of discrete cosine transformation basis to the design matrix to filter out the low frequency noises present in the BOLD signal.

We also spatially smoothed the data.

3.3 Generalized linear regression results

(add)

3.4 Validation

We use Shapiro-Wilk's test to validate our normality assumption on the error term. (More)

4 Lessons and obstacles

The hardest part was to connect the theoretical models with fMRI data to make appropriate inferences.

References

- [1] G. REPOVS AND D. M. BARCH, *Brain network connectivity in individuals with schizophrenia and their siblings*, Biological Psychiatry, 69 (2011), pp. 967–973.
- [2] ———, *Working memory related brain network connectivity in individuals with schizophrenia and their siblings*, Frontiers in Human Neuroscience, 6 (May 2012), pp. 1–15.