Association of Neural Activity with Response to Dietary Regimine for Treatment of Depression

Maime Guan, Lingge Li, Duy Ngo,
Fulya Ozcan, Dustin Pluta, Yuxiao Wang

November 8, 2016

Motivating Questions

- Is SPECT brain imaging predictive of response to treatment of depression through a dietary regimine?
- Are there specific brain regions associated with response to treatment?
- How does the association of brain activation with treatment response differ across to demographic subgroups?

Project Goals

- Build a model to predict treatment response from SPECT imaging and clinical covariates.
- Test the strength of association between brain activation and treatment response for selected scientifically relevant brain regions.
- Identify patient subgroups for which the association of activation with response differs

Background

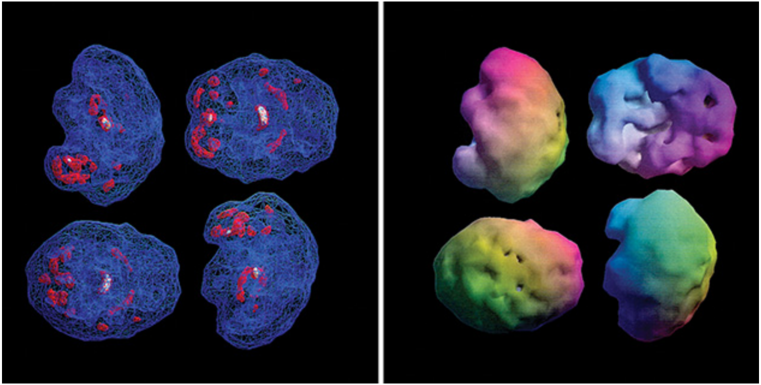
- SPECT uses a radioactive tracer to measure cerebral blood flow.
- A 3D map of the brain is produced, indicating areas of high activity and low activity.
- Results of previous studies using SPECT:
  - Distinguish traumatic brain injury (TBI) cases from post-traumatic stress disorder cases with high accuracy (Raji 2015).
  - Distinguish NFL players with repeated head trauma from healthy population (Amen 2016).
  - Distinguish autism spectrum disorder from healthy subjects (Amen 2016).

Background

A SPECT Scanner

Background

A SPECT Image



Data Description

Patients

1. 828 patients diagnosed with depression. Most patients had a history of various mental illnesses or disorders, and were typically not responsive to traditional treatments before being referred to Dr. Amen's clinic.
2. Basic demographic information was collected for all patients.
3. Patients were given preliminary clinical exams to assess any medical conditions, and any co-diagnoses recorded.
4. Single Photon Emission Computed Tomography (SPECT) scans were also collected for all patients during the initial exam.

Data Description

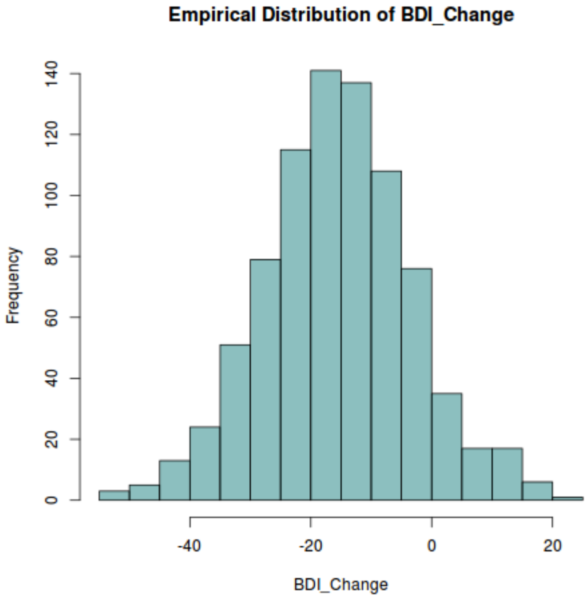
Response Values

1. During the initial exam, patients were given the Beck Depression Inventory (BDI) survey to measure level of depression.
2. Each patient was asked to follow a dietary regimine of supplements and vitamins designed to naturally treat depression and other disorders.
3. After 5-7 months of dietary treatment patients were given a follow-up BDI survey.
4. "Raw" response values are taken as

$$\Delta BDI = Pre\_BDI - Post\_BDI$$

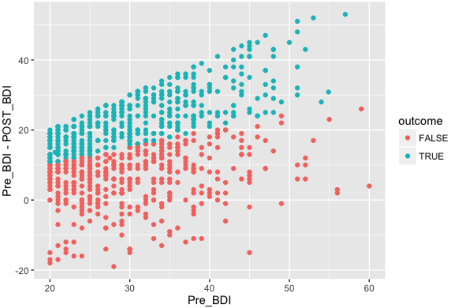
5. These values were then dichotomized into *Responders* and *Non-responders*.

Data Description



Who are the responders

In the original data, patients with BDI less than 25 prior to treatment were all labeled as non-responders. Our models were able to get above 60% accuracy but they were actually only finding patients with low BDI. Therefore, we changed the labels as one would expect.

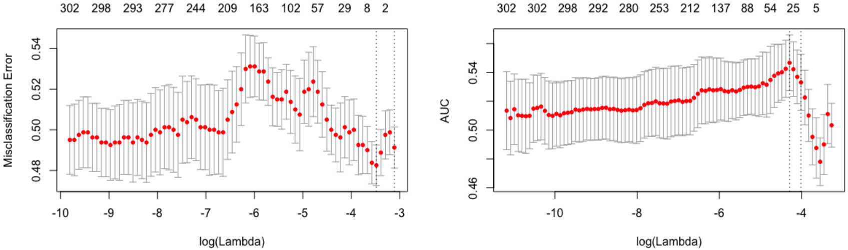


Data Description

Covariate	N (%)	Mean BDI_Change
Gender		
Male	460 (0.56)	-15.26
Female	368 (0.44)	-14.86
Age Group		
Pediatric	14 (0.02)	-14.68
Adult	753 (0.91)	-15.13
Geriatric	61 (0.07)	-14.02
Co-diagnosis		
ADHD	439	-14.98
Anxiety	631	-14.89
Bipolar	71	-15.35

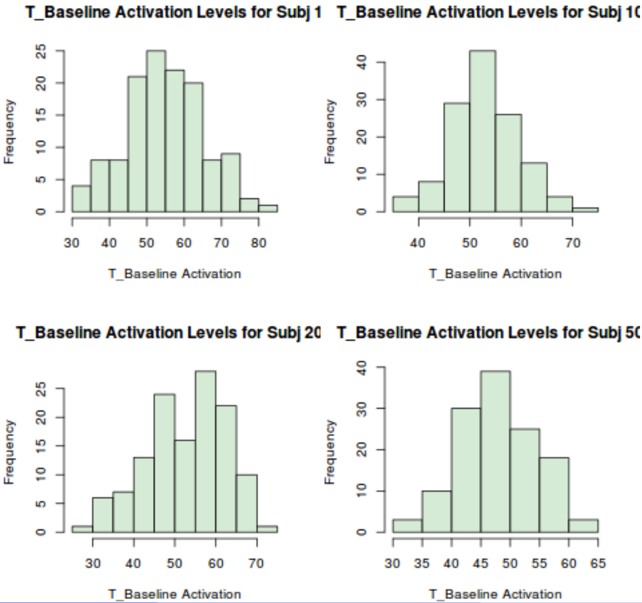
Basic Predictive Modelling

We tried logistic regression with $L1$ penalty, support vector machine, and boosted trees. However, they performed barely better than tossing a coin.

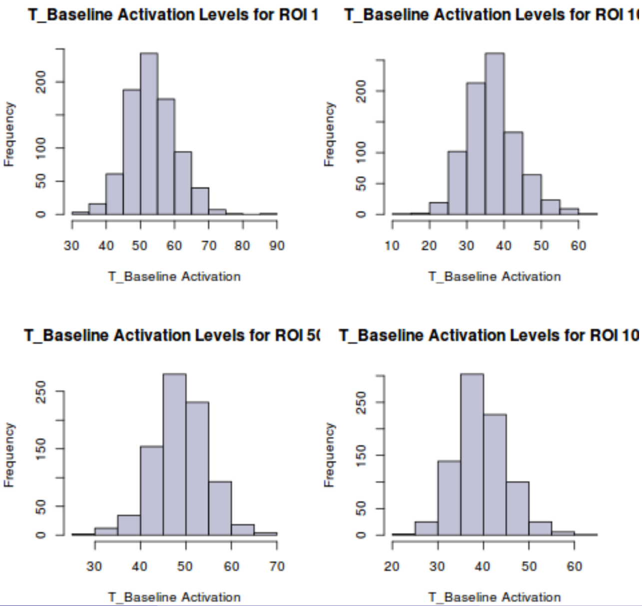


Cross validation LASSO results

Exploratory Analysis

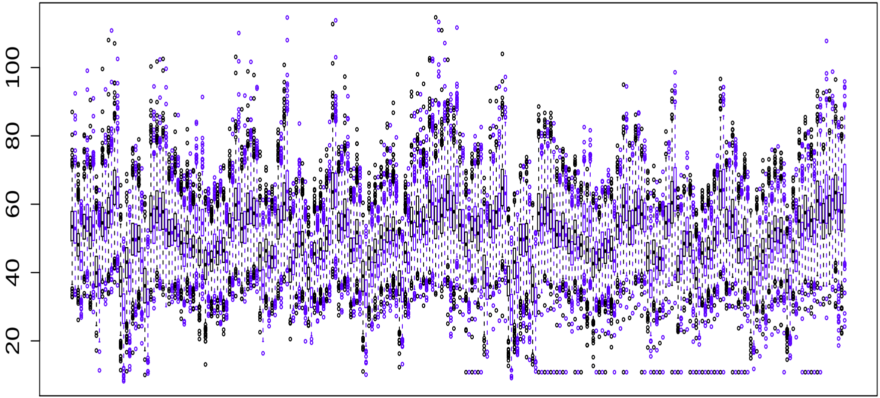


Exploratory Analysis



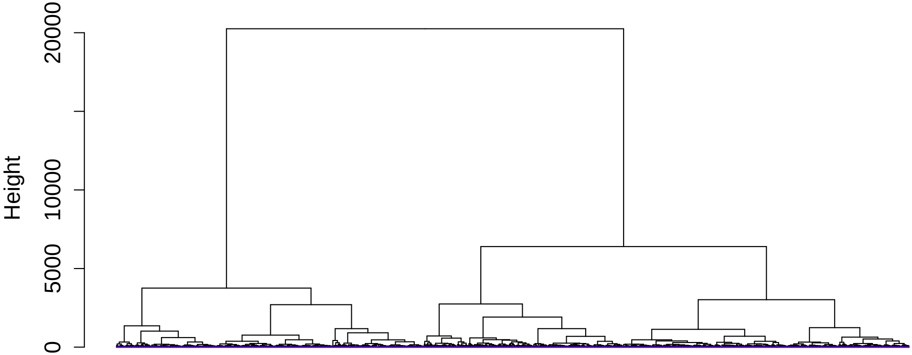
Exploratory Analysis

- Number of subject: 1041. Number of brain region: 256.



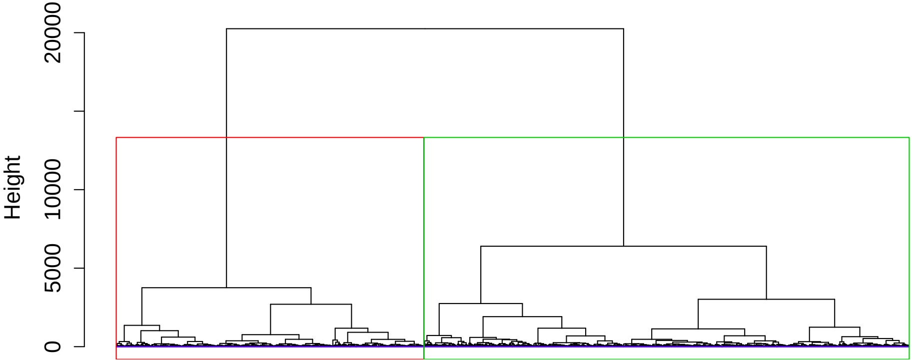
Clustering Analysis: Hierarchical Tree

- Hierarchical tree (dendrogram) for all subjects.



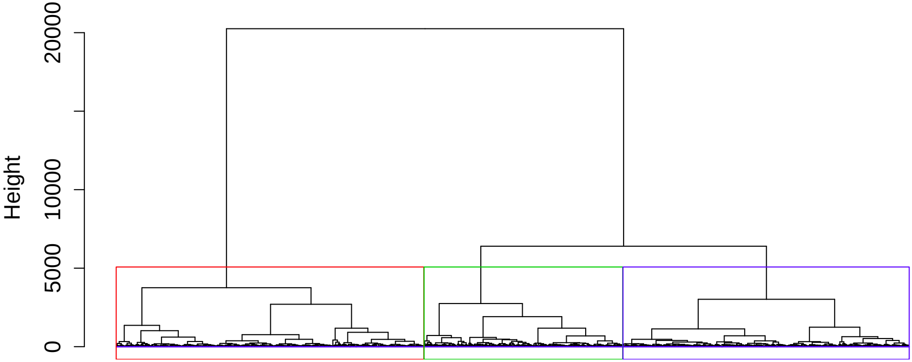
Clustering Analysis: Hierarchical Tree

- Number of cluster = 2



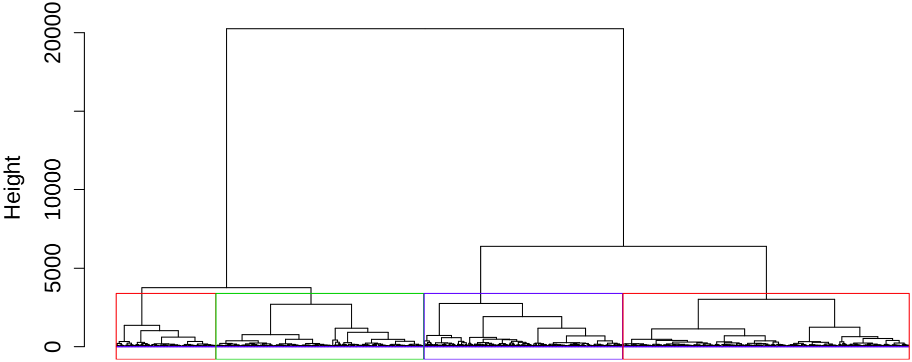
Clustering Analysis: Hierarchical Tree

- Number of cluster = 3



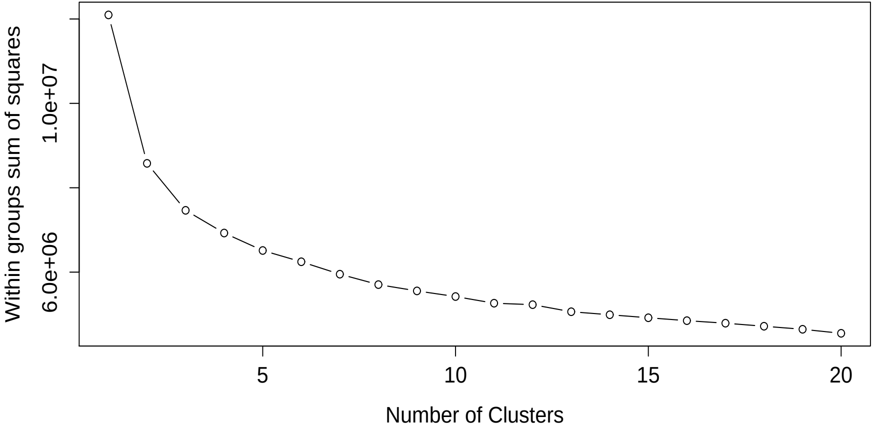
Clustering Analysis: Hierarchical Tree

- Number of cluster = 4 and so on.



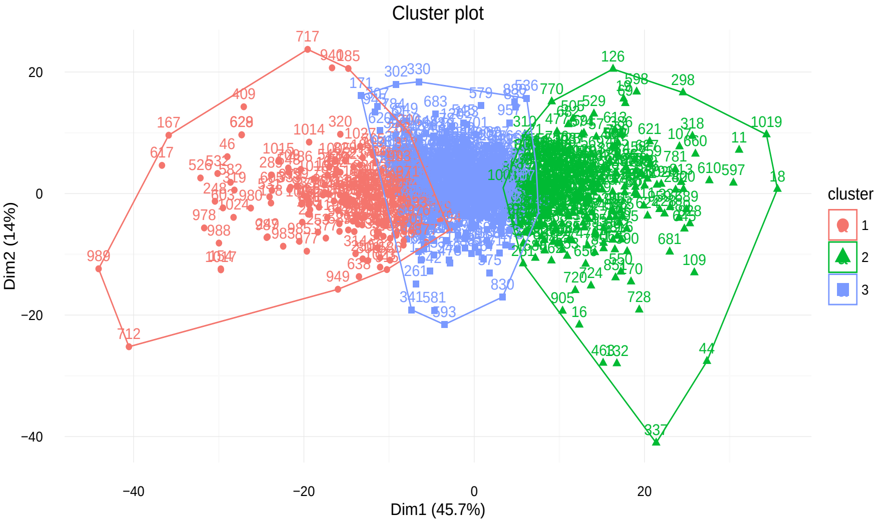
Clustering Analysis: K-Mean

- Decide number of cluster based on the percentage of variance explained as a function of the number of clusters



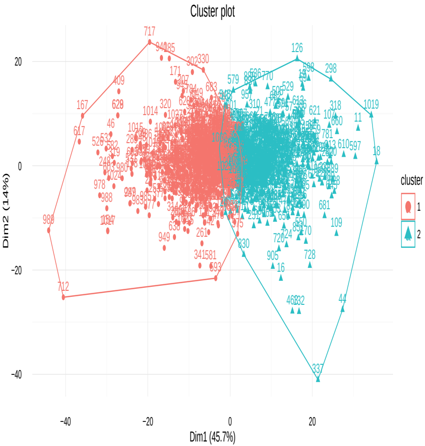
Clustering Analysis: K-Mean

- Number of cluster = 3.

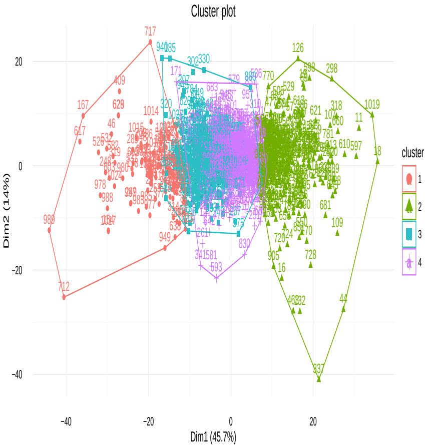


Clustering Analysis: K-Mean

Number of cluster = 2



Number of cluster = 4



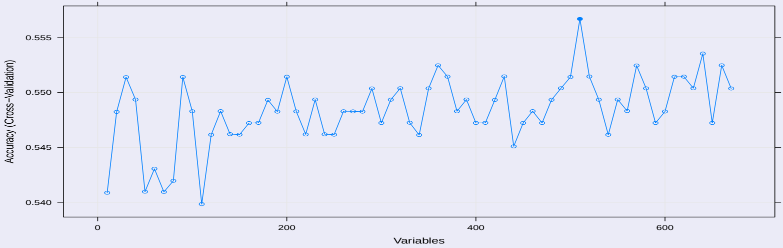
Feature Reduction

Challenges

- 954 observations, 1654 possible features, 11 main feature groups
- Missing data (especially for certain group of features)

Step-0: Feature Reduction Using All Possible Features

- Top 5 features:
  "Pre_GSI", "PRE_PST", "Pre_PSDI", "Pre_BDI", "Baseline_location"



Feature Reduction

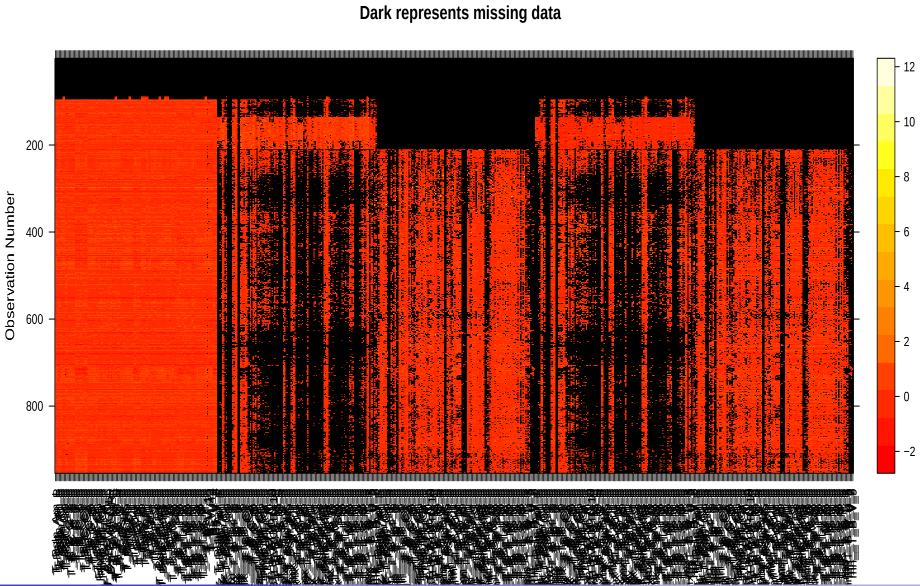
Step-1: Grouping Features

- 11 feature groups:
  - Demographics, Comorbidities, T Baseline, T Concentration, Baseline, Concentration, max cluster size, min cluster size, max cluster T Baseline, min cluster T Baseline

Step-2: Imputation

1. Remove columns with more than a certain threshold of missing values
2. Impute the remaining missing values using Gelman's **mi** package

Feature Reduction



Feature Reduction

Step-3: Imputation

1. Remove columns with more than a certain threshold of missing values
2. Impute the remaining missing values using Gelman's **mi** package

Step-4: K-means within feature groups and outcome

1. K-means on comorbidities
2. K-means on groups

Feature Reduction

Step-5: Overall Feature Reduction-Recursive Feature Elimination

1. Constructs an Learning Vector Quantization (LVQ)
2. Variable importance is estimated for each group
3. Most important variables are merged in a new feature matrix
4. Number of features reduced from 1655 to 652

Step-6: Logistic Regression to find significant variables

- Number of features reduced from 652 to 94

Feature Reduction

[1] "T_Concentration_Cerebellum_7b_R"
[2] "T_Concentration_Fusiform_R"
[3] "T_Concentration_Paracentral_Lobule_L"
[4] "T_Concentration_ParaHippocampal_L"
[5] "T_Concentration_Supp_Motor_Area_R"
[6] "T_Concentration_Temporal_Mid_Post_L"
[7] "T_Concentration_Temporal_Mid_Post_R"
[8] "T_Concentration_Vermis_8"
[9]"max_cluster_size_T_Concentration_cerebellum_6_L"
[10] "max_cluster_size_T_Concentration_Thalamus_L"
[11] "max_cluster_T_Concentration_Putamen_R"
[12] "km_gr3$cluster"
[13] "Activation_Amygdala_R"
[14] "Activation_Cerebellum_10_L"
[15] "Activation_Cerebellum_10_R"
[16] "Activation_Cerebellum_6_L"
[17] "Activation_Cerebellum_7b_L"
[18] "Activation_Cerebellum_7b_R"
[19] "Activation_Cerebellum_Crus1_L"
[20] "Activation_Cingulum_Mid_L"
[21] "Activation_Cingulum_Post_R"
[22] "Activation_Frontal_Inf_Orb_R"
[23] "Activation_Frontal_Mid_L"
[24] "Activation_Frontal_Mid_Orb_R_10"
[25] "Activation_Frontal_Sup_L"
[26] "Activation_Frontal_Sup_Medial_L"
[27] "Activation_Frontal_Sup_Orb_R"
[28] "Activation_Frontal_Sup_R"
[29] "Activation_Fusiform_R"
[30] "Activation_Hippocampus_R"
[31] "Activation_Insula_L"
[32] "Activation_Occipital_Sup_L"
[33] "Activation_Pallidum_R"
[34] "Activation_Parietal_Sup_L"
[35] "Activation_Postcentral_R"
[36] "Activation_Precuneus_L"
[37] "Activation_Precuneus_R"
[38] "Activation_Rectus_R"
[39] "Activation_SupraMarginal_R"
[40] "Activation_Temporal_Inf_Ant_L"
[41] "Activation_Temporal_Inf_Mid_R"
[42] "Activation_Temporal_Mid_Ant_L"
[43] "Activation_Temporal_Mid_Ant_R"
[44] "Activation_Temporal_Mid_Post_L"
[45] "Activation_Temporal_Mid_Post_R"
[46] "Activation_Temporal_Pole_Sup_L"
[47] "Activation_Temporal_Pole_Sup_R"
[48] "Activation_Thalamus_R"
[49] "Activation_Vermis_8"
[50] "T_Baseline_Calcarine_L"
[51] "T_Baseline_Caudate_R"
[52] "T_Baseline_Cerebellum_4_5_R"
[53] "T_Baseline_Cerebellum_6_R"
[54] "T_Baseline_Cerebellum_7b_R"
[55] "T_Baseline_Frontal_Inf_Tri_L"
[56] "T_Baseline_Frontal_Mid_Orb_L"
[57] "T_Baseline_Frontal_Sup_Orb_R"
[58] "T_Baseline_Fusiform_R"
[59] "T_Baseline_Lingual_L"
[60] "T_Baseline_Occipital_Inf_L"
[61] "T_Baseline_Pallidum_L"
[62] "T_Baseline_Pallidum_R"
[63] "T_Baseline_Paracentral_Lobule_L"
[64] "T_Baseline_Parietal_Sup_R"
[65] "T_Baseline_Postcentral_R"
[66] "T_Baseline_Precentral_L"
[67] "T_Baseline_Precuneus_L"
[68] "T_Baseline_Precuneus_R"
[69] "T_Baseline_Putamen_R"
[70] "T_Baseline_Rectus_R"
[71] "T_Baseline_Supp_Motor_Area_L"
[72] "T_Baseline_Supp_Motor_Area_R"
[73] "T_Baseline_SupraMarginal_L"
[74] "T_Baseline_SupraMarginal_R"
[75] "T_Baseline_Temporal_Inf_Ant_L"
[76] "T_Baseline_Temporal_Mid_Ant_L"
[77] "T_Baseline_Temporal_Pole_Mid_L"
[78] "T_Baseline_Temporal_Sup_Ant_L"
[79] "T_Baseline_Thalamus_L"
[80] "T_Baseline_Thalamus_R"
[81] "T_Baseline_Vermis_3"
[82] "T_Baseline_Vermis_8"
[83] "T_Baseline_Vermis_9"
[84] "max_cluster_size_T_Baseline_cerebellum_6_L"
[85] "max_cluster_size_T_Concentration_cerebellum_Crus1_I
[86] "max_cluster_size_T_Concentration_Putamen_L"
[87] "max_cluster_size_T_Concentration_Thalamus_L"
[88] "min_cluster_size_T_Baseline_Postcentral_R"
[89] "min_cluster_size_T_Concentration_cerebellum_Crus2_F
[90] "max_cluster_T_Baseline_Calcarine_L"
[91] "max_cluster_T_Baseline_cerebellum_6_R"
[92] "max_cluster_T_Baseline_cerebellum_Crus1_L"
[93] "T_Baseline_Vermis_9"
[94] "km_gr11$cluster"

Bayesian Logistic Regression

Still has the same components:

1. Random: $Y_i \sim Bernoulli\left(p_i\right)$
2. Systematic and link: $\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots$

To estimate the β_k's in a Bayesian framework, we define priors for the coefficients of the predictors. Using a multivariate normal prior for β_k:

$$\beta_k \sim MVN\left(\mu = 0, 1/\sigma^2 = .0001\right), k = 1, 2, \ldots$$

Can be implemented with 'MCMCpack' or bayesglm (from 'arm' library) in R, or with JAGS.

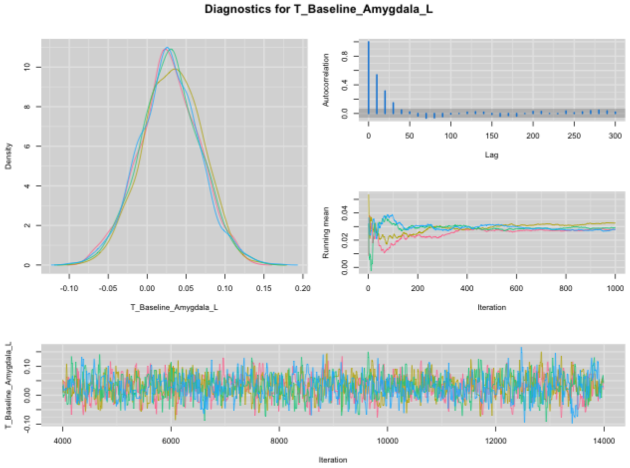
Bayesian Logistic Regression

Models tried so far:

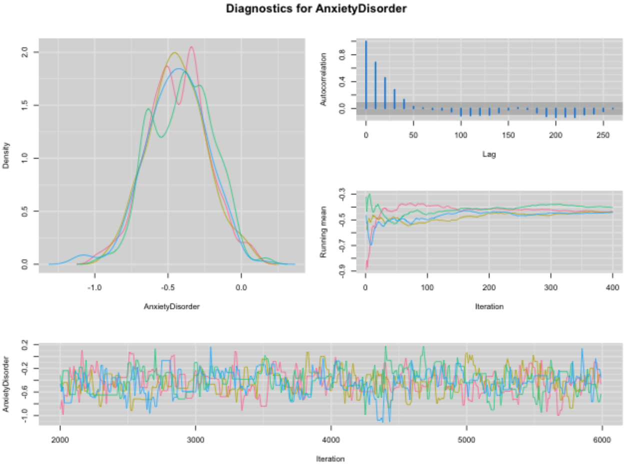
- Using *all 128 brain regions* as predictors (chains did not converge).
- Using brain regions known to be *associated with depression*: amygdala, hippocampus, thalamus, and insula. (prediction = 52.5%)
- Using some *comorbidities* of interest (anxiety, ADHD, substance abuse, frontal lobe dysfunction, brain trauma, and PTSD) with *compliance score*. (prediction = 51.5%)
- Using some comorbodities, compliance score, and brain regions associated with depression. (prediction = 54.5%)

Cross-validation with 828 patients in training set and 200 patients in test set.

This is quite representative of the findings for the brain regions: posterior samples of coefficients mostly right about 0 . . .

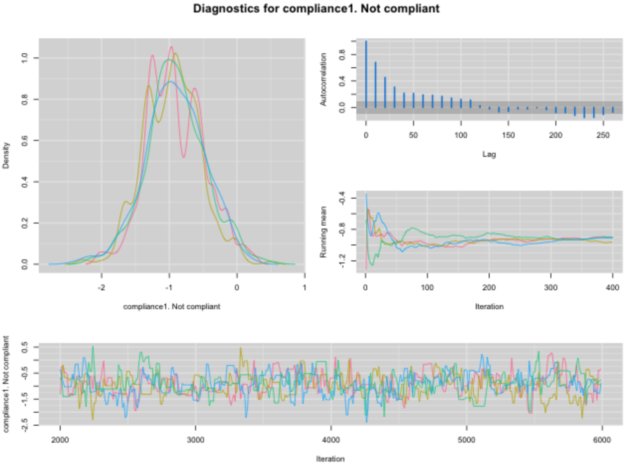


However, some coefficients may be more interesting ...



About a predicted 40% decrease in the odds of responding to treatment if a patient has anxiety as well as depression, compared to not having anxiety.

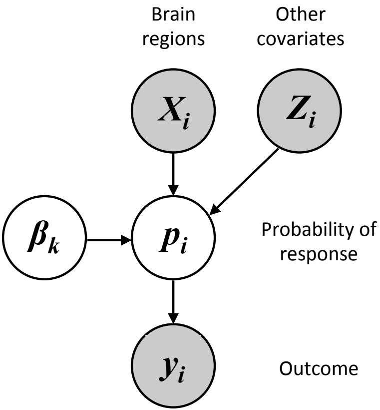
However, some coefficients may be more interesting . . .



About a predicted 63% decrease in the odds of responding to treatment if a patient self-reported being "not compliant", compared to not answering.

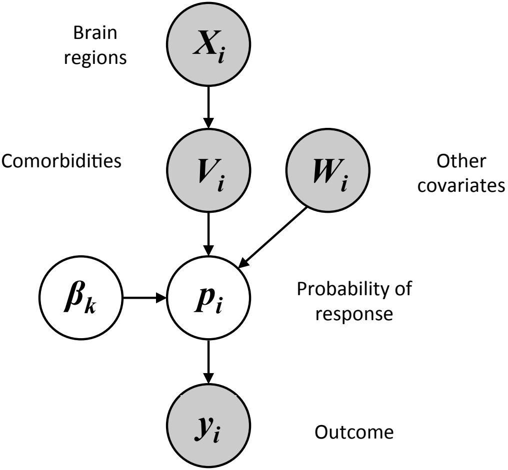
Next Steps

So far, we have only tried Bayesian logistic regression, where all the predictors are directly related to the odds of response to treatment.



Next Steps

We can also try different underlying structures, defining other possible relationships between variables. For example:



Next Steps

1. Select a dimension reduction method for SPECT covariates.
2. Build, validate, and interpret models for hypothesis testing.
3. Tune and compare prediction models.

Summary

- Exploratory analyses and preliminary modeling suggest that separating Responders from Non-responders is difficult from the available covariates.
  - This is a more challenging setting than previous studies since we are classifying within Depression patient subgroup, instead of separating cases from controls.
- There is some weak clustering by SPECT imaging, but these clusters aren't strongly associated with change in BDI. Supervised dimension reduction methods may give better groupings.

Thanks!