Results Summary

Protocol Title: Cranial electric stimulation (CES) to modify suicide risk factors in psychiatric inpatients

Study No.: HP-00065640, Date January 4th 2018

Principal Investigator: Nithin Krishna MD, 410-328-6610, 410-328-5386 (24 hour access)

ClinicalTrials.gov Identifier: NCT02846740

Christopher E. Stamper and Christopher A. Lowry Statisticians.

Data Preprocessing:

Random forest analysis does not support missing values. The following experimental variables were removed prior to analysis because there were more than 2 missing values.

HAMAT1, HAMAT2, MADRA_T2, MADRA_T3, PSQI_T2, PSQI_T3, T2time, compositetotalT2, granulocytespercent, granulocyteabsolute, lymphocyteabsloute, lymphocytepercent, monocyteabsolute, monocytepercent, basophileabs, basophiperc, eosinophilabd, Eosinphilsperc, nucleatedrbcabs, nucleatedrbcperc, MCH, MCHC, Immaturegransabs, Immaturegranspec, Magnesium, Phosphorus, TotalProtein, Albumin, AST, ALT, TotalBilirubin, AlkPhos, TSH, VITB12, Folate

If the number of missing values in an experimental variable was less than or equal to 2 the value was changed to the median of the column and documented below.

The following data points were altered:

Level of confidence for participant CES005 was changed from blank to 3 (the median of the column). This was the only missing value in this column.

wbc for participant CES030 was changed from blank to 7.8 (the median of the column). This was the only missing value in this column.

hb for participant CES030 was changed from blank to 13 (the median of the column). This was the only missing value in this column.

htc for participant CES030 was changed from blank to 40.1 (the median of the column). This was the only missing value in this column.

platelets for participant CES030 was changed from blank to 285 (the median of the column). This was the only missing value in this column.

MCV for participant CES030 was changed from blank to 87.2 (the median of the column). This was the only missing value in this column.

MPV for participant CES019 and CES030 was changed from blank to 9.85 (the median of the column). These were the only missing values in this column.

RDW for participant CES030 was changed from blank to 13.6 (the median of the column). This was the only missing value in this column.

Sodium for participant CES019 and CES030 was changed from blank to 139 (the median of the column). These were the only missing values in this column.

Potassium for participant CES019 and CES030 was changed from blank to 3.95 (the median of the column). These were the only missing values in this column.

Chloride for participant CES030 was changed from blank to 104 (the median of the column). This was the only missing value in this column.

CO_2 for participant CES019 and CES030 was changed from blank to 27 (the median of the column). These were the only missing values in this column.

AnionGap for participant CES019 and CES030 was changed from blank to 9 (the median of the column). These were the only missing values in this column.

BUN for participant CES019 and CES030 was changed from blank to 12.5 (the median of the column). These were the only missing values in this column.

Creatinine for participant CES019 and CES030 was changed from blank to .785 (the median of the column). These were the only missing values in this column.

Glucose for participant CES019 and CES030 was changed from blank to 94 (the median of the column). These were the only missing values in this column.

Calcium for participant CES019 and CES030 was changed from blank to 9 (the median of the column). These were the only missing values in this column.

The following experimental variables were removed because the values were identical to the treatment codes.
Actual_Devise_Used_ACTIVE_vs._Sham

The following experimental variables were added to examine the change between the last observation and the baseline observation for the primary target variables.
MADRA_Delta: MADRALAST_OBS – MADRA_Baseline_T0
HAMAT_Delta: HAMLAST_OBS – HAMAT0
PSQI_Delta: PSQILAST_OB – PSQI_T1

See the attached csv file, "Alpha_stim_data" for the data that were used in the analysis.

Method
Analysis done in R Version 3.4.1

Random forest results:
All random forest models were run to create 2001 trees.

*Random forest model predicting treatment group (Active vs Sham)*

OOB (out-of-bag) error rate: 60%
- This means that the model was wrong 60% of the time and correct 40% of the time.

Here is a plot showing how the model is converging as the number of trees increases. Overall this is the inverse of what you would want to see. Optimally you want the error to start high and decrease to a convergence or a plateau as the number of trees increases. The convergence is settling at an error of ~0.6 ,which is your OOB error rate.

rf
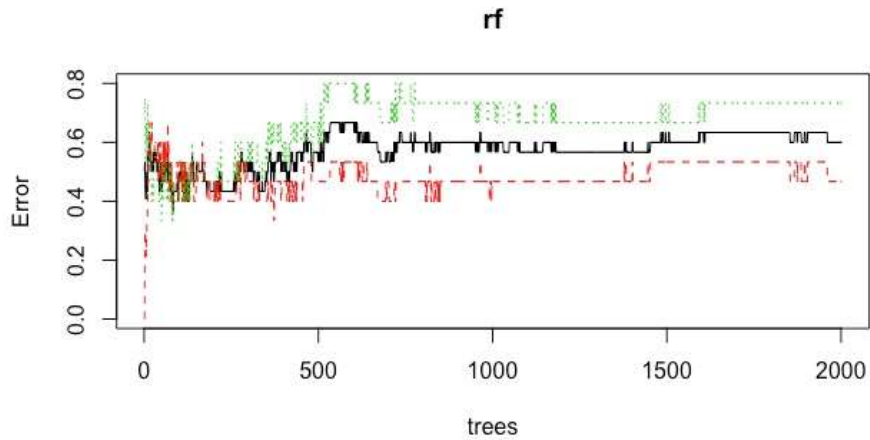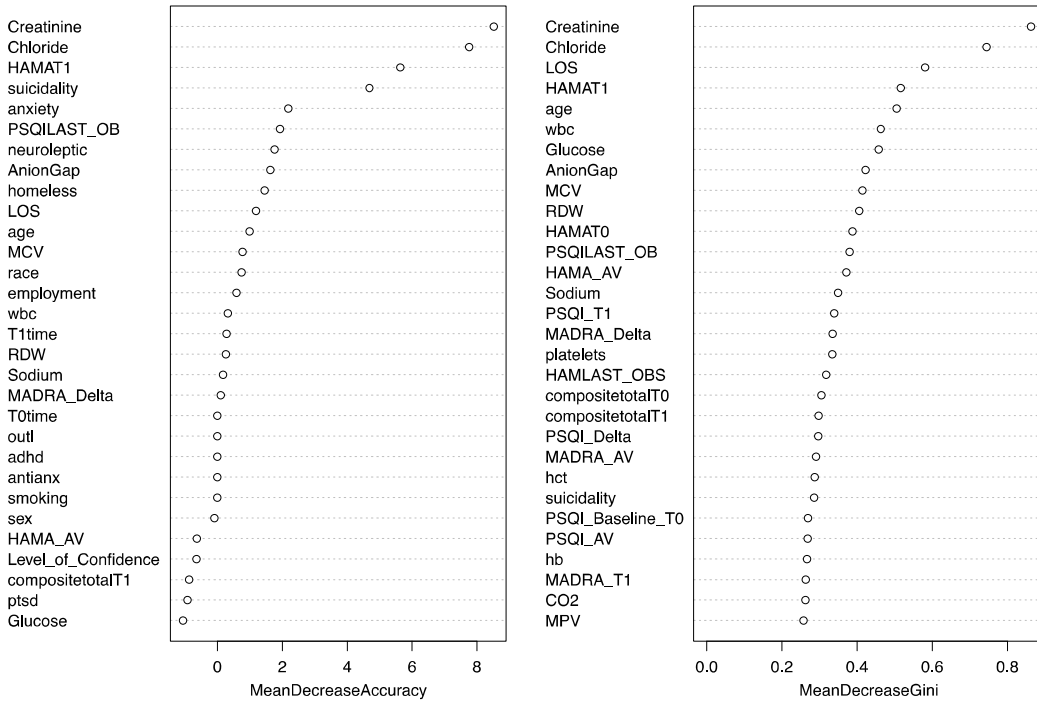

For reference this is what a more favorable plot looks like. This plot has an OOB error rate of 5.73%.

rf_H
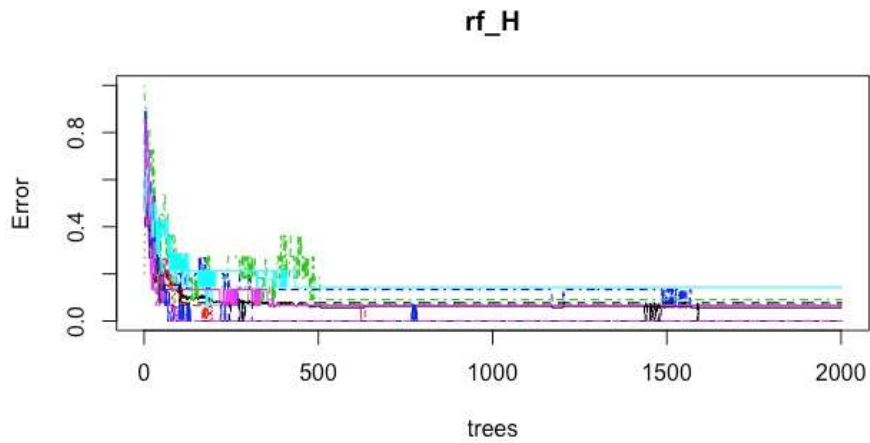
Back to your data, here are the experimental variables that are most predictive of treatment group. There are 2 accuracy metrics here that are used interchangeably based on the preference of the investigator. The accuracy numbers are a bit arbitrary and there is no threshold of significance per se. These are simply the most predictive variables of the target experimental variable.

rf

MeanDecreaseAccuracy

Creatinine
Chloride
HAMAT1
suicidality
anxiety
PSQILAST_OB
neuroleptic
AnionGap
homeless
LOS
age
MCV
race
employment
wbc
T1time
RDW
Sodium
MADRA_Delta
T0time
outl
adhd
antianx
smoking
sex
HAMA_AV
Level_of_Confidence
compositetotalT1
ptsd
Glucose

0   2   4   6   8
MeanDecreaseAccuracy

MeanDecreaseGini

Creatinine
Chloride
LOS
HAMAT1
age
wbc
Glucose
AnionGap
MCV
RDW
HAMAT0
PSQILAST_OB
HAMA_AV
Sodium
PSQI_T1
MADRA_Delta
platelets
HAMLAST_OBS
compositetotalT0
compositetotalT1
PSQI_Delta
MADRA_AV
hct
suicidality
PSQI_Baseline_T0
PSQI_AV
hb
MADRA_T1
CO2
MPV

0.0   0.2   0.4   0.6   0.8
MeanDecreaseGini

Random forest model predicting Montgomery-Asberg Depression Scale change between the final observation and baseline (MADRA_Delta)

The MADRA_Delta score is continuous; therefore, the model produces a % variation explained rather than an OOB error rate.
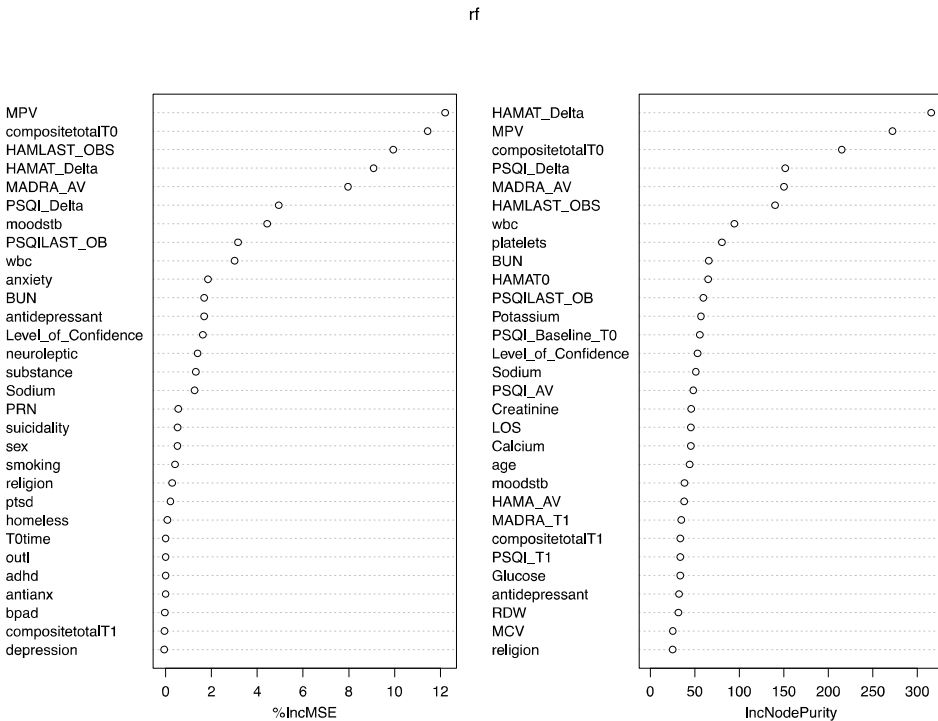
% variation explained: 14.14%
- Meaning that ~85% of the variation is unexplained

Here is the plot showing how the model is converging. Again the model is not converging very well.
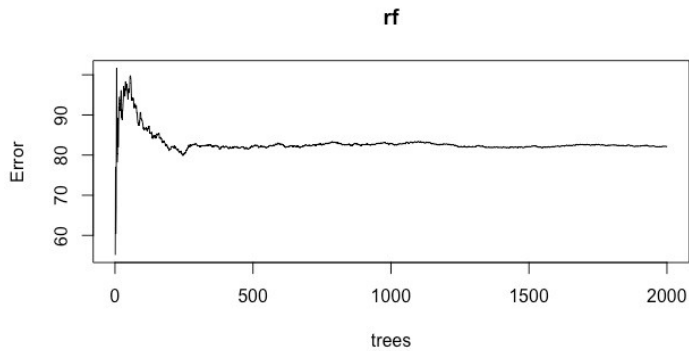
Here are the experimental variables that are explaining the most variation or are the most predictive of the MADRA_Delta score. It is encouraging that HAMAT_Delta score and PSQI-Delta score are two of the best predictors of MADRA_Delta score.
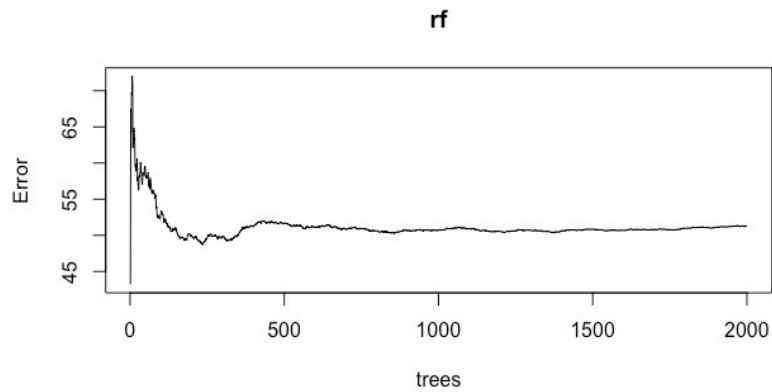
Random forest model predicting Hamilton Depression Score change between the final observation and baseline (HAMAT_Delta)
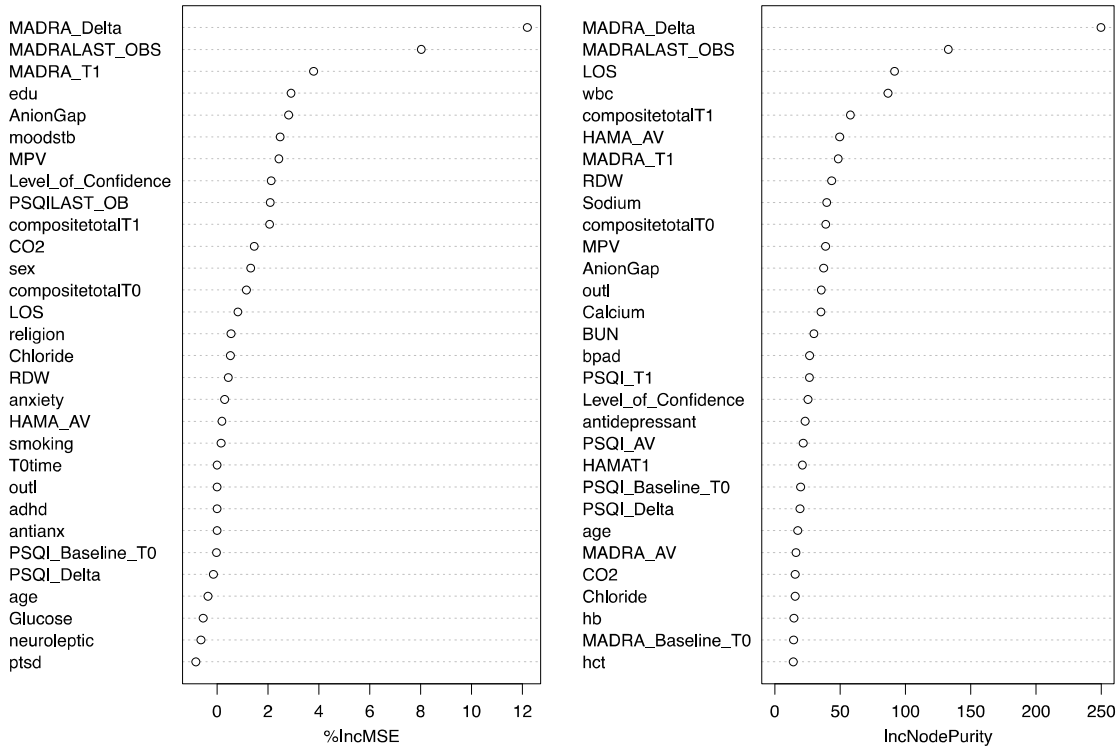% variation explained: 0.83%

- Meaning that ~99% of the variation is unexplained

Here is the plot showing how the model is converging. Again the model is not converging very well.

rf

Here are the experimental variables that are explaining the most variation or are the most predictive of the HAMAT_Delta score. It is encouraging that MADRA_Delta score is one of the best predictors of HAMAT_Delta score.
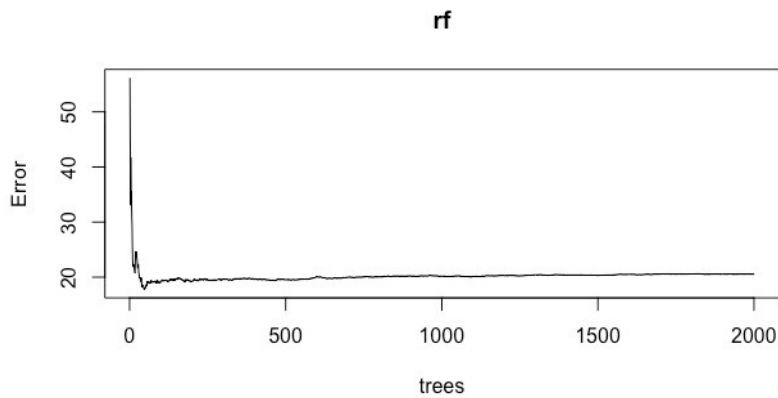
Random forest model predicting Pittsburg Sleep Quality Index score change between the final observation and baseline (PSQI_Delta).
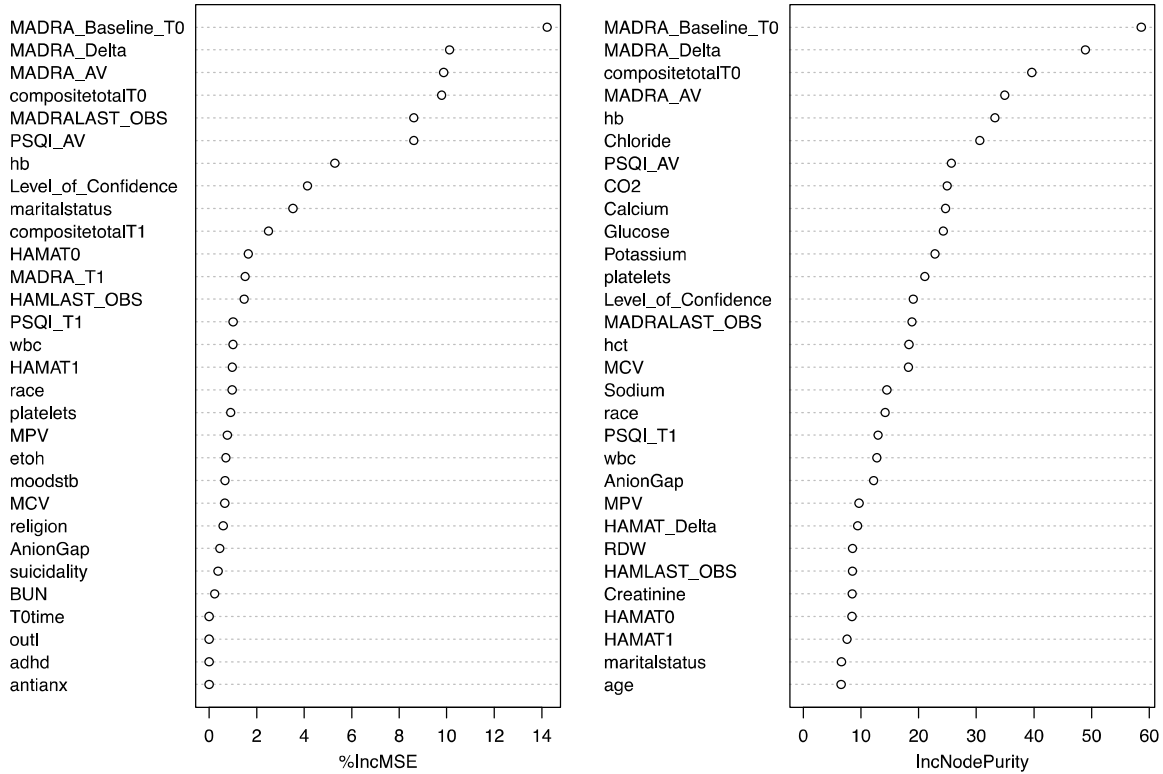% variation explained: 13.01%

- Meaning that ~87% of the variation is unexplained

Here is the plot showing how the model is converging. Better.

Here are the experimental variables that are explaining the most variation or are the most predictive of the PSQI_Delta score. It is encouraging that MADRA_Delta score is one of the best predictors of PSQI_Delta score.
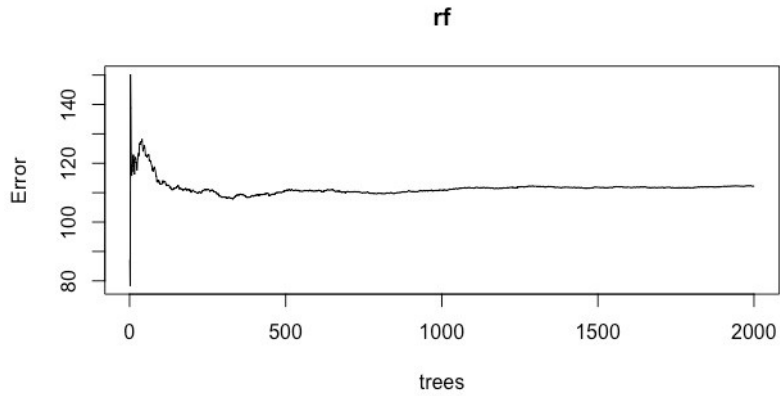
rf

Random forest model predicting Length of Stay (LOS).

% variation explained: 18.7%

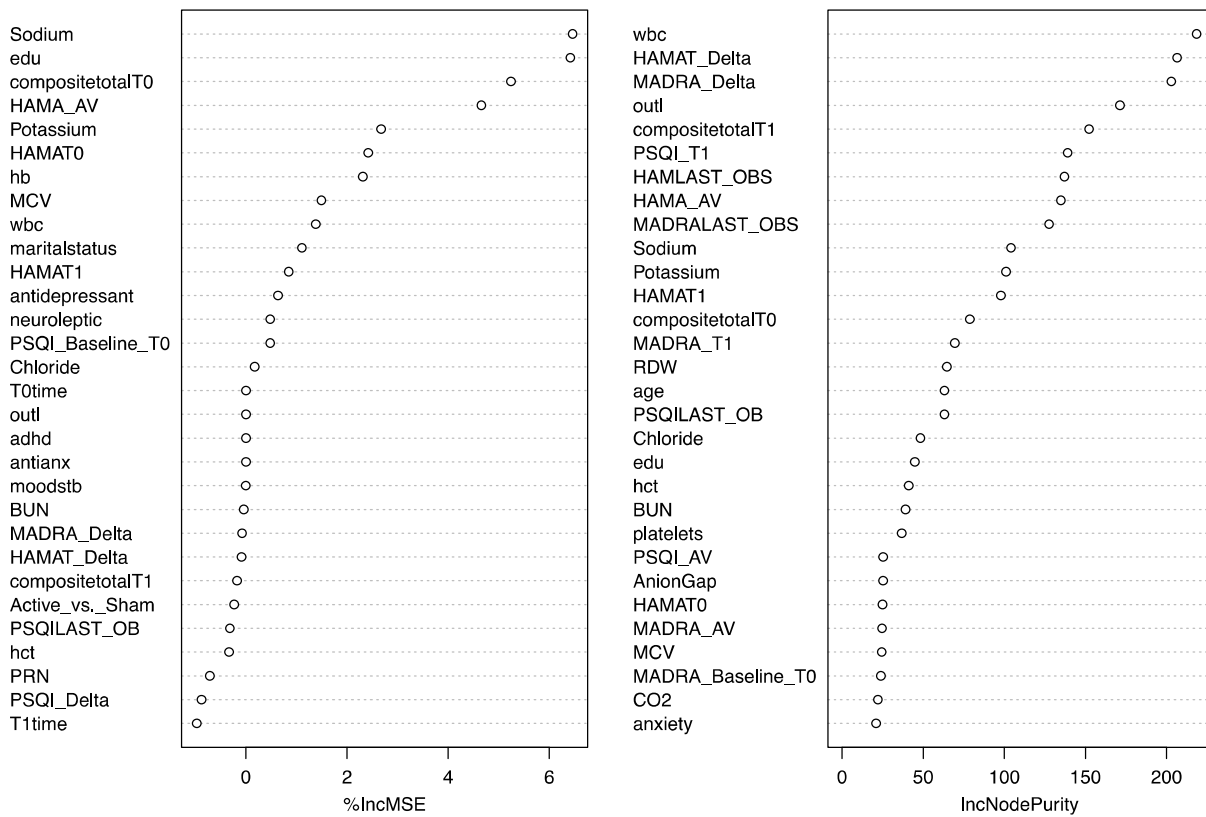- Meaning that ~81% of the variation is unexplained

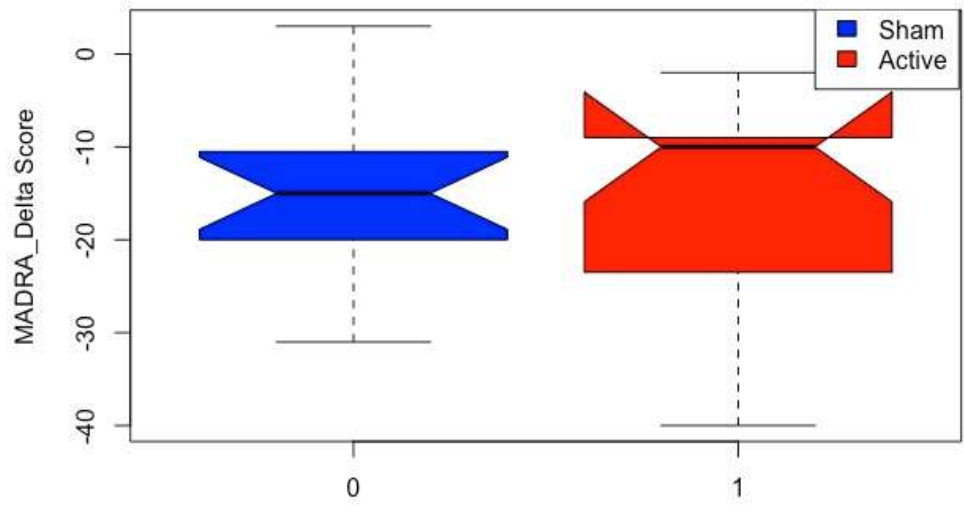Here is the plot showing how the model is converging.

rf

Here are the experimental variables that are explaining the most variation or are the most predictive of the LOS.
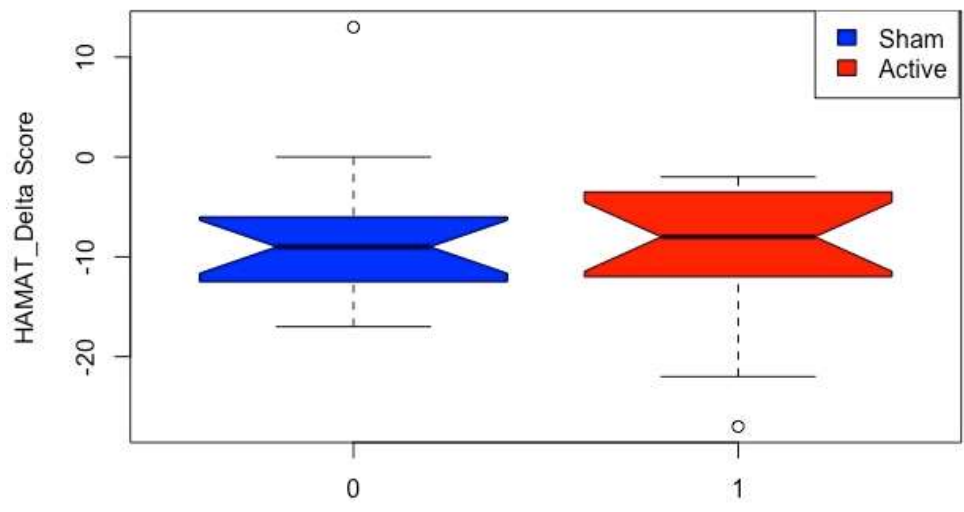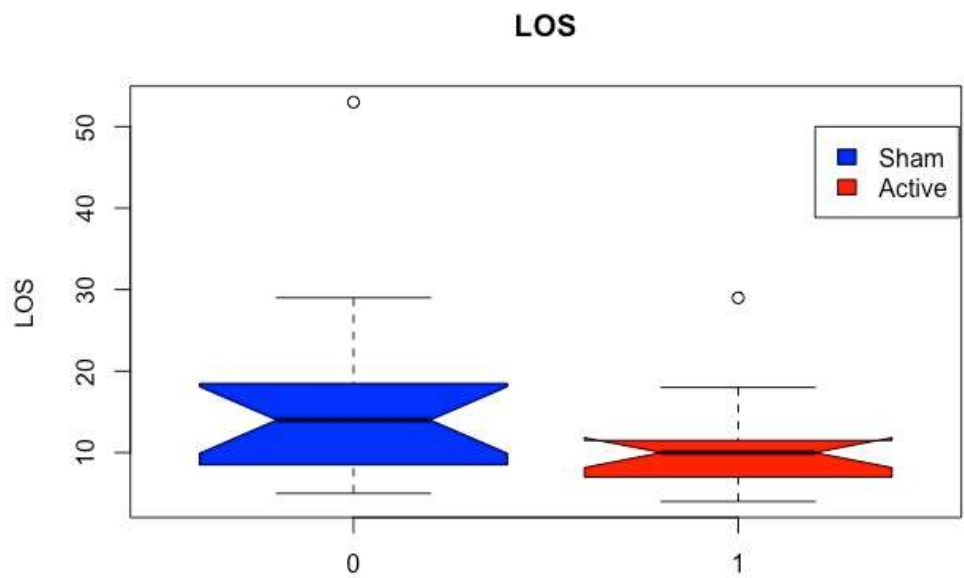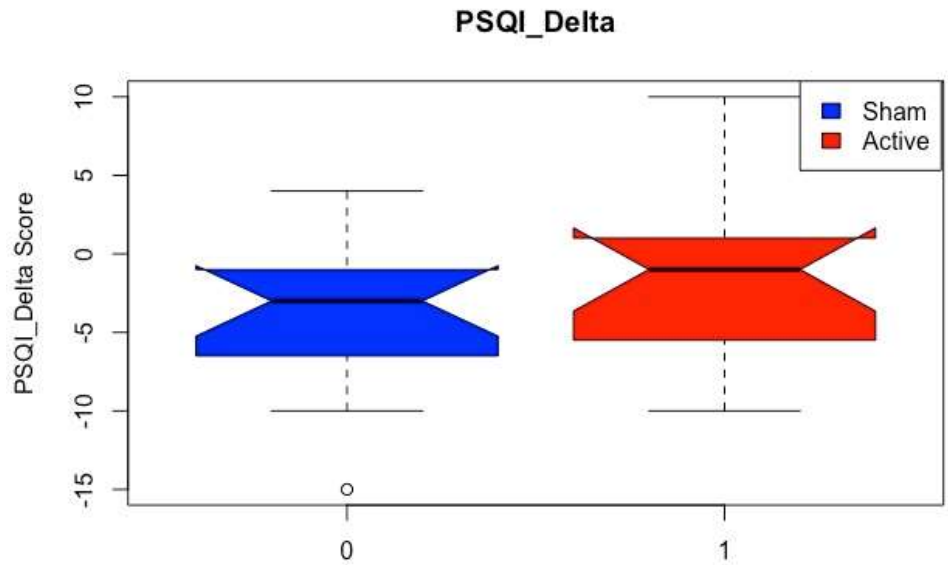
rf

I made some box plots of the target variables with notches. If the notches are overlapping that means that the difference is not significant.

MADRA_Delta

HAMAT_Delta

PSQI_Delta

LOS

I also ran some linear models on the target variables to see if there was a significant proportion of the variation explained by the treatment. Here are the various models. A summary of the results is that none of the target variables explained a significant proportion of the variance in the treatment variable. In addition the treatment variable did not explain a significant proportion of the variance in any of the target variables.

Active vs Sham

```
Call:
lm(formula = dat$Active_vs._Sham ~ dat$MADRA_Delta + dat$PSQI_Delta +
    dat$LOS + dat$HAMAT_Delta)

Residuals:
    Min      1Q  Median      3Q     Max
-0.7116 -0.4783  0.1233  0.4296  0.6952

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      1.680905   0.272114   6.177 1.85e-06 ***
dat$MADRA_Delta -0.008970   0.014086  -0.637    0.530
dat$PSQI_Delta   0.024970   0.021857   1.142    0.264
dat$LOS         -0.014567   0.011225  -1.298    0.206
dat$HAMAT_Delta  0.006887   0.019592   0.352    0.728
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5104 on 25 degrees of freedom
Multiple R-squared:  0.1317,    Adjusted R-squared:  -0.007216
F-statistic: 0.9481 on 4 and 25 DF,  p-value: 0.4529
```

MADRA_Delta

```
Call:
lm(formula = dat$MADRA_Delta ~ dat$HAMAT_Delta + dat$PSQI_Delta +
    dat$Active_vs._Sham + dat$LOS)

Residuals:
    Min      1Q  Median      3Q     Max
-19.1440  -4.4362   0.3207   4.8098  13.4379

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)             -4.75714    4.17875  -1.138 0.265741
dat$HAMAT_Delta          0.84571    0.21889   3.864 0.000703 ***
dat$PSQI_Delta           0.65285    0.28752   2.271 0.032044 *
dat$Active_vs._Sham1    -1.77929    2.79430  -0.637 0.530071
dat$LOS                 -0.07574    0.16263  -0.466 0.645431
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.188 on 25 degrees of freedom
Multiple R-squared:  0.5498,    Adjusted R-squared:  0.4778
F-statistic: 7.633 on 4 and 25 DF,  p-value: 0.0003661
```

HAMAT_Delta

```
Call:
lm(formula = dat$HAMAT_Delta ~ dat$MADRA_Delta + dat$PSQI_Delta +
    dat$Active_vs._Sham + dat$LOS)

Residuals:
     Min       1Q   Median       3Q      Max
-10.2459  -2.5605   0.1305   1.8374  11.9851

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)             -5.84565    2.86952  -2.037 0.052353 .
dat$MADRA_Delta          0.44207    0.11442   3.864 0.000703 ***
dat$PSQI_Delta          -0.05453    0.22805  -0.239 0.812978
dat$Active_vs._Sham1     0.71411    2.03158   0.352 0.728154
dat$LOS                  0.26534    0.10549   2.515 0.018688 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.197 on 25 degrees of freedom
Multiple R-squared:  0.5646,    Adjusted R-squared:  0.4949
F-statistic: 8.104 on 4 and 25 DF,  p-value: 0.000247
```

PSQI_Delta

```
Call:
lm(formula = dat$PSQI_Delta ~ dat$MADRA_Delta + dat$HAMAT_Delta +
    dat$Active_vs._Sham + dat$LOS)

Residuals:
    Min      1Q   Median      3Q     Max
-11.3494  -2.1282   0.1626   1.9709   9.8305

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)             0.65390    2.71122   0.241    0.811
dat$MADRA_Delta         0.26188    0.11534   2.271    0.032 *
dat$HAMAT_Delta        -0.04184    0.17500  -0.239    0.813
dat$Active_vs._Sham1    1.98695    1.73926   1.142    0.264
dat$LOS                -0.04182    0.10311  -0.406    0.688
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.553 on 25 degrees of freedom
Multiple R-squared:  0.2684,    Adjusted R-squared:  0.1513
F-statistic: 2.293 on 4 and 25 DF,  p-value: 0.08759
```

LOS
```
Call:
lm(formula = dat$LOS ~ dat$MADRA_Delta + dat$PSQI_Delta + dat$Active_vs._Sham +
    dat$HAMAT_Delta)

Residuals:
    Min      1Q   Median      3Q     Max
-12.844   -4.353   -2.341    2.845   23.249

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)            20.0404     3.3876   5.916 3.57e-06 ***
dat$MADRA_Delta        -0.1136     0.2438  -0.466   0.6454
dat$PSQI_Delta         -0.1563     0.3854  -0.406   0.6885
dat$Active_vs._Sham1   -4.3329     3.3387  -1.298   0.2062
dat$HAMAT_Delta         0.7611     0.3026   2.515   0.0187 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.802 on 25 degrees of freedom
Multiple R-squared:  0.3168,    Adjusted R-squared:  0.2075
F-statistic: 2.898 on 4 and 25 DF,  p-value: 0.04239
```