

2020-02-18 14:48:27

Time to summarize all results so far. We start with a sample of 393 subjects that have PRS. Some details about the sample:

```
> data0 =
readRDS('~/.data/baseline_prediction/prs_start/complete_massagedResids_0205
2020.rds')
> table(data0$sex)

Female    Male
   130    263
> mean(data0$base_age)
[1] 8.238728
> sd(data0$base_age)
[1] 2.64162
> mean(data0$last_age)
[1] 13.08527
> sd(data0$last_age)
[1] 2.979007
> table(data0$thresh0.00_inatt_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      159      115       73         46
> table(data0$thresh0.50_hi_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      159      76      112         46
```

There are 6 data domains:

- neuropsych: 'FSIQ', 'VMI.beery', 'SSB.wisc', 'SSF.wisc', 'DSF.wisc', 'DSB.wisc', 'DS.wj', 'VM.wj'
- demographic: 'base_age', 'sex', 'SES'
- genomics: PRS scores for the entire cohort (i.e. not the European-only PRS)
- DTI: FA values for JHU tracts, collapsed to reduce variables
- anatomy: thickness data for collapsed Freesurfer lobar regions
- clinics: 'internalizing', 'externalizing', 'base_sx'

Analysis 1: univariate results

Within each data domain, check which variables are significantly predicted in the model that uses the linear relationship between the 4 groups as the main predictor. In other words:

```
model = lme(myvar ~ ordered + covariates, ~1|FAMID)
```

And we collected the p-value and betas for the linear model associated with the variable **ordered**. The order of the groups is always ('nv012', 'notGE6adhd', 'imp', 'nonimp'). The covariates change per data domain:

- neuropsych: base_age
- demographic: sex
- genomics: population PCs, base_age, sex
- DTI: "meanX.trans", "meanY.trans", "meanZ.trans", "meanX.rot", "meanY.rot", "meanZ.rot", "goodVolumes", age_at_scan, sex
- anatomy: "mprage_score", "ext_avg", "int_avg", age_at_scan, sex
- clinics: sex

Those were the initial covariates in each domain, but I used stepAIC to select which covariates we should keep for the best model possible.

Also, the number of observations varies per domain, as there was no imputation in this analysis. So, the final number per domain is as follows:

```
> idx=!is.na(data0[, 'FSIQ'])
> sum(idx)
[1] 390
> table(data0[idx,]$thresh0.00_inatt_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      157      114       73       46
> table(data0[idx,]$thresh0.50_hi_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      157      75      112       46
> idx=!is.na(data0[, 'VMI.beery'])
> sum(idx)
[1] 315
> table(data0[idx,]$thresh0.00_inatt_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      124      91      61      39
> table(data0[idx,]$thresh0.50_hi_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      124      60      92      39
> idx=!is.na(data0[, 'SSB.wisc'])
> sum(idx)
[1] 245
> table(data0[idx,]$thresh0.00_inatt_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      91      74      42      38
> table(data0[idx,]$thresh0.50_hi_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      91      49      67      38
> idx=!is.na(data0[, 'DS.wj'])
```

```

> sum(idx)
[1] 339
> table(data0[idx,]$thresh0.00_inatt_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      137      102       57       43
> table(data0[idx,]$thresh0.50_hi_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      137      66       93       43
> idx=!is.na(data0[, 'CC_fa'])
> sum(idx)
[1] 180
> table(data0[idx,]$thresh0.00_inatt_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      74      48       31       27
> table(data0[idx,]$thresh0.50_hi_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      74      32       47       27
> idx=!is.na(data0[, 'parietal'])
> sum(idx)
[1] 286
> table(data0[idx,]$thresh0.00_inatt_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      124      78       47       37
> table(data0[idx,]$thresh0.50_hi_GE6_wp05)

      nv012      imp      nonimp notGE6adhd
      124      52       73       37

```

Finally, the data used for anat and DTI are already the result of excluding outliers using the same methods we used in the heritability of change paper, after removing any scans outside the 95th percentile. Specifically, all longitudinal scans for the 393 subjects with PRS were taken into consideration. After removing the outliers, I picked the first scan and only kept it if it was acquired within 1 year of the baseline clinical assessment.

Results

So, the complete list of variables tested was:

```

VMI.beery
VM.wj
FSIQ
externalizing
IFO_fa
DS.wj
ADHD_PRS0.001000 (only the best PRS was taken into FDR)
OFC
ATR_fa

```

```

CST_fa
base_age
cingulate
DSF.wisc
UNC_fa
SSB.wisc
temporal
CC_fa
SSF.wisc
parietal
ILF_fa
CIN_fa
frontal
insula
sensorimotor
DSB.wisc
internalizing
occipital
SES
SLF_fa

```

Our chosen approach was to do an FDR across all domains. Then it's a matter of choosing what q level to go with:

```

> hi_p[p2<.05,'var']
[1] VMI.beery      VM.wj      FSIQ      externalizing
[5] IF0_fa        DS.wj      ADHD_PRS0.001000  OFC
[9] ATR_fa
> hi_p[p2<.1,'var']
[1] VMI.beery      VM.wj      FSIQ      externalizing
[5] IF0_fa        DS.wj      ADHD_PRS0.001000  OFC
[9] ATR_fa        CST_fa     base_age    cingulate
[13] DSF.wisc
> inatt_p[p2<.05,'var']
[1] FSIQ      VMI.beery      VM.wj      base_age
[5] externalizing  ADHD_PRS0.000500
> inatt_p[p2<.1,'var']
[1] FSIQ      VMI.beery      VM.wj      base_age
[5] externalizing  ADHD_PRS0.000500 DSF.wisc    IF0_fa
[9] DS.wj

```

Or we could choose to not include base_age, because it's not really a predictor and more of a covariate anyways (like sex):

```

> hi_p[p2<.05,'var']
[1] VMI.beery      VM.wj      FSIQ      externalizing
[5] IF0_fa        DS.wj      ADHD_PRS0.001000  OFC
[9] ATR_fa
> hi_p[p2<.1,'var']

```

```

[1] VMI.beery      VM.wj      FSIQ      externalizing
[5] IF0_fa        DS.wj      ADHD_PRS0.001000 OFC
[9] ATR_fa        CST_fa     cingulate  DSF.wisc
> inatt_p[p2<.05, 'var']
[1] FSIQ      VMI.beery  VM.wj      externalizing
[5] ADHD_PRS0.000500
> inatt_p[p2<.1, 'var']
[1] FSIQ      VMI.beery  VM.wj      externalizing
[5] ADHD_PRS0.000500 DSF.wisc   IF0_fa     DS.wj

```

Analysis 2: big model

This analysis had the goal of checking how well we can model the groups by combining all the "good" univariate variables from analysis 1.

As we cannot deal with NAs here, we decided to impute the data using as the base the 180 kids who have both PRS and DTI.

The model is a multinomial logistic regression that ignores the family term. It is also not ordered, because it performed better (i.e. higher ROC AUC) than the ordered model:

```
group ~ good_vars + covars
```

The good_vars came from analysis 1:

```

hi_vars = c('VMI.beery', 'VM.wj', 'FSIQ', 'externalizing', 'IF0_fa',
'DS.wj',
            'ADHD_PRS0.001000', 'OFC', 'ATR_fa', 'CST_fa', 'cingulate',
            'DSF.wisc')
inatt_vars = c('FSIQ', 'VMI.beery', 'VM.wj', 'externalizing',
               'ADHD_PRS0.000500', 'DSF.wisc', 'IF0_fa', 'DS.wj')
covars = c('base_age', 'sex')

```

Also, note that every domain has already been residualized within domain (for example, PRS was residualized based on PCs, etc). We can evaluate the models based on ROC AUC, and check how important each variable was in the prediction (just the sum of the absolute value of the coefficients across the different categories).

The results are quite disproportional, as one would expect since externalizing is very unbalanced:

```

[1] "inatt"

Overall
FSIQ      0.7862872
VMI.beery 1.6524547
VM.wj     1.1801762
externalizing1 49.3133069
ADHD_PRS0.000500 0.8268296

```

```
DSF.wisc      0.4345088
IFO_fa        0.9686923
DS.wj         1.3378331
base_age      1.7978418
sexMale       1.0346242
Multi-class area under the curve: 0.7411
```

```
[1] "hi"
```

```
Overall
VMI.beery     1.7175702
VM.wj         0.7886613
FSIQ          0.9107406
externalizing1 47.7411726
IFO_fa        0.5858586
DS.wj         1.5779318
ADHD_PRS0.001000 0.6740715
OFC           0.9341405
ATR_fa        0.4345321
CST_fa        0.7608150
cingulate     0.2842507
DSF.wisc      0.5487167
base_age      1.2516826
sexMale       1.2585796
Multi-class area under the curve: 0.7578
```

How does it look when we add in base_sx as one of the predictors?

```
[1] "inatt"
```

```
Overall
FSIQ          0.6806480
VMI.beery     1.3805707
VM.wj         1.2073122
externalizing1 47.3177951
ADHD_PRS0.000500 1.2894672
DSF.wisc      2.7361367
IFO_fa        1.5614350
DS.wj         0.6733051
base_inatt    27.3284253
base_age      0.9617274
sexMale       2.6653584
Multi-class area under the curve: 0.9138
```

```
[1] "hi"
```

```
Overall
VMI.beery     1.8009884
VM.wj         3.2346574
FSIQ          1.3900451
externalizing1 23.7156506
IFO_fa        1.6400383
DS.wj         0.4748365
ADHD_PRS0.001000 2.0801047
OFC           1.5064806
```

```

ATR_fa      0.5491814
CST_fa      2.7601051
cingulate   1.1449211
DSF.wisc    0.5819857
base_hi     24.2712899
base_age    3.8420027
sexMale     2.6573462
Multi-class area under the curve: 0.9169

```

Big jump in AUC, which makes sense as base_sx is used in defining the groups. But interesting that externalizing is almost as good, maybe even better...

We can also check the increase in AUC as we add the different domains in a step-wise fashion:

```

[1] "inatt"
      demo      clin      gen      dti neuropsych
0.6134810 0.8804639 0.8919944 0.8922737 0.9137963
[1] "hi"
      demo      clin      gen      dti      anat neuropsych
0.6065227 0.8345617 0.8455017 0.8616758 0.8765961 0.9169098

```

Out of curiosity, what are our numbers if we use the exact same framework, but trying to distinguish just between the two clinical groups? Results are also not that bad...

```

[1] "inatt"
      demo      clin      gen      dti neuropsych
0.6871640 0.8185484 0.8319892 0.8319892 0.8830645
[1] "hi"
      demo      clin      gen      dti      anat neuropsych
0.6702128 0.8071809 0.8091755 0.8138298 0.8244681 0.8590426

```

Now, of course the externalizing distribution is very sparse:

```

> table(data$externalizing, data[, phen])

      nv012 imp nonimp notGE6adhd
0         74  27     42          25
1          0   5      5           2

```

So, what happens to the weights if I either don't run that variable in the 4 group comparison? Or don't include nv012, and only run 3 groups?

```

[1] "inatt"
      Overall
FSIQ      0.6713690

```

```

VMI.beery      1.8008210
VM.wj          1.1684526
ADHD_PRS0.000500 0.7708455
DSF.wisc       0.4803280
IFO_fa        1.1708453
DS.wj         1.0548323
base_age      1.6850936
sexMale       0.9483320
Multi-class area under the curve: 0.7229

```

```
[1] "hi"
```

```

Overall
VMI.beery      1.8268554
VM.wj          0.8664441
FSIQ           0.9230327
IFO_fa        0.7473989
DS.wj         1.2511845
ADHD_PRS0.001000 0.6451152
OFC           0.8566875
ATR_fa        0.5427812
CST_fa        0.7208566
cingulate     0.4296081
DSF.wisc       0.6122018
base_age      1.1866683
sexMale       1.1602818
Multi-class area under the curve: 0.747

```

Overall, not as bad, and we didn't lose much in terms of AUC. Let's see how it looks after we add in base_sx:

```
[1] "inatt"
```

```

Overall
FSIQ           0.5190794
VMI.beery      1.2832258
VM.wj          1.1350190
ADHD_PRS0.000500 1.1784921
DSF.wisc       3.0009173
IFO_fa        1.6415447
DS.wj         0.7241186
base_inatt     28.0924540
base_age       0.8509839
sexMale       2.2257656
Multi-class area under the curve: 0.9115

```

```
[1] "hi"
```

```

Overall
VMI.beery      1.8090450
VM.wj          3.2007661
FSIQ           1.4060350
IFO_fa        1.6473047
DS.wj         0.5276963
ADHD_PRS0.001000 2.0131734
OFC           1.5049995

```



```

ATR_fa          0.5186103
CST_fa          2.7554278
cingulate       1.1505311
DSF.wisc        0.5744980
base_hi         24.1625875
base_age        3.8087047
sexMale         2.6791743
Multi-class area under the curve: 0.9163

```

We still see the big jump from before, again expected as base_sx is used to define the groups. What if we run 3 groups only?

```

[1] "inatt"
Overall
FSIQ          1.3936495
VMI.beery     0.4464261
VM.wj         0.7960408
externalizing1 2.4373028
ADHD_PRS0.000500 1.3741709
DSF.wisc      0.5311392
IFO_fa        0.4959420
DS.wj         0.4958580
base_age      1.0964769
sexMale       1.1486840
Multi-class area under the curve: 0.7569

```

```

[1] "hi"
Overall
VMI.beery     0.4864423
VM.wj         0.5471967
FSIQ          1.6412003
externalizing1 2.5931840
IFO_fa        0.6310506
DS.wj         0.4103319
ADHD_PRS0.001000 1.1681734
OFC           1.0591294
ATR_fa        0.2912992
CST_fa        0.9043042
cingulate     0.2246797
DSF.wisc      0.5846238
base_age      1.0156251
sexMale       1.6160933
Multi-class area under the curve: 0.7823

```

It still carries weight, but at least it's not over the top. AUC is still decent... And if we include base_sx?

```

[1] "inatt"
Overall
FSIQ          0.7149091

```

```

VMI.beery      0.2881959
VM.wj          1.1430780
externalizing1 2.7247453
ADHD_PRS0.000500 1.1872987
DSF.wisc       0.1431103
IFO_fa         0.8085379
DS.wj          0.5387884
base_inatt     5.3156872
base_age       0.7218521
sexMale        2.4801733
Multi-class area under the curve: 0.8596

```

```

[1] "hi"

Overall
VMI.beery      0.7125827
VM.wj          0.6703544
FSIQ           1.7043603
externalizing1 1.1160982
IFO_fa         1.0440388
DS.wj          0.3050194
ADHD_PRS0.001000 1.9647090
OFC            1.2091195
ATR_fa         0.5853600
CST_fa         1.3282894
cingulate      0.2831388
DSF.wisc       1.2255681
base_hi        4.9314779
base_age       3.7619075
sexMale        1.8857750
Multi-class area under the curve: 0.8786

```

Yep, make sense. Lastly, what happens when we add in medication_status_at_observation? First, the distribution is very skewed, like externalizing:

```
> table(data$medication_status_at_observation,
data$thresh0.50_hi_GE6_wp05)
```

	nv012	imp	nonimp	notGE6adhd
none	159	52	78	37
nonstim	0	3	2	1
stim	0	21	32	8

So, of course the linear fit for this variable by itself (similar to analysis #1) is very significant. Adding it to the big model as is will just give results very similar to when I was adding externalizing. So, let's just run the model with 3 groups instead. First, without base_sx:

```

[1] "inatt"

Overall
FSIQ           1.5794703

```

```

VMI.beery 0.4064579
VM.wj 0.7626587
externalizing1 2.2779352
ADHD_PRS0.000500 1.4058793
DSF.wisc 0.8553022
IFO_fa 0.5188099
DS.wj 0.4738631
base_age 1.0790133
sexMale 0.9700791
medication_status_at_observationnonstim 2.5387123
medication_status_at_observationstim 2.5825405
Multi-class area under the curve: 0.766

```

```
[1] "hi"
```

```

Overall
VMI.beery 0.3949537
VM.wj 0.6577070
FSIQ 1.7894836
externalizing1 2.1609333
IFO_fa 0.6791764
DS.wj 0.4973218
ADHD_PRS0.001000 1.2240894
OFC 1.0448133
ATR_fa 0.2579573
CST_fa 0.9364096
cingulate 0.2619426
DSF.wisc 0.8888433
base_age 0.9071128
sexMale 1.9337588
medication_status_at_observationnonstim 17.1836361
medication_status_at_observationstim 1.9955208
Multi-class area under the curve: 0.8002

```

The medication variable dominates the predictors in the HI case. It's not a considerable improvement in AUC though. What if we add base_sx?

```
[1] "inatt"
```

```

Overall
FSIQ 0.7414561
VMI.beery 0.3028156
VM.wj 1.0315119
externalizing1 2.7496942
ADHD_PRS0.000500 1.1781947
DSF.wisc 0.2350047
IFO_fa 0.7335736
DS.wj 0.4558638
base_inatt 5.2208425
base_age 0.7560045
sexMale 2.3631864
medication_status_at_observationnonstim 1.6377845
medication_status_at_observationstim 0.8823375

```

```
Multi-class area under the curve: 0.8641

[1] "hi"

Overall
VMI.beery      0.8857438
VM.wj          0.4973983
FSIQ           1.8330103
externalizing1 0.7141466
IFO_fa         0.8766953
DS.wj          0.2795922
ADHD_PRS0.001000 1.9459321
OFC            1.1504444
ATR_fa         0.6752563
CST_fa         1.5048279
cingulate      0.1098872
DSF.wisc       1.1715570
base_hi        5.0591415
base_age       3.6486695
sexMale        2.3601317
medication_status_at_observationnonstim 15.6503299
medication_status_at_observationstim    2.6920838
Multi-class area under the curve: 0.9
```

We see the usual bump in AUC, but medication is still extremely important.

Analysis 3: ML

The idea here is to take all variables that were analyzed in the univariate analysis (#1), but instead of taking them individually we take them all together (after any within-domain residualizing procedures).

There is no imputation because the classifiers are trained within domain. We separate for testing everyone but one participant in the same family. Some of the testing cases will have data only for some of the domains, similarly to what we will have in the training data. Note that the testing data is never used during training, but it's not a clean cross-validation: the test data is not independent from the training data because of the family component, and also because of the residualizing procedure that uses the entire dataset for robustness.

So, we train the best classifier we can within each domain. We also train an ensemble classifier that learns to combine the "vote" for each domain. In other words, each domain votes (with a probability) what group a given participant belongs to, and the ensemble classifier learns how to best consider each vote (i.e. trust/take into consideration some domains more than others). When there is no data for a given domain it either votes NA, or just the class probability deducted from the training data. I tried it both ways, the difference being that if voting NA we need to use an ensembler that takes that in (i.e. any GLM/weighted majority voting won't work).

The training itself is a 10-fold repeated cross validation (10 times), which happens only within the training set. For this analysis, we can not only assess variable importance within domain, but also how important each domain was in the ensemble classifier.

I only ran this for the 3 and 2 class scenarios, as I didn't think it'd be fair to run externalizing and medication variables in the 4-class case. The results using 3-classes is just an average of the 3 different ROC curves. Still, not very impressive: .62 for inatt and .5 for hi.

For the 2-class case, we get up to .75 for inatt and .72 for hi.

These are the training/testing splits in each domain (neuropsych was further divided to avoid additional imputation):

```
"Training iq_vmi on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 103 participants"
[1] "Training wisc on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 72 participants"
[1] "Training wj on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 106 participants"
[1] "Training demo on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 133 participants"
[1] "Training clin on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 133 participants"
[1] "Training gen on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 133 participants"
[1] "Training dti on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 56 participants"
[1] "Training anat on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 84 participants"
[1] "iq_vmi"
[1] "Testing on 49 participants"
[1] "wisc"
[1] "Testing on 43 participants"
[1] "wj"
[1] "Testing on 50 participants"
[1] "demo"
[1] "Testing on 55 participants"
[1] "clin"
[1] "Testing on 55 participants"
[1] "gen"
[1] "Testing on 55 participants"
[1] "dti"
[1] "Testing on 23 participants"
[1] "anat"
[1] "Testing on 41 participants"
```

```
[1] "Training iq_vmi on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 126 participants"
[1] "Training wisc on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 93 participants"
[1] "Training wj on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 130 participants"
[1] "Training demo on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 160 participants"
[1] "Training gen on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 160 participants"
[1] "Training dti on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 71 participants"
```

```

[1] "Training anat on thresh0.00_inatt_GE6_wp05 (sx=inatt, model=lda)"
[1] "Training on 105 participants"
[1] "iq_vmi"
[1] "Testing on 65 participants"
[1] "wisc"
[1] "Testing on 60 participants"
[1] "wj"
[1] "Testing on 69 participants"
[1] "demo"
[1] "Testing on 74 participants"
[1] "gen"
[1] "Testing on 74 participants"
[1] "dti"
[1] "Testing on 35 participants"
[1] "anat"
[1] "Testing on 57 participants"

```

If we spit out the importance of the variables in each domain, and the ensemble, we get (inattention first, then hi):

```

      Importance
FSIQ      100
      Importance
SSB.wisc  100.00
DSB.wisc   36.89
DSF.wisc   13.33
SSF.wisc    0.00
      Importance
VM.wj      100
DS.wj       0
      Importance
base_age   100.00
SES        39.41
sex         0.00
      Importance
base_inatt 100.000
externalizing 1.290
medication_status_at_observation 1.207
internalizing 0.000
      Importance
ADHD_PRS0.000100 1.000e+02
ADHD_PRS0.000050 8.856e+01
ADHD_PRS0.000500 7.214e+01
ADHD_PRS0.001000 3.651e+01
ADHD_PRS0.005000 3.226e+01
ADHD_PRS0.005000 1.393e+01
ADHD_PRS0.010000 1.364e+01
ADHD_PRS0.020000 9.238e+00
ADHD_PRS0.010000 8.504e+00
ADHD_PRS0.030000 2.053e+00
ADHD_PRS0.040000 7.107e-14
ADHD_PRS0.050000 0.000e+00
      Importance

```

CIN_fa	100.00
UNC_fa	63.89
ATR_fa	59.72
CC_fa	56.25
IFO_fa	38.89
ILF_fa	15.97
CST_fa	10.42
SLF_fa	0.00

Importance

insula	100.000
parietal	92.473
occipital	74.194
cingulate	69.892
OFC	35.484
sensorimotor	20.430
temporal	6.452
frontal	0.000

Overall

gen	100.00
clin	100.00
anat	33.83
iq_vmi	30.08
wisc	25.56
demo	0.00
wj	0.00
dti	0.00

Importance

VMI.beery	100
FSIQ	0

Importance

SSF.wisc	100.00
DSB.wisc	82.55
DSF.wisc	26.17
SSB.wisc	0.00

Importance

VM.wj	100
DS.wj	0

Importance

base_age	100.00
sex	63.91
SES	0.00

Importance

base_hi	100.000
medication_status_at_observation	4.985
internalizing	2.085
externalizing	0.000

Importance

ADHD_PRS0.000100	100.000
ADHD_PRS0.000500	78.663
ADHD_PRS0.010000	75.835

ADHD_PRS0.001000	73.265
ADHD_PRS0.005000	68.638
ADHD_PRS0.000050	66.067
ADHD_PRS0.300000	12.596
ADHD_PRS0.400000	5.398
ADHD_PRS0.500000	4.113
ADHD_PRS0.100000	3.085
ADHD_PRS0.200000	2.314
ADHD_PRS0.050000	0.000

Importance

CST_fa	100.000
ATR_fa	93.333
ILF_fa	54.444
UNC_fa	41.111
SLF_fa	22.222
CC_fa	17.778
CIN_fa	8.889
IFO_fa	0.000

Importance

occipital	100.00
OFC	94.17
sensorimotor	71.67
cingulate	62.50
parietal	37.50
insula	23.33
frontal	23.33
temporal	0.00

Overall

clin	100.0
dti	78.2
demo	0.0
anat	0.0
wj	0.0
gen	0.0
wisc	0.0
iq_vmi	0.0