

Appendices

Data Processing

The subjects were scanned at seven different acquisition sites as shown in Table 4. A few subjects did not pass the quality control thresholds, so their imaging data were not included in the study. The additional information about acquisition parameters can be found at http://fcon_1000.projects.nitrc.org/indi/adhd200/. A standard preprocessing pipeline was applied to the resting-state functional magnetic resonance imaging (Rs-fMRI) data using the Data Processing Assistant for Resting-State fMRI Toolbox (DPARSF) as described in Yan and Zang (2010). The pipeline included various steps such as removing the first five volumes, correcting for slice timing, realigning volumes to account for head motion, co-registering the anatomical image to the functional image, and performing nuisance variable regression. Nuisance variable regression involved linear detrending, accounting for mean global signal, white matter and cerebrospinal fluid signals, as well as six motion parameters. Following nuisance variable regression, the data were normalized to the MNI template. The blood-oxygen-level-dependent (BOLD) time series from each voxel in the brain were deconvolved using blind deconvolution to estimate voxel-specific hemodynamic response functions and obtain latent neural signals (Wu et al., 2013). A bandpass filter with a bandwidth of 0.01-0.1 Hz was then applied to temporally filter the data. Mean time series were extracted from 190 functionally homogeneous brain regions based on the CC200 template (Craddock et al., 2012). Finally, functional connectivity between these 190 regions was calculated using Pearson’s correlation. Pearson correlation represents the covariance of two time series normalized for their individual autovariance. Importantly, PCC remains invariant under separate changes in location and scale for the variables, showing that it is not influenced by amplitude fluctuations. It does not distinguish between active or inactive states of regions A and B but rather captures synchronization of activity, distinct from activation. As the majority of the correlations were noise, a t-test was used with a liberal threshold ($p < 0.05$ uncorrected) to select the 1179 features. Therefore, these correlations delivered the features used as predictors of ADHD.

Table 4: The site distribution for ADHD data across the seven imaging sites used in our study.

Imaging site	Control	ADHD	Total
Peking University	143	102	245
Kennedy Krieger Institute	69	25	94
NeuroIMAGE Sample	37	36	73
New York University Child Study Center	110	147	257
Oregon Health & Science University	70	43	113
University of Pittsburgh	94	4	98
Washington University	50	0	50
Total Subjects	573	357	930

LOGIT Results

Table 5: Proportion (%) of SWAG models that contain each feature (second column); proportion (%) of times the coefficient of the feature is positive (third column); median value of coefficients for the feature (fourth column).

Feature	% Models	% Positive β	Median β
'Insula-R' \leftrightarrow 'Paracentral-Lobule-R'	1	0	-0.533
'Frontal-Sup-R' \leftrightarrow 'Frontal-Mid-R'	1	1	0.675
'Temporal-Mid-L' \leftrightarrow 'Lingual-L'	1	1	0.508
'Insula-L' \leftrightarrow 'Frontal-Mid-L'	1	0	-0.977
'Frontal-Sup-R' \leftrightarrow 'Frontal-Inf-Orb-R'	0.989	0.003	-0.313
'Supp-Motor-Area-L' \leftrightarrow 'SupraMarginal-L'	0.989	1	0.357
'Frontal-Sup-R' \leftrightarrow 'ParahippocampaGyrus'	0.953	0	-2.081
'Vermis-1-2' \leftrightarrow 'Temporal-Mid-L-4'	0.953	1	0.781
'Vermis-4-5' \leftrightarrow 'Frontal-Sup-Orb-R'	0.853	0	-1.099
'Frontal-Sup-Medial-L' \leftrightarrow 'Cingulum-Mid-L'	0.745	0	-1.007
'Temporal-Mid-L' \leftrightarrow 'Pallidum-R'	0.738	0.979	0.222
'Occipital-Mid-L' \leftrightarrow 'Cerebelum-Crus2-L'	0.70	1	2.261
'Cingulum-Ant-L' \leftrightarrow 'Paracentral-Lobule-R'	0.677	0	-1.104
'Temporal-Pole-Sup-R' \leftrightarrow 'Frontal-Mid-L'	0.617	0	-1.072
'Frontal-Mid-L' \leftrightarrow 'Caudate-R'	0.543	1	1.162
'Frontal-Mid-L' \leftrightarrow 'Frontal-Sup-L'	0.528	1	0.939

Feature	% Models	% Positive β	Median β
'Temporal-Mid-L' ↔ 'Temporal-Sup-L'	0.451	1	0.671
'Frontal-Sup-R' ↔ 'Precuneus-R'	0.391	0	-1.042
'Insula-L' ↔ 'Postcentral-R'	0.22	0.56	0.005
'Frontal-Mid-L' ↔ 'Cingulum-Ant-L'	0.192	1	0.778
'Frontal-Sup-Medial-R' ↔ 'Cingulum-Mid-L'	0.173	0.303	-0.077
'Cingulum-Post-L' ↔ 'Postcentral-R'	0.165	0.651	0.073
'Supp-Motor-Area-L' ↔ 'SupraMarginal-L'	0.144	1	0.143
'Frontal-Sup-L' ↔ 'ParahippocampaGyrus'	0.134	0	-0.3
'Caudate-R' ↔ 'Paracentral-Lobule-R'	0.134	0	-0.551
'Frontal-Sup-Medial-R' ↔ 'Frontal-Inf-Tri-L'	0.121	0.174	-0.031
'Insula-L' ↔ 'Paracentral-Lobule-R'	0.118	0.956	0.36
'Cingulum-Post-L' ↔ 'Paracentral-Lobule-R'	0.113	0.326	-0.046
'Parietal-Sup-R' ↔ 'Cuneus-L'	0.097	1	0.275
'Cingulum-Ant-R' ↔ 'Frontal-Sup-R'	0.097	0	-0.124
'Parietal-Inf-L' ↔ 'Caudate-Head-L'	0.089	0	-0.562
'Frontal-Sup-Medial-L' ↔ 'Frontal-Inf-Tri-L'	0.073	0.964	0.094
'Insula-R' ↔ 'Pallidum-R'	0.066	1	0.865
'Temporal-Mid-L' ↔ 'Putamen-L'	0.06	0.217	-0.228
'Insula-L' ↔ 'Cingulum-Ant-L'	0.06	0.304	-0.037
'Caudate-Head-L' ↔ 'Cingulum-Ant-L'	0.055	1	0.336
'Frontal-Med-Or-R' ↔ 'Frontal-Inf-Tri-L'	0.052	0.3	-0.119
'Vermis-4-5' ↔ 'Frontal-Sup-Medial-L'	0.047	0	-0.693
'Frontal-Sup-R' ↔ 'Cingulum-Ant-L'	0.034	0.308	-0.087
'Cingulum-Mid-R' ↔ 'Frontal-Sup-L'	0.031	0	-0.243
'Frontal-Med-Orb-R' ↔ 'Frontal-Mid-R'	0.024	0.778	0.071
'Rolandic-Oper-L' ↔ 'Temporal-Inf-L'	0.024	1	0.994
'Lingual-L' ↔ 'Frontal-Sup-L'	0.021	0	-1.862
'Cingulum-Ant-R' ↔ 'Frontal-Mid-L'	0.018	1	0.665
'Cingulum-Ant-R' ↔ 'Cingulum-Mid-L'	0.018	0.571	0.014
'Cingulum-Ant-L' ↔ 'Precuneu-R'	0.013	0	-1.212
'Pallidum-R' ↔ 'Temporal-Mid-R'	0.008	0	-0.034
'Caudate-Head-L' ↔ 'Precuneus-R'	0.008	0	-0.582
'Cingulum-Ant-L' ↔ 'Caudate-R'	0.003	1	0.022
'Temporal-Inf-L' ↔ 'Lingual-L'	0.003	1	0.915
'Temporal-Sup-L' ↔ 'Temporal-Mid-L'	0.003	0	-0.169
'Vermis-6' ↔ 'Temporal-Inf-L'	0.003	0	-0.525
'Cingulum-Ant-R' ↔ 'Paracentral-Lobule-R'	0.003	0	-0.097

Sensitivity and Specificity

Table 6: Sensitivity for the three classifiers with and without SWAG. The second column shows the sensitivities of the original models without SWAG. The range of sensitivities of the SWAG models is shown in the last column.

Classifiers	Without SWAG	Classifiers	With SWAG
Lasso Logistic	0.16	Logistic	(0.16,0.29)
Linear SVM	0.45	SVM-L	(0,0.23)
Radial SVM	0.34	SVM-R	(0.22,0.34)

Table 7: Specificity for the three classifiers with and without SWAG. The second column shows the specificities of the original models without SWAG. The range of specificities of the SWAG models is shown in the last column.

Classifiers	Without SWAG	Classifiers	With SWAG
Lasso Logistic	0.87	Logistic	(0.81,0.89)
Linear SVM	0.66	SVM-L	(0.85,1)
Radial SVM	0.86	SVM-R	(0.82,0.89)

SVM Results

SWAG produced a total of 114 models of dimensions 18, 19, and 20 using SVM-Linear learning mechanism. Table 8 reports the most frequently appeared brain regions that are associated with the selected features. Table 10 shows the raw names of the most frequent features selected by SWAG using SVM-L.

Table 8: Frequency table for SVM-L models. Features are aggregated into brain regions (left column) with corresponding frequency (right column).

Brain Region	Frequency
Frontal	527
Temporal	334
Thalamus	299
Vermis	228
Fusiform	228
Parietal	220
Insula	114

SWAG produced a total of 146 models of dimensions 17, 18, 19 and 20 using the radial kernel SVM learning mechanism. Table 9 reports the brain regions most frequently represented in the selected features. Table 11 shows the raw names of the most frequent features selected by SWAG using SVM-R.

Table 9: Frequency table for SVM-R models. Features are aggregated into brain regions (left column) with corresponding frequency (right column).

Brain Region	Frequency
Frontal	1271
Temporal	557
Cingulum	434
Fusiform	419
Parietal	283
Rolandic	146
Insula	146

Table 10: Frequency table with the raw feature names for SVM-L models

Feature	Frequency
'Insula-R'↔'Paracentral-Lobule-R'	114
'Frontal-Sup-L'↔'Lingual-L'	114
'Thalamus-L'↔'Frontal-Sup-Medial-L'	114
'Olfactory'↔'Frontal-Mid-R'	114
'Vermis-4-5'↔'Frontal-Sup-Orb-R'	114
'Temporal-Sup-L'↔'Thalamus-L'	114
'Parietal-Inf-R'↔'Precentral-R'	114
'Vermis-1-2'↔'Temporal-Mid-L'	114
'Angular-R'↔'Fusiform-L'	114
'Fusiform-R'↔'Cingulum-Ant-R'	114
'Cerebelum-6-L'↔'Cerebelum-Crus2-L'	114
'Parietal-Inf-L'↔'Temporal-Pole-Mid-L'	106
'Thalamus-L'↔'Frontal-Sup-L'	71
'Vermis-1-2'↔'Pallidum-R'	67
'Cuneus-R'↔'Frontal-Inf-Orb-L'	66
'Frontal-Sup-Medial-L'↔'Precentral-L'	66
'Frontal-Mid-R'↔'Cingulum-Ant-L'	66
'Parietal-Inf-L'↔'Fusiform-R'	56
'Frontal-Sup-Orb-L'↔'Cuneus-L'	40
'Angular-R'↔'Cerebelum-Crus2-L'	39
'Rolandic-Oper-L'↔'Frontal-Inf-Orb-L'	28
'Midbrain'↔'Temporal-Sup-L'	27
'Temporal-Pole-Sup-R'↔'Frontal-Inf-Orb-R'	24
'Cingulum-Ant-R'↔'Temporal-Mid-R'	22
'SupraMarginal-L'↔'Cingulum-Mid-L'	16
'Parietal-Inf-L'↔'Cingulum-Ant-L'	15
'Frontal-Sup-R'↔'Putamen-L'	12
'Precentral-L'↔'Occipital-Mid-R'	12
'Temporal-Mid-L'↔'Frontal-Inf-Orb-L'	11
'Parietal-Inf-L'↔'Precuneus-L'	10
'Frontal-Inf-Oper-R'↔'SupraMarginal-R'	10

Feature	Frequency
'Occipital-Mid-R' ↔ 'Lingual-R'	9
'Occipital-Sup-R' ↔ 'Parietal-Sup-R'	9
'Thalamus-R' ↔ 'Frontal-Sup-Medial-L'	8
'Temporal-Mid-L' ↔ 'Frontal-Mid-R'	8
'Frontal-Mid-L' ↔ 'Frontal-Sup-L'	7
'Cingulum-Ant-R' ↔ 'Frontal-Inf-Orb-R'	7
'Insula-L' ↔ 'Brainstem'	6
'Calcarine-L' ↔ 'Cerebelum-8-R'	6
'Frontal-Med-Orb-R' ↔ 'Frontal-Mid-L'	6
'Rolandic-Oper-L' ↔ 'Temporal-Pole-Sup-R'	6
'Frontal-Inf-Orb-L' ↔ 'Putamen-L'	5
'Cingulum-Mid-R' ↔ 'Frontal-Mid-L'	5
'Temporal-Pole-Sup-R' ↔ 'Supp-Motor-Area-R'	5
'Cerebelum-8-R' ↔ 'Cerebelum-6-L'	5
'Rectus-L' ↔ 'Brainstem'	5
'Rectus-L' ↔ 'Cingulum-Mid-L'	5
'Rectus-L' ↔ 'Cerebelum-6-L'	5
'Cingulum-Ant-R' ↔ 'Temporal-Mid-R'	5
'Frontal-Inf-Oper-R' ↔ 'Frontal-Inf-Tri-R'	5
'Cuneus-L' ↔ 'Precuneus-L'	4
'Thalamus-L' ↔ 'Frontal-Sup-Medial-R'	4
'Frontal-Mid-L' ↔ 'Frontal-Sup-Medial-L'	4
'Occipital-Sup-R' ↔ 'Precuneus-R'	4
'Frontal-Inf-Oper-R' ↔ 'Cingulum-Mid-R'	4
'Midbrain' ↔ 'Frontal-Inf-Tri-L'	4
'Cerebelum-6-L' ↔ 'Temporal-Inf-R'	4
'Frontal-Sup-R' ↔ 'Frontal-Mid-Orb-R'	3
'Insula-L' ↔ 'Supp-Motor-Area-R'	3
'Insula-L' ↔ 'Postcentral-R'	3
'Caudate-R' ↔ 'Temporal-Pole-Mid-R'	3
'Precentral-L' ↔ 'Insula-L'	3
'Temporal-Sup-L' ↔ 'Cingulum-Mid-R'	3
'Supp-Motor-Area-L' ↔ 'Angular-L'	3
'Precuneus-L' ↔ 'Precuneus-L'	3
'RectalGyrus' ↔ 'Frontal-Inf-Orb-R'	3
'Cuneus-L' ↔ 'Cingulum-Ant-L'	2
'Occipital-Mid-R' ↔ 'Temporal-Mid-R'	2
'Frontal-Sup-Medial-R' ↔ 'Caudate-R'	2
'Angular-R' ↔ 'Frontal-Sup-Orb-R'	2
'Supp-Motor-Area-L' ↔ 'SupraMarginal-L'	2
'RectalGyrus' ↔ 'Cerebelum-8-R'	2
'Paracentral-Lobule-L' ↔ 'Postcentral-R'	2
'Paracentral-Lobule-L' ↔ 'Cingulum-Mid-L'	2
'Temporal-Mid-L' ↔ 'Cingulum-Ant-L'	2
'Parietal-Inf-L' ↔ 'Angular-L'	2
'Cingulum-Ant-R' ↔ 'Cingulum-Mid-L'	2

Feature	Frequency
'Cuneus-L'↔'Putamen-L'	1
'Insula-L'↔'Temporal-Mid-L'	1
'Frontal-Mid-L'↔'Temporal-Inf-R'	1
'Cingulum-Mid-R'↔'Precentral-L'	1
'Parietal-Sup-L'↔'Precentral-R'	1
'Cingulum-Mid-L'↔'Frontal-Inf-Orb-L'	1
'Temporal-Inf-L'↔'Frontal-Sup-L'	1
'Frontal-Sup-R'↔'Paracentral-Lobule-R'	1
'Vermis-9'↔'Frontal-Mid-R'	1
'Temporal-Mid-L'↔'Frontal-Sup-R'	1
'Postcentral-R'↔'Insula-R'	1
'Insula-L'↔'Parietal-Sup-L'	1
'Frontal-Mid-Orb-L'↔'Cingulum-Mid-L'	1
'Frontal-Sup-Medial-L'↔'Frontal-Inf-Orb-R'	1
'Temporal-Sup-L'↔'Postcentral-R'	1
'Frontal-Med-Orb-R'↔'Angular-R'	1
'Supp-Motor-Area-L'↔'Frontal-Mid-R'	1
'Temporal-Pole-Sup-R'↔'Temporal-Mid-L'	1
'Temporal-Pole-Sup-R'↔'Frontal-Sup-Orb-L'	1
'SupraMarginal-L'↔'Frontal-Mid-Orb-L'	1
'Rolandic-Oper-L'↔'Cingulum-Ant-L'	1
'Rolandic-Oper-L'↔'Frontal-Mid-L'	1
'Cerebelum-3-R'↔'Thalamus-R'	1
'Postcentral-L'↔'Postcentral-R'	1
'Precentral-L'↔'Cingulum-Ant-R'	1
'RectalGyrus'↔'Insula-L'	1
'Parietal-Inf-L'↔'Frontal-Mid-L'	1
'Occipital-Sup-L'↔'Occipital-Mid-R'	1
'Frontal-Inf-Orb-R'↔'Vermis-6'	1
'Fusiform-R'↔'Caudate-R'	1
'Parietal-Inf-L'↔'Paracentral-Lobule-R'	1
'Temporal-Pole-Sup-R'↔'Thalamus-R'	1
'Temporal-Mid-L'↔'Temporal-Inf-L'	1

Table 11: Frequency table with the raw feature names for SVM-R models

Feature	Frequency
'Frontal-Sup-L' ↔ 'ParahippocampaGyrus'	146
'Temporal-Inf-L' ↔ 'Frontal-Sup-Medial-L'	146
'Frontal-Mid-L' ↔ 'Cingulum-Ant-L'	146
'Temporal-Mid-L' ↔ 'Frontal-Inf-Orb-L'	146
'Rolandic-Oper-L' ↔ 'Insula-L'	146
'Parietal-Inf-L' ↔ 'Fusiform-R'	146
'Cingulum-Ant-R' ↔ 'Supp-Motor-Area-L'	146
'Cingulum-Ant-L' ↔ 'Precuneus-R'	142
'Frontal-Sup-L' ↔ 'Lingual-L'	142
'Frontal-Mid-L' ↔ 'Frontal-Inf-Tri-L'	142
'Fusiform-R' ↔ 'Parietal-Sup-L'	137
'Fusiform-R' ↔ 'Frontal-Mid-L'	136
'Frontal-Mid-R' ↔ 'Frontal-Sup-Orb-R'	135
'Temporal-Inf-L' ↔ 'Lingual-R'	133
'Pallidum-R' ↔ 'Temporal-Mid-R'	132
'Midbrain' ↔ 'Frontal-Sup-L'	132
'Cerebelum-Crus2-R' ↔ 'Temporal-Pole-Sup-R'	67
'Fusiform-R' ↔ 'Frontal-Sup-L'	53
'Frontal-Sup-L' ↔ 'Putamen-L'	49
'Frontal-Sup-L' ↔ 'Frontal-Inf-Tri-L'	44
'Precentral-R' ↔ 'Precentral-R'	40
'Vermis-9' ↔ 'Temporal-Pole-Mid-R'	31
'Frontal-Mid-L' ↔ 'Insula-R'	27
'Temporal-Pole-Sup-R' ↔ 'Frontal-Med-Orb-R'	27
'Cingulum-Ant-R' ↔ 'Vermis-4-5'	26
'Frontal-Sup-R' ↔ 'Cingulum-Ant-L'	25
'Frontal-Mid-L' ↔ 'Cingulum-Ant-L'	22
'Cingulum-Ant-R' ↔ 'Paracentral-Lobule-R'	22
'Insula-L' ↔ 'Cingulum-Ant-L'	19
'Frontal-Sup-R' ↔ 'ParahippocampaGyrus'	14
'Lingual-L' ↔ 'Frontal-Sup-L'	14
'Frontal-Inf-Oper-L' ↔ 'Frontal-Sup-L'	14
'Cingulum-Ant-R' ↔ 'Precuneus-R'	10
'Cingulum-Ant-R' ↔ 'Caudate-Head-L'	10
'Insula-L' ↔ 'Cingulum-Mid-L'	9
'Insula-L' ↔ 'Frontal-Mid-L'	9
'Cingulum-Ant-L' ↔ 'Paracentral-Lobule-R'	8
'Frontal-Inf-Orb-R' ↔ 'Caudate-R'	6
'Vermis-4-5' ↔ 'Frontal-Sup-Orb-R'	6
'Frontal-Sup-L' ↔ 'Cerebelum-6-L'	5
'Frontal-Sup-Medial-L' ↔ 'Cingulum-Mid-L'	5
'Rolandic-Oper-L' ↔ 'Temporal-Mid-L'	5
'Temporal-Inf-L' ↔ 'Putamen-L'	4
'Frontal-Sup-Medial-R' ↔ 'Cingulum-Mid-L'	4
'Precentral-L' ↔ 'Cingulum-Ant-R'	4
'Temporal-Mid-L' ↔ 'Rolandic-Oper-L'	3
'Fusiform-L' ↔ 'Fusiform-R'	1

Comparison of Results

Table 12: Features that are common in all the three methods along with their frequencies in the respective methods (after post-processing)

Feature	LOGIT	SVM-L	SVM-R
'Insula-R' ↔ 'Paracentral-Lobule-R'	381	114	
'Supp-Motor-Area-L' ↔ 'SupraMarginal-L'	377	2	
'Vermis-1-2' ↔ 'Temporal-Mid-L'	363	114	
'Vermis-4-5' ↔ 'Frontal-Sup-Orb-R'	325	114	6
'Insula-L' ↔ 'Frontal-Mid-L'	381		9
'Frontal-Sup-R' ↔ 'ParahippocampaGyrus'	363		14
'Frontal-Sup-Medial-L' ↔ 'Cingulum-Mid-L'	284		5
'Cingulum-Ant-L' ↔ 'Paracentral-Lobule-R'	258		8
'Frontal-Mid-L' ↔ 'Cingulum-Ant-L'	73		146
'Frontal-Sup-Medial-R' ↔ 'Cingulum-Mid-L'	66		4
'Frontal-Sup-L' ↔ 'ParahippocampaGyrus'	51		146
'Insula-L' ↔ 'Cingulum-Ant-L'	23		19
'Frontal-Sup-R' ↔ 'Cingulum-Ant-L'	13		25
'Lingual-L' ↔ 'Frontal-Sup-L'	8		14
'Cingulum-Ant-L' ↔ 'Precuneus-R'	5		142
'Pallidum-R' ↔ 'Temporal-Mid-R'	3		132
'Cingulum-Ant-R' ↔ 'Paracentral-Lobule-R'	1		22
'Frontal-Sup-L' ↔ 'Lingual-L'		114	142
'Parietal-Inf-L' ↔ 'Fusiform-R'		56	146
'Precentral-L' ↔ 'Cingulum-Ant-R'		1	4

Table 13: Frequency of connections between brain regions for each method (after post-processing).

Feature	LOGIT	SVM-L	SVM-R
Frontal↔Frontal	1062	26	335
Temporal ↔Vermis	364	114	31
Caudate↔Frontal	207	2	6
Rolandic↔Temporal	9	6	8
Cingulum↔Frontal	492	84	202
Frontal↔Vermis	343	116	6
Cingulum↔Paracentral	302	2	30
Frontal↔Temporal	235	47	319
Frontal↔Lingual	8	114	156
Insula↔Paracentral	426	114	
Temporal↔Temporal	173	2	
Insula↔Postcentral	84	4	
Cingulum↔Cingulum	7	2	
Lingual↔Temporal	382		133
Frontal↔ParahippocampaGyrus	414		160
Pallidum↔Temporal	284		132
Frontal↔Insula	381		36
Putamen↔Temporal	23		4
Cingulum↔Insula	23		28
Caudate↔Cingulum	22		10
Fusiform↔Parietal		56	283
Frontal↔Putamen		17	49
Cerebelum↔Temporal		4	67
Frontal↔Midbrain		4	132
Cingulum↔Precentral		2	4