More similarities than differences between ADHD and ASD in functional brain connectivity

Joe Bathelt[1,2], Matthan W.A. Caan[2,3], and Hilde M. Geurts[1,2]

[1] Dutch Autism & ADHD Research Center
  Department of Psychology
  University of Amsterdam

[2] Amsterdam Brain and Cognition (ABC)
  University of Amsterdam

[3] Department of Biomedical Engineering & Physics
  Amsterdam UMC
  University of Amsterdam

Abstract

Attention deficit hyperactivity disorder (ADHD) and autism spectrum disorder (ASD) are highly comorbid neurodevelopmental conditions. There is an ongoing debate regarding the nature of their overlap. Behavioral symptoms and cognitive profiles indicate differences between the conditions, but genetic studies and neuroimaging investigations suggest at least some shared etiology. The current study investigated if functional connectivity can be used to distinguish ADHD and ASD using a machine-learning approach. Towards this aim, we trained a machine learning algorithm to distinguish ASD and ADHD cases from each other and from comparison cases in a total sample of 805 cases, comprising of 243 ASD cases, 164 ADHD cases, and 398 comparison cases between 7 and 21 years of age. We compared the performance of the best performing machine learning algorithm (l2-regularised support vector classification) when classifying unseen cases of ADHD, ASD, and CMP. The results indicated lower classification performance when distinguishing ADHD from ASD compared to classifying diagnostic groups vs a typical comparison group. The model trained to distinguish ASD and comparison cases performed equally well when tasked with classifying ADHD vs CMP. A Bayesian analysis gave strong evidence for similarity ADHD and ASD. The ADHD and ASD group showed overlap in connections of the right ventral attention network, the salience network, and the default mode network. In sum, these results suggest a substantial overlap in functional brain connectivity between ADHD and ASD. We discuss the implications of these findings for the quest to identify functional neuroimaging biomarkers and provide recommendation for future research.

Keywords: Attention Deficit Hyperactivity Disorder, Autism Spectrum Disorder, machine-learning classification, resting-state functional connectivity

Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is the most commonly diagnosed mental disorder in children with a lifetime prevalence between 5 and 7%[1]. ADHD is characterized by inattention, hyperactivity, and impulsivity, which encompass difficulties to stay on task, sit still, and wait for one's turn. In contrast, autism spectrum disorder (ASD) is a rarer condition with a prevalence of around 2%[2]. The behavioral characteristics of ASD are difficulties with social communication and interaction alongside restricted, repetitive, and/or stereotyped behaviors and narrow interests (RRBIs). Despite these seemingly different behavioral presentations, clinical experts have long recognized that there is considerable overlap between ADHD and ASD. Formal comparisons indicated that between 20 and 80% of children of a diagnosis of ASD also meet DSM-IV criteria for ADHD[3,4]. In fact, ADHD is the most common comorbid problem in children with ASD[5]. In turn, 30-60% of those with ADHD showed clinically significant symptoms of ASD[3,6]. Notably, these problems encompass all characteristics of ASD, including problems with social interaction/communication and RRBIs. This overlap has been recognized in the latest revision of the Diagnostic and Statistical Manual of Mental Disorder (DSM). While ADHD and ASD were mutually exclusive diagnoses in previous iterations of the DSM, DSM-5 allows for comorbid ADHD and ASD diagnoses[7]. However, this does not settle the debate regarding the nature of the overlap. The overlap could merely reflect a superficial similarity in the presentation of behavioral symptoms, or it could point towards a deeper shared etiology of both conditions.

Genetic studies indicate considerable overlap between ADHD and ASD. Regarding inheritance, a large population study indicated a higher probability to be diagnosed with ADHD for relatives of

individuals with ASC, which was substantially higher for monozygotic than dizygotic twins[8]. This overlap was explained by genetic effects [9]. Further, a large number of known copy number variants (CNVs) and chromosomal abnormalities are associated with a higher likelihood for both ADHD and ASD, including 2q11.2 deletions, duplications of X and Y chromosomes among others (see [7] for a recent review). A cross-syndrome genome-wide association study (GWAS) also indicated a shared likelihood for ADHD and ASD [10]. However, there may be partially distinct genetic influences as indicated by a recent study that reported a weak negative correlation between genetic factors associated with ADHD and ASD [11]. Environmental influences indicate a similar pattern of shared vulnerability alongside some condition-specific effects.

Studies of cognition indicate an overlap between ADHD and ASD in difficulties with executive function and social cognition. Problems in executive function have been identified in both ADHD and ASD in children and adults [12–14]. While both ADHD and ASD show difficulties with multiple aspects of executive function, the executive dysfunction profiles may be partially distinct. Children with ASD were found to show more problems with flexibility and planning, while children with ADHD had more difficulties with inhibition [13–15]. Problems in social cognition have also been identified in both ADHD and ASD. Difficulties with emotion recognition and theory of mind have been identified in both ADHD and ASD, but the problems are milder in ADHD and do not seem to be present in adults with ADHD [16].

Studies that investigated the neural basis of ADHD and ASD suggest similarities in the brain systems involved. Both conditions are increasingly characterized as disorders of brain connectivity that involve differences in the interplay between large-scale brain systems. The most prominent differences in ADHD have been observed in the integration of the default mode network (DMN), the salience network (SN), the ventral attention network (VAN), and the frontoparietal attention network (FAN) [17,18]. Further, the basal ganglia, particularly frontostriatal circuits, have been consistently implicated in ADHD [19–21]. The most consistent differences in connectivity in ASD have been identified in the DMN [22–25]. Some studies also reported hypoconnectivity in the SN [26] and connections between the medial and orbital prefrontal cortex (PFC) with the amygdala [27]. Several attempts have been made to derive biomarkers that distinguish ADHD or ASD by combining functional neuroimaging with machine learning classification. While the initial results indicated very high accuracy (>80%)[28,29], later studies that compared results across samples from different acquisition sites and employed held-out data for validation suggest that classification accuracy of around 60-70% is more realistic[30–32].

Two studies examined similarities and differences in ADHD and ASD in functional brain organization. Kernbach and colleagues identified brain-level phenotypes in the connectivity of the DMN, SN, and dorsal attention network (DAN) in ADHD and ASD[33]. The results indicated that ADHD was characterized by reduced connectivity between the DMN and DAN, and ASD by reduced DMN-SN connectivity. Di Martino et al. identified hypoconnectivity in the precuneus that is common to both ADHD and ASD, and showed that connectivity of the basal ganglia distinguished ASD participants with or without comorbid ADHD symptoms[34]. While these studies provided insight into the potential shared and unique brain connectivity in both diagnostic categories, the question remains if the behaviorally-defined diagnostic categories are meaningful at brain-level. If this was the case, it should be possible to distinguish the diagnostic categories using brain-level markers. The current study aimed to address this in a large sample of ADHD, ASD, and neurotypical cases (N>800).

We employed a machine learning approach that allowed us to pinpoint the connections that together predict the diagnosis of unseen participants on the basis of their functional connectivity profile. This ensures that all reliable information for diagnostic classification that is present in functional connectivity profiles is utilized. Further, it provided us with a measure of classification performance to compare how well functional connectivity can distinguish between diagnostic and comparison groups and between diagnostic groups. Based on the literature that indicates a high overlap between ADHD and ASD at the behavioral, genetic, and neural level, we expected worse performance when classifying ADHD vs ASD compared to the comparisons between diagnostic and neurotypical groups. In the case of substantial similarity between ADHD and ASD, we expected better

performance when grouping both conditions together into a common atypical group. However, worse performance is expected if both groups had unique features that distinguished them from neurotypical individuals. In order to determine the probability of either of these findings, we add a Bayesian approach.

Materials & Methods

Participants

The current analysis was based on data from the Autism Brain Imaging Data Exchange (ABIDE)[35] and the ADHD-200 dataset[36]. Both the ABIDE and ADHD-200 data were collected across multiple sites without prior harmonization of the imaging protocol. The current analysis made use of the publicly available preprocessed version of this dataset provided by the Preprocessed Connectome Project to ensure the replicability of the findings (http://preprocessed-connectomes-project.org/datasets.html). Specifically, we used the fMRI data preprocessed through *fMRIPrep* v1.2.5 [10.1038/s41592-018-0235-4], a state-of-the-art processing pipeline for fMRI[37]. In order to reduce heterogeneity in the sample we excluded some participants. Because of the broad age-range and skewed distribution of age in both datasets, we first selected the 7-to-21-year age range were most data was available. Secondly, we excluded females, because there were far fewer female than male participants in the datasets. Thirdly, we excluded datasets outside the recommended range of image quality indices (frame-wise displacement [FD] > 0.5: n=30 [8.31%]; spatial standard deviation of the data [DVARS] > 5%: n=13 [3.60%])[38]. Lastly, we removed acquisition sites that had only one participant per category (ADHD, ASD, CMP) after the selection for age, gender, and quality control were applied (see **Figure 1** for a breakdown of exclusions). The final sample consisted of 164 participants with ADHD (age: mean=12.11, std=3.215, range [min-max]: 7.24-20.89 [all values in years]), 243 participants with ASD (age: mean=13.4, std=3.1, range: 7.1-20.8), 140 comparison cases from the ADHD-200 dataset (age: mean=12.0, std=3.0, range: 7.2-20.5), and 258 comparison cases from the ABIDE dataset (age: mean=13.7, std=3.2, range: 7.2-20.7). There was no significant difference in age between the ADHD group and their comparison group (Welch-corrected t-test: t(299.51)=0.42, p=0.677, d=0.04), nor between the ASD group and their comparison group (t(498.31)=-1.34, p=0.181, d=-0.12). The ADHD group was younger than the ASD group (t(342.72)=-3.91, p<0.001, d=-0.41). Please note that case-control comparisons were conducted with comparison cases of the same dataset, i.e. ADHD vs ADHD-CMP, ASD vs ASD-CMP. We provide a full list of the included participant IDs for each study in the supplementary materials.
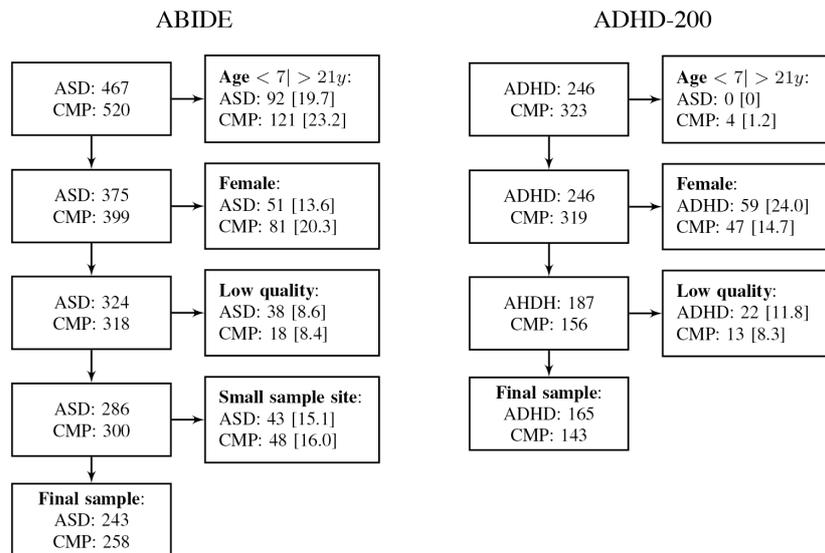
Figure 1 Breakdown of exclusions that lead to the final sample size for the ABIDE dataset (left) and the ADHD-200 dataset (right). The values in brackets show the percentage of excluded cases.

fMRI processing

3.2.1. Preprocessing

The pre-processing of the T1-weighted and fMRI sequences was carried out using *fMRIPrep* v1.2.5 (PRID:SCR_016216). The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) using *N4BiasFieldCorrection*[40] (ANTs 2.2.0) and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped using *antsBrainExtraction.sh* (ANTs 2.2.0, RRID:SCR_004757), using *OASIS* as the target template. Brain surfaces were reconstructed using *recon-all* [41](FreeSurfer 6.0.1, RRID:SCR_001847), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentation of the cortical grey-matter of *Mindboggle*[42] (RRID:SCR_002438). Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c[43] (RRID:SCR_008796) was performed through nonlinear registration with antsRegistration (ANTs 2.2.0, RRID:SCR_004757), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white matter (WM) and grey-matter (GM) was performed on the brain-extracted T1w using FAST[44] (FSL 5.0.9, RRID:SCR_002823).

For each BOLD run per subject (across all sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. The BOLD reference was then co-registered to the T1w reference using *bbregister* (FreeSurfer) which implements boundary-based registration[45]. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using *mcflirt*[46] (FSL v5.0.9). The BOLD time-series were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as preprocessed BOLDpreprocessed BOLD. The BOLD time-series were resampled to MNI152NLin2009cAsym standard space, generating a preprocessed BOLD run in that space. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three

region-wise global signals. FD and DVARS were calculated for each functional run, both using their implementations in Nipype (following the definitions by 10.1016/j.neuroimage.2013.08.048[3]). Global signals were extracted within a CSF and a WM mask defined in the anatomical image for confound regression. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*[47]). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (*tCompCor*) and anatomical (*aCompCor*). Six tCompCor components are then calculated from the top 5% variable voxels within a mask covering the subcortical regions. This subcortical mask is obtained by heavily eroding the brain mask, which ensures it does not include cortical GM regions. For *aCompCor*, six components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). To account for movement-related and physiological artefacts, we regressed the CSF and WM signals and rotational and translations, their temporal derivative, squared term, and squared temporal derivative (36 parameters) from the time series of each voxel [39]. We further regressed the time course of the first 5 spatial and temporal noise components identified through *CompCorr*. This regression was first carried out at the voxel level and then at the ROI level following the protocol described in a study that systematically compared classification accuracy in the ABIDE dataset[30]. In addition, we applied a band-pass filter (0.009-0.1Hz) and spatial smoothing (5mm FWHM) to the resting-state data.

3.2.2. Functional connectome construction

We derived functional connectomes as the correlation between time series within the MSDL atlas. The MSDL atlas is a probabilistic atlas based on multi-subject dictionary learning (MSDL)[40] that demonstrated superior performance in classifying cases and controls in the ABIDE dataset[30]. For greater comparability across samples and studies, the current study employed the pre-computed, labelled MSDL atlas that is distributed with the Nilearn package (https://nilearn.github.io/). Functional connectivity was calculated as the Pearson correlation between the time series of ROIs. The correlation values were transformed using Fisher's r-to-z transformation. To account for differences between sites and changes with age, we regressed the effect of study site and age from the functional connectome edges. Control analyses that consisted of training machine-learning algorithm to predict the age or study site from the regressed data confirmed that the regression successfully removed the influence of age and acquisition site (see Supplementary Materials). The code for the analyses can be found on the Open Science Framework website (osf.io/jekbs).
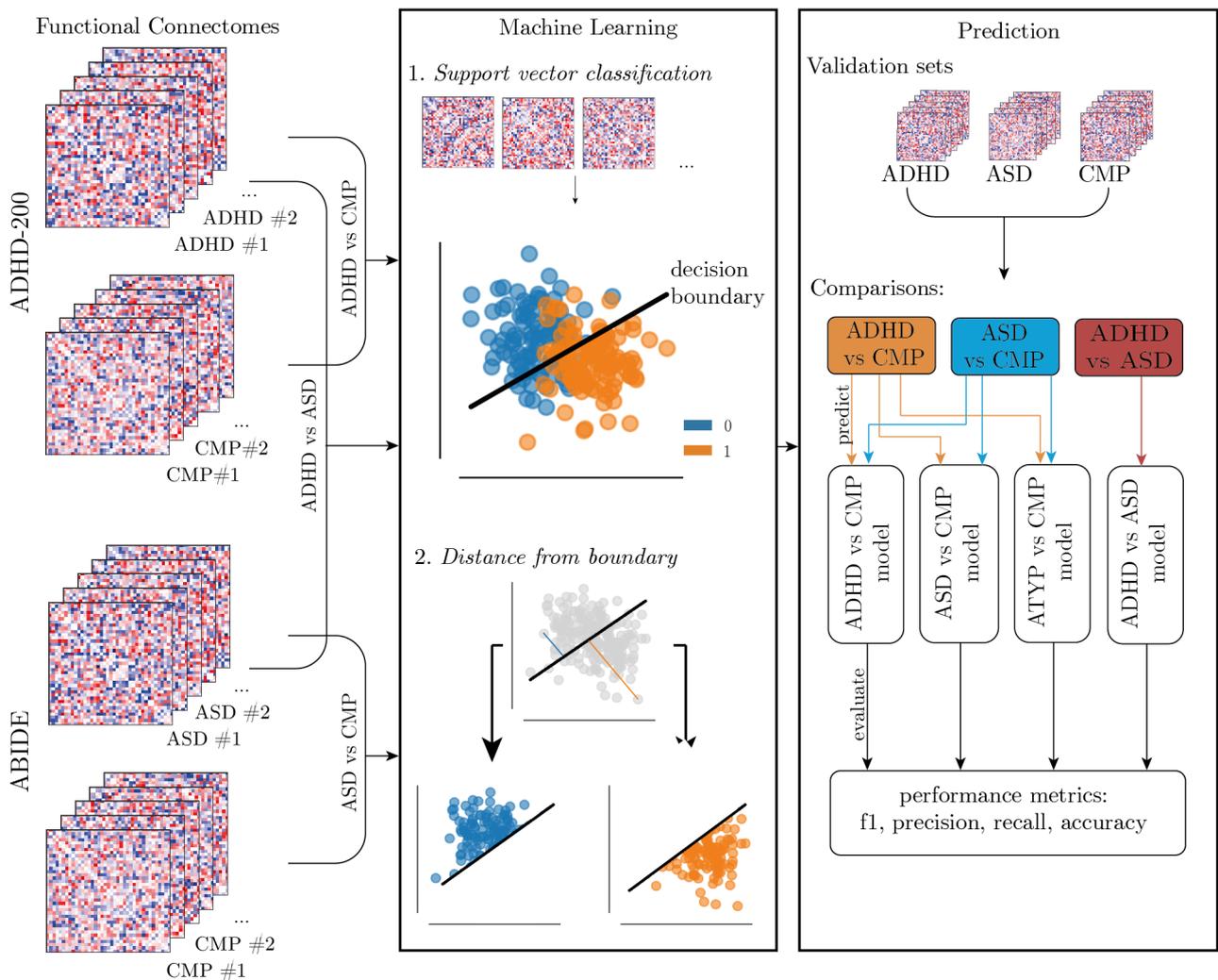
Figure 2 Overview of the machine learning pipeline. The machine learning pipeline created a model for each binary comparison (ADHD vs CMP, ASD vs CMP, ATYP vs CMP, ADHD vs ASD) by identifying a decision boundary using support vector classification. These models were then used to predict group membership in the held-out validation data. The performance of the models was then evaluated and compared between the binary comparisons. The ADHD and ASD groups were combined to form the ATYP group in the ATYP vs CMP comparison. Please note that only two features are shown for the classification approach for illustration purposes. The actual classification was based on a high-dimensional feature space.

3.2.3. Machine learning for classification

The aim of our machine learning approach was two-fold. First, we wanted to compare the classification accuracy between ADHD and ASD to identify potential similarities or differences between the conditions. In addition, we wanted to obtain interpretable brain differences that we could relate to either condition or their overlap. For the machine learning, we employed the same algorithms that were used in a recent authoritative analysis of classification performance in ASD[30], i.e. l1- and l2-regularised support vector classification (l1-/l2-SVC), ridge regression, random forest classification. For model tuning and evaluation, we split the data into a training (80%) and a validation set (20%). Within the training set, we used cross-validation to tune the optimal parameters in a grid search, i.e. regularization parameter for the SVC and ridge regression, minimum samples to split for the random forest. We employed a stratified shuffle split, whereby 60% of the training sample were randomly selected to tune the parameters and 40% was used as an internal test set to avoid overfitting. This procedure was repeated 10 times with different random splits. We applied stratification to maintain equal proportions of the diagnostic and comparison group across splits (see Figure S1 for an illustration). In addition, we used out-of-bag prediction for the random forest classification to

minimize overfitting. We further used the inverse ratio of diagnostic and comparison groups as class weights to account for imbalances in their relative proportion, i.e. assigning a proportionally higher weight to the smaller group. We used the f1-score as the objective to find the optimal balance between precision and recall. Finally, we evaluated the tuned model using the held-out validation set. To compare the prediction accuracy of the models across diagnostic groups, we used the validation set from another comparison, i.e., only data not used in the tuning of any model was used for model comparison. For this cross-prediction, we treated the new data as if it had the label that the model was originally trained on. For instance, we fed ASD and CMP cases into the model trained to distinguish ADHD and CMP and counted ASD cases classified as ADHD by this model as correct. We obtained confidence intervals for the prediction metrics through bootstrap resampling of the validation data with 100 repetitions. We compared each model's classification performance to chance performance with reference to a binomial distribution using SciPy[48] v1.2.1.

3.2.4 Comparison of SVC distances

To further characterize the ability of each model to distinguish the groups, we assessed the distance of each data point from the decision boundary identified by the SVC. To this end, we used the decision boundaries identified in each training set and calculated the distance from this boundary for the whole dataset. We compared differences in this distance using an analysis of variance (ANOVA) model with group (ADHD, ASD, CMP) as a fixed factor. For follow-up comparisons, we employed Tukey's Honest Significance Test (HSD). In addition, we applied a Bayesian ANOVA to assess if any absence of a significant difference in the frequentist analysis is indicative of equivalence between the groups. We used the null model as a reference for the Bayesian analysis and report log-scaled Bayes factor in support of the alternative hypothesis, i.e., $\log(BF_{10})$. We used the Jeffreys scale for interpreting the Bayes factors, i.e. $\log(BF10) > 2$: decisive, $> 1.5\text{-}2$: very strong, $1\text{-}1.5$: strong, $0.5\text{-}1$: substantial, $0\text{-}1/2$: barely worth mentioning, < 0: supports alternative hypothesis and the descriptors for alternative hypothesis [$\log(BF_01)$] are used). For the follow-up contrasts, we report log-scaled posterior odds that were adjusted for multiple comparisons[49] by fixing the prior probability to 0.5. These statistical analyses were carried out using JASP v0.1[50].

Results

Differences in classification performance between ADHD, ASD, and CMP

The comparison of the classification algorithms showed that support vector classification (SVC) with l2-regularisation performed best (see

Table 1). Using this algorithm, we examined if the trained models could predict other classes. The model trained to distinguish ADHD and CMP performed significantly above chance when predicting unseen cases ADHD or CMP cases (see

Table 2). The ADHD vs CMP model also performed above chance when entering cases from the combined ADHD/ASD group (ATYP). The ASD vs CMP model performed above chance for all classifications, i.e. when classifying 'ADHD vs CMP', 'ASD vs CMP', and 'ATYP vs CMP' (see **Table 2**). The model trained to distinguish the combined ADHD/ASD group (ATYP) and the CMP group also performed above chance for all classification, i.e. 'ADHD vs CMP', 'ASD vs CMP', 'ATYP vs CMP'.

To further characterize the separation between classes indicated by the SVC model, we calculated the distance of all data points from the decision boundary identified in the training data and compared the distances between the groups (ADHD, ASD, CMP). For the model trained to distinguish ADHD and CMP, there was a significant difference in these distances between the groups. Post-hoc comparisons indicated significant differences between all groups with an ordering of ADHD < ASD < CMP (see Figure 3). The distances indicated by the model trained to distinguish ASD and CMP also showed a significant difference between groups. Follow-up contrasts indicated differences between all three groups with an order of ASD < ADHD < CMP. There was no significant group effect for the model trained to distinguish ASD and ADHD. Further, there was very strong evidence for no difference between the groups according to the Bayesian analysis.

Table 1 Overview of the performance of alternative machine learning pipelines. Abbreviations: l1-SVC – l1-regularised support vector classification, l2-SVC – l2-regularised support vector classification

	accuracy			
	ADHD vs CMP	ASD vs CMP	ADHD vs ASD	ATYP vs CMP
l1-SVC	0.56±0.006	0.60±0.005	0.58±0.006	0.62±0.006
l2-SVC	0.66±0.005	0.67±0.007	0.66±0.006	0.66±0.006
Ridge Regression	0.61±0.006	0.61±0.006	0.59±0.006	0.62±0.006
Random Forest	0.60±0.006	0.59±0.005	0.48±0.007	0.52±0.008
	f1			
l1-SVC	0.58±0.007	0.63±0.006	0.61±0.006	0.64±0.006
l2-SVC	0.70±0.005	0.70±0.007	0.68±0.007	0.69±0.006
Ridge Regression	0.57±0.007	0.56±0.008	0.55±0.008	0.58±0.008
Random Forest	0.44±0.010	0.50±0.007	0.14±0.008	0.33±0.010
	precision			
l1-SVC	0.52±0.008	0.55±0.007	0.53±0.008	0.57±0.008
l2-SVC	0.60±0.007	0.61±0.008	0.59±0.008	0.60±0.008
Ridge Regression	0.57±0.008	0.56±0.010	0.55±0.009	0.58±0.010
Random Forest	0.61±0.011	0.55±0.008	0.29±0.016	0.47±0.014
	recall			
l1-SVC	0.67±0.008	0.75±0.008	0.75±0.007	0.75±0.008
l2-SVC	0.86±0.006	0.83±0.008	0.82±0.008	0.82±0.007
Ridge Regression	0.58±0.009	0.58±0.010	0.55±0.009	0.59±0.010
Random Forest	0.35±0.010	0.47±0.009	0.10±0.006	0.26±0.009

Table 2 Cross-prediction performance. The left column indicates what the model was trained on, i.e. the 'adhd' model was trained to distinguish ADHD and CMP, the 'asd' model to distinguish ASD and CMP, the 'atyp' model to distinguish ASD + ADHD and CMP, and the 'adhd-asd' model to distinguish ADHD and ASD. The group columns show the data that were entered for cross-prediction. The mean ± standard deviation (SD) of classification performance metrics across 500 bootstrap permutations are shown. The p-value is based on comparison of the correct classification to a binomial distribution. Classification performance significantly above chance is marked with an asterisk.

| model	group1	group2	accuracy	f1	precision	recall	p	
adhd	ADHD	CMP	0.68±0.075	0.66±0.061	0.58±0.075	0.67±0.058	<0.001	*
	ASD	CMP	0.56±0.063	0.55±0.058	0.55±0.063	0.57±0.057	0.137	
	ATYP	CMP	0.63±0.039	0.62±0.033	0.58±0.039	0.63±0.032	0.001	*
asd	ADHD	CMP	0.57±0.066	0.57±0.053	0.51±0.066	0.57±0.053	0.027	*
	ASD	CMP	0.59±0.062	0.59±0.053	0.57±0.062	0.59±0.053	0.045	*
	ATYP	CMP	0.75±0.050	0.75±0.040	0.69±0.050	0.75±0.040	<0.001	*
atyp	ADHD	CMP	0.87±0.071	0.87±0.040	0.78±0.071	0.87±0.040	<0.001	*
	ASD	CMP	0.80±0.061	0.80±0.045	0.74±0.061	0.81±0.044	<0.001	*
	ATYP	CMP	0.58±0.049	0.58±0.039	0.54±0.049	0.58±0.039	0.011	*
asd-adhd	ADHD	ASD	0.47±0.064	0.49±0.065	0.58±0.064	0.50±0.064	0.953	
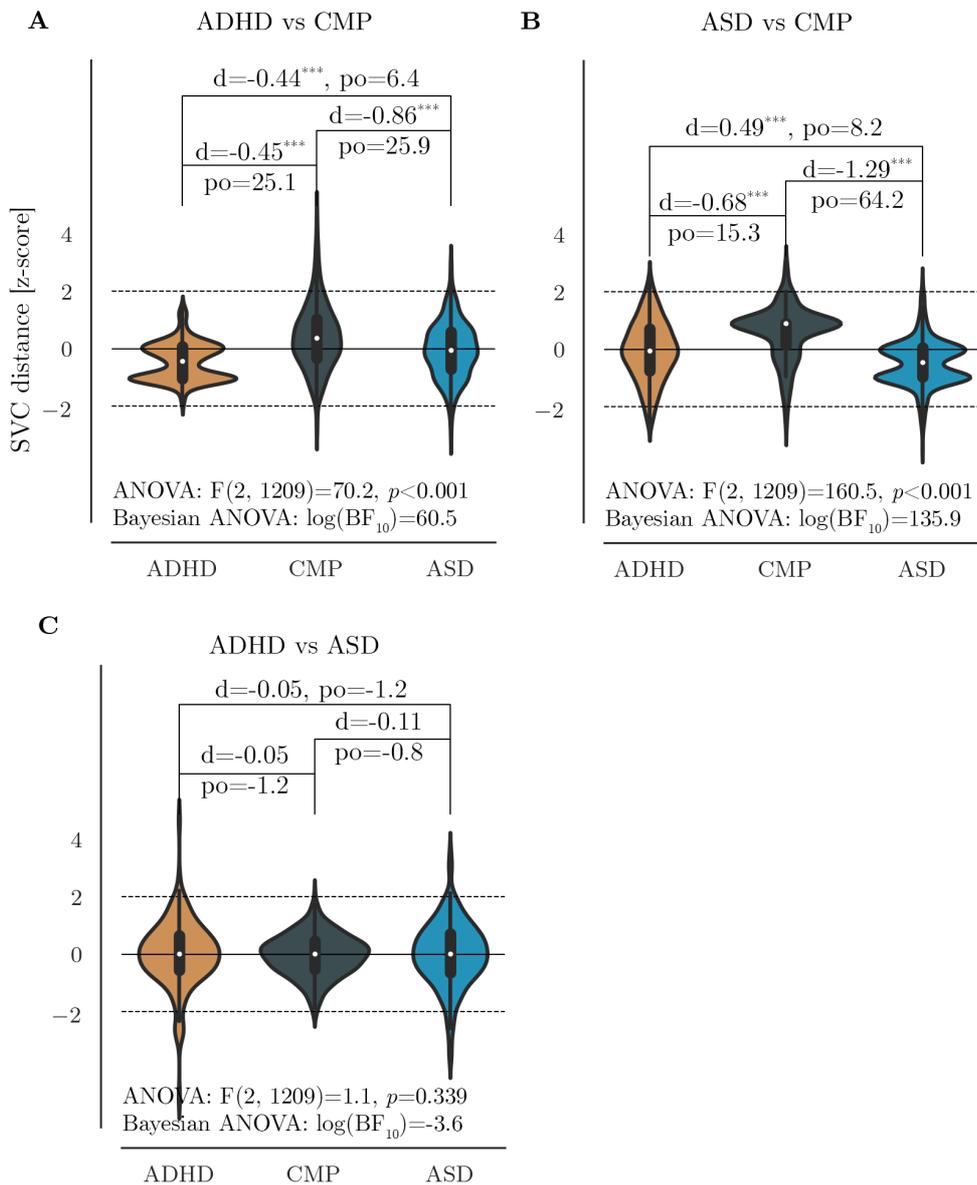
Figure 3 Comparison of the SVC distances indicated by each model for each group. A) Model trained to distinguish ADHD and CMP cases, B) Model trained to distinguish ASD and CMP cases, c) Model trained to distinguish ADHD and ASD cases. The x-axis shows the label of the cases that were fed into the model for prediction. The y-axis shows the distance of cases from the decision boundary in the model. The annotation underneath each plot indicates the group effect according indicated by a frequentist (Cohens' d) and a Bayesian ANOVA (po). The null model was used as a reference for the Bayesian ANOVA. The annotations show the results of post-hoc contrasts. For the frequentist analysis, Cohen's d is shown. The asterisks indicate a p-value < 0.001 (Tukey HSD). For the Bayesian analysis, the log(posterior odds) (po) adjusted for multiple comparisons are shown.

Shared and distinct functional connectivity in ADHD and ASD

In an additional exploratory analysis, we investigated the functional connections that distinguish the groups. To that end, we fitted the l2-SVC model to the whole dataset with the optimized parameters indicated in the training set. The l2-SVC algorithm assigns a coefficient to each feature, in this case functional connectivity edge. To make the results more interpretable, we introduced sparsity through a bootstrap procedure. The sparse solution suggested that the ADHD vs CMP model loaded most strongly on connections between the cingulo-insular network (Cing-Ins) and default mode network

(DMN), the salience network and right ventral attention network (R V Att), and the auditory (Aud) and secondary visual network (Vis Sec, see **Figure** *4*).

The model that was trained to distinguish ASD and CMP had the highest coefficients for connections between the auditory (Aud) and salience network, the DMN and the dorsal posterior cingulate cortex (Dors PCC), the right dorsal attention network (D Att) and the salience network, and the left ventral attention network (L V Att) and the language network (Lang, see **Figure** *4*).

The model that was trained to distinguish the ADHD and ASD group for connections between the secondary visual network (Vis Sec) and the default mode network (DMN), between Vis Sec and the language network, the dorsal attention network (D Att) and the language network, and between the left ventral attention network (L V Att) and the language network.

The coefficients of some connections were included across models. The 'ADHD vs CMP' and the 'ASD vs CMP model both loaded on connections between the ventral anterior cingulate cortex (V ACC) and the posterior occipital cortex (Occ post), the right posterior temporal cortex (R Post Temp) and the left DMN, and between the left DMN and the right temporo-parietal junction (R TPJ). The 'ADHD vs CMP' and the 'ADHD vs ASD' model between the right anterior intra parietal sulcus (R Ant IPS) and the right DMN. The 'ASD vs CMP' and the 'ADHD vs ASD' model both loaded on connections between the left lateral occipital complex (LOC) and the left auditory cortex, and between the left LOC and the parietal region of the left ventral attention network (L Par).
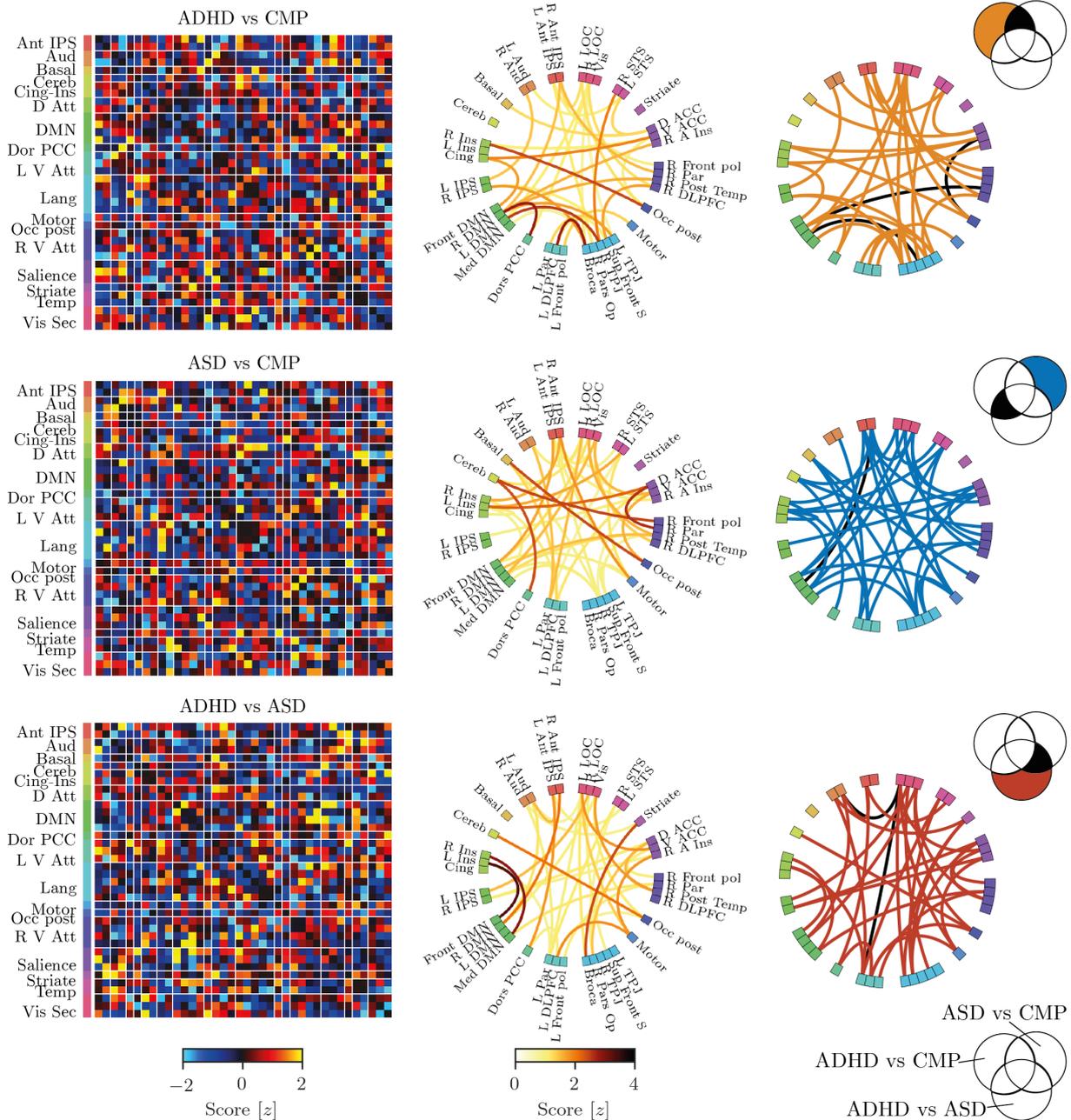
Figure 4 Comparison of the SVC solution for the different models. The model trained to distinguish ADHD and CMP is shown at the top, the ASD vs CMP model in the middle, and the ADHD vs ASD model at the bottom. The left panel shows the full SVC coefficient matrix. The colour bar at the right of each matrix indicates the canonical functional network for each ROI. The boundaries between canonical networks are marked with white lines in the matrix. Warmer colours indicate a more positive coefficient, while cooler colours indicate a more negative coefficient. The middle panel shows the edges that were significant at p<0.05 compared to a null model with scrambled diagnoses. Only positive coefficients are shown because only positive coefficients were indicated to be significant. Darker red colour indicates a higher coefficient. The diagrams on the right show the unique connections in each model and the shared connections. The Venn diagrams indicate which unique and shared connections are shown. Please note that the whole sample was used to generate these plots, i.e., training and validation set.

Discussion

The current analysis investigated similarities and differences in functional brain organization between ADHD and ASD. Towards this end, we trained machine learning models to distinguish each diagnostic group from a typical comparison group, and another model to distinguish the diagnostic groups from each other. The results suggested that a model trained to distinguish ASD from

13

neurotypical cases also performed well when classifying ADHD and comparison cases, but a model trained to distinguish ADHD and comparison cases did not perform well when classifying ASD and comparison cases. A model trained to distinguish atypical cases (ADHD, ASD) from comparison cases performed well when classifying unseen cases from either diagnostic group. In contrast, the model trained to distinguish the groups from each other performed substantially less well when tasked with distinguishing the groups from typical comparison cases. A Bayesian analysis suggested strong evidence that the model comparing the diagnostic groups could not distinguish between them. The functional connectivity profiles showed overlap between ADHD and ASD in connections of the right ventral attention network, the salience network, and the default mode network.

The results of the current analysis agree with previous empirical studies and theoretical work that suggests a substantial overlap between ADHD and ASD. Specifically, our finding of low performance when trying to distinguish ADHD and ASD and the strong Bayesian evidence for similarity suggest that functional brain connectivity is either too similar between both groups, or too variable within each group for reliable classification using functional connectivity. The modest performance that can be achieved when distinguishing individuals with a diagnosis from comparison individuals[30] may suggest that it is the latter. An alternative view is that ADHD and ASD form a continuum with typical and autistic individuals at the extreme and ADHD in the middle [51,52]. If this were the case, one would expect that it is easier to distinguish autistic from neurotypical individuals compared to ADHD vs neurotypical, because autistic and neurotypical individuals are more dissimilar according to this dimensional account. However, the current analysis indicated similar performance for both comparisons. In other words, autistic individuals are not more dissimilar from neurotypical individuals than ADHD individuals, at least in functional brain connectivity. Further, one would expect lower performance when combining ADHD and ASD in the atypical vs typical comparison if the dimensional account would hold. That is because the supposed intermediate ADHD group is included in this comparison. However, the current analysis indicated even better performance in the combined comparisons, i.e. atypical vs comparison, as in the separate comparisons, i.e. ADHD vs CMP, ASD vs CMP. Further, the comparison of model weights suggest that ADHD is closer to ASD rather than in the middle between CMP and ASD. However, the cross-prediction indicated that the model trained to distinguish ADHD and CMP cases did not generalize when tasked to classify ASD vs CMP, while the ASD vs CMP model performed well when classifying ADHD cases. This could suggest a dimensional ordering that makes it easier to create a classification model from the more extreme exemplars. However, the ADHD model was also trained with fewer cases which might hamper its generalizability.

Arguably, the results of the current analysis are most consistent with a theoretical account that suggests a high degree of overlap between condition categories without assuming a dimensional order. According to this view, all dimensions of psychopathology load onto a common factor p [53], akin to the general intelligence factor g in cognition. The dominant p factor would produce a high degree of overlap between the diagnostic categories and would distinguish them from the typical comparison group. Any condition-specific contribution is assumed to be much smaller and, therefore, harder to distinguish. This is consistent with the low performance of the model that was trained to distinguish the diagnostic groups relative to the other comparisons.

It is important to keep in mind some limitations of the current analysis. This type of analysis requires a large data set to produce robust and meaningful results and had to rely on large, publicly available datasets to reach a sufficient sample size. Even though these datasets are unprecedented in their ambition and contribution to the science of neurodevelopmental disorder, they are also limited. Firstly, there was no previous harmonization between the sites which resulted in differences in sample recruitment, diagnosis, behavioral assessment, and neuroimaging. Our attempt to remove this variation by regressing sample site as a confounding factor and comparing cases and control within each sample was probably not sufficient to fully account for these differences. Large-scale studies with comprehensive harmonization between participating acquisition sites are already underway and will pave the way for more in-depth characterization of neurodevelopment. Secondly, there were no

14

behavioral measures that were common across the ADHD and ASD groups. Future studies that incorporate detailed measures of brain, behavior, and cognition will need to investigate how the brain-level differences relate to behavioral and cognitive similarities in ASD and ADHD. Thirdly, the current analysis relied on data collected with case-control comparisons in mind. However, the behavioral and cognitive traits associated with ASD and ADHD may be continuously distributed across the population[54]. The differences identified with the data used in the current analysis may appear more extreme because individuals with intermediate phenotypes were not well represented in the comparison group. Future studies that measure ADHD and ASD traits in broader samples that span the typical and atypical spectrum will be better able to identify these associations.

The results of the current analysis highlight the challenges of identifying brain-level markers of behaviorally-defined diagnostic groups. A growing body of literature employs ever-more sophisticated machine learning approaches to distinguished diagnostic groups like ASD or ADHD from typical comparison groups. The modest performance of 60-70% accuracy that can be achieved demonstrates that functional or structural neuroimaging markers are of limited utility for clinical purposes. Furthermore, as the current study demonstrates, classifying separate clinical groups is even more challenging resulting in close-to-chance performance even with large datasets and state-of-the-art machine learning algorithms. This provides an important note of caution for studies that aim to identify biomarkers from case-control comparison studies. The markers identified in these studies are likely to reflect general markers of "atypicality" but cannot be regarded as true biomarkers of the condition if they cannot distinguish the diagnostic group from their common comorbid conditions, as shown here for ADHD and ASD. Future studies are likely to be of more clinical utility if they focused on the neurobiological correlates of specific symptom dimensions across the population or if comorbid diagnostic groups are included to establish the specificity of functional neuroimaging biomarkers.

References

1. Polanczyk GV, Willcutt EG, Salum GA, Kieling C, Rohde LA. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. *International Journal of Epidemiology* 2014; **43**: 434–442.
2. Zablotsky B, Black LI, Maenner MJ, Schieve LA, Danielson ML, Bitsko RH *et al.* Prevalence and Trends of Developmental Disabilities among Children in the United States: 2009–2017. *Pediatrics* 2019; **144**: e20190811.
3. Ronald A, Simonoff E, Kuntsi J, Asherson P, Plomin R. Evidence for overlapping genetic influences on autistic and ADHD behaviors in a community twin sample. *Journal of Child Psychology and Psychiatry* 2008; **49**: 535–542.
4. Matson JL, Rieske RD, Williams LW. The relationship between autism spectrum disorders and attention-deficit/hyperactivity disorder: An overview. *Research in Developmental Disabilities* 2013; **34**: 2475–2484.
5. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G. Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *Journal of the American Academy of Child & Adolescent Psychiatry* 2008; **47**: 921–929.
6. Clark T, Feehan C, Tinline C, Vostanis P. Autistic symptoms in children with attention deficit-hyperactivity disorder. *European Child & Adolescent Psychiatry* 1999; **8**: 50–55.
7. Antshel KM, Zhang-James Y, Wagner KE, Ledesma A, Faraone SV. An update on the comorbidity of ADHD and ASD: a focus on clinical management. *Expert Review of Neurotherapeutics* 2016; **16**: 279–293.
8. Ghirardi L, Brikell I, Kuja-Halkola R, Freitag CM, Franke B, Asherson P *et al.* The familial co-aggregation of ASD and ADHD: a register-based cohort study. *Molecular Psychiatry* 2017; **23**: 257–262.
9. Ghirardi L, Pettersson E, Taylor MJ, Freitag CM, Franke B, Asherson P *et al.* Genetic and environmental contribution to the overlap between ADHD and ASD trait dimensions in young adults: a twin study. *Psychological Medicine* 2018; **49**: 1713–1721.
10. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet* 2013; **381**: 1371–1379.
11. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR *et al.* An atlas of genetic correlations across human diseases and traits. *Nature Genetics* 2015; **47**: 1236–1241.

12. Nydén A, Niklasson L, Stahlberg O, Anckarsater H, Wentz E, Rastam M et al. Adults with autism spectrum disorders and ADHD neuropsychological aspects. *Research in Developmental Disabilities* 2010; **31**: 1659–1668.

13. Geurts HM, Verte S, Oosterlaan J, Roeyers H, Sergeant JA. How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology and Psychiatry* 2004; **45**: 836–854.

14. Happé F, Booth R, Charlton R, Hughes C. Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition* 2006; **61**: 25–39.

15. Ozonoff S, Jensen J. *Journal of Autism and Developmental Disorders* 1999; **29**: 171–177.

16. Bora E, Pantelis C. Meta-analysis of social cognition in attention-deficit/hyperactivity disorder (ADHD): comparison with healthy controls and autistic spectrum disorder. *Psychological Medicine* 2015; **46**: 699–716.

17. Cortese S, Kelly C, Chabernaud C, Proal E, Di Martino A, Milham MP et al. Toward Systems Neuroscience of ADHD: A Meta-Analysis of 55 fMRI Studies. *American Journal of Psychiatry* 2012; **169**: 1038–1055.

18. Sripada C, Kessler D, Fang Y, Welsh RC, Prem Kumar K, Angstadt M. Disrupted network architecture of the resting brain in attention-deficit/hyperactivity disorder. *Human Brain Mapping* 2014; **35**: 4693–4705.

19. Frodl T, Skokauskas N. Meta-analysis of structural MRI studies in children and adults with attention deficit hyperactivity disorder indicates treatment effects. *Acta Psychiatrica Scandinavica* 2011; **125**: 114–126.

20. Ellison-Wright I, Ellison-Wright Z, Bullmore E. Structural brain change in Attention Deficit Hyperactivity Disorder identified by meta-analysis. *BMC Psychiatry* 2008; **8**. doi:10.1186/1471-244x-8-51.

21. Nakao T, Radua J, Rubia K, Mataix-Cols D. Gray Matter Volume Abnormalities in ADHD: Voxel-Based Meta-Analysis Exploring the Effects of Age and Stimulant Medication. *American Journal of Psychiatry* 2011; **168**: 1154–1163.

22. Bathelt J, Koolschijn PC, Geurts HM. Age-variant and age-invariant features of functional brain organization in middle-aged and older autistic adults. *Molecular Autism* 2020; **11**. doi:10.1186/s13229-020-0316-y.

23. Lynch CJ, Uddin LQ, Supekar K, Khouzam A, Phillips J, Menon V. Default Mode Network in Childhood Autism: Posteromedial Cortex Heterogeneity and Relationship with Social Deficits. *Biological Psychiatry* 2013; **74**: 212–219.

24. Padmanabhan A, Lynch CJ, Schaer M, Menon V. The Default Mode Network in Autism. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 2017; **2**: 476–486.

25. Glerean E, Pan RK, Salmi J, Kujala R, Lahnakoski JM, Roine U et al. Reorganization of functionally connected brain subnetworks in high-functioning autism. *Human Brain Mapping* 2015; **37**: 1066–1079.

26. Hull JV, Dokovna LB, Jacokes ZJ, Torgerson CM, Irimia A, Van Horn JD. Resting-State Functional Connectivity in Autism Spectrum Disorders: A Review. *Frontiers in Psychiatry* 2017; **7**. doi:10.3389/fpsyt.2016.00205.

27. Gotts SJ, Simmons WK, Milbury LA, Wallace GL, Cox RW, Martin A. Fractionation of social brain circuits in autism spectrum disorders. *Brain* 2012; **135**: 2711–2725.

28. Anderson JS, Nielsen JA, Froehlich AL, DuBray MB, Druzgal TJ, Cariello AN et al. Functional connectivity magnetic resonance imaging classification of autism. *Brain* 2011; **134**: 3742–3754.

29. Wang H, Chen C, Fushing H. Extracting Multiscale Pattern Information of fMRI Based Functional Brain Connectivity with Application on Classification of Autism Spectrum Disorders. *PLoS ONE* 2012; **7**: e45502.

30. Abraham A, Milham MP, Di Martino A, Craddock RC, Samaras D, Thirion B et al. Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. *NeuroImage* 2017; **147**: 736–745.

31. Brown MRG, Sidhu GS, Greiner R, Asgarian N, Bastani M, Silverstone PH et al. ADHD-200 Global Competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements. *Frontiers in Systems Neuroscience* 2012; **6**. doi:10.3389/fnsys.2012.00069.

32. Plitt M, Barnes KA, Martin A. Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *NeuroImage: Clinical* 2015; **7**: 359–366.

33. Kernbach JM, Satterthwaite TD, Bassett DS, Smallwood J, Margulies D, Krall S et al. Shared endo-phenotypes of default mode dysfunction in attention deficit/hyperactivity disorder and autism spectrum disorder. *Translational Psychiatry* 2018; **8**. doi:10.1038/s41398-018-0179-6.

34. Di Martino A, Zuo X-N, Kelly C, Grzadzinski R, Mennes M, Schvarcz A et al. Shared and Distinct Intrinsic Functional Network Centrality in Autism and Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry* 2013; **74**: 623–632.

35. Di Martino A, Yan C-G, Li Q, Denio E, Castellanos FX, Alaerts K et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry* 2013; **19**: 659–667.

36. Consortium T. The ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in Systems Neuroscience* 2012; **6**. doi:10.3389/fnsys.2012.00062.

37. Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods* 2018; **16**: 111–116.

38. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* 2012; **59**: 2142–2154.

39. Satterthwaite TD, Wolf DH, Loughead J, Ruparel K, Elliott MA, Hakonarson H et al. Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth. *NeuroImage* 2012; **60**: 623–632.

40. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, and Gee JC. N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging* 2010; **29**(6): 1310–20.

41. Dale AM, Fischl B, and Sereno MI. Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage* 1999; **9**(2): 179–94.

42. Klein A, Ghosh SS, Bao FS, Giard J, Häme Y, Stavsky E, Lee N, et al. Mindboggling morphometry of human brains. *PLOS Computational Biology* 2017; **13**(2): e1005350.

43. Fonov VS, Evans AC, McKinstry RC, Almli RC, Collins DL. Unbiased Nonlinear Average Age-Appropriate Brain Templates from Birth to Adulthood." *NeuroImage*, Organization for human brain mapping 2009 annual meeting 2009; **47**: S102

44. Zhang Y, Brady M, Smith S. Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm. *IEEE Transactions on Medical Imaging* 2001; **20**(1): 45–57.

45. Greve DN, Fischl B. Accurate and Robust Brain Image Alignment Using Boundary-Based Registration. *NeuroImage* 2009; **48**(1): 63–72.

46. Jenkinson M, Bannister P, Brady M, Smith, S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* 2002; **17**(2): 825–41.

47. Behzadi Y, Restom K, Liau J, Liu TT. A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI. *NeuroImage* 2007; **37**(1): 90–101.

48. Oliphant TE. Python for Scientific Computing. *Computing in Science & Engineering* 2007; **9**: 10-20.

49. Westfall PH, Johnson WO, Utts JM. A Bayesian perspective on the Bonferroni adjustment. *Biometrika* 1997; **84**(2): 419-427.

50. JASP Team. JASP (Version 0.11.1) 2019 [Computer software].

51. Rommelse N, Buitelaar JK, Hartman CA. Structural brain imaging correlates of ASD and ADHD across the lifespan: a hypothesis-generating review on developmental ASD–ADHD subtypes. *Journal of Neural Transmission* 2016; **124**: 259–271.

52. Kern JK, Geier DA, Sykes LK, Geier MR, Deth RC. Are ASD and ADHD a Continuum? A Comparison of Pathophysiological Similarities Between the Disorders. *Journal of Attention Disorders* 2012; **19**: 805–827.

53. Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S *et al.* The p Factor. *Clinical Psychological Science* 2013; **2**: 119–137.

54. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K *et al.* Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *American Journal of Psychiatry* 2010; **167**: 748–751.