Using the Excitation/Inhibition Ratio to Optimize the Classification of Autism Spectrum Disorder and Schizophrenia

Lavinia Carmen Uscătescu[1], Christopher J. Hyatt[1], Jack Dunn[8], Martin Kronbichler[6,7], Vince Calhoun[3], Silvia Corbera[4], Kevin Pelphrey[5], Brian Pittman[2], Godfrey Pearlson[1,2] and Michal Assaf[1,2].

1 Olin Neuropsychiatry Research Center, Institute of Living, Hartford, CT, USA

2 Yale University, School of Medicine, Department of Psychiatry, New Haven, CT, USA

3 Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS) Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, GA, USA

4 Central Connecticut State University, Department of Psychological Science, New Britain, CT, USA

5 University of Virginia, Department of Neurology, Charlottesville, VA, USA

6 Centre for Cognitive Neuroscience & Department of Psychology, Paris-Lodron University of Salzburg, Salzburg, Austria

7 Neuroscience Institute, Christian-Doppler Medical University Hospital, Paracelsus Medical University, Salzburg, Austria

8 Interpretable AI, Cambridge, MA, USA

Abstract

The excitation/inhibition (E/I) ratio has been shown to be elevated in both autism spectrum disorder (ASD) and schizophrenia (SZ), relative to neurotypical controls. However, the degree of E/I imbalance overlap and differentiation between SZ and ASD is not known. Our main objectives were therefore (1) to quantify group differences in the E/I ratio between controls, ASD and SZ, and (2) to assess the potential of the E/I ratio for differential diagnosis. We collected resting state fMRI (rsfMRI) and phenotypic data from 55 controls, 30 ASD, and 39 SZ, ages 18 to 35 (IQ>80). For each participant, we computed the Hurst exponent (H), an indicator of the E/I ratio, for the timecourses of 53 independent components covering the entire brain. Next, using Optimal Classification Trees (OCT), we ran a classification analysis on the two clinical groups using five incremental feature sets (i.e., models): (1) Positive and Negative Syndrome Scale (PANSS) and the Autism Diagnostic Observation Schedule (ADOS) only; (2) PANSS, ADOS, Bermond–Vorst Alexithymia Questionnaire (BVAQ), Empathy Quotient (EQ), and IQ; (3) H only; (4) H, PANSS and ADOS; (5) H, PANSS, ADOS, BVAQ, EQ and IQ. We observed decreased H (i.e., increase in E/I ratio) in ASD and SZ compared to controls, and in SZ compared to ASD in the Cerebellar, Sensorimotor, Visual and Cognitive Control networks. The

OCT classification showed a consistent increase in discrimination accuracy across models between ASD and SZ, suggesting that the E/I ratio in combination with phenotypic measures can contribute to differential diagnosis in adults.

1. Introduction

Autism spectrum disorder (ASD) and schizophrenia (SZ) have been historically conceptualized as a continuum of the same disorder, when Eugen Bleuler posited autism to be a type of impairment in SZ, alongside affectivity, association, and ambivalence (Bleuler, 1911). In the 1970s they were recognized as independent diagnoses and have remained so until now (*Diagnostic and Statistical Manual of Mental Disorders* 5th ed.; DSM–5; American Psychiatric Association, 2013). While ASD is primarily characterized by impairments in social communication skills and by repetitive behaviors, a SZ diagnosis consists of positive (e.g., hallucinations, delusions) and negative (e.g., social withdrawal) symptoms. The heterogeneity of both diagnostic categories (Benkarim et al., 2022; Segal et al., 2022) and their phenotypic overlap (Kästner et al., 2015) can hinder accurate diagnosis. More precisely, ASD and SZ co-occur in approximately 4% of cases (Lai et al., 2019), and share both social (Oliver et al., 2020) and sensory-motor deficits (Du et al., 2021). Common clinical observational or self-report tools, such as the Autism Diagnostic Observation Schedule (ADOS) and the Positive and Negative Syndrome Scale (PANSS), do not have good specificity (Bastiaansen et al., 2011; Trevisan et al., 2020), thus efforts have been made to find discriminative neuroimaging biomarkers (Horien et al., 2022). A recent international machine learning challenge aimed to classify ASD and controls showed that fMRI data can yield a diagnosis accuracy of ~ 80% (Traut et al., 2022). However, the challenge is even greater when tackling the overlap of different disorders, such as ASD and SZ, that share both genetic and neuroimaging commonalities (Moreau et al., 2021).

One explanatory hypothesis for the observed sensory deficits in ASD has been that of increased excitation/inhibition (E/I) ratio, first proposed by Rubenstein and Merzenich (2003). Among the evidence they cite is the fact that parietal and cerebellar areas of the ASD brain show ~50% less glutamic acid decarboxylase (GAD), the enzyme that synthesizes the inhibitory neurotransmitter γ -aminobutyric acid (GABA), compared to controls (Fatemi et al., 2002). Additionally, in ASD, cortical minicolumns, which are functional units composed of GABAergic and glutamatergic neurons processing thalamic inputs, are smaller and more numerous in ASD compared to controls (Casanova et al., 2002). A more recent summary specifically points towards the relationship between reduced inhibition and its negative impact on cortical and hippocampal functioning in ASD (Sohal & Rubenstein, 2019). Whether the E/I imbalance (hereafter, we use the term E/I imbalance to indicate an elevated E/I ratio in patients vs. controls) is mainly due to excessive excitatory activity, or deficient inhibitory activity, is not entirely clear (Dickinson, Jones & Milne, 2016; Ford & Crewther, 2016), but recent evidence points to the E/I imbalance in ASD being caused by concomitant increase in excitation and

decrease in inhibition (for an excellent summary of findings see Canitano & Palumbi, 2021).

Post-mortem and genetic evidence point to the presence of E/I imbalance in SZ as well (Anticevic & Lisman, 2017), and computational modeling revealed that this imbalance causes hyperconnectivity in association brain areas (Yang et al., 2016). In addition, a review has shown a link between an E/I imbalance and aberrant internal sensory processing in SZ, such as hallucinations (Jardri et al., 2016). Finally, dopamine appears to be crucial in maintaining the E/I balance by modulating the excitability of glutamate and GABAergic neurons (Purves-Tyson et al., 2021), with direct impact on memory function and prefrontal dysconnectivity (see Winterer and Weinberger, 2004, for an extended account).

Supportive evidence in favor of the E/I imbalance in ASD and SZ has been corroborated by animal models and post-mortem human studies, as well as by experimental, genetic and magnetic resonance spectroscopy studies (MRS) (for a comprehensive review, see Dickinson, Jones & Milne, 2016). It has been proposed that there are common neuronal pathways underlying E/I imbalance in both ASD and SZ (Canitano & Pallagrosi, 2017; Foss-Feig et al., 2017), and that this relies in turn on shared genotype (Gao & Penzes, 2015). Additionally, a recent meta-analysis also showed that comorbidity of ASD and SZ has a 4% prevalence (Lai et al., 2019). However, given the substantial heterogeneity in both SZ and ASD (e.g., Segal et al., 2022), it is difficult to ascertain to which extent and in which brain areas or networks there is overlap between the two diagnoses.

In recent years, numerous machine learning approaches have been employed to improve differential diagnosis of mental disorders. Among these, interpretable models, such as classification trees, have become increasingly popular due to their transparency, as opposed to the traditional “black box” methods, such as deep learning models (Murdoch et al., 2019; Rudin, 2019). Interpretable machine learning has been successfully used in ASD research across multiple independent samples and has shown good accuracy (~78%) when differentiating ASD from typically developed individuals (TD) using brain imaging derived features (Supekar et al., 2022). Similarly, decision trees with a feature set comprising items from several psychosis rating tools, have been used to distinguish between schizophrenia, schizo-affective disorder, depression and bipolar disorder with remarkable performance (Gibbons et al., 2021). Likewise, a decision tree approach yielded high accuracy in diagnosing ASD based on a large collection of features documenting family history, longitudinal information of symptom onset and severity, and clinical ratings (Hassan & Mokhtar, 2019). Finally, two decision tree approaches have been used to distinguish between ASD and SZ. One was based on structural features, and its accuracy did not exceed 65% (Yassin et al., 2020), while the other one used neuropsychological and self-report measures with accuracy of ~ 83% (Demetriou et al., 2020).

Traditional decision tree algorithms have however been shown to suffer from various limitations which can cause the classification performance to reflect local rather than exhaustive characteristics of the dataset (for a more detailed explanation, see

section 2.4 as well as Bertsimas & Dunn, 2017). In the present study, we therefore used a novel and improved decision tree algorithm, Optimal Classification Trees (OCT), developed by Bertsimas & Dunn (2017). Classifications based on decision tree algorithms traditionally operate through successive binary decision splits according to feature values along decision branches. Previous classification trees all share a common shortcoming in that they are biased towards finding local optimal solutions which may not accurately characterize the best global model fit. The main advantage of OCT is that it allows to overcome this limitation by fitting the whole tree in one optimized step, as opposed to using successive splits. The advantage of this approach is that it optimizes the trade-off between model complexity and accuracy, thus enabling us to compare the performance of different models containing a varying number of features. Finally, OCT also allows for a multivariate classification, where each split is decided upon a combination of features instead of one feature at a time. Although multivariate splits were previously proposed by Murthy, Kasif and Salzberg (1994), this was still in the context of a greedy optimization approach (i.e., using as many variables as possible in the first split and then proceeding with many successive splits) which could lead to extreme overfitting. OCT prevents this shortcoming by aiming to achieve a best global fit with fewer variables in the first split. For this study, OCT enabled us to leverage the combined predictive power of all the phenotypic and neuroimaging-based information and find the features that lead to the most generalizable classification performance.

To classify ASD and SZ patients, we used five distinct sets of features comprising phenotypic and clinical assessment scores, the E/I ratio (as indexed by the H exponent) of multiple brain areas, or both. For each model, we also assessed the relative importance of each feature included in that respective model. To quantify the E/I imbalance based on resting state functional magnetic resonance (rsfMRI) timeseries, one approach is to compute the Hurst exponent (H) of predefined functional brain areas. The H exponent has been refined as a reliable computational approximation of synaptic E/I based on extensive physiological and *in silico* studies, and conventionally takes values between 0 and 1 (e.g., Trakoshis et al., 2020). However, Trakoshis et al. (2020) advise that the upper bound should not be limited to 1 in order to avoid ceiling effects and thus risk not capturing individual differences. We have opted for the same approach in the current paper, as further detailed in section 2.4. below. For the clinical features we focused on core symptoms assessments: ADOS – measuring social and communication impairments, and PANSS – measuring positive (e.g. delusions, hallucinations) and negative (e.g., social withdrawal) symptoms and general psychopathology (e.g., attention deficits), the IQ estimate, and two social cognitive measures: EQ – measuring empathy, and BVAQ – measuring alexithymia, both of which have been shown to be impaired in ASD and SZ (van't Wout et al., 2007; Warrier et al., 2018; Kinnaird, Stewart & Tchanturia, 2019).

2. Methods

2.1. Participants

Participants were recruited via the Olin Neuropsychiatry Research Center (ONRC) and the Yale University School of Medicine and underwent resting state fMRI scanning for the current study. After discarding datasets displaying head motion > 10 mm, our dataset contained 58 TD, 39 ASD, and 41 SZ. Of these, some were subsequently excluded to incomplete phenotypic assessment information, thus resulting in the following final samples: 55 TD, 30 ASD, and 39 SZ. As this dataset has been previously used by Hyatt et al. (2020, 2021) and Rabany et al. (2019), the exclusion criteria were the same, namely: intellectual disability (i.e., estimated IQ < 80), neurological disorder (e.g. epilepsy), current drug use as indicated by pre-scanning interview and urine test, incompatibility with MRI safety measures (e.g., metal implants), and a history of psychiatric diagnoses in HC.

2.2. Clinical and phenotypical assessment

The clinical assessment targeted the severity of psychotic and autistic symptoms, while phenotypic measures quantified empathy, alexithymia and IQ. The severity of psychotic symptoms was assessed using the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987) in both the ASD and SZ groups. The PANSS scores can be interpreted along three subscales: positive symptoms, reflecting the severity of hallucinations and delusions; negative symptoms, reflecting the severity of blunted affect and anhedonia, and a general subscale quantifying other psychopathology such as poor attention and lack of insight. The ADOS, module 4 (Lord et al., 2000) was administered to all participants and the total score was used in this study, to confirm/rule out ASD diagnosis. Intelligence Quotient (IQ) was calculated for the entire sample using the Vocabulary and Block Design subtests of the *Wechsler Scale of Adult Intelligence-III* (WAIS-III; Wechsler, 1997; Sattler and Ryan, 1999). Additionally, all participants were asked to complete the Empathizing Quotient (EQ; Wakabayashi, Baron-Cohen & Wheelwright, 2006) which measures general empathy including both the affective and the cognitive empathy components, and the Bermond–Vorst Alexithymia Questionnaire (BVAQ; Vorst & Bermond, 2001). BVAQ subscores are computed along five distinct dimensions: “verbalizing” reflects one’s propensity to talk about one’s feelings; “identifying” reflect the extent to which one is able to accurately define one’s emotional states; “analyzing” quantifies the extent to which one seeks to understand the reason for one’s emotions; “fantasizing” quantifies one’s tendency to day-dream, and “emotionalizing” reflects the extent to which a person is emotionally aroused by emoting inducing events. Means and standard deviations, as well as group comparison tests of the above-mentioned instruments are given in Table 1 below. The Structured clinical interview for DSM-IV-TR axis I disorders (SCID; First et al., 2002) was additionally used to confirm SZ diagnosis and the absence of any Axis I diagnoses in HC.

Table 1. Means and standard deviations (in parentheses) of demographics, phenotypic and clinical instrument scores for all three groups. est. IQ = estimated Intelligence Quotient; EQ = Empathy Quotient; ADOS = Autism Diagnostic Observation Schedule module 4; BVAQ = Bermond–Vorst Alexithymia Questionnaire; FD = framewise displacement. Group statistics are shown in the last four columns. Pairwise comparisons were performed using Welch two-samples t test. Both uncorrected (i.e., p) and false discovery rate corrected (i.e., p_{FDR}) p values are shown.

	TD	**ASD**	**SZ**	**ASD v. SZ v. TD**	**ASD > TD**	**ASD > SZ**	**TD > SZ**
Females/Males	29/26	5/25	8/31	$\chi^2(2) = 15.8$, <.000			
				F(2, 121), *p*	**t(df), *p*, p_{FDR}**	**t(df), *p*, p_{FDR}**	**t(df), *p*, p_{FDR}**
FD	0.08 (0.03)	0.09 (0.04)	0.11 (0.1)	**5.23, .007**	2.52 (66.6), .014, .021	-0.772 (69.6), .443, .443	-2.71 (54.3), .009, .021
Age	24 (3.73)	22 (3.74)	26 (3.58)		-1.98 (79.8), .051, .051	-4.06 (76.9), <.000, <.000	-2.46 (88), .016, .024
est. IQ	112.26 (14.62)	109.1 (15.21)	99.41 (13.34)	**9.366, <.000**	-1.33 (82.8), .186, .186	**3.26 (77), .002, .003**	**4.97 (90.5), <.000, <.000**
EQ	49 (10.28)	33.57 (11.1)	39.8 (12.26)	**22.37, <.000**	**-6.73 (72.2), <.000, <.000**	**-2.50 (73.9), .015, .015**	**3.65 (72.2), <.000, <.000**
ADOS	1.87 (1.45)	10.1 (2.61)	8.41 (5.26)	**78.26, < .000**	**17.5 (52.2), <.000, <.000**	1.67 (58.8), 1.72, 1.72	**-7.62 (44.2), <.000, <.000**
BVAQ Verbalizing	18.55 (5.51)	22.2 (4.54)	22 (5.03)	**7.02, .001**	**3.07 (81.4), .003, .005**	-0.0571 (75.1), .955, .955	**-3.22 (90.4), .002, .005**
BVAQ Fantasizing	19.69 (5.12)	17.9 (6.03)	21.36 (5.48)	**4, .02**	-1.17 (73.5), .247, .247	**-2.72 (75.2), .008, .024**	-1.88 (82.8), .064, .096
BVAQ Identifying	15.53 (4.82)	18.57 (6.4)	20.41 (4.96)	**10.43, <.000**	**2.59 (63), .012, .018**	-1.29 (68.2), .203, .203	**-4.79 (85.2), <.000, <.000**
BVAQ Emotionalizing	22.38 (3.74)	21.83 (3.51)	22.56 (4.2)	0.6, .55	-0.810 (81.4), .421, .632	-1.07 (74), .289, .632	-0.389 (77.9), .698, .698
BVAQ Analyzing	17.91 (4.43)	18.47 (4.27)	19.9 (4.5)	2.81, 0.6	0.576 (81.9), .566, .566	-1.62 (76), .11, .165	**-2.27 (85.4), .026, .078**
PANSS Positive		12.1 (2.86)	15.36 (4.86)			**-3.88 (66.96), <.000, 0**	

PANSS Negative	15.57 (4.7)	19.26 (6.2)	**-2.644 (72.75), .01, 1**
PANSS General	26.7 (5.62)	31.59 (6.98)	**-3.716 (73), <.000, 0**

2.3. Imaging data acquisition and preprocessing

Resting state fMRI scans lasted 7.5 min and were collected using a Siemens Skyra 3 T scanner at the ONRC. Participants lay still, with eyes open, while fixating a centrally presented cross. Blood oxygenation level dependent (BOLD) signal was obtained with T2*-weighted echo planar imaging (EPI) sequence: TR/TE = 475/30 msec, flip-angle = 60 deg, 48 slices, multiband (8), interleaved slice order, 3 mm3 voxels.

Neuroimaging data were preprocessed using SPM8 (www.fil.ion.ucl.ac.uk/spm/software/spm8/). Each dataset was realigned to the first T2* image using the INRIAlign toolbox (https://www-sop.inria.fr/epidaure/Collaborations/IRMf/INRIAlign.html), coregistered to their corresponding high signal-to-noise single-band reference image (sbREF; (Glasser et al., 2013)), spatially normalized to the Montreal Neurological Institute (MNI) standard template (Friston et al., 1995), and spatially smoothed (6 mm3). Finally, framewise displacement (FD) motion parameters were computed according to the FSL library algorithm (Jenkinson et al., 2012) and used the mean FD value for each run as a covariate in group analyses.

2.4. Data analysis

We ran a fully automated independent component analysis (ICA) on the preprocessed fMRI data, using the Group ICA for fMRI Toolbox (GIFT v4.0c; https://trendscenter.org/software/gift/; Calhoun et al., 2001) to define functional brain regions. The 53 replicable independent component (IC) templates from the NeuroMark pipeline (Du et al., 2020) were used to estimate participant-specific spatially-independent components using a spatially-constrained ICA algorithm (Du et al., 2018). A complete list of the NeuroMark IC templates, arranged into seven functional domains, and peak MNI coordinates for each IC template are given in Supplement Table 1 and illustrated in Supplement Figure 1. After detrending and despiking using 3dDespike (AFNI, 1995), the resulting 53 IC timecourses for each participant were used to compute the Hurst exponent.

The Hurst exponent (H) was estimated for each component and participant using the nonfractal MATLAB toolbox (https://github.com/wonsang/nonfractal; Wonsang et al., 2012). Specifically, we used the function *bfn_mfin_ml.m* with the “filter” argument set

to “haar” and the “ub” and “lb” arguments set to [1.5,10] and [-0.5,0], respectively, as previously recommended by Trakoshis et al. (2020).

Finally, we classified the ASD and SZ groups using the OCT procedure (Bertsimas & Dunn, 2017), as implemented in the Interpretable Artificial Intelligence (IAI) toolbox (https://www.interpretable.ai/) and accessible through the R 5.2 interface (R Core Team, 2018). Five models were used for the OCT classification of SZ and ASD, containing the following features: (a) PANSS 3 factor scores and ADOS total scores; (b) PANSS, ADOS, EQ, BVAQ and IQ scores; (c) the H only values of the 53 brain areas; (d) the 53 H values plus the PANSS and ADOS scores, and (e) the 53 H values plus the PANSS, ADOS, EQ, BVAQ and IQ scores. Of each of the two clinical groups, 80% of the participants were allocated to the training group, while the rest 20% were used to test the generalizability of the classification.

Other statistical analyses, performed with R 5.2., included ANCOVA with Age, Sex, IQ and FD as covariates, and two-sided two-sample Welch t tests, which were used to further compare groups to each-other.

3. Results

3.1. Group differences in demographic, clinical and phenotypic assessment

The demographics such as age and sex were significantly different between groups (see Table 1). Therefore, these variables (together with average FD, as described above) were included as covariates in all further group-comparison analyses. With respect to the clinical assessment, the two patient groups did not significantly differ in their social and communication skills, as indicated by the ADOS scores, but the PANSS scores on all three domains (i.e., positive and negative symptom severity and general psychopathology) were significantly elevated in SZ compared to ASD. Finally, with respect to phenotypic measures, the alexithymia scores indicating fantasizing were significantly decreased in ASD compared to SZ. Estimated IQ also significantly differed between groups, and was therefore used as a covariate alongside FD, age and sex. Empathy was significantly decreased in ASD compared to both SZ and TD, and in SZ compared to TD.

3.2. Group differences in H

Overall group comparisons of the H values from each component are given in Table 2 below. Due to concerns that Age, Sex, IQ and FD might act as confounded variables when assessing group differences in H, we ran ANCOVA analyses using Age, Sex, IQ and FD as covariates (see Table 2). Two-sided two-sample Welch t tests were used to further compare groups to each-other (see Table 3).

The areas most sensitive to overall group differences, after controlling for Age, Sex, IQ and FD belonged to the Sensorimotor (i.e., left and light postcentral gyrus and paracentral lobule – IC no. 9, 11, and 10 respectively), Visual (i.e., calcarine gyrus, middle occipital gyrus, cuneus, right middle occipital gyrus, fusiform gyrus, inferior

occipital gyrus and lingual gyrus – IC no. 17, 18, 20, 21, 22, 23 and 24) and the Cerebellar domains – IC no. 50 - 53). In addition, the supplementary motor area and inferior frontal gyrus (IC no. 34, 35), associated with the Cognitive Control domain, also yielded significant group differences (see Table 2).

Table 2. Group differences in Hurst exponent (H) per component, after controlling for Age, Sex, IQ and FD. The effect size was calculated using ηp2.

		H mean (sd)						
IC no.		TD	ASD	SZ	F(2, 117)	*p*	p_{fdr}	ηp2
	SUBCORTICAL DOMAIN							
1	Caudate 1	1.05 (0.15)	0.99 (0.12)	0.96 (0.15)	2.55	.08	0.18	.04
2	Subthalamus/hypothalamus	0.8 (0.14)	0.84 (0.12)	0.75 (0.12)	2.38	.1	0.2	.04
3	Putamen	0.98 (0.14)	0.98 (0.12)	0.91 (0.16)	1	.36	0.56	.02
4	Caudate 2	0.98 (0.13)	0.97 (0.13)	0.95 (0.14)	0.01	.6	0.78	.01
5	Thalamus	0.94 (0.18)	0.93 (0.15)	0.84 (0.16)	2.8	.07	0.18	.05
	AUDITORY DOMAIN							
6	Superior temporal gyrus	0.99 (0.16)	0.99 (0.14)	1.01 (0.16)	0.99	.37	0.56	.02
7	Middle temporal gyrus	0.91 (0.14)	0.95 (0.14)	0.92 (0.12)	0.33	.72	0.86	.01
	SENSORIMOTOR DOMAIN							
8	Postcentral gyrus 1	1.37 (0.17)	1.32 (0.13)	1.23 (0.17)	2.55	.08	0.18	.04
9	**Left postcentral gyrus**	**1.29 (0.18)**	**1.24 (0.15)**	**1.15 (0.17)**	**6.18**	**.003**	**0.02**	**.1**

10	**Paracentral lobule 1**	**1.18 (0.17)**	**1.15 (0.19)**	**1.05 (0.18)**	**6.25**	**.003**	**0.02**	**.1**
11	**Right postcentral gyrus**	**1.3 (0.19)**	**1.28 (0.16)**	**1.15 (0.19)**	**4.57**	**.01**	**0.05**	**.7**
12	Superior parietal lobule	1.2 (0.15)	1.19 (0.18)	1.09 (0.17)	2.8	.07	0.18	.05
13	Paracentral lobule 2	1.11 (0.15)	1.09 (0.18)	0.99 (0.15)	3	.05	0.16	.05
14	Precentral gyrus	1.22 (0.16)	1.23 (0.12)	1.12 (0.14)	2.78	.07	0.18	.05
15	Superior parietal lobule	1.26 (0.13)	1.24 (0.13)	1.17 (0.18)	0.5	.61	0.78	.008
16	Postcentral gyrus 2	1.29 (0.13)	1.26 (0.14)	1.17 (0.19)	1.93	.15	0.28	.03
	VISUAL DOMAIN							
17	**Calcarine gyrus**	**1.31 (0.16)**	**1.22 (0.16)**	**1.15 (0.15)**	**6.23**	**.003**	**0.02**	**.1**
18	**Middle occipital gyrus**	**1.19 (0.16)**	**1.13 (0.16)**	**1.01 (0.17)**	**5.76**	**.004**	**0.03**	**.09**
19	**Middle temporal gyrus**	**1.2 (0.16)**	**1.15 (0.13)**	**1.03 (0.14)**	**6.33**	**.003**	**0.02**	**.1**
20	Cuneus	1.28 (0.17)	1.21 (0.17)	1.12 (0.16)	3.79	.03	0.13	.06
21	**Right middle occipital gyrus**	**1.18 (0.14)**	**1.12 (0.14)**	**1.01 (0.14)**	**8.46**	**<.000**	**0**	**.13**
22	Fusiform gyrus	1.03 (0.14)	1.01 (0.1)	0.93 (0.13)	3	.05	0.16	.05
23	**Inferior occipital gyrus**	**1.11 (0.14)**	**1.06 (0.12)**	**0.94 (0.15)**	**8.25**	**<.000**	**0**	**.12**
24	**Lingual gyrus**	**1.25 (0.18)**	**1.17 (0.17)**	**1.05 (0.14)**	**7.9**	**<.000**	**0**	**.12**

25	Middle temporal gyrus	1.04 (0.15)	1.02 (0.12)	0.93 (0.14)	2.69	.7	0.86	.04

COGNITIVE CONTROL DOMAIN

26	Inferior parietal lobule 1	1.39 (0.1)	1.37 (0.12)	1.35 (0.15)	0.29	.75	0.86	.01
27	Insula	1.04 (0.12)	1.03 (0.12)	0.97 (0.14)	0.7	.5	0.73	.01
28	Superior medial frontal gyrus	1.16 (0.14)	1.14 (0.13)	1.13 (0.16)	0.24	.78	0.86	.004
29	Inferior frontal gyrus	1.14 (0.1)	1.11 (0.14)	1.07 (0.15)	0.18	.83	0.88	.003
30	Right inferior frontal gyrus	1.21 (0.12)	1.18 (0.13)	1.14 (0.12)	0.63	.54	0.73	.01
31	Middle frontal gyrus	1.2 (0.12)	1.18 (0.12)	1.1 (0.16)	0.48	.62	0.78	.01
32	Inferior parietal lobule 2	1.37 (0.11)	1.33 (0.1)	1.29 (0.15)	1.63	.2	0.33	.03
33	Left inferior parietal lobule	1.31 (0.14)	1.29 (0.13)	1.23 (0.14)	0.08	.93	0.93	.001
34	**Supplementary motor area**	**1.26 (0.11)**	**1.2 (0.14)**	**1.22 (0.12)**	**5.1**	**.008**	**0.05**	**.08**
35	Superior frontal gyrus	1.22 (0.1)	1.19 (0.1)	1.2 (0.14)	1.4	.25	0.4	.02
36	Middle frontal gyrus 1	1.11 (0.13)	1.1 (0.13)	1.04 (0.17)	.21	.81	0.88	.004
37	Hippocampus 1	0.77 (0.14)	0.79 (0.11)	0.76 (0.12)	.13	.88	0.91	.002
38	Left inferior parietal lobule	1.29 (0.11)	1.23 (0.13)	1.22 (0.14)	2.62	.08	0.18	.04
39	Middle cingulate cortex	1 (0.13)	0.98 (0.14)	0.9 (0.15)	2.37	.1	0.2	.04

40	Inferior frontal gyrus	1.12 (0.12)	1.06 (0.12)	1.01 (0.15)	3.37	.04	0.15	.06
41	Middle frontal gyrus 2	1.06 (0.14)	1.06 (0.14)	0.96 (0.14)	1.73	.18	0.32	.03
42	Hippocampus 2	0.83 (0.12)	0.87 (0.14)	0.81 (0.18)	0.29	.75	0.86	.01
	DEFAULT MODE DOMAIN							
43	Precuneus 1	1.25 (0.1)	1.2 (0.11)	1.15 (0.15)	1.7	.19	0.32	.03
44	Precuneus 2	1.16 (0.12)	1.13 (0.08)	1.09 (0.12)	2	.14	0.27	.03
45	Anterior cingulate cortex 1	1.1 (0.11)	1.09 (0.1)	1.02 (0.15)	0.62	.54	0.73	.01
46	Posterior cingulate cortex 1	0.96 (0.12)	0.98 (0.13)	0.97 (0.16)	0.26	.77	0.86	.004
47	Anterior cingulate cortex 2	1.04 (0.14)	1.03 (0.13)	1.01 (0.18)	0.12	.89	0.91	.002
48	Precuneus 3	1.18 (0.14)	1.17 (0.11)	1.08 (0.13)	1.8	.17	0.31	.03
49	Posterior cingulate cortex 2	1.36 (0.11)	1.37 (0.12)	1.31 (0.14)	0.63	.54	0.73	.01
	CEREBELLAR DOMAIN							
50	Cerebellum 1	0.98 (0.17)	0.95 (0.16)	0.87 (0.17)	2.71	.07	0.18	.04
51	Cerebellum 2	1.1 (0.14)	1.04 (0.16)	0.98 (0.15)	3.49	.03	0.13	.06
52	Cerebellum 3	0.89 (0.16)	0.91 (0.17)	0.8 (0.17)	3.25	.04	0.15	.05

53	Cerebellum 4	1.02 (0.14)	0.98 (0.18)	0.91 (0.14)	2.99	.05	0.16	.05

Direct group comparisons revealed that ASD exhibited significantly smaller H values (i.e., reduced E/I ratio) compared to TD in areas of the Visual network (i.e., calcarine gyrus, right middle occipital gyrus, and lingual gyrus), Cognitive Control network (i.e., supplementary motor area, left inferior parietal lobule, and inferior frontal gyrus), and Default Mode network (i.e., precuneus). Significantly smaller H values (i.e., reduced E/I ratio) were also found in SZ compared to TD in areas of the Subcortical domain (i.e., caudate, putamen, and thalamus), the entire Sensorimotor domain (i.e., precentral and postcentral gyrus, paracentral lobule, superior parietal lobule, and middle occipital gyrus), the entire Visual domain (i.e., calcarine, fusiform, inferior occipital and lingual gyrus, cuneus, middle occipital and temporal gyrus), some areas of the Cognitive Control domain (i.e., insula, inferior and middle frontal gyrus, inferior parietal lobule, and middle cingulate cortex), some areas of the Default Mode domain (i.e., precuneus and anterior cingulate cortex), and the entire Cerebellar domain. Finally, significantly larger H values (i.e., reduced E/I ratio) were found in ASD compared to SZ in areas of the Subcortical network (putamen and thalamus), the entire Sensorimotor domain (i.e., precentral and postcentral gyrus, paracentral lobule, superior parietal lobule, and middle occipital gyrus), the Visual domain (fusiform, inferior occipital and lingual gyrus, cuneus, middle occipital and temporal gyrus), some areas of the Cognitive Control domain (middle and inferior frontal gyrus, and middle cingulate gyrus), some areas of the Default Mode domain (precuneus and anterior cingulate cortex), and two areas of the Cerebellar domain. The statistical values for all the above-listed findings are listed in Table 3.

Table 3. Group differences in Hurst exponent (H) per component, as calculated using two-sided two-sample Welch t-tests. Both uncorrected and false discovery rate corrected p values are provided, as well as Hedge's g effect sizes. Degrees of freedom (df) are provided in parentheses, next to the t values.

		TD > ASD				**TD > SZ**				**ASD > SZ**			
IC no.		t(df)	p_{unc}	p_{fdr}	g	t(df)	$p_{unc.}$	p_{fdr}	g	t(df)	$p_{unc.}$	p_{fdr}	g
SUBCORTICAL DOMAIN													
1	**Caudate 1**	1.88 (69.83)	.06	0.35	0.4	**2.82 (83.2)**	**.006**	**0.01**	**0.58**	0.92 (66.32)	.36	0.44	0.22
2	Subthalamus/hypothalamus	-1.33 (65.54)	.19	0.5	-0.29	1.83 (87.67)	.07	0.09	0.37	2.96 (62)	.004	0.02	0.71
3	**Putamen**	-0.2 (65.16)	.85	0.88	-0.04	**2**	**.05**	**0.07**	**0.43**	**2.03 (66.95)**	**.05**	**0.09**	**0.47**

						(74.75)							
4	Caudate 2	0.43 (59.17)	.67	0.84	0.1	1 (76.95)	.33	0.38	0.21	0.47 (64.85)	.64	0.71	0.11
5	**Thalamus**	0.39 (70.27)	.7	0.84	0.08	**2.76 (88.28)**	**.007**	**0.01**	**0.56**	**2.24 (64.1)**	**.03**	**0.06**	**0.53**
	AUDITORY DOMAIN												
6	Superior temporal gyrus	0.18 (65.02)	.86	0.88	0.04	-0.36 (80.57)	.72	0.73	-0.08	-0.49 (65.73)	.63	0.71	-0.1
7	Middle temporal gyrus	-1.27 (61.56)	.21	0.53	-0.28	-0.54 (88.77)	.59	0.64	-0.11	0.81 (58.29)	.42	0.49	0.2
	SENSORIMOTOR DOMAIN												
8	**Postcentral gyrus 1**	1.5 (73.27)	.14	0.47	0.31	**3.83 (80.59)**	**<.000**	**0**	**0.8**	**2.4 (67)**	**.02**	**0.05**	**0.56**
9	**Left postcentral gyrus**	1.42 (66.98)	.16	0.47	0.31	**3.98 (83.22)**	**<.000**	**0**	**0.82**	**2.36 (65.42)**	**.02**	**0.05**	**0.56**
10	**Paracentral lobule 1**	0.76 (53.5)	.45	0.77	0.18	**3.73 (78.64)**	**<.000**	**0**	**0.78**	**2.32 (60.4)**	**.02**	**0.05**	**0.56**
11	**Right postcentral gyrus**	0.54 (68.6)	.59	0.79	0.12	**3.91 (82.34)**	**<.000**	**0**	**0.81**	**3.17 (66.24)**	**.003**	**0.02**	**0.75**
12	**Superior parietal lobule**	0.26 (53.33)	.8	0.87	0.06	**3.22 (76.31)**	**.002**	**0**	**0.68**	**2.39 (61.74)**	**.02**	**0.05**	**0.58**
13	**Paracentral lobule 2**	0.5 (53.33)	.6	0.79	0.12	**3.87 (84.4)**	**<.000**	**0**	**0.8**	**2.57 (56.05)**	**.02**	**0.05**	**0.63**
14	**Precentral gyrus**	-0.35 (74.22)	.73	0.84	-0.73	**3.15 (86.1)**	**.002**	**0**	**0.65**	**3.45 (66.6)**	**<.000**	**0**	**0.81**
15	**Superior parietal lobule**	0.58 (59.58)	.56	0.79	0.13	**2.77 (65.3)**	**.007**	**0.01**	**0.61**	**2.04 (66.83)**	**.05**	**0.09**	**0.47**
16	**Postcentral gyrus 2**	0.97 (57.13)	.34	0.77	0.22	**3.39 (63.43)**	**.001**	**0**	**0.75**	**2.26 (66.88)**	**.03**	**0.06**	**0.52**
	VISUAL DOMAIN												

17	**Calcarine gyrus**	**2.64 (61.68)**	**.01**	**0.35**	**0.59**	**5 (87.27)**	**<.000**	**0**	**1**	1.77 (59.89)	.08	0.14	0.43
18	**Middle occipital gyrus**	1.47 (60.96)	.15	0.47	0.33	**4.93 (78.86)**	**<.000**	**0**	**1**	**3 (64.83)**	**.004**	**0.02**	**0.71**
19	**Middle temporal gyrus**	1.59 (68.33)	.12	0.47	0.34	**5.66 (87.42)**	**<.000**	**0**	**1**	**3.7 (63.74)**	**<.000**	**0**	**0.88**
20	**Cuneus**	1.89 (58.58)	.06	0.35	0.43	**4.88 (85.75)**	**<.000**	**0**	**1**	**2.27 (59.1)**	**.03**	**0.06**	**0.55**
21	**Right middle occipital gyrus**	**1.98 (57.13)**	**.05**	**0.35**	**0.45**	**6 (80.41)**	**<.000**	**0**	**1.26**	**3.24 (61.88)**	**.002**	**0.02**	**0.78**
22	**Fusiform gyrus**	0.78 (75.36)	.44	0.77	0.16	**3.58 (85.74)**	**<.000**	**0**	**0.73**	**2.83 (66.86)**	**.006**	**0.03**	**0.66**
23	**Inferior occipital gyrus**	1.73 (65.92)	.09	0.47	0.38	**5.67 (79.31)**	**<.000**	**0**	**1.19**	**3.66 (66.36)**	**<.000**	**0**	**0.86**
24	**Lingual gyrus**	**2.04 (62.82)**	**.05**	**0.35**	**0.45**	**6.1 (91)**	**<.000**	**0**	**1.21**	**3.18 (55.96)**	**.002**	**0.02**	**0.78**
25	**Middle temporal gyrus**	0.8 (71.27)	.43	0.77	0.17	**3.67 (85.35)**	**<.000**	**0**	**0.75**	**2.76 (66.05)**	**.007**	**0.03**	**0.65**

COGNITIVE CONTROL DOMAIN

26	Inferior parietal lobule 1	0.7 (49.47)	.49	0.79	0.17	1.16 (61.1)	.25	0.3	0.26	0.41 (66.59)	.69	0.71	0.1
27	**Insula**	0.53 (61.49)	.6	0.79	0.12	**2.45 (74.89)**	**.02**	**0.03**	**0.52**	1.73 (66.36)	.09	0.14	0.41
28	Superior medial frontal gyrus	0.54 (61.29)	.6	0.79	0.12	0.94 (72.17)	.35	0.39	0.2	0.39 (66.82)	.7	0.71	0.1
29	**Inferior frontal gyrus**	0.88 (47)	.39	0.77	0.22	**2.26 (61.9)**	**.03**	**0.04**	**0.5**	1.1 (65.25)	.28	0.37	0.26
30	**Right inferior frontal gyrus**	1.21 (55.76)	.23	0.55	0.28	**2.59 (80.7)**	**.01**	**0.01**	**0.54**	1 (60.74)	.32	0.41	0.24
31	**Middle frontal gyrus**	0.83 (60.78)	.41	0.77	0.19	**3.35 (66.88)**	**.001**	**0**	**0.73**	**2.37 (66.89)**	**.02**	**0.05**	**0.55**
32	**Inferior parietal lobule 2**	1.46 (61.83)	.15	0.47	0.32	**2.84 (65.08)**	**.006**	**0.01**	**0.62**	1.47 (66.39)	.15	0.2	0.34
33	**Left inferior parietal lobule**	0.81 (64.79)	.42	0.77	0.18	**2.75 (81.92)**	**.008**	**0.01**	**0.57**	1.72 (65.12)	.09	0.14	0.41

34	**Supplementary motor area**	**2.26 (50.17)**	**.03**	**0.35**	**0.54**	1.68 (79.22)	.1	0.12	0.36	-0.81 (57.13)	.42	0.49	-0.2
35	Superior frontal gyrus	1.6 (58.8)	.11	0.47	0.36	1 (65.29)	.3	0.35	0.22	-0.39 (66.92)	.7	0.71	-0.0
36	**Middle frontal gyrus 1**	0.28 (60.63)	.78	0.86	0.06	**2.2 (66.52)**	**.03**	**0.04**	**0.48**	1.77 (66.86)	.08	0.14	0.41
37	Hippocampus 1	-0.55 (69)	.58	0.79	-0.12	0.43 (86.83)	.67	0.7	0.09	0.94 (64.45)	.35	0.44	0.22
38	**Left inferior parietal lobule**	**2.3 (53.29)**	**.03**	**0.35**	**0.54**	**2.73 (69.5)**	**.008**	**0.01**	**0.6**	0.32 (65.23)	.75	0.75	0.08
39	**Middle cingulate cortex**	0.57 (53.56)	.57	0.79	0.13	**3.41 (73.54)**	**.001**	**0**	**0.73**	**2.31 (63.48)**	**.02**	**0.05**	**0.55**
40	**Inferior frontal gyrus**	**2.1 (59.5)**	**.05**	**0.35**	**0.47**	**3.8 (68.83)**	**<.000**	**0**	**0.82**	1.67 (66.97)	.1	0.15	0.39
41	**Middle frontal gyrus 2**	0.16 (61.37)	.88	0.88	0.035	**3.27 (81.17)**	**.002**	**0**	**0.68**	**2.73 (63.92)**	**.008**	**0.04**	**0.65**
42	Hippocampus 2	-1.4 (54.7)	.17	0.47	-0.33	0.43 (61.34)	.67	0.7	0.1	1.46 (66.9)	.15	0.2	0.34

DEFAULT MODE DOMAIN

43	**Precuneus 1**	**1.97 (56.16)**	**.05**	**0.35**	**0.45**	**3.39 (62.92)**	**.001**	**0**	**0.75**	1.45 (66.93)	.15	0.2	0.34
44	**Precuneus 2**	1.48 (79.1)	.14	0.47	0.3	**2.86 (79.23)**	**.005**	**0.01**	**0.6**	**1.66 (65.1)**	**.01**	**0.04**	**0.38**
45	**Anterior cingulate cortex 1**	0.4 (64.7)	.69	0.84	.09	**2.85 (69.47)**	**.006**	**0.01**	**0.62**	**2.32 (66.74)**	**.02**	**0.05**	**0.54**
46	Posterior cingulate cortex 1	-0.72 (54.78)	.48	0.79	-0.17	-0.31 (67.64)	.76	0.76	-0.07	0.33 (66.4)	.7	0.71	0.08
47	Anterior cingulate cortex 2	0.15 (66.27)	.88	0.88	0.03	0.91 (70.6)	.37	0.41	0.2	0.73 (66.67)	.47	0.54	0.17
48	**Precuneus 3**	0.45 (73.98)	.66	0.84	.09	**3.49 (85.08)**	**<.000**	**0**	**0.72**	**3 (66.76)**	**.004**	**0.02**	**0.7**
49	Posterior cingulate cortex 2	-0.28 (54.68)	.78	0.86	-0.07	1.72 (68.94)	.09	0.11	0.37	1.71 (66)	.09	0.14	0.4

CEREBELLAR DOMAIN

50	**Cerebellum 1**	0.76 (60.9)	.45	0.77	0.18	**3.12 (80.53)**	**.003**	**0.01**	**0.65**	**2 (64)**	**.05**	**0.09**	**0.49**
51	**Cerebellum 2**	1.66 (52.69)	.1	0.47	0.39	**3.83 (77.64)**	**<.000 0**	**0**	**0.81**	1.56 (60.42)	.13	0.19	0.38
52	**Cerebellum 3**	-0.36 (57.13)	.72	0.84	-0.08	**2.72 (80.74)**	**.008**	**0.01**	**0.57**	**2.63 (61.67)**	**.01**	**0.04**	**0.63**
53	**Cerebellum 4**	0.94 (48)	.35	0.77	0.23	**3.75 (82.95)**	**<.000 0**	**0**	**0.76**	1.82 (52.26)	.07	0.13	0.45

3.3. Classification accuracy

Using the “misclassification” criterion, we obtained the following classification accuracy on the test set for each of the five models: (a) 50%; (b) 64%; (c) 78%; (d) 78%, and (e) 85%. A comparison of classification accuracy across models is illustrated in Figure 1 below, displaying the receiver operating characteristic (ROC) for each model.

The best learner classification trees produced during training for each of the five models are also illustrated in Supplement Figure 2. Each hyperplane shows the features that most contributed to achieving a maximal training performance, as well as their respective weights.

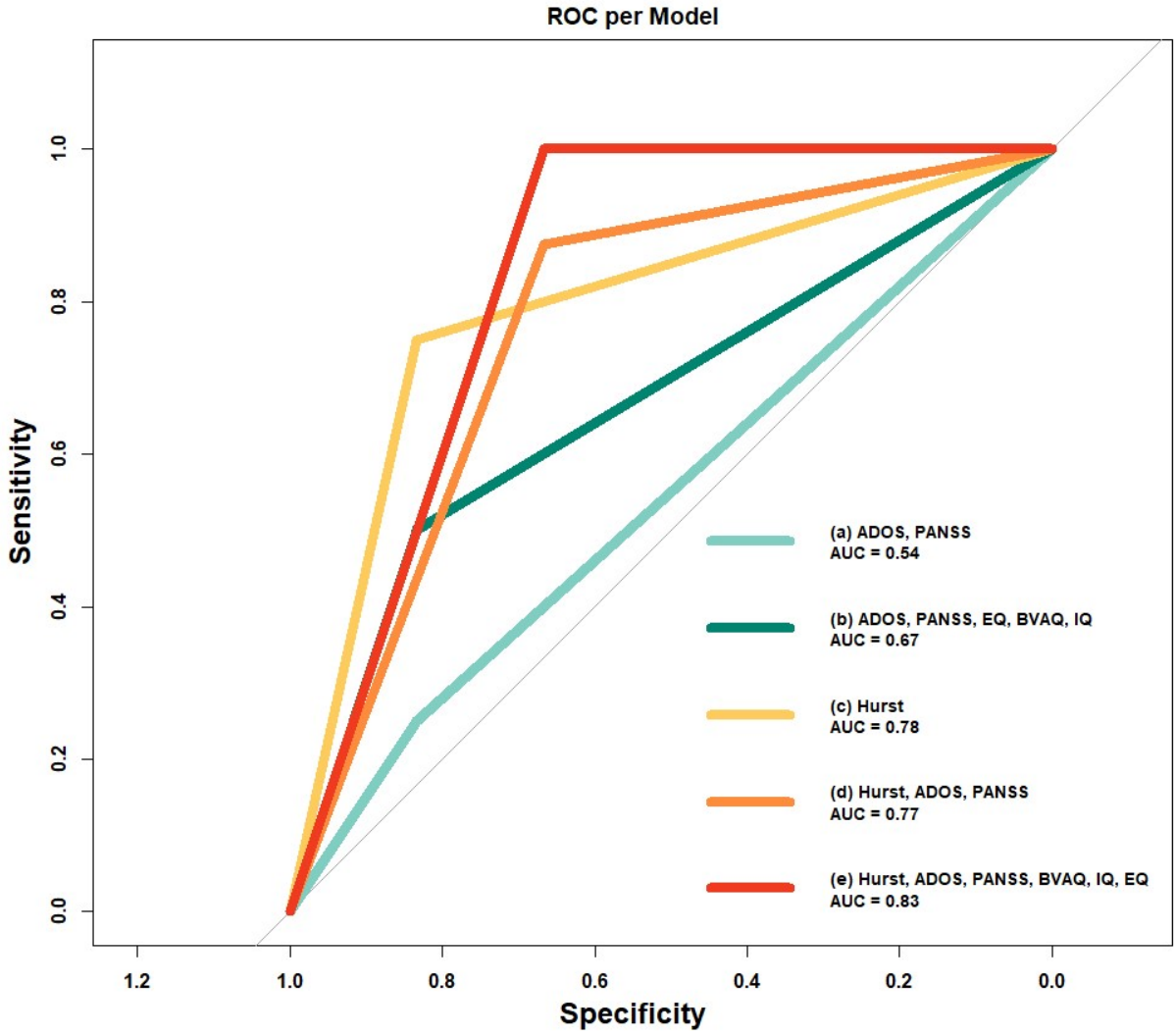

Figure 1. Receiver operating characteristic (ROC) for each model. Note that there is only one inflection point in each ROC curve because for all five classification instances the decision tree achieved its best performance by splitting only once. AUC: area under the curve; ADOS: Autism Diagnostic Observation Schedule; PANSS: Positive and Negative Syndrome Scale; EQ: Empathy Quotient; BVAQ: Bermond–Vorst Alexithymia Questionnaire; IQ: Intelligence Quotient. [could you do this for more thresholds to get smoother curves?

For each classification instance, we also inspected the contribution of each feature included in that respective model. More specifically, for each model, a ranking reflects each feature's overall contribution, or 'importance', to the classification accuracy during training. These feature importance values per model are summarized in Supplement Table 2. For an overview of feature importance across all five models, we refer the reader to Figure 2. In order to account for the different number of features included in each model we scaled the feature importance values in Figure 2 by multiplying raw importance values for each feature (i.e., in Supplement Table 2) by the number of features included in that model. We chose this approach to improve interpretability of importance values when all five models are visualized together in the stacked bar plots of Figure 2.

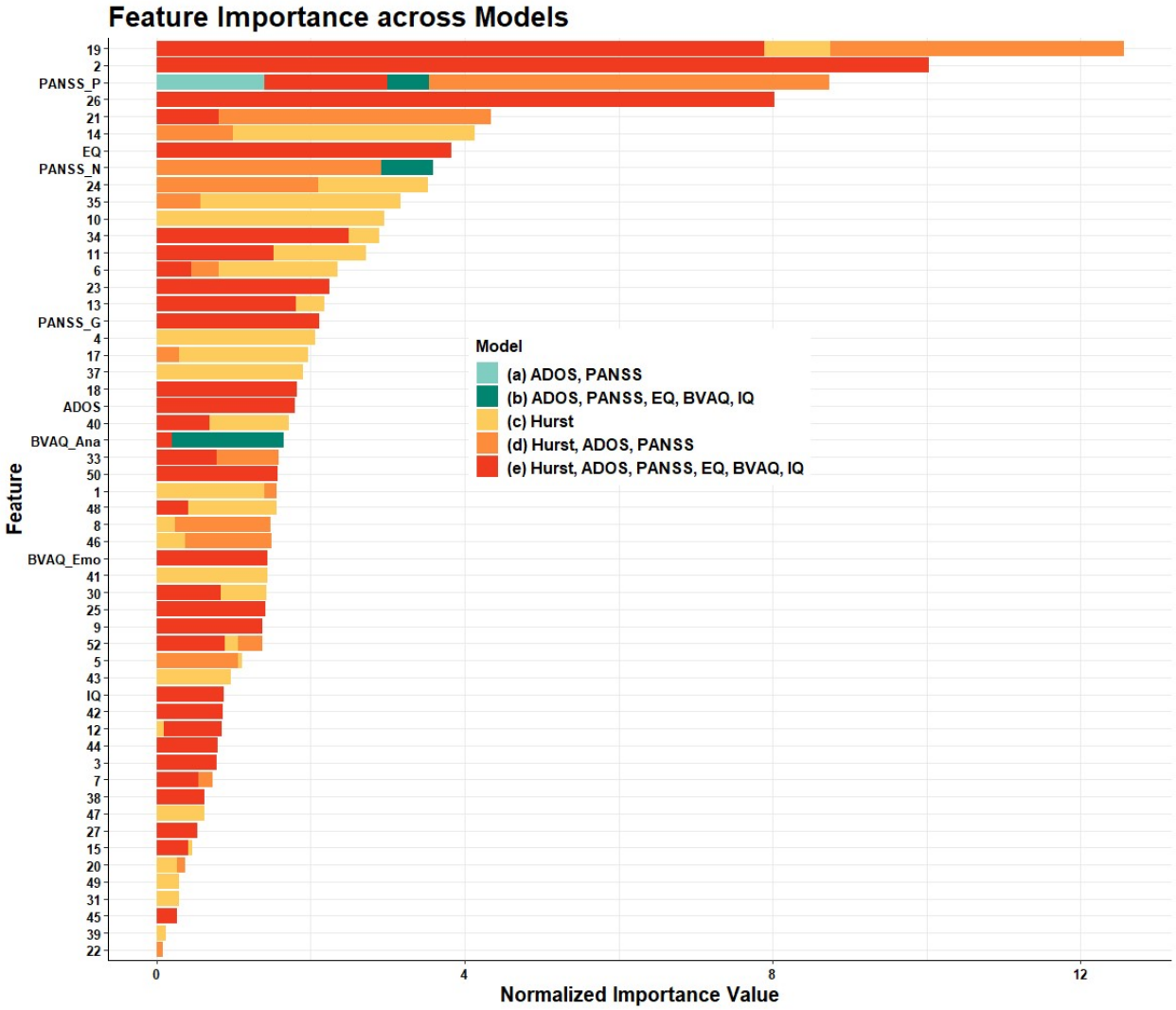

Figure 2. Scaled importance values of features across all five models. The x axis displays *scaled importance values* (i.e., obtained by multiplying raw importance values from Table 5 by the number of features in the model from which the values were obtained). On the y axis, feature names are displayed; the ones represented by numbers reflect the NeuroMark components from which the H values were computed (with corresponding labels given in Table 2). The other abbreviated feature labels on the y axis reflect the following phenotypic measures: PANSS_P = PANSS positive total; PANSS_N = PANSS negative total; PANSS_G = PANSS general total; BVAQ_Ana = BVAQ analyzing; BVAQ_Emo = BVAQ emotionalizing. Features which did not have a contribution > 0 are not shown in this plot, but are listed in Table 5 in the Supplement.

Discussion

The current paper assessed the feasibility of using the E/I ratio, as estimated by the H exponent, to distinguish between ASD and SZ. Of primary interest was to compare the classification accuracy of the two patient groups when different sets of clinical, phenotypic and imaging features were combined. We obtained an increasingly higher classification accuracy (see Figure 1) as we integrated the neuroimaging and phenotypic features, with the lowest performance being given by model (a), containing ADOS and PANSS scores only (50% testing accuracy) and the highest by model (e), containing H, ADOS, PANSS, BVAQ, EQ and IQ (85% testing accuracy). The second to last classification accuracy was achieved by model (b), containing PANSS, ADOS, BVAQ, EQ and IQ (64% testing accuracy). Next, a classification accuracy of 78% testing accuracy was achieved by both model (c), using H only, and by model (d), combining H with ADOS and PANSS. We note that comparing

classification models of different complexity (i.e., number of features) can be problematic, especially with conventional classification tree algorithms, due to their greedy nature (i.e., maximizing local, but not global performance). With OCT, however, this limitation is contained given the optimized trade-off between model accuracy and complexity by fitting the tree within one global step.

Despite the SZ group showing significantly higher PANSS symptom severity,BVAQ fantasizing and EQ scores as well as lower IQ scores compared to the ASD group, using clinical and phenotypic measures alone allowed only chance-level classification accuracy. Previous data has shown that PANSS and ADOS do not excel at distinguishing between ASD and SZ, but positive symptom severity does tend to perform better (Bastiaansen et al., 2011; Trevisan et al., 2020). Indeed, in our analysis PANSS-positive scores contributed more than PANSS-negative and general scores to the group discriminability across the models including these features (i.e., a, b, d and e). Additionally, a recent comprehensive meta-analysis comparing social cognitive performance between SZ and ASD revealed no significant group differences with respect to either theory of mind or emotion processing (Oliver et al., 2020). Finally, the BVAQ has recently been shown to have a sub-optimal factor structure upon which improvements can be made (Williams & Gotham, 2021). In our sample, only the fantasizing dimension of the BVAQ reflected significant group differences between the patient samples, while the analyzing and the emotionalizing dimensions contributed, albeit modestly, to the classification of the ASD and SZ samples. Empathy scores also contributed relatively highly to distinguishing between the two clinical groups, and similarly to previously reported results (Pepper et al., 2019), the ASD in our sample scored significantly lower on self-reported empathy compared to both SZ and TD.

The group differences that were statistically significant reflected decreased H values (i.e., increased E/I ratio) in ASD compared to TD, which is in line with previous findings (e.g., Lai et al., 2010; Trakoshis et al., 2020). Likewise, we found decreased H values in SZ compared to TD, indicating an increased E/I ratio, as was previously reported by Sokunbi et al. (2014). This finding supports previous results showing a disruption of inhibitory transmission in SZ (Purves-Tyson et al., 2021; Taylor & Tso, 2015). Finally, in our sample, ASD exhibited increased H (i.e., reduced E/I ratio) compared to SZ. To the best of our knowledge, no previous studies have directly compared the E/I ratio, as estimated via H, between these two clinical samples.

The brain networks whose E/I ratio contributed the most to patient classification were the Visual network (the middle temporal gyrus and the right middle occipital gyrus, corresponding to ICs 19 and 21; see Figure 2), the Subcortical network (thalamus/hypothalamus — IC 2), the Cognitive Control domain (the inferior parietal lobule — IC 26, and the Sensorimotor domain (the precentral gyrus, corresponding to IC 14). These areas have been previously documented to reflect effects (e.g., dysconnectivity, decrease in cognitive performance) linked to E/I imbalance in both ASD and SZ. In ASD, a study combining fMRI with Proton Magnetic Resonance Spectroscopy (MRS) showed a negative correlation between increased levels of excitatory metabolites and decreased functional connectivity between the dorsal anterior cingulate cortex and the insula, parietal and limbic regions, while this

association was positive in controls (Siegel-Ramsay et al., 2021). A link between altered GABAergic neurotransmission in visual areas and diminished binocular rivalry in ASD has also been established with the help of MRS (e.g., Robertson et al., 2016). Increased levels of glutamatergic metabolites in the dorsal anterior cingulate cortex have previously been linked to poorer social functioning in ASD compared to TD (Cochran et al., 2015). In SZ, Dempster et al. (2015) showed that increased levels of excitatory metabolites in the dorsal anterior cingulate cortex and the thalamus, as measured with MRS, correlated negatively with executive functioning performance. Increased levels of excitatory metabolites in these same areas have also been shown to positively correlate with increased severity of negative symptoms and decreased global functioning (Egerton et al., 2012). Additionally, disrupted thalamo-cortical inhibitory control has been shown to negatively impact sustained attention and thus lead to increased distractibility (John et al., 2018).

A limitation inherent to the top-down approach we took in this study is that diagnosis overlap between ASD and SZ cannot be fully characterized given the a priori reliance on clinically pre-defined group labels (Moreau et a., 2021), especially due to the heterogeneity of ASD and SZ (Port, Oberman & Roberts, 2019; Benkarim et al., 2022). For example, it has been shown that the relationship between the E/I balance and the cerebro-cerebellar functional connectivity in ASD is not uniform across samples (Hegarty et al., 2018), therefore a bottom-up approach in a larger sample would be extremely valuable to inform on different E/I-based clinical subtypes. In addition, due to the small number of female participants, we were unable to assess sex differences in the current study. This is a particularly relevant aspect to be addressed in future studies, as sex-specific effects of GABA concentrations have been reported in ASD, with higher concentrations found in males compared to females (Fung et al., 2021). At the same time, rsfMRI studies have reported decreased H values in ASD males, but not females, in prefrontal areas, compared to controls (Trakoshis et al., 2020). Finally, it has been shown that H values differentially change in various brain areas as a natural consequence of healthy aging. For example, Wink et al., (2016) showed that with age H increased in the hippocampus, while Dong et al. (2018) showed a decrease of H in the left insula, bilateral parahippocampal gyrus, left superior temporal gyrus, and bilateral superior temporal pole, concomitant with an increase of H in the left angular gyrus and the left superior parietal gyrus. At the same time, interactions between age and sex have not been definitively established (for a summary, see Dong et al., 2018). Future studies should investigate the long-term changes in H in patient groups, some of which, such as SZ, have been shown to exhibit premature brain aging (e.g., Chen et al., 2020). Finally, since the co-occurrence of epilepsy and ASD is considerable (> 4 %; Strasser et al., 2017), we chose to exclude participants with a history of epilepsy. However, since the E/I ratio has been shown to be perturbed in ASD populations who have epilepsy (Bozzi, Provenzano and Casarosa, 2017), it would be interesting, as a future direction, to see how the H can be used to capture even more specific E/I group differences in ASD with and without epilepsy.

In conclusion, in the current study, we explored how the E/I ratio, as estimated via H, can contribute to the differential diagnosis of ASD and SZ. In addition, we used a

novel interpretable machine learning approach, OCT, to classify the two groups based on various sets of features. We showed that H alone is superior to traditional observational clinical tools with respect to classification accuracy, and that differential diagnosis can be best achieved when neuroimaging features are combined with certain traditional clinical scores. In summary, our study highlights the potential for the E/I ratio and H to serve in developing personalized psychiatric tools for differential diagnosis.

Acknowledgments

We thank James McPartland, PhD, for his helpful comments on the manuscript.

References

AFNI (1995): http://afni.nimh.nih.gov/afni.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Anticevic, A., & Lisman, J. (2017). How can global alteration of excitation/inhibition balance lead to the local dysfunctions that underlie schizophrenia?. *Biological Psychiatry*, *81*(10), 818-820.

Bastiaansen, J. A., Meffert, H., Hein, S., Huizinga, P., Ketelaars, C., Pijnenborg, M., ... & De Bildt, A. (2011). Diagnosing autism spectrum disorders in adults: the use of Autism Diagnostic Observation Schedule (ADOS) module 4. *Journal of autism and developmental disorders*, *41*(9), 1256-1266.

Benkarim, O., Paquola, C., Park, B. Y., Kebets, V., Hong, S. J., Vos de Wael, R., ... & Bzdok, D. (2022). Population heterogeneity in clinical cohorts affects the predictive accuracy of brain imaging. *PLoS biology*, *20*(4), e3001627.

Bertsimas, D., & Dunn, J. (2017). Optimal classification trees. *Machine Learning*, *106*(7), 1039-1082.

Bozzi, Y., Provenzano, G., & Casarosa, S. (2018). Neurobiological bases of autism–epilepsy comorbidity: a focus on excitation/inhibition imbalance. *European Journal of Neuroscience*, *47*(6), 534-548.

Casanova MF, Buxhoeveden DP, Switala AE & Roy E (2002) Minicolumnar pathology in autism. Neurology 58, 428–432.

Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J., 2001. A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapp*. 14, 140–151.

Canitano, R., & Pallagrosi, M. (2017). Autism spectrum disorders and schizophrenia spectrum disorders: excitation/inhibition imbalance and developmental trajectories. *Frontiers in psychiatry*, *8*, 69.

Canitano, R., & Palumbi, R. (2021). Excitation/Inhibition Modulators in Autism Spectrum Disorder: Current Clinical Research. Frontiers in Neuroscience, 15.

Chen, C. L., Hwang, T. J., Tung, Y. H., Yang, L. Y., Hsu, Y. C., Liu, C. M., ... & Tseng, W. Y. I. (2020). Multifaceted brain age measures reveal premature brain aging and associations with clinical manifestations in schizophrenia. medRxiv. https://doi.org/10.1101/2020.11.09.20228064

Cochran, D. M., Sikoglu, E. M., Hodge, S. M., Edden, R. A., Foley, A., Kennedy, D. N., ... & Frazier, J. A. (2015). Relationship among glutamine, γ-aminobutyric acid, and social cognition in autism spectrum disorders. *Journal of child and adolescent psychopharmacology*, *25*(4), 314-322.

Demetriou, E. A., Park, S. H., Ho, N., Pepper, K. L., Song, Y. J., Naismith, S. L., ... & Guastella, A. J. (2020). Machine Learning for Differential Diagnosis Between Clinical Conditions With Social Difficulty: Autism Spectrum Disorder, Early Psychosis, and Social Anxiety Disorder. *Frontiers in psychiatry*, *11*, 545.

Dempster, K., Norman, R., Théberge, J., Densmore, M., Schaefer, B., & Williamson, P. (2015). Glutamatergic metabolite correlations with neuropsychological tests in first episode schizophrenia. *Psychiatry Research: Neuroimaging*, *233*(2), 180-185.

Dickinson, A., Jones, M., & Milne, E. (2016). Measuring neural excitation and inhibition in autism: different approaches, different findings and different interpretations. *Brain Research*, *1648*, 277-289.

Du, Y., Fryer, S. L., Lin, D., Sui, J., Yu, Q., Chen, J., ... & Mathalon, D. H. (2018). Identifying functional network changing patterns in individuals at clinical high-risk for psychosis and patients with early illness schizophrenia: a group ICA study. *NeuroImage: Clinical*, *17*, 335-346.

Du, Y., Fu, Z., Sui, J., Gao, S., Xing, Y., Lin, D., ... & Alzheimer's Disease Neuroimaging Initiative. (2020). NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders. *NeuroImage: Clinical*, 28, 102375.

Du, Y., Fu, Z., Xing, Y., Lin, D., Pearlson, G., Kochunov, P., ... & Calhoun, V. D. (2021). Evidence of shared and distinct functional and structural brain signatures in schizophrenia and autism spectrum disorder. Communications biology, 4(1), 1-16.

Egerton, A., Brugger, S., Raffin, M., Barker, G. J., Lythgoe, D. J., McGuire, P. K., & Stone, J. M. (2012). Anterior cingulate glutamate levels related to clinical status following treatment in first-episode schizophrenia. *Neuropsychopharmacology*, *37*(11), 2515-2521.

Fatemi S, Halt A, Stary J, Kanodia R, Schulz S & Realmuto G (2002) Glutamic acid decarboxylase 65 and 67 kDa proteins are reduced in autistic parietal and cerebellar cortices. Biol Psychiatry 52, 805.

Ford, T. C., & Crewther, D. P. (2016). A comprehensive review of the 1H-MRS metabolite spectrum in autism spectrum disorder. Frontiers in Molecular Neuroscience, 9, 14.

Foss-Feig, J. H., Adkinson, B. D., Ji, J. L., Yang, G., Srihari, V. H., McPartland, J. C., ... & Anticevic, A. (2017). Searching for cross-diagnostic convergence: neural mechanisms governing excitation and inhibition balance in schizophrenia and autism spectrum disorders. Biological psychiatry, 81(10), 848-861.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. (1995). Spatial registration and normalization of images. *Human brain mapping*, *3*(3), 165-189.

Fung, L. K., Flores, R. E., Gu, M., Sun, K. L., James, D., Schuck, R. K., ... & Hardan, A. Y. (2021). Thalamic and prefrontal GABA concentrations but not GABAA receptor densities are altered in high-functioning adults with autism spectrum disorder. *Molecular psychiatry*, *26*(5), 1634-1646.

Gao, R., & Penzes, P. (2015). Common mechanisms of excitatory and inhibitory imbalance in schizophrenia and autism spectrum disorders. *Current molecular medicine*, *15*(2), 146-167.

Gibbons, R. D., Chattopadhyay, I., Meltzer, H. Y., Kane, J. M., & Guinart, D. (2021). Development of a computerized adaptive diagnostic screening tool for psychosis. *Schizophrenia Research*.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., ... & Wu-Minn TDP Consortium. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*, *80*, 105-124.

Hassan, M. M., & Mokhtar, H. M. (2019). Investigating autism etiology and heterogeneity by decision tree algorithm. *Informatics in Medicine Unlocked*, *16*, 100215.

Hegarty, J. P., Weber, D. J., Cirstea, C. M., & Beversdorf, D. Q. (2018). Cerebro-cerebellar functional connectivity is associated with cerebellar excitation–inhibition balance in autism spectrum disorder. Journal of autism and developmental disorders, 48(10), 3460-3473.

Horien, C., Floris, D. L., Greene, A. S., Noble, S., Rolison, M., Tejavibulya, L., ... & Constable, R. T. (2022). Functional connectome-based predictive modelling in autism. *Biological Psychiatry*.

Hyatt, C. J., Calhoun, V. D., Pittman, B., Corbera, S., Bell, M. D., Rabany, L., ... & Assaf, M. (2020). Default mode network modulation by mentalizing in young adults with autism spectrum disorder or schizophrenia. *NeuroImage: Clinical*, *27*, 102343.

Hyatt, C. J., Wexler, B. E., Pittman, B., Nicholson, A., Pearlson, G. D., Corbera, S., ... & Assaf, M. (2021). Atypical Dynamic Functional Network Connectivity State Engagement during Social–Emotional Processing in Schizophrenia and Autism. *Cerebral Cortex*.

Jardri, R., Hugdahl, K., Hughes, M., Brunelin, J., Waters, F., Alderson-Day, B., ... & Denève, S. (2016). Are hallucinations due to an imbalance between excitatory and inhibitory influences on the brain?. Schizophrenia bulletin, 42(5), 1124-1134.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage* 62(2), 782–790.

John, Y. J., Zikopoulos, B., Bullock, D., & Barbas, H. (2018). Visual attention deficits in schizophrenia can arise from inhibitory dysfunction in thalamus or cortex. *Computational Psychiatry (Cambridge, Mass.)*, *2*, 223.

Kay, S. R., Opler, A., Fiszbein, A., Ramirez, P. M., & White, L. (1987). The Positive and Négative Syndrome Scale for schizophrénia. *Schizophr Bull*, *3*, 26-76.

Kästner, A., Begemann, M., Michel, T. M., Everts, S., Stepniak, B., Bach, C., ... & Ehrenreich, H. (2015). Autism beyond diagnostic categories: characterization of autistic phenotypes in schizophrenia. *BMC psychiatry*, *15*(1), 1-12.

Kinnaird, E., Stewart, C., & Tchanturia, K. (2019). Investigating alexithymia in autism: A systematic review and meta-analysis. *European Psychiatry*, *55*, 80-89.

Lai, M. C., & Baron-Cohen, S. (2015). Identifying the lost generation of adults with autism spectrum conditions. *The Lancet Psychiatry*, *2*(11), 1013-1027.

Lai, M. C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., ... & Ameis, S. H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis. *The Lancet Psychiatry*, *6*(10), 819-829.

Lai, M. C., Lombardo, M. V., Chakrabarti, B., Sadek, S. A., Pasco, G., Wheelwright, S. J., ... & MRC AIMS Consortium. (2010). A shift to randomness of brain oscillations in people with autism. Biological psychiatry, 68(12), 1092-1099.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, *30*(3), 205-223.

Moreau, C. A., Raznahan, A., Bellec, P., Chakravarty, M., Thompson, P. M., & Jacquemont, S. (2021). Dissecting autism and schizophrenia through neuroimaging genomics. Brain, 144(7), 1943-1957.

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. Proceedings of the National Academy of Sciences, 116(44), 22071-22080.

Murthy, S. K., Kasif, S., & Salzberg, S. (1994). A system for induction of oblique decision trees. *Journal of artificial intelligence research*, *2*, 1-32.

Oliver, L. D., Moxon-Emre, I., Lai, M. C., Grennan, L., Voineskos, A. N., & Ameis, S. H. (2021). Social cognitive performance in schizophrenia spectrum disorders compared with autism spectrum disorder: a systematic review, meta-analysis, and meta-regression. JAMA psychiatry, 78(3), 281-292.

Pepper, K. L., Demetriou, E. A., Park, S. H., Boulton, K. A., Hickie, I. B., Thomas, E. E., & Guastella, A. J. (2019). Self-reported empathy in adults with autism, early psychosis, and social anxiety disorder. *Psychiatry Research*, *281*, 112604.

Port, R. G., Oberman, L. M., & Roberts, T. P. (2019). Revisiting the excitation/inhibition imbalance hypothesis of ASD through a clinical lens. The British journal of radiology, 92(1101), 20180944.

Purves-Tyson, T. D., Brown, A. M., Weissleder, C., Rothmond, D. A., & Weickert, C. S. (2021). Reductions in midbrain GABAergic and dopamine neuron markers are linked in schizophrenia. Molecular Brain, 14(1), 1-19.

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rabany, L., Brocke, S., Calhoun, V. D., Pittman, B., Corbera, S., Wexler, B. E., ... & Assaf, M. (2019). Dynamic functional connectivity in schizophrenia and autism spectrum disorder: Convergence, divergence and classification. *NeuroImage: Clinical*, *24*, 101966.

Robertson, C. E., Ratai, E. M., & Kanwisher, N. (2016). Reduced GABAergic action in the autistic brain. *Current Biology*, *26*(1), 80-85.

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, *2*(5), 255-267.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215.

Sattler, J.M., Ryan, J.J., (1999). Assessment of Children: WAIS-III Supplement (Rev. and Updated, 3rd ed. Jerome M. Sattler Publisher, Inc, La Mesa, CA.

Segal, A., Parkes, L., Aquino, K., Kia, S.M., …, Fornito, A. (2022). Regional, circuit, and network heterogeneity of brain abnormalities in psychiatric disorders. MedRxiv. doi: https://doi.org/10.1101/2022.03.07.22271986

Siegel-Ramsay, J. E., Romaniuk, L., Whalley, H. C., Roberts, N., Branigan, H., Stanfield, A. C., ... & Dauvermann, M. R. (2021). Glutamate and functional connectivity-support for the excitatory-inhibitory imbalance hypothesis in autism spectrum disorders. Psychiatry Research: Neuroimaging, 313, 111302.

Sohal, V. S., & Rubenstein, J. L. (2019). Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. Molecular psychiatry, 24(9), 1248-1257.

Sokunbi, M. O., Gradin, V. B., Waiter, G. D., Cameron, G. G., Ahearn, T. S., Murray, A. D., ... & Staff, R. T. (2014). Nonlinear complexity analysis of brain FMRI signals in schizophrenia. Plos one, 9(5), e95146.

Strasser, L., Downes, M., Kung, J., Cross, J. H., & De Haan, M. (2018). Prevalence and risk factors for autism spectrum disorder in epilepsy: A systematic review and meta-analysis. *Developmental Medicine & Child Neurology*, *60*(1), 19-29.

Supekar, K., Ryali, S., Yuan, R., Kumar, D., de los Angeles, C., & Menon, V. (2022). Robust, generalizable, and interpretable AI-derived brain fingerprints of autism and social-communication symptom severity. Biological Psychiatry.

Taylor, S. F., & Tso, I. F. (2015). GABA abnormalities in schizophrenia: a methodological review of in vivo studies. Schizophrenia research, 167(1-3), 84-90.

Trakoshis, S., Rocchi, F., Canella, C., You, W., Chakrabarti, B., Ruigrok, A. N., ... & MRC AIMS Consortium. (2020). Intrinsic excitation-inhibition imbalance affects medial prefrontal cortex differently in autistic men versus women. *Elife*, 9, e55684.

Traut, N., Heuer, K., Lemaître, G., Beggiato, A., Germanaud, D., Elmaleh, M., ... & Varoquaux, G. (2022). Insights from an autism imaging biomarker challenge: promises and threats to biomarker discovery. NeuroImage, 119171.

Trevisan, D. A., Foss-Feig, J. H., Naples, A. J., Srihari, V., Anticevic, A., & McPartland, J. C. (2020). Autism spectrum disorder and schizophrenia are better differentiated by positive symptoms than negative symptoms. Frontiers in psychiatry, 11, 548.

Van't Wout, M., Aleman, A., Bermond, B., & Kahn, R. S. (2007). No words for feelings: alexithymia in schizophrenia patients and first-degree relatives. *Comprehensive psychiatry*, *48*(1), 27-33.

Vorst, H. C., & Bermond, B. (2001). Validity and reliability of the Bermond–Vorst alexithymia questionnaire. *Personality and individual differences*, *30*(3), 413-434.

Warrier, V., Toro, R., Chakrabarti, B., Børglum, A. D., Grove, J., Hinds, D. A., ... & Baron-Cohen, S. (2018). Genome-wide analyses of self-reported empathy: correlations with autism, schizophrenia, and anorexia nervosa. *Translational psychiatry*, *8*(1), 1-10.

Wakabayashi, A., Baron-Cohen, S., & Wheelwright, S. (2006). Individual and gender differences in Empathizing and Systemizing: Measurement of individual differences by the Empathy Quotient (EQ) and the Systemizing Quotient (SQ). *Japanese Journal of Psychology*.

Wechsler III, D. S. WAIS-III, Wechsler Adult Intelligence Scale--Administration and Scoring Manual (1997). *San Antonio, TX: Psychological Corporation*.

Williams, Z. J., & Gotham, K. O. (2021). Improving the measurement of alexithymia in autistic adults: a psychometric investigation of the 20-item Toronto Alexithymia Scale and generation of a general alexithymia factor score using item response theory. *Molecular autism*, *12*(1), 1-24.

Wink, A. M., Bernard, F., Salvador, R., Bullmore, E., & Suckling, J. (2006). Age and cholinergic effects on hemodynamics and functional coherence of human hippocampus. *Neurobiology of aging*, *27*(10), 1395-1404.

Winterer, G., & Weinberger, D. R. (2004). Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. *Trends in neurosciences*, *27*(11), 683-690.

Yang, G. J., Murray, J. D., Wang, X. J., Glahn, D. C., Pearlson, G. D., Repovs, G., ... & Anticevic, A. (2016). Functional hierarchy underlies preferential connectivity disturbances in schizophrenia. *Proceedings of the National Academy of Sciences*, *113*(2), E219-E228.

Yassin, W., Nakatani, H., Zhu, Y., Kojima, M., Owada, K., Kuwabara, H., ... & Koike, S. (2020). Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. Translational psychiatry, 10(1), 1-11.