# From pattern classification to stratification: towards conceptualizing the heterogeneity of Autism Spectrum Disorder

Thomas Wolfers[1,2], Dorothea L. Floris[1,2], Richard Dinga[3], Daan van Rooij[1,2], Christina Isakoglou[1,2], Seyed Mostafa Kia[1,2], Mariam Zabihi[1,2], Alberto Llera[1,2], Rajanikanth Chowdanayaka[4], Vinod J. Kumar[5], Han Peng[1,20], Charles Laidi[6], Dafnis Batalle[7,8], Ralica Dimitrova[7,8], Tony Charman[9], Eva Loth[7], Meng-Chuan Lai[10,11,12], Emily Jones[13], Sarah Baumeister[14], Caroline Moessnang[14], Tobias Banaschewski[14], Christine Ecker[15,7], Guillaume Dumas[16], Jonathan O'Muircheartaigh[7,8,17], Declan Murphy[7,8,17], Jan K. Buitelaar[1,2,18 #], Andre F. Marquand[1,2,19 #], Christian F. Beckmann[1,2,20 #]

[#]shared last authors

1 Donders Center for Brain, Cognition and Behavior, Radboud University Nijmegen, Nijmegen, The Netherlands

2 Department for Cognitive Neuroscience, Radboud University Medical Center Nijmegen, Nijmegen, The Netherlands

3 Department of Psychiatry, Amsterdam UMC, Amsterdam, The Netherlands.

4 Department of Genetics and Genomics, University of Mysore, Mysuru, India.

5 Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

6 Institut National de la Santé et de la Recherche Médicale (INSERM), U955, Institut Mondor de Recherche Biomédicale, Pôle de Psychiatrie, Assistance Publique–Hôpitaux de Paris (AP-HP), Faculté de Médecine de Créteil, DHU PePsy, Hôpitaux Universitaires Mondor, Créteil, France.

7 Department of Forensic and Neurodevelopmental Science, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom.

8 Centre for the Developing Brain, School of Biomedical Engineering & Imaging Sciences, King's College London, London, United Kingdom.

9 Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom.

10 Centre for Addiction and Mental Health and The Hospital for Sick Children, Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada.

11 Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom.

12 Department of Psychiatry, National Taiwan University Hospital and College of Medicine, Taipei, Taiwan.

13 Centre for Brain and Cognitive Development, Birkbeck, University of London, London, United Kingdom.

14 Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany.

15 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt am Main, Goethe-University, Frankfurt am Main, Germany.

16 Human Genetics and Cognitive Functions, Institut Pasteur, Université Paris Diderot, Sorbonne Paris Cité, CNRS UMR3571 / USR 3756, Paris, France.

17 MRC Centre for Neurodevelopmental Disorders, King's College London, London, United Kingdom.

18 Karakter Child and Adolescent Psychiatry University Center, Radboud University Medical Center, Nijmegen, The Netherlands

19 Department of Neuroimaging, Institute of Psychiatry, King's College London, London, United Kingdom

20 Centre for Functional MRI of the Brain (FMRIB), University of Oxford, Oxford, United Kingdom

**Abstract**

Pattern classification and stratification approaches have increasingly been used in research on Autism Spectrum Disorder (ASD) over the last ten years with the goal of translation towards clinical applicability. Here, we present an extensive scoping literature review on those two approaches. We screened a total of 635 studies, of which 57 pattern classification and 19 stratification studies were included. We observed large variance across pattern classification studies in terms of predictive performance from about 60% to 98% accuracy, which is among other factors likely linked to sampling bias, different validation procedures across studies, the heterogeneity of ASD and differences in data quality. Stratification studies were less prevalent with only two studies reporting replications and just a few showing external validation. While some identified strata based on cognition and intelligence reappear across studies, biology as a stratification marker is clearly underexplored. In summary, mapping biological differences at the level of the individual with ASD is a major challenge for the field now. Conceptualizing those mappings and individual trajectories that lead to the diagnosis of ASD, will become a major challenge in the near future.

**Keywords**

Autism Spectrum Disorder, Machine learning, Pattern Recognition, Classification, Clustering,

Stratification, Biotypes, Precision Medicine

## Introduction

Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental disorder with an estimated worldwide prevalence of about one percent (Elsabbagh et al., 2012). The diagnosis of ASD is based on behavioral symptoms such as impairments in social communication and interaction, and restricted and repetitive behaviors, interests and activities. The most recent DSM-5 (American Psychiatric Association, 2013) revision of the diagnostic criteria has dropped previously defined subtypes. Today, we have no effective pharmacological treatments for the core symptoms of ASD and most clinical trials fail (LeClerc and Easley, 2015). Two of the major reasons for this are that the biology(ies) of ASD are poorly understood (i.e. we lack treatment targets), and prior trials used an 'all comers' approach (i.e. gave the same treatment to all ASD individuals - though they likely vary considerably). Hence there has been increasing focus on the identification of biologically meaningful subcategories within the ASD phenotype (Amaral et al., 2008; Coleman, 2005; Ecker et al., 2015; Lombardo et al., 2019; Masi et al., 2017; Simonoff et al., 2008; Tang et al., 2018). There are three main types of heterogeneity that impact studies. First, different clinical symptom profiles can lead to the diagnosis of ASD (clinical heterogeneity). Second, different biological mechanisms may converge onto a common set of symptoms for ASD (biological heterogeneity). Third, different environmental factors, may modulate the expression of ASD (environmental heterogeneity). Largely due to these forms of heterogeneity, no current theory captures all aspects of ASD based on biological processes (Sanders, 2015). Therefore initiatives are essential, comprising increasingly larger cohorts of individuals with ASD to capture the variation in the biological and clinical characteristics across individuals (Di Martino et al., 2014; Loth et al., 2017; Szatmari et al., 2007).

In order to take advantage of increasingly large datasets, pattern recognition and machine learning methods are gaining more importance. Generally speaking these approaches

can be summarized under the umbrella terms supervised and unsupervised learning. In the former, labels -for instance clinical diagnoses- are known and utilized to find an optimal decision rule. In the latter, the algorithm infers a decision on class membership by relying exclusively on the inherent structure of the unlabeled input data (Bishop, 2007; Hastie et al., 2009). Here, we use the term pattern classification to refer to supervised approaches that integrate biological and/or behavioral measures in order to extract a predictive pattern corresponding to the diagnosis of ASD. In contrast, we use the term stratification to refer to unsupervised approaches that use different sources of information to find meaningful substructures within the ASD phenotype.

A number of reviews on mental disorders have focused on one of these two approaches -pattern classification or stratification- (Andrews et al., 2018; Arbabshirani et al., 2017; Hahn et al., 2017; Huys et al., 2016; Marquand et al., 2016b; Orrù et al., 2012; Varoquaux, 2017; Wolfers et al., 2015). However, none has synthesized the utility of both approaches systematically. Therefore, we surveyed the literature on pattern classification and stratification methods in ASD using our previously published methods (Marquand et al., 2016b; Wolfers et al., 2015). In this scoping review, we contribute to the literature by focusing on pattern classification studies based on behavioral, neuroimaging, and other biological readouts in ASD. Earlier work mostly focused on one modality, usually brain imaging. In addition, we also included stratification studies and compared the two approaches on their utility to shape the future of ASD research. In doing so we i) provide an in-depth review of both approaches, ii) identify important common outcomes, iii) outline opportunities and shortcomings, and iv) present potential future directions for pattern classification and stratification approaches in ASD research and clinical practice.

## Methods

*Scoping review*

We conducted a literature search on all studies that used pattern classification as well as stratification approaches in ASD. We defined pattern classification studies as those that predict ASD clinical diagnostic status either cross-sectionally or longitudinally on the basis of biology, cognition and/or behavior. Importantly, we only included studies that report out-of-sample predictions. In other words, a predictive model was trained on one part of the data and tested on another. Other statistical approaches that are validated within the same sample may describe biological factors underlying ASD but are not predictive at the individual level. Stratification studies were defined as those that aimed to identify meaningful clusters within ASD on the basis of biological, cognitive, behavioral or symptom measures. The search terms[1][2], inclusion criteria and the number of studies that were reviewed are depicted in Figure 1. The search was concluded on the 10th of April 2019.

[insert Figure 1]

*Pattern classification and stratification*

Pattern classification approaches on the basis of quantifiable biological readouts and behavior started gaining prominence in the ASD literature about 10 years ago. Since then many studies have been performed with the goal to predict ASD diagnosis. Similarly, stratifications on the basis of primarily behavior have gained more attention in the ASD literature in the last

[1] Search term - pattern classification:
(Autism OR Autism spectrum disorder) AND (pattern classification OR machine learning)
[2] Search term - stratification:
(Autism OR Autism spectrum disorder) AND (subtyping OR stratification OR clustering)

decade. In this section, we briefly introduce the main categories by which we described and classified the existing literature. This classification scheme is kept throughout the text, tables, and figures.

*Modality and features*

In the present review, we classified studies based on the modalities they used to extract features from for their predictions. Modality refers to the type of biological readout. We cannot easily measure brain structure or function directly. Therefore, we rely on indirect measures, such as electroencephalography (EEG), magnetoencephalography (MEG), magnetic resonance imaging (MRI), cognitive and behavioral assessments and genetic measures. All these measures have unique advantages and disadvantages in assessing biological or psychological states. While, for instance, functional MRI allows us to image the brain even in deep subcortical nuclei of the brain, it has a poor temporal resolution (Dale and Halgren, 2001). In contrast, EEG has exceptional temporal, but poor spatial, resolution.

A feature is a characteristic that can be extracted from the data generated with a certain measurement modality and is used as input to any kind of algorithm. In the last decade, the predominant measures used for classification were biological features while stratification was primarily based on symptoms and cognition. Importantly, while most clustering algorithms require the measurements to be continuous or ordinal, some algorithms can deal with different types of variables measured on different kinds of non-continuous scales (Bishop, 2007). The engineering of novel features that can be extracted from all kinds of biological measurements is an important research topic as the biological relevance and reliability of this step often determines the predictability of ASD more substantially than the classifier itself.

*Classifiers and stratification algorithms*

Classifiers used for pattern classification range from simple linear models to more sophisticated nonlinear variants such as multilayer neural networks or Gaussian processes (Bishop, 2007). All these approaches were used in the reviewed literature. In simple terms, a classifier learns a rule, which separates the classes. The algorithms thus differ with regard to the method that determines this rule. In the following we shortly introduce the main algorithms. A Linear Discriminant Classifier (LDC), a classical linear model, is used to separate classes by maximizing the ratio of between-class to within-class variance. A Logistic Regression Classifier (LRC) is a probabilistic discriminant model that aims to learn an optimal decision rule by modelling the log-odds ratio as a linear combination of predictor variables. Under the Gaussian assumption (assuming that individuals within each class are distributed according to a Gaussian distribution), LDC and LRC are equivalent (Hastie et al., 2009). Both methods yield probabilistic predictions that a new case corresponds to a particular class and can be transformed into a class label. The Support Vector Machine (SVM) is an algorithm designed for binary classification that maximizes the margin between classes in a high dimensional space. Mathematically, the discriminant function is defined by a weight vector orthogonal to the decision boundary, which can be uniquely specified by the samples that lie closest to the decision boundary, referred to as support vectors. The decision boundary represents the rule for classification of new examples. A Gaussian Process Classifier (GPC) is a Bayesian extension of LRC, a probabilistic model often described as a distribution over functions. In contrast to SVM, the predicted class is augmented by an estimate of the certainty of the prediction. Based on Bayes's rule, the posterior distribution of functions on the training data is found by maximizing the negative log-likelihood. This posterior distribution is then used to classify new examples according to the rules of probability. Previous approaches typically utilize linear techniques for class boundary definitions, though non-linear extensions are possible. For example, Artificial Neural Networks (NN) are a broad class of algorithms that

are inspired from biological neural networks (Rosenblatt, 1958). Generally, NNs consist of a set of artificial neurons that are trained by adjusting the weights connecting them and can be used for a range of pattern recognition tasks including classification. The learned relation of these artificial neurons represents the decision rule which is applied to make predictions for new examples. Deep learning is one prominent extension of the neural networks theory, which is characterized by many layers of artificial neurons (Lecun et al., 2015). In many fields, these networks outperform other algorithms (Silver et al., 2016), provided that sufficient training data is available (Jia Deng et al., 2009). This is a problem in neuroscience in which big-data is just emerging as the acquisition of a large number of samples is very challenging and costly (Miller et al., 2016).

With regard to stratification, there are two related types of approaches that have been used in ASD research, i) clustering and ii) finite mixture models (FMM). Clustering methods explore the data with the goal to identify clusters or subtypes within a dataset. The goal is to identify a partitioning of the data such that the samples comprising each cluster are more similar to one another than to samples assigned to the other clusters. Given this goal, these algorithms require a measure of similarity or distance to be defined (Xu and Tian, 2015). A simple clustering algorithm is the iterative 'K-means' approach (MacQueen, 1967). It involves two steps: (i) for each data point, find the closest cluster center according to for example the squared Euclidean distance and (ii) replace each cluster center with the coordinate-wise mean of all points assigned to it. These steps are iterated until the cluster assignments do not change. K-means is just one of the many methods that have been used in the literature (Hartigan and Wong, 1979). Generally, these algorithms differ in the way they operationalize class similarity or class difference (Xu and Tian, 2015; Xu and Wunsch, 2005). It is important to note that the results of the clustering method heavily depend on the assumptions of the algorithm chosen and the particular notion of similarity or distance that each algorithm implies (von Luxburg et

al., 2012). Unfortunately, despite a proliferation of different algorithms for clustering, there is no way to tell unambiguously whether an algorithm performs better than any other as there is no universal metric that can adjudicate this. Finite mixture models are a broad class of probabilistic stratification approaches (Bishop, 2007). The models partition the data into a mixture of a given number of parametric distributions. For instance, a Gaussian mixture model describes the input data as a mixture of Gaussians (Bishop, 2007). Generally, these models work by estimating the number of components that are represented by a certain probability distribution and the algorithm estimates the mixing coefficients that determine the proportion that each component contributing to the mixture along with the parameters of each component distribution. A number of different models belong to the class of finite mixture models (Bratchell, 1987; Jobson, 1992; Vermunt and Magidson, 2002), that all have their own advantages and disadvantages. The performance evaluation of pattern classifications and stratification is generally quite different. Therefore, we discuss both approaches separately in the following two sections.

*Performance evaluation for pattern classification*

To estimate model performance on unseen data, usually, some form of data splitting is performed. The simplest, but least efficient way (Steyerberg et al., 2007), is to split the data into two parts, a training set used to develop a model, and a test set that is used for model validation. However, this limits the number of available samples for evaluation and the results are too dependent on the initial split. Hence, cross-validation is usually performed. As part of this, data are split into a number of parts, and the training and validation steps are repeated, each time leaving a different partition out as a test set. A special case where one subject is left out each time is called leave-one-out cross-validation. This method has an intuitive appeal (since most of the subjects can used for model training). However, it is less reliable than

alternatives and so it is now not recommended (Varoquaux, 2017). Another method, that is less common in psychiatry research, is called bootstrapping. Here, the training samples are repeatedly selected with replacement from the original data.

In medicine, it is recommended that both diagnostic and prognostic models are probabilistic and are evaluated based on their calibration and discrimination (Collins et al., 2015). Calibration is defined as degree of agreement between the predicted probabilities of an event and its observed frequency (e.g. good calibration means that events that are predicted with 0.6 probability, if they happen 60% of the time). In contrast, discrimination is usually evaluated using the area under the receiver operating characteristic curve (ROC), where we compare the false-positive rate of a classifier relative to the true positive rate and which is equivalent to concordance probability in the case of a binary outcome. It can therefore be interpreted as the probability that a randomly selected participant from a specific group will be predicted to belong to this group with higher probability than a randomly selected person from another group. Evaluating models based solely on thresholded categorical predictions (i.e. accuracy, balanced accuracy, sensitivity, specificity), is usually not recommended, because categorical predictions are very crude and hide potentially clinically important information. It also moves the decision making from 'stakeholders' (e.g. clinicians or patients) to the data analyst - thus assuming that the appropriate decision threshold is known and constant across all situations the model can be applied to. It is important to note, however, that measures like ROC and balanced accuracy are not affected by relative class frequency or disease prevalence. Therefore, it is possible to have what is seemingly a high performing model that still produces a high number of false positive predictions. The reported performance measures varied across articles we have reviewed here. Despite its problems, accuracy was the most commonly reported performance measure. Therefore, we focused our review on accuracy. In studies where accuracy was not reported, we report the area under the curve.

*Performance evaluation for stratification*

Generally, stratification requires the user of most clustering algorithms to determine the number of clusters or components a priori. This has a substantial effect on the outcome of the algorithm as it determines its flexibility. Therefore, different heuristics and test procedures have been developed that can be used to determine whether a certain cluster solution is better than another, or if it is appropriate in the first place (Bratchell, 1987; Jobson, 1992; Vermunt and Magidson, 2002). However, all these methods rely on assumptions on the nature of similarity and they apply heuristics that can (and often do) fail. Therefore, there is no way of determining the optimal number of clusters with certainty (Bratchell, 1987). Consequentially, it is essential that identified stratifications are replicated (e.g. to assess the stability of the cluster solution) and validated (e.g. to assess clinical plausibility) to be meaningful. This is not straightforward and specific steps need to be taken. Moreover, it is important to note that those steps (while important) might not be sufficient for all kinds of stratification(s). First, the number of clusters should be replicated in an independent dataset that includes the same kind of variables. If this replication is successful, the clusters should then be validated against variables that are of interest but have not been part of the clustering procedure. In the case of clinical validation, one would be interested in a certain outcome measure that is predictive of long-term assessment of quality of life for instance. In an ideal world emerging clusters boil down to reliable subtypes and are externally validated against those clinical measures. In the real world, however, this is often impossible (as the number of datasets available with identical measures is limited). In the present article we reviewed the literature on those validation criteria.

**Results**

*Pattern classification of ASD*

For the section of this review, we characterized the pattern classification literature on ASD in detail based on biological, cognitive and behavioral factors. We included a total of 57 ASD studies for detailed review and condensed the most salient features of those studies in Figure 2. In the following, we review some of the most prominent findings of those studies.

One of the first studies in which ASD was predicted was based on structural brain measures, used a SVM approach, and reported promising results with accuracies of up to 86% (Ecker et al., 2010b, 2010a). Subsequently, several other studies also reported accuracies that were up to 90% and higher, and these were based on structural MRI, resting state MRI, and other imaging modalities (Ahmadlou et al., 2010; Jiao et al., 2010; Uddin et al., 2011). These initial results were also backed up by follow-up publications showing similarly high accuracies, indicating that it was possible to discriminate ASD from healthy individuals (Table1), and in some cases from those with other neurodevelopmental disorders (such as ADHD). After about three years, the sample size of pattern classification studies increased from around 50 individuals with ASD to more than 300. One study that represents a shift in the literature performed a classification of 325 individuals with ASD using structural MRI (Sabuncu and Konukoglu, 2014). They reported, however, that ASD could only be predicted with an accuracy of 60%. Since this was a multisite study, scanner 'noise' might be an important factor influencing the lower accuracies. Nonetheless, the prior smaller studies might have been prone to cross-validation failure (Varoquaux, 2017) and a stronger publication bias. For example, a non-surprising finding in a small sample is less likely to find its place in the literature than the same finding in a large sample. Furthermore, larger samples also imply that more of the intrinsic heterogeneity of both non-autistic and individuals with ASD is sampled. This results in a larger overlap between groups and lower accuracies during classification. In one of the

largest multi-site studies on ASD to date (using the ABIDE sample) showed that with increasing sample sizes, inter-site prediction approached intra-site predictions with the highest accuracies of around 67% (Abraham et al., 2017). Therefore, increasingly larger sample sizes may allow for the identification of more robust decision functions. While diagnostic predictions were important in determining the predictability of ASD on the basis of biological measures, there are increasing reports from longitudinal studies of 'at-risk' infants (Hazlett et al., 2017; Shen et al., 2017) with promising (but mixed) predictive accuracies ranging from 94% and 69% for the development of ASD. Although, this performance is too low and unreliable for clinical translation at this stage, those studies hold the potential for the identification of signatures indicative for the development of ASD very early in life.

In summary, predictive accuracies are variable across studies from about 60% to 98% dependent on features, cross-validation, and sampling. Structural and functional MRI predictions of ASD are over-represented in comparison with diffusion MRI, EEG, behavior and multimodal data-based classifications (Figure 2). The accuracies dropped with sample size and the predictions are usually not calibrated by the base-rate of the diagnosis with ASD. Cross-sectional prediction studies are the most prevalent, and only recently have longitudinal prediction studies allowed the assessment of ASD trajectories.

[insert Figure 2]

[insert Table 1]

*Stratification of ASD*

For this section of this review, we inspected the stratification literature on ASD in detail based on biological and behavioral factors. We included a total of 19 studies and reported the most salient features of those studies in Figure 2. In the following, we review some of the most prominent findings of those studies.

About ten years ago the first research on stratification of ASD was performed (Bitsika et al., 2008; Munson et al., 2008). One of the most prominent clustering approaches is hierarchical clustering. Generally, the number of identified clusters was independent of the respective features on which stratification was based and ranged from three to six across all papers reviewed (Table 2, Figure 2). There was no clear association between the applied stratification method and the number of resulting subgroups. The sample size ranged from about 100 participants to studies that included more than 4000 individuals with ASD (Table 2). The measures that were mostly used for stratification were either symptom scores or cognitive measures, while only a few included sensory processing or biological variables (Lane et al., 2010; Sacco et al., 2012). While the etiology of ASD is likely in large parts biological (Abrahams and Geschwind, 2008) the inclusion of biological information for stratification appeared in 2018 for the first time. Some studies also included a range of different measures for the purpose of stratification; however, in general this did not include biological information such as genetic or brain imaging measures. Of the 19 reviewed studies only two studies replicated the identified clusters in independent samples (Lombardo et al., 2016; Veatch et al., 2014) and about half of the studies did not validate the identified clusters externally. While it is difficult to synthesize common results across stratification studies due to the factors discussed above, there are a few common outcomes. First, a large proportion of the reviewed studies show at least one subgroup that is characterized by decreased cognitive performance or intelligence (Table 2). More importantly, there are only a few studies that report biological subtypes and these subtypes do not seem to converge (Table 2). This highlights the need for more systematic investigation into biological subtypes of ASD. Further, while many samples include individuals with comorbidities on top of a primary diagnosis with ASD, the majority of studies neglected other disorders. This is a limitation for the identification of transdiagnostic clusters that may map better onto biology than clusters constrained by boundaries due to current

psychiatric classification which is mostly based on symptoms. Finally, while with increasing sample size the accuracy reported in pattern classifications studies decrease, in stratification studies the reported number of clusters increase. With increasing sample size, one is more likely to sample the large variety of individuals with ASD, so that in larger samples the number of strata or subtypes would increase. This review provides first evidence for this observation (Figure 2).

In summary, the body of literature on stratification approaches is considerably smaller than that on pattern classification approaches on ASD. While the number of identified subgroups differs between studies, the measures utilized for stratifications focused primarily on symptom scores or cognition. Biological measures were largely neglected. Most studies did not independently replicate their findings and about half did not validate their results externally (Figure 2). While these limitations are the core theme across studies, a large proportion of studies also report a subgroup of individuals with ASD characterized by low cognitive performance or intelligence. In line with the debate on heterogeneity of the ASD phenotype we provide first evidence that with increasing sample size the number of clusters reported for ASD increases.

[insert Table 2]

**Discussion**

In this scooping review, we surveyed the literature on pattern classification and stratification studies on ASD. With increasingly larger samples being made available for analysis, these methods will determine whether we can eventually translate 'big data approaches' into clinical practice. We observed variable accuracies across studies dependent on the selection of features, cross-validation strategy, sampling and differences in data quality. Structural and functional MRI predictions are overrepresented in comparison with diffusion

MRI, EEG, behavior, and multimodal classifications. Cross-sectional prediction studies are the most prevalent, and only recently longitudinal studies on high-risk samples allow for the assessment of ASD trajectories in a predictive framework. With respect to stratification methods applied to ASD, the number of identified ASD subgroups differ substantially and symptom scores or cognition were usually the basis of these approaches. This may be a consequence of the high heterogeneity of the ASD phenotype or sensitivity of resulting clusters on arbitrary user-defined clustering parameters. Therefore, clustering methods require further development and need to be more user independent. Generally, the sample size of most stratification studies is large, suggesting that the intrinsic clinical, biological and environmental heterogeneity of the ASD phenotype was captured. That said, biological measures gained momentum in those studies only very recently. This is important, as a stratification on the basis of symptoms in the case of ASD is not sufficient. The emerging subgroups are too heterogeneous and often not reproducible. Therefore, we need to be able to identify clusters that map better onto biology. Cluster approaches based on genetics are being adapted for the stratification for ASD and are expected to gain importance as they did in other medical disciplines (Hofree et al., 2013; Kim et al., 2018). In the following, we discuss in detail the evolution of those two approaches and what we can learn from for the past and present for the future of these approaches in the context of ASD research.

*The past of pattern classification and stratification in ASD research*

ASD has been investigated using both approaches. Especially, in the early days pattern classification approaches received a lot of attention when on the basis of biological measures, the prediction of ASD was possible with relatively high accuracies of more than 80% (Table 1). Considerable resources have since been spent on identifying biological signatures. However, over time it has become increasingly apparent that earlier studies may have overestimated the 'real' predictability of ASD. This was likely largely due to (for example)

sampling biases (Wolfers et al., 2015), cross-validation failures (i.e. cross-validation overestimates the generalizability of an algorithm in comparison to using a test set (Varoquaux, 2017), less heterogeneity in smaller studies, and potentially also publication bias. With increasing sample size, the reported performance drops across studies (Figure 2). Nonetheless, a recent study reported an increase samples size that led to better learning (Abraham et al., 2017) - suggesting that larger samples will allow for the estimation of more complex decision functions that allow us to capture the complex phenotype of ASD with a single decision function. While this debate is not over, we suggest that the more important challenge is to parse the heterogeneity within ASD.

While pattern classification algorithms have been applied to ASD for more than ten years, stratification studies are much less prevalent in ASD research. This may be due to the fact that stratification is a more difficult problem. For instance, replication and external validations become more important in unsupervised learning problems (Marquand et al., 2016b; Schwenker and Trentin, 2014). In other words, subgroups must be validated against an external estimate such as the course of ASD and these variables are usually not readily available. Furthermore, many studies were simply not similar enough in terms of acquisition procedures, protocols, and variables to allow for replication. Last, the biology of potentially emerging subtypes needs to be mapped. While pattern classification was often based on biological factors, stratification studies largely neglect this information. Instead, symptom counts or cognitive measures were fed into a clustering schema. This is in line with the long-standing stratification efforts of ASD based on symptom profiles, which however has shown limited success. Therefore, it may be possible to improve clustering efforts by including biological variables into the stratifications and at the same time using a more systematic way of replicating and validating emerging subgroups. Generally, these improvements require the

acquisition of large samples which can be validated against external measures such as course and developmental outcomes.

In summary, i) pattern classification approaches show promise but clinical applicability has not been reached, ii) stratification approaches have not yet robustly detected subgroups for ASD and/or shown how well they map onto underlying biology(ies).

*The present of pattern classification and stratification in ASD research*

Pattern classification and stratification approaches for ASD are affected by general trends in the field today. In the following, we discuss major developments and how pattern classification and stratification approaches can contribute.

The analysis of increasingly larger samples is a general mantra across different fields. Therefore, we expect the acquisition of even larger samples and pooling of data across studies (Van Rooij et al., 2018). In genetics, for instance, this has led to the identification of common risk variants for many major mental disorders, among others ASD (Lee et al., 2013). While those studies report robust group level differences with small effect sizes, it is unclear if those differences translate toward individual predictions. After all we treat and care for individuals, not averages or groups. Piling up data and building more advanced classifiers, might allow for the learning of very complex decision functions potentially resulting in more optimal classifications of individuals with ASD (Abraham et al., 2017). Therefore, we see considerable investments into more advanced classification approaches based on, for instance, deep learning (Lecun et al., 2015). Earlier studies using deep learning with neuroimaging data have shown good classification accuracy in small samples (Heinsfeld et al., 2018; Vieira et al., 2017), however, there is no confirmed benchmark in large datasets. A recent prediction challenge of the complex phenotype intelligence suggests that the benefit of deep learning on neuroimage

data for improving performance is marginal or non-existent[3], pointing towards a degree of overfitting in smaller imaging studies on complex traits such as ASD. Considering that we are predicting a highly heterogenous and comorbid disorder, predicting diagnostic status is not sufficient. Therefore, the prediction of the course of ASD which is clinically more meaningful gains further momentum. Research on neonates at risk for development of ASD (Hazlett et al., 2017; Shen et al., 2017) might in particular benefit from applications of pattern classification approaches that can predict their developmental trajectories.

While these developments will remain important, we anticipate that the larger challenge concerns the heterogeneity of ASD. We can clearly observe that classification of ASD tends to decrease with increasing sample size (Figure 2), although the reasons for it are not entirely clear. It may be that the biological overlap of ASD with a healthy population increases when both are sampled more representatively. Therefore, we expect that stratification approaches become more important. The identification of meaningful subgroups within the ASD diagnosis is a very challenging task and requires larger samples that optimally reflect the biological diversity of patients. While this is an important step, it is still only the first, replication of emerging strata is equally important as is validation. Today, most studies do not provide replications, especially when the identified strata are based on biology. Only one study has successfully replicated the cognitive subgroups of ASD (Lombardo et al., 2016) in an independent sample. Note that biology did not play any role in this study. Therefore, we anticipate that replications and biological stratifications of ASD will become important.

We identified three major trends today: i) increasingly larger samples might allow for the training of more complex algorithms. ii) Pattern classification is more focused on predicting the course of individuals at risk for ASD rather than the diagnosis itself. iii) Stratification methods are gaining more importance in comparison with pattern classification. While

[3] https://sibis.sri.com/abcd-np-challenge/

systematic replication of potentially emerging subtypes is still ahead of us, we anticipate that the identification of meaningful subtypes will gain much more momentum in the near future as well as assembling appropriate datasets for replication and validation.

*The future of pattern classification and stratification in ASD research*

Here, we outline major trends in the field that might shape our understanding of ASD suggesting potential novel applications for pattern classification and stratification approaches.

First, increasingly larger samples are and have been acquired with the goal to improve our understanding of ASD. For instance, longitudinal large-scale studies are emerging that capture the heterogeneity of the ASD phenotype to a fuller extent (Di Martino et al., 2014; Loth et al., 2017). These samples contain biological and behavioral readouts acquired with novel methods and provide information across a large number of individuals and biological layers. In order to integrate all this information in a meaningful way, novel methods that are geared toward big-data will become more important (Calhoun and Sui, 2016; Groves et al., 2011; Wolfers et al., 2017). In a recent study, we show that such an integration approach can identify multimodal brain structures relevant to describe complex behaviors biologically (Arenas et al., under review), which are relevant in ASD. Generally, the field is characterized by the emergence of novel methods that allow us to extract features across biological readouts. In this context deep learning might gain more prominence as it allows for the automatic construction of features (Lecun et al., 2015). We expect that those techniques introduce features that have a clearer biological foundation and, therefore, might shape a better understanding of ASD in the near future.

Second, we expect that the limitations of clustering approaches become more apparent and that novel approaches such as normative modelling (Marquand et al., 2016a; Wolfers et al., 2018; Zabihi et al., 2019) will gain more momentum. As mentioned earlier, clustering of

the ASD phenotype is primarily based on behavioral data, symptom profiles and only recently includes genetics, and/or brain imaging data as the basis to identify subgroups. All reported studies have similar limitations, such as the predefined number of clusters, which is usually an arbitrary choice, or a lack of external validation of the prospective subtypes. Further, clustering algorithms always give a result, and they usually do not test the null hypothesis that there may be no clusters in the data at all (Liu et al., 2008). Only a few studies have performed extensive out-of-sample validations. One study that generated considerable attention mapped symptom counts on resting state data (Drysdale et al., 2017) in depression. In this way, the researchers identified two dimensions, which formed the basis for hierarchical clustering. The researchers identified four subgroups for depression, which were subsequently validated extensively. While the number of external validations was impressive, the clusters identified seem arbitrary, as individuals might simply be described along two identified continuous dimensions without utilizing clustering at all. In a recent attempt to reproduce the main results of the paper in an independent sample, limitation of this approach became apparent (Dinga et al., 2019). Those pertain the utilization of clustering and the application of statistical methods as well as the replicability of the cluster solutions. In line with this observation, our results, obtained on ASD, ADHD, Bipolar Disorder and Schizophrenia (Wolfers et al., 2019, 2018; Zabihi et al., 2019), show that inter-individual differences between patients with the same diagnosis are extreme. With respect to ASD, we mapped its heterogeneity in terms of brain structure at the level of the individual (Zabihi et al., 2019). In this way we showed that ASD is more variable than anticipated and that many individuals with this disorder have patient-specific deviations from the healthy range, suggesting that each person with this diagnosis is quite different from one another. Note that only a subset of patients showed deviations from an expected normal pattern, and many other patients had a neuroanatomical profile that overlapped with the healthy range. The deviating participants were in many cases quite extreme, which suggests a possible reason

for inconsistencies in case-control studies (Bethlehem et al., 2018). Therefore, the description of patients on the group level is certainly not sufficient, a cluster level description may not be refined enough to capture the complexity of ASD, which may, in fact, be relatively patient specific. For these reasons, we expect that approaches which can describe the individual patient will gain moment in the near future.

Third, in line with the previous arguments, we expect that research initiatives, which suggest an approach to investigate ASD through systematic research across cognitive domains and levels of biological description will gain further relevance. A prominent approach would be the Research Domain Criteria (RDoC) approach (Insel et al., 2010). While RDoC has a number of problems (Weinberger and Goldberg, 2014), we think that a systematic characterization of individuals across different behavioral, cognitive and biological domains is important in addition to moving beyond a classical clustering and subtyping approach. Concretely, we expect that phenotypic instruments capturing different aspects of biology from large populations cohorts will be used to describe biological processes in the general population. These processes can subsequently be captured in a normative modelling framework to build reference panels across biological, cognitive and behavioral domains to map individual differences in ASD. Provided that it makes sense to stratify ASD based on biology, we can chart variation in brain systems. For this effort to be successful we need to acquire samples of patients that captures the full heterogeneity of the ASD and related disorders and that can be placed with respect to variation in population reference samples. If clinical studies included only individuals with ASD without co-occurring psychiatric disorders/symptoms these stratification efforts would inherently be limited by the categorial divisions that are imposed through the current psychiatric (i.e. DSM) classification scheme. Therefore, stratification studies might miss biologically meaningful groups, because they start off with categorical boundaries that may have limited biological relevance as can also be observed in large scale

cross-disorder work, showing considerable biological overlap across disorders (Lee et al., 2013). A way forward would be to analyze individuals with multiple different disorders together by circulating those individuals, thus not groups, around a population reference (Marquand et al., 2019). In this way we do not preselect any group neither the healthy nor the disordered to uncover biologically meaningful strata independent of current psychiatric classifications. We expect that these developments will take effect in the near future.

In summary, we expect that the future of ASD research with respect to pattern classification and stratification approaches will be characterized by a few main developments. Pattern classification approaches remain very important for the integration of information in order to extract novel predictive biological signatures. This is important especially for outcome predictions in babies and toddlers at risk for ASD and the prediction of the developmental course of the individual. While studies will further increase in sample size we anticipate that clinical utility of diagnostic predictions based on biology will not be reached and instead stratification approaches will gain further importance. Furthermore, we foresee that the field will move beyond stratification approaches towards the conceptualization of heterogeneity within ASD and across other cooccurring disorders at level of the individual.

**Conclusion**

Pattern classification approaches have extensively been used in research on ASD in the last ten years. While initial studies showed promising results, it has not been possible to predict ASD to a degree that translated to clinical practice. This is probably at least partly due to its heterogeneity. While larger samples might allow us to improve our predictions further, we foresee that the predictions at the individual level will remain challenging. Instead, in the future, we will see more efforts to disentangle the heterogeneity of ASD. Further, we expect that it will become increasingly important to predict the developmental trajectories of

individuals with ASD, especially in preverbal infants who are at risk of developing ASD. Therefore, there will be more effort directed to those at-risk populations. We anticipate that the biological foundations even in those restricted groups are large and that stratification approaches will be vital here as well. Based on recently emerging work we think that ASD is going to be best described mechanistically at the level of the individual. Therefore, we expect that mapping individual differences using, for instance, normative models will be an important step toward precision medicine in ASD research and eventually clinical practice.

**Acknowledgments**

This work has been primarily supported by the EU-AIMS (European Autism Interventions) program, which receives support from Innovative Medicines Initiative Joint Undertaking Grant No. 115300 and No. 777394, the resources of which are composed of financial contributions from the European Union's Seventh Framework Program (Grant No. FP7/2007-2013), from the European Federation of Pharmaceutical Industries and Associations companies' in-kind contributions. In addition, it has received funding from the AIMS-2-TRIALS program, which receives support from the Innovative Medicines Initiative Joint Undertaking under grant agreement No 777394. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and EFPIA and AUTISM SPEAKS, Autistica and SFARI. Furthermore, this work was supported by the Niels Stensen Fellowship (to TW) and the Netherlands Organization for Scientific Research VIDI Grant Nos. 016.156.415 (to AFM) and 864.12.003 (to CFB).

**References**


Abraham, A., Milham, M.P., Di Martino, A., Craddock, R.C., Samaras, D., Thirion, B., Varoquaux, G., 2017. Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. Neuroimage 147, 736–745. https://doi.org/10.1016/j.neuroimage.2016.10.045

Abrahams, B.S., Geschwind, D.H., 2008. Advances in autism genetics: On the threshold of a new neurobiology. Nat. Rev. Genet. 9, 341–355. https://doi.org/10.1038/nrg2346

Ahmadlou, M., Adeli, H., Adeli, A., 2012. Fuzzy Synchronization Likelihood-wavelet methodology for diagnosis of autism spectrum disorderAhmadlou, M., Adeli, H. and Adeli, A. (2012) 'Fuzzy Synchronization Likelihood-wavelet methodology for diagnosis of autism spectrum disorder', Journal of Neurosc. J. Neurosci. Methods 211, 203–209. https://doi.org/10.1016/j.jneumeth.2012.08.020

Ahmadlou, M., Adeli, H., Adeli, A., 2010. Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder. J. Clin. Neurophysiol. 27, 328–333. https://doi.org/10.1097/WNP.0b013e3181f40dc8

Amaral, D.G., Schumann, C.M., Nordahl, C.W., 2008. Neuroanatomy of autism. Trends Neurosci. 31, 137–145. https://doi.org/10.1016/j.tins.2007.12.005

American Psychiatric Association, 2013. DSM-V, American Journal of Psychiatry. https://doi.org/10.1176/appi.books.9780890425596.744053

Anderson, J.S., Nielsen, J.A., Froehlich, A.L., Dubray, M.B., Druzgal, T.J., Cariello, A.N., Cooperrider, J.R., Zielinski, B.A., Ravichandran, C., Fletcher, P.T., Alexander, A.L., Bigler, E.D., Lange, N., Lainhart, J.E., 2011. Functional connectivity magnetic resonance imaging classification of autism. Brain 134, 3739–3751. https://doi.org/10.1093/brain/awr263

Andrews, D.S., Marquand, A., Ecker, C., McAlonan, G., 2018. Using Pattern Classification to Identify Brain Imaging Markers in Autism Spectrum Disorder. pp. 289–320. https://doi.org/10.1007/7854

Arbabshirani, M.R., Plis, S., Sui, J., Calhoun, V.D., 2017. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. Neuroimage 145, 137–165. https://doi.org/10.1016/j.neuroimage.2016.02.079

Ausderau, K.K., Sideris, J., Little, L.M., Furlong, M., Bulluck, J.C., Baranek, G.T., 2016. Sensory subtypes and associated outcomes in children with autism spectrum disorders. Autism Res. 9, 1316–1327. https://doi.org/10.1002/aur.1626

Bethlehem, R.A., Seidlitz, J., Romero-Garcia, R., Dumas, G., Lombardo, M.V., 2018. Normative age modelling of cortical thickness in autistic males. Bioarchive.

Bishop, C.M., 2007. Pattern Recognition and Machine Learning, 1st ed, Journal of Electronic Imaging. Springer. https://doi.org/10.1117/1.2819119

Bitsika, V., Sharpley, C.F., Orapeleng, S., 2008. An exploratory analysis of the use of cognitive, adaptive and behavioural indices for cluster analysis of ASD subgroups. J. Intellect. Disabil. Res. 52, 973–985. https://doi.org/10.1111/j.1365-2788.2008.01123.x

Bratchell, N., 1987. Cluster analysis. Chemom. Intell. Lab. Syst. https://doi.org/10.1016/0169-7439(87)80054-0

Calderoni, S., Retico, A., Biagi, L., Tancredi, R., Muratori, F., Tosetti, M., 2012. Female children with autism spectrum disorder: An insight from mass-univariate and pattern classification analyses. Neuroimage 59, 1013–1022. https://doi.org/10.1016/j.neuroimage.2011.08.070

Calhoun, V.D., Sui, J., 2016. Multimodal Fusion of Brain Imaging Data: A Key to Finding the Missing Link(s) in Complex Mental Illness. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 1, 230–244. https://doi.org/10.1016/j.bpsc.2015.12.005

Chanel, G., Pichon, S., Conty, L., Berthoz, S., Chevallier, C., Grèzes, J., 2016. Classification of autistic individuals and controls using cross-task characterization of fMRI activity. NeuroImage Clin. 10, 78–88. https://doi.org/10.1016/j.nicl.2015.11.010

Chen, C.P., Keown, C.L., Jahedi, A., Nair, A., Pflieger, M.E., Bailey, B.A., Müller, R.A., 2015. Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism. NeuroImage Clin. 8, 238–245. https://doi.org/10.1016/j.nicl.2015.04.002

Chen, Heng, Duan, X., Liu, F., Lu, F., Ma, X., Zhang, Y., Uddin, L.Q., Chen, Huafu, 2016. Multivariate classification of autism spectrum disorder using frequency-specific resting-state functional connectivity-A multi-center study. Prog. Neuro-Psychopharmacology Biol. Psychiatry 64, 1–9. https://doi.org/10.1016/j.pnpbp.2015.06.014

Cheng, H., Yu, J., Xu, L., Li, J., 2019. Power spectrum of spontaneous cerebral homodynamic oscillation shows a distinct pattern in autism spectrum disorder. Biomed. Opt. Express 10, 1383. https://doi.org/10.1364/boe.10.001383

Cholemkery, H., Medda, J., Lempp, T., Freitag, C.M., 2016. Classifying Autism Spectrum Disorders by ADI-R: Subtypes or Severity Gradient? J. Autism Dev. Disord. 46, 2327–2339. https://doi.org/10.1007/s10803-016-2760-2

Coleman, M., 2005. Advances in autism research. Dev. Med. Child Neurol. 47, 148. https://doi.org/10.1017/S0012162205000277

Collins, G.S., Reitsma, J.B., Altman, D.G., Moons, K.G.M., 2015. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur. Urol. 67, 1142–1151. https://doi.org/10.1016/j.eururo.2014.11.025

Crippa, A., Salvatore, C., Perego, P., Forti, S., Nobile, M., Molteni, M., Castiglioni, I., 2015. Use of Machine Learning to Identify Children with Autism and Their Motor

Abnormalities. J. Autism Dev. Disord. 45, 2146–2156. https://doi.org/10.1007/s10803-015-2379-8

Dale, A.M., Halgren, E., 2001. Spatiotemporal mapping of brain activity by integration of multiple imaging modalities. Curr. Opin. Neurobiol. 11, 202–208. https://doi.org/10.1016/S0959-4388(00)00197-5

Deshpande, G., Libero, L.E., Sreenivasan, K.R., Deshpande, H.D., Kana, R.K., 2013. Identification of neural connectivity signatures of autism using machine learning. Front. Hum. Neurosci. 7, 670. https://doi.org/10.3389/fnhum.2013.00670

Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D.A., Gallagher, L., Kennedy, D.P., Keown, C.L., Keysers, C., Lainhart, J.E., Lord, C., Luna, B., Menon, V., Minshew, N.J., Monk, C.S., Mueller, S., Müller, R.A., Nebel, M.B., Nigg, J.T., O'Hearn, K., Pelphrey, K.A., Peltier, S.J., Rudie, J.D., Sunaert, S., Thioux, M., Tyszka, J.M., Uddin, L.Q., Verhoeven, J.S., Wenderoth, N., Wiggins, J.L., Mostofsky, S.H., Milham, M.P., 2014. The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol. Psychiatry 19, 659–667. https://doi.org/10.1038/mp.2013.78

Dinga, R., Schmaal, L., Penninx, B.W.J.H., van Tol, M.J., Veltman, D.J., van Velzen, L., Mennes, M., van der Wee, N.J.A., Marquand, A.F., 2019. Evaluating the evidence for biotypes of depression: Methodological replication and extension of Drysdale et al. (2017). NeuroImage Clin. 22, 101796. https://doi.org/10.1016/j.nicl.2019.101796

Doshi-Velez, F., Ge, Y., Kohane, I., 2014. Comorbidity Clusters in Autism Spectrum Disorders: An Electronic Health Record Time-Series Analysis. Pediatrics 133, e54–e63. https://doi.org/10.1542/peds.2013-0819

Drysdale, A.T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R.N.,

Zebley, B., Oathes, D.J., Etkin, A., Schatzberg, A.F., Sudheimer, K., Keller, J., Mayberg, H.S., Gunning, F.M., Alexopoulos, G.S., Fox, M.D., Pascual-leone, A., Voss, H.U., Casey, B.J., Dubin, M.J., Liston, C., 2017. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat. Med. 23, 264. https://doi.org/10.1038/nm0217-264d

Duffy, F.H., Als, H., 2019. Autism, spectrum or clusters? An EEG coherence study. BMC Neurol. 19, 1–13. https://doi.org/10.1186/s12883-019-1254-1

Duffy, F.H., Als, H., 2012. A stable pattern of EEG spectral coherence distinguishes children with autism from neuro- typical controls - a large case control study A stable pattern of EEG spectral coherence distinguishes children with autism from neuro- typical controls - a large ca. BMC Med. 10, 64. https://doi.org/10.1186/1741-7015-10-64

Easson, A., Fatima, Z., McIntosh, A.R., 2018. Functional connectivity-based subtypes of individuals with and without autism spectrum disorder. Netw. Neurosci. 1–32. https://doi.org/10.1162/netn_a_00067

Ecker, C., Bookheimer, S.Y., Murphy, D.G.M., 2015. Neuroimaging in autism spectrum disorder: Brain structure and function across the lifespan. Lancet Neurol. 14, 1121–1134. https://doi.org/10.1016/S1474-4422(15)00050-2

Ecker, C., Marquand, A., Mourao-Miranda, J., Johnston, P., Daly, E.M., Brammer, M.J., Maltezos, S., Murphy, C.M., Robertson, D., Williams, S.C., Murphy, D.G.M., 2010a. Describing the Brain in Autism in Five Dimensions--Magnetic Resonance Imaging-Assisted Diagnosis of Autism Spectrum Disorder Using a Multiparameter Classification Approach. J. Neurosci. 30, 10612–10623. https://doi.org/10.1523/JNEUROSCI.5413-09.2010

Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Miranda, J., Marquand, A., Daly, E.M., Brammer, M.J., Murphy, C., Murphy, D.G., 2010b. Investigating the predictive value of

whole-brain structural MR scans in autism: A pattern classification approach. Neuroimage 49, 44–56. https://doi.org/10.1016/j.neuroimage.2009.08.024

Eldridge, J., Lane, A.E., Belkin, M., Dennis, S., 2014. Robust features for the automatic identification of autism spectrum disorder in children. J. Neurodev. Disord. 6, 1–12. https://doi.org/10.1186/1866-1955-6-12

Elsabbagh, M., Divan, G., Koh, Y.J., Kim, Y.S., Kauchali, S., Marcín, C., Montiel-Nava, C., Patel, V., Paula, C.S., Wang, C., Yasamy, M.T., Fombonne, E., 2012. Global Prevalence of Autism and Other Pervasive Developmental Disorders. Autism Res. 5, 160–179. https://doi.org/10.1002/aur.239

Emerson, R.W., Adams, C., Nishino, T., Hazlett, H.C., Wolff, J.J., Zwaigenbaum, L., Constantino, J.N., Shen, M.D., Swanson, M.R., Elison, J.T., Kandala, S., Estes, A.M., Botteron, K.N., Collins, L., Dager, S.R., Evans, A.C., Gerig, G., Gu, H., McKinstry, R.C., Paterson, S., Schultz, R.T., Styner, M., Infant Brain Imaging Study (IBIS) network, Schlaggar, B.L., Pruett Jr., J.R., Piven, J., Autism, 2017. Functional neuroimaging of high-risk 6- month-old infants predicts a diagnosis of autism at 24 months of age Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. Sci. Transl. Med. 2882, 1–8. https://doi.org/10.1126/scitranslmed.aag2882

Feczko, E., Balba, N.M., Miranda-Dominguez, O., Cordova, M., Karalunas, S.L.L., Irwin, L., Demeter, D.V. V., Hill, A.P.P., Langhorst, B.H.H., Grieser Painter, J., Van Santen, J., Fombonne, E.J.J., Nigg, J.L.T., Fair, D.A.A., 2018. Subtyping cognitive profiles in Autism Spectrum Disorder using a random forest algorithm. Neuroimage 172, 674–688. https://doi.org/10.1016/j.neuroimage.2017.12.044

Fountain, C., Winter, A.S., Bearman, P.S., 2012. Six Developmental Trajectories Characterize Children With Autism. Pediatrics 129, e1112–e1120.

https://doi.org/10.1542/peds.2011-1601

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, L., Bryson, S., Fombonne, E., Volden, J., Mirenda, P., Smith, I., Roberts, W., Vaillancourt, T., Waddell, C., Bennett, T., Thompson, A., 2013. Investigating phenotypic heterogeneity in children with autism spectrum disorder: A factor mixture modeling approach. J. Child Psychol. Psychiatry Allied Discip. 54, 206–215. https://doi.org/10.1111/j.1469-7610.2012.02588.x

Ghafouri-Fard, S., Taheri, M., Omrani, M.D., Daaee, A., Mohammad-Rahimi, H., Kazazi, H., 2019. Application of Single-Nucleotide Polymorphisms in the Diagnosis of Autism Spectrum Disorders: A Preliminary Study with Artificial Neural Networks. J. Mol. Neurosci. 1–7. https://doi.org/10.1007/s12031-019-01311-1

Ghiassian, S., Greiner, R., Jin, P., Brown, M.R.G., 2016. Using Functional or Structural Magnetic Resonance Images and Personal Characteristic Data to Identify ADHD and Autism. PLoS One 11, e0166934. https://doi.org/10.1371/journal.pone.0166934

Gori, I., Giuliano, A., Muratori, F., Saviozzi, I., Oliva, P., Tancredi, R., Cosenza, A., Tosetti, M., Calderoni, S., Retico, A., 2015. Gray Matter Alterations in Young Children with Autism Spectrum Disorders: Comparing Morphometry at the Voxel and Regional Level. J. Neuroimaging 25, 866–874. https://doi.org/10.1111/jon.12280

Grossi, E., Olivieri, C., Buscema, M., 2017. Computer Methods and Programs in Biomedicine Diagnosis of autism through EEG processed by advanced computational algorithms : A pilot study. Comput. Methods Programs Biomed. 142, 73–79. https://doi.org/10.1016/j.cmpb.2017.02.002

Groves, A.R., Beckmann, C.F., Smith, S.M., Woolrich, M.W., 2011. Linked independent component analysis for multimodal data fusion. Neuroimage 54, 2198–2217. https://doi.org/10.1016/j.neuroimage.2010.09.073

Guo, X., Dominick, K.C., Minai, A.A., Li, H., Erickson, C.A., Lu, L.J., 2017. Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method. Front. Neurosci. 11, 1–19. https://doi.org/10.3389/fnins.2017.00460

Hahn, T., Nierenberg, A.A., Whitfield-Gabrieli, S., 2017. Predictive analytics in mental health: Applications, guidelines, challenges and perspectives. Mol. Psychiatry 22, 37–43. https://doi.org/10.1038/mp.2016.201

Hartigan, J.A., Wong, M.A., 1979. A K-Means Clustering Algorithm. J. R. Stat. Soc. 28, 100–108.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, 2nd ed. Springer. https://doi.org/10.1007/b94608

Hazlett, H.C., Gu, H., Munsell, B.C., Kim, S.H., Styner, M., Wolff, J.J., Elison, J.T., Swanson, M.R., Zhu, H., Botteron, K.N., Collins, D.L., Constantino, J.N., Dager, S.R., Estes, A.M., Evans, A.C., Fonov, V.S., Gerig, G., Kostopoulos, P., McKinstry, R.C., Pandey, J., Paterson, S., Pruett, J.R., Schultz, R.T., Shaw, D.W., Zwaigenbaum, L., Piven, J., 2017. Early brain development in infants at high risk for autism spectrum disorder. Nature 542, 348–351. https://doi.org/10.1038/nature21369

Heinsfeld, A.S., Franco, A.R., Craddock, R.C., Buchweitz, A., Meneguzzi, F., Sólon, A., Rosa, A., Craddock, R.C., Buchweitz, A., Meneguzzi, F., 2018. NeuroImage: Clinical Identification of autism spectrum disorder using deep learning and the ABIDE dataset. NeuroImage Clin. 17, 16–23. https://doi.org/10.1016/j.nicl.2017.08.017

Hofree, M., Shen, J.P., Carter, H., Gross, A., Ideker, T., 2013. Network-based stratification of tumor mutations. Nat. Methods 10, 1108–1118. https://doi.org/10.1038/nmeth.2651

Hong, S.-J., Valk, S.L., Di Martino, A., Milham, M.P., Bernhardt, B.C., 2017. Multidimensional Neuroanatomical Subtyping of Autism Spectrum Disorder. Cereb.

Cortex 1–11. https://doi.org/10.1093/cercor/bhx229

Hu, V.W., Steinberg, M.E., 2009. Novel clustering of items from the autism diagnostic interview-revised to define phenotypes within autism spectrum disorders. Autism Res. 2, 67–77. https://doi.org/10.1002/aur.72

Huys, Q.J.M., Maia, T. V., Frank, M.J., 2016. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat. Neurosci. 19, 404–413. https://doi.org/10.1038/nn.4238

Iidaka, T., 2015. Resting state functional magnetic resonance imaging and neural network classified autism and control. Cortex 63, 55–67. https://doi.org/10.1016/j.cortex.2014.08.011

Ingalhalikar, M., Parker, D., Bloy, L., Roberts, T.P.L., Verma, R., 2011. Diffusion based abnormality markers of pathology: Toward learned diagnostic prediction of ASD. Neuroimage 57, 918–927. https://doi.org/10.1016/j.neuroimage.2011.05.023

Ingalhalikar, M., Parker, W.A., Bloy, L., Roberts, T.P.L., Verma, R., 2014. Creating multimodal predictors using missing data: Classifying and subtyping autism spectrum disorder. J. Neurosci. Methods 235, 1–9. https://doi.org/10.1016/j.jneumeth.2014.06.030

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., Sanislow, C., Wang, P., 2010. Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. Am. J. Psychiatry 167, 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379

Jahedi, A., Nasamran, C.A., Faires, B., Fan, J., Müller, R.-A., 2017. Distributed Intrinsic Functional Connectivity Patterns Predict Diagnostic Status in Large Autism Cohort. Brain Connect. 7, 515–525. https://doi.org/10.1089/brain.2017.0496

Jiao, Y., Chen, R., Ke, X., Chu, K., Lu, Z., Herskovits, E.H., 2010. Predictive models of autism spectrum disorder based on brain regional cortical thickness. Neuroimage 50,

589–599. https://doi.org/10.1016/j.neuroimage.2009.12.047

Jobson, J.D., 1992. Cluster Analysis and Multidimensional Scaling, in: Applied Multivariate Data Analysis. https://doi.org/10.1007/978-1-4612-0921-8_5

Jun, E., Kang, E., Choi, J., Suk, H. Il, 2019. Modeling regional dynamics in low-frequency fluctuation and its application to Autism spectrum disorder diagnosis. Neuroimage 184, 669–686. https://doi.org/10.1016/j.neuroimage.2018.09.043

Just, M.A., Cherkassky, V.L., Buchweitz, A., Keller, T.A., Mitchell, T.M., 2014. Identifying autism from neural representations of social interactions: Neurocognitive markers of autism. PLoS One 9, 1–22. https://doi.org/10.1371/journal.pone.0113879

Kazeminejad, A., Sotero, R.C., 2019. Topological Properties of Resting-State fMRI Functional Networks Improve Machine Learning-Based Autism Classification. Front. Neurosci. 12, 1–10. https://doi.org/10.3389/fnins.2018.01018

Kim, S.H., Macari, S., Koller, J., Chawarska, K., 2016. Examining the phenotypic heterogeneity of early autism spectrum disorder: Subtypes and short-term outcomes. J. Child Psychol. Psychiatry Allied Discip. 57, 93–102. https://doi.org/10.1111/jcpp.12448

Kim, Y.M., Poline, J.B., Dumas, G., 2018. Experimenting with reproducibility: A case study of robustness in bioinformatics. Gigascience 7, 1–8. https://doi.org/10.1093/gigascience/giy077

Lane, A.E., Young, R.L., Baker, A.E.Z., Angley, M.T., 2010. Sensory processing subtypes in autism: Association with adaptive behavior. J. Autism Dev. Disord. 40, 112–122. https://doi.org/10.1007/s10803-009-0840-2

LeClerc, S., Easley, D., 2015. Pharmacological therapies for autism spectrum disorder: a review. P T 40, 389–97.

Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Lee, S.H., Ripke, S., Neale, B.M., Faraone, S. V., Purcell, S.M., Perlis, R.H., Mowry, B.J., Thapar, A., Goddard, M.E., Witte, J.S., Absher, D., Agartz, I., Akil, H., Amin, F., Andreassen, O.A., Anjorin, A., Anney, R., Anttila, V., Arking, D.E., Asherson, P., Azevedo, M.H., Backlund, L., Badner, J.A., Bailey, A.J., Banaschewski, T., Barchas, J.D., Barnes, M.R., Barrett, T.B., Bass, N., Battaglia, A., Bauer, M., Bayés, M., Bellivier, F., Bergen, S.E., Berrettini, W., Betancur, C., Bettecken, T., Biederman, J., Binder, E.B., Black, D.W., Blackwood, D.H.R., Bloss, C.S., Boehnke, M., Boomsma, D.I., Breen, G., Breuer, R., Bruggeman, R., Cormican, P., Buccola, N.G., Buitelaar, J.K., Bunney, W.E., Buxbaum, J.D., Byerley, W.F., Byrne, E.M., Caesar, S., Cahn, W., Cantor, R.M., Casas, M., Chakravarti, A., Chambert, K., Choudhury, K., Cichon, S., Cloninger, C.R., Collier, D.A., Cook, E.H., Coon, H., Cormand, B., Corvin, A., Coryell, W.H., Craig, D.W., Craig, I.W., Crosbie, J., Cuccaro, M.L., Curtis, D., Czamara, D., Datta, S., Dawson, G., Day, R., De Geus, E.J., Degenhardt, F., Djurovic, S., Donohoe, G.J., Doyle, A.E., Duan, J., Dudbridge, F., Duketis, E., Ebstein, R.P., Edenberg, H.J., Elia, J., Ennis, S., Etain, B., Fanous, A., Farmer, A.E., Ferrier, I.N., Flickinger, M., Fombonne, E., Foroud, T., Frank, J., Franke, B., Fraser, C., Freedman, R., Freimer, N.B., Freitag, C.M., Friedl, M., Frisén, L., Gallagher, L., Gejman, P. V., Georgieva, L., Gershon, E.S., Geschwind, D.H., Giegling, I., Gill, M., Gordon, S.D., Gordon-Smith, K., Green, E.K., Greenwood, T.A., Grice, D.E., Gross, M., Grozeva, D., Guan, W., Gurling, H., De Haan, L., Haines, J.L., Hakonarson, H., Hallmayer, J., Hamilton, S.P., Hamshere, M.L., Hansen, T.F., Hartmann, A.M., Hautzinger, M., Heath, A.C., Henders, A.K., Herms, S., Hickie, I.B., Hipolito, M., Hoefels, S., Holmans, P.A., Holsboer, F., Hoogendijk, W.J., Hottenga, J.J., Hultman, C.M., Hus, V., Ingason, A., Ising, M., Jamain, S., Jones, E.G., Jones, I., Jones, L., Tzeng, J.Y., Kähler, A.K., Kahn, R.S., Kandaswamy, R., Keller, M.C., Kennedy, J.L., Kenny, E., Kent, L., Kim, Y., Kirov,

G.K., Klauck, S.M., Klei, L., Knowles, J.A., Kohli, M.A., Koller, D.L., Konte, B., Korszun, A., Krabbendam, L., Krasucki, R., Kuntsi, J., Kwan, P., Landén, M., Långström, N., Lathrop, M., Lawrence, J., Lawson, W.B., Leboyer, M., Ledbetter, D.H., Lee, P.H., Lencz, T., Lesch, K.P., Levinson, D.F., Lewis, C.M., Li, J., Lichtenstein, P., Lieberman, J.A., Lin, D.Y., Linszen, D.H., Liu, C., Lohoff, F.W., Loo, S.K., Lord, C., Lowe, J.K., Lucae, S., Macintyre, D.J., Madden, P.A.F., Maestrini, E., Magnusson, P.K.E., Mahon, P.B., Maier, W., Malhotra, A.K., Mane, S.M., Martin, C.L., Martin, N.G., Mattheisen, M., Matthews, K., Mattingsdal, M., Mccarroll, S.A., Mcghee, K.A., Mcgough, J.J., Mcgrath, P.J., Mcguffin, P., Mcinnis, M.G., Mcintosh, A., Mckinney, R., Mclean, A.W., Mcmahon, F.J., Mcmahon, W.M., Mcquillin, A., Medeiros, H., Medland, S.E., Meier, S., Melle, I., Meng, F., Meyer, J., Middeldorp, C.M., Middleton, L., Milanova, V., Miranda, A., Monaco, A.P., Montgomery, G.W., Moran, J.L., Moreno-De-Luca, D., Morken, G., Morris, D.W., Morrow, E.M., Moskvina, V., Muglia, P., Mühleisen, T.W., Muir, W.J., Müller-Myhsok, B., Murtha, M., Myers, R.M., Myin-Germeys, I., Neale, M.C., Nelson, S.F., Nievergelt, C.M., Nikolov, I., Nimgaonkar, V., Nolen, W.A., Nöthen, M.M., Nurnberger, J.I., Nwulia, E.A., Nyholt, D.R., O'dushlaine, C., Oades, R.D., Olincy, A., Oliveira, G., Olsen, L., Ophoff, R.A., Osby, U., Owen, M.J., Palotie, A., Parr, J.R., Paterson, A.D., Pato, C.N., Pato, M.T., Penninx, B.W., Pergadia, M.L., Pericak-Vance, M.A., Pickard, B.S., Pimm, J., Piven, J., Posthuma, D., Potash, J.B., Poustka, F., Propping, P., Puri, V., Quested, D.J., Quinn, E.M., Ramos-Quiroga, J.A., Rasmussen, H.B., Raychaudhuri, S., Rehnström, K., Reif, A., Ribasés, M., Rice, J.P., Rietschel, M., Roeder, K., Roeyers, H., Rossin, L., Rothenberger, A., Rouleau, G., Ruderfer, D., Rujescu, D., Sanders, A.R., Sanders, S.J., Santangelo, S.L., Sergeant, J.A., Schachar, R., Schalling, M., Schatzberg, A.F., Scheftner, W.A., Schellenberg, G.D., Scherer, S.W., Schork, N.J., Schulze, T.G., Schumacher, J.,

Schwarz, M., Scolnick, E., Scott, L.J., Shi, J., Shilling, P.D., Shyn, S.I., Silverman, J.M., Slager, S.L., Smalley, S.L., Smit, J.H., Smith, E.N., Sonuga-Barke, E.J.S., St. Clair, D., State, M., Steffens, M., Steinhausen, H.C., Strauss, J.S., Strohmaier, J., Stroup, T.S., Sutcliffe, J.S., Szatmari, P., Szelinger, S., Thirumalai, S., Thompson, R.C., Todorov, A.A., Tozzi, F., Treutlein, J., Uhr, M., Van Den Oord, E.J.C.G., Van Grootheest, G., Van Os, J., Vicente, A.M., Vieland, V.J., Vincent, J.B., Visscher, P.M., Walsh, C.A., Wassink, T.H., Watson, S.J., Weissman, M.M., Werge, T., Wienker, T.F., Wijsman, E.M., Willemsen, G., Williams, N., Willsey, A.J., Witt, S.H., Xu, W., Young, A.H., Yu, T.W., Zammit, S., Zandi, P.P., Zhang, P., Zitman, F.G., Zöllner, S., Devlin, B., Kelsoe, J.R., Sklar, P., Daly, M.J., O'donovan, M.C., Craddock, N., Sullivan, P.F., Smoller, J.W., Kendler, K.S., Wray, N.R., 2013. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat. Genet. 45, 984–994. https://doi.org/10.1038/ng.2711

Li, B., Sharma, A., Meng, J., Purushwalkam, S., Gowen, E., 2017. Applying machine learning to identify autistic adults using imitation: An exploratory study. PLoS One 12, 1–19. https://doi.org/10.1371/journal.pone.0182652

Libero, L.E., DeRamus, T.P., Lahti, A.C., Deshpande, G., Kana, R.K., 2015. Multimodal neuroimaging based classification of autism spectrum disorder using anatomical, neurochemical, and white matter correlates. Cortex 66, 46–59. https://doi.org/10.1016/j.cortex.2015.02.008

Lim, L., Marquand, A., Cubillo, A.A., Smith, A.B., Chantiluke, K., Simmons, A., Mehta, M., Rubia, K., 2013. Disorder-Specific Predictive Classification of Adolescents with Attention Deficit Hyperactivity Disorder (ADHD) Relative to Autism Using Structural Magnetic Resonance Imaging. PLoS One 8, e63660. https://doi.org/10.1371/journal.pone.0063660

Liu, W., Li, M., Yi, L., 2016. Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. Autism Res. 9, 888–898. https://doi.org/10.1002/aur.1615

Liu, Y., Hayes, D.N., Nobel, A., Marron, J.S., 2008. Statistical significance of clustering for high-dimension, low-sample size data. J. Am. Stat. Assoc. 103, 1281–1293. https://doi.org/10.1198/016214508000000454

Lombardo, M. V., Lai, M.C., Auyeung, B., Holt, R.J., Allison, C., Smith, P., Chakrabarti, B., Ruigrok, A.N.V., Suckling, J., Bullmore, E.T., Ecker, C., Craig, M.C., Murphy, D.G.M., Happé, F., Baron-Cohen, S., 2016. Unsupervised data-driven stratification of mentalizing heterogeneity in autism. Sci. Rep. 6, 1–15. https://doi.org/10.1038/srep35333

Lombardo, M. V., Lai, M.C., Baron-Cohen, S., 2019. Big data approaches to decomposing heterogeneity across the autism spectrum. Mol. Psychiatry. https://doi.org/10.1038/s41380-018-0321-0

Loth, E., Charman, T., Mason, L., Tillmann, J., Jones, E.J.H., Wooldridge, C., Ahmad, J., Auyeung, B., Brogna, C., Ambrosino, S., Banaschewski, T., Baron-Cohen, S., Baumeister, S., Beckmann, C., Brammer, M., Brandeis, D., Bölte, S., Bourgeron, T., Bours, C., De Bruijn, Y., Chakrabarti, B., Crawley, D., Cornelissen, I., Acqua, F.D., Dumas, G., Durston, S., Ecker, C., Faulkner, J., Frouin, V., Garces, P., Goyard, D., Hayward, H., Ham, L.M., Hipp, J., Holt, R.J., Johnson, M.H., Isaksson, J., Kundu, P., Lai, M.C., D'Ardhuy, X.L., Lombardo, M. V., Lythgoe, D.J., Mandl, R., Meyer-Lindenberg, A., Moessnang, C., Mueller, N., O'Dwyer, L., Oldehinkel, M., Oranje, B., Pandina, G., Persico, A.M., Ruigrok, A.N.V., Ruggeri, B., Sabet, J., Sacco, R., Cáceres, A.S.J., Simonoff, E., Toro, R., Tost, H., Waldman, J., Williams, S.C.R., Zwiers, M.P., Spooren, W., Murphy, D.G.M., Buitelaar, J.K., 2017. The EU-AIMS Longitudinal

European Autism Project (LEAP): Design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. Mol. Autism 8, 1–19. https://doi.org/10.1186/s13229-017-0146-8

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations, in: In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. pp. 281–297. https://doi.org/10.1007/s11665-016-2173-6

Marquand, A.F., Kia, S.M., Zabihi, M., Wolfers, T., Buitelaar, J.K., Beckmann, C.F., 2019. Conceptualizing mental disorders as deviations from normative functioning. Mol. Psychiatry. https://doi.org/10.1038/s41380-019-0441-1

Marquand, A.F., Rezek, I., Buitelaar, J., Beckmann, C.F., 2016a. Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. Biol. Psychiatry 80, 552–561. https://doi.org/10.1016/j.biopsych.2015.12.023

Marquand, A.F., Wolfers, T., Mennes, M., Buitelaar, J., Beckmann, C.F., 2016b. Beyond Lumping and Splitting: A Review of Computational Approaches for Stratifying Psychiatric Disorders. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 1, 433–447. https://doi.org/10.1016/j.bpsc.2016.04.002

Masi, A., DeMayo, M.M., Glozier, N., Guastella, A.J., 2017. An Overview of Autism Spectrum Disorder, Heterogeneity and Treatment Options. Neurosci. Bull. 33, 183–193. https://doi.org/10.1007/s12264-017-0100-y

Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L.R., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19, 1523–1536. https://doi.org/10.1038/nn.4393

Munson, J., Faja, S., Meltzoff, A., Abbott, R., Dawson, G., 2008. Neurocognitive predictors of social and communicative developmental trajectories in preschoolers with autism spectrum disorders. J. Int. Neuropsychol. Soc. 14, 956–66. https://doi.org/10.1017/S1355617708081393

Murdaugh, D.L., Shinkareva, S. V., Deshpande, H.R., Wang, J., Pennick, M.R., Kana, R.K., 2012. Differential Deactivation during Mentalizing and Classification of Autism Based on Default Mode Network Connectivity. PLoS One 7, e50064. https://doi.org/10.1371/journal.pone.0050064

Orrù, G., Pettersson-Yeo, W., Marquand, A.F., Sartori, G., Mechelli, A., 2012. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. Neurosci. Biobehav. Rev. 36, 1140–1152. https://doi.org/10.1016/j.neubiorev.2012.01.004

Parikh, M.N., Li, H., He, L., 2019. Enhancing Diagnosis of Autism With Optimized Machine Learning Models and Personal Characteristic Data. Front. Comput. Neurosci. 13, 1–5. https://doi.org/10.3389/fncom.2019.00009

Payabvash, S., Palacios, E., Owen, J.P., Wang, M.B., Tavassoli, T., Gerdes, M.R., Brandes Aitken, A., Cuneo, D., Marco, E., Mukherjee, P., 2019. White Matter Connectome Edge Density in Children with Autism Spectrum Disorders: Potential Imaging Biomarkers Using Machine Learning Models. Brain Connect. 9, 209–220. https://doi.org/10.1089/brain.2018.0658

Plitt, M., Anne, K., Martin, A., 2015. NeuroImage : Clinical Functional connectivity classi fi cation of autism identi fi es highly predictive brain features but falls short of biomarker standards. Ynicl 7, 359–366. https://doi.org/10.1016/j.nicl.2014.12.013

Rapin, I., Dunn, M.A., Allen, D.A., Stevens, M.C., Fein, D., 2009. Subtypes of language disorders in school-age children with autism. Dev. Neuropsychol. 34, 66–84.

https://doi.org/10.1080/87565640802564648

Rosenblatt, F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. Psychol. Rev. https://doi.org/10.1037/h0042519

Sabuncu, M.R., Konukoglu, E., 2014. Clinical Prediction from Structural Brain MRI Scans: A Large-Scale Empirical Study. Neuroinformatics 13, 31–46. https://doi.org/10.1007/s12021-014-9238-1

Sacco, R., Lenti, C., Saccani, M., Curatolo, P., Manzi, B., Bravaccio, C., Persico, A.M., 2012. Cluster analysis of autistic patients based on principal pathogenetic components. Autism Res. 5, 137–147. https://doi.org/10.1002/aur.1226

Sadeghi, M., Khosrowabadi, R., Bakouie, F., Mahdavi, H., Eslahchi, C., Pouretemad, H., 2017. Screening of autism based on task-free fMRI using graph theoretical approach. Psychiatry Res. - Neuroimaging 263, 48–56. https://doi.org/10.1016/j.pscychresns.2017.02.004

Sanders, S.J., 2015. First glimpses of the neurobiology of autism spectrum disorder. Curr. Opin. Genet. Dev. 33, 80–92. https://doi.org/10.1016/j.gde.2015.10.002

Schwenker, F., Trentin, E., 2014. Pattern classification and clustering: A review of partially supervised learning approaches. Pattern Recognit. Lett. 37, 4–14. https://doi.org/10.1016/j.patrec.2013.10.017

Segovia, F., Holt, R., Spencer, M., GÃ³rriz, J.M., RamÃ-rez, J., Puntonet, C.G., Phillips, C., Chura, L., Baron-Cohen, S., Suckling, J., 2014. Identifying endophenotypes of autism: a multivariate approach. Front. Comput. Neurosci. 8, 1–8. https://doi.org/10.3389/fncom.2014.00060

Sen, B., Borle, N.C., Greiner, R., Brown, M.R.G., 2018. A general prediction model for the detection of ADHD and Autism using structural and functional MRI. PLoS One 13, e0194856. https://doi.org/10.1371/journal.pone.0194856

Shen, M.D., Kim, S.H., McKinstry, R.C., Gu, H., Hazlett, H.C., Nordahl, C.W., Emerson, R.W., Shaw, D., Elison, J.T., Swanson, M.R., Fonov, V., Gerig, G., Dager, S., Botteron, K.N., Paterson, S., Schultz, R.T., Evans, A.C., Estes, A., Zwaigenbaum, L., Styner, M.A., Amaral, D.G., Piven, J., Chappell, C., Botteron, K., McKinstry, R., Constantino, J., Pruett, J., Schultz, R., Elison, J., Collins, D.L., Pike, G.B., Kostopoulos, P., Das, S., Styner, M., 2017. Increased Extra-axial Cerebrospinal Fluid in High-Risk Infants Who Later Develop Autism. Biol. Psychiatry 82, 186–193. https://doi.org/10.1016/j.biopsych.2017.02.1095

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Driessche, G. Van Den, Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., 2016. Mastering the game of Go with deep neural networks and tree search. Nature 529, 484–489. https://doi.org/10.1038/nature16961

Simões, M., Monteiro, R., Andrade, J., Mouga, S., França, F., Oliveira, G., Carvalho, P., Castelo-Branco, M., 2018. A novel biomarker of compensatory recruitment of face emotional imagery networks in autism spectrum disorder. Front. Neurosci. 12, 1–15. https://doi.org/10.3389/fnins.2018.00791

Simonoff, E., Pickels, A., Charman, T., Chandler, S., Loucas, T., Baird, G., 2008. Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. J. Am. Acad. Child Adolecent Psychiatry 55, 104–112. https://doi.org/10.1097/CHI.0b013e318179964f

Soussia, M., Rekik, I., 2018. Unsupervised Manifold Learning Using High-Order Morphological Brain Networks Derived From T1-w MRI for Autism Diagnosis. Front. Neuroinform. 12, 1–10. https://doi.org/10.3389/fninf.2018.00070

Steyerberg, E.W., Harrell, F.E., Borsboom, G.J.J.M., Eijkemans, M.J.C.R., Vergouwe, Y.,

Habbema, J.D.F., 2007. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. Proc. ICED 2007, 16th Int. Conf. Eng. Des. DS 42, 774–781. https://doi.org/10.1016/S0895-4356(01)00341-9

Subbaraju, V., Suresh, M.B., Sundaram, S., Narasimhan, S., 2017. Identifying differences in brain activities and an accurate detection of autism spectrum disorder using resting state functional-magnetic resonance imaging : A spatial filtering approach. Med. Image Anal. 35, 375–389. https://doi.org/10.1016/j.media.2016.08.003

Szatmari, P., Paterson, A.D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X.Q., Vincent, J.B., Skaug, J.L., Thompson, A.P., Senman, L., Feuk, L., Qian, C., Bryson, S.E., Jones, M.B., Marshall, C.R., Scherer, S.W., Vieland, V.J., Bartlett, C., Mangin, L.V., Goedken, R., Segre, A., Pericak-Vance, M.A., Cuccaro, M.L., Gilbert, J.R., Wright, H.H., Abramson, R.K., Betancur, C., Bourgeron, T., Gillberg, C., Leboyer, M., Buxbaum, J.D., Davis, K.L., Hollander, E., Silverman, J.M., Hallmayer, J., Lotspeich, L., Sutcliffe, J.S., Haines, J.L., Folstein, S.E., Piven, J., Wassink, T.H., Sheffield, V., Geschwind, D.H., Bucan, M., Brown, W.T., Cantor, R.M., Constantino, J.N., Gilliam, T.C., Herbert, M., LaJonchere, C., Ledbetter, D.H., Lese-Martin, C., Miller, J., Nelson, S., Samango-Sprouse, C.A., Spence, S., State, M., Tanzi, R.E., Coon, H., Dawson, G., Devlin, B., Estes, A., Flodman, P., Klei, L., McMahon, W.M., Minshew, N., Munson, J., Korvatska, E., Rodier, P.M., Schellenberg, G.D., Smith, M., Spence, M.A., Stodgell, C., Tepper, P.G., Wijsman, E.M., Yu, C.E., Rogé, B., Mantoulan, C., Wittemeyer, K., Poustka, A., Felder, B., Klauck, S.M., Schuster, C., Poustka, F., Bölte, S., Feineis-Matthews, S., Herbrecht, E., Schmötzer, G., Tsiantis, J., Papanikolaou, K., Maestrini, E., Bacchelli, E., Blasi, F., Carone, S., Toma, C., Van Engeland, H., De Jonge, M., Kemner, C., Koop, F., Langemeijer, M., Hijimans, C., Staal, W.G., Baird, G., Bolton, P.F., Rutter, M.L., Weisblatt, E., Green, J., Aldred, C., Wilkinson, J.A., Pickles, A., Le Couteur, A.,

Berney, T., McConachie, H., Bailey, A.J., Francis, K., Honeyman, G., Hutchinson, A., Parr, J.R., Wallace, S., Monaco, A.P., Barnby, G., Kobayashi, K., Lamb, J.A., Sousa, I., Sykes, N., Cook, E.H., Guter, S.J., Leventhal, B.L., Salt, J., Lord, C., Corsello, C., Hus, V., Weeks, D.E., Volkmar, F., Tauber, M., Fombonne, E., Shih, A., 2007. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. Nat. Genet. 39, 319–328. https://doi.org/10.1038/ng1985

Tang, S., Sun, N., Floris, D.L., Zhang, X., Di Martino, A., Yeo, B.T.T., 2018. Reconciling Dimensional and Categorical Models of Autism Heterogeneity: a Brain Connectomics & Behavioral Study. Under Rev. 1–32.

Tariq, Q., Daniels, J., Schwartz, J.N., Washington, P., Kalantarian, H., Wall, D.P., 2018. Mobile detection of autism through machine learning on home video: A development and prospective validation study. PLoS Med. 15, 1–20. https://doi.org/10.1371/journal.pmed.1002705

Tomchek, S.D., Little, L.M., Myers, J., Dunn, W., 2018. Sensory Subtypes in Preschool Aged Children with Autism Spectrum Disorder. J. Autism Dev. Disord. 48, 2139–2147. https://doi.org/10.1007/s10803-018-3468-2

Uddin, L.Q., Menon, V., Young, C.B., Ryali, S., Chen, T., Khouzam, A., Minshew, N.J., Hardan, A.Y., 2011. Multivariate searchlight classification of structural magnetic resonance imaging in children and adolescents with autism. Biol. Psychiatry 70, 833–841. https://doi.org/10.1016/j.biopsych.2011.07.014

Uddin, L.Q., Supekar, K., Lynch, C.J., Khouzam, A., Phillips, J., Feinstein, C., Ryali, S., Menon, V., 2013. Salience network-based classification and prediction of symptom severity in children with autism. JAMA Psychiatry 70, 869–879. https://doi.org/10.1001/jamapsychiatry.2013.104

Van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G.F.,

Calderoni, S., Daly, E., Deruelle, C., Di Martino, A., Dinstein, I., Duran, F.L.S., Durston, S., Ecker, C., Fair, D., Fedor, J., Fitzgerald, J., Freitag, C.M., Gallagher, L., Gori, I., Haar, S., Hoekstra, L., Jahanshad, N., Jalbrzikowski, M., Janssen, J., Lerch, J., Luna, B., Martinho, M.M., McGrath, J., Muratori, F., Murphy, C.M., Murphy, D.G.M., O'Hearn, K., Oranje, B., Parellada, M., Retico, A., Rosa, P., Rubia, K., Shook, D., Taylor, M., Thompson, P.M., Tosetti, M., Wallace, G.L., Zhou, F., Buitelaar, J.K., 2018. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: Results from the ENIGMA ASD working group. Am. J. Psychiatry 175, 359–369. https://doi.org/10.1176/appi.ajp.2017.17010100

Varoquaux, G., 2017. Cross-validation failure: Small sample sizes lead to large error bars. Neuroimage 1–10. https://doi.org/10.1016/j.neuroimage.2017.06.061

Veatch, O.J., Veenstra-Vanderweele, J., Potter, M., Pericak-Vance, M.A., Haines, J.L., 2014. Genetically meaningful phenotypic subgroups in autism spectrum disorders. Genes, Brain Behav. 13, 276–285. https://doi.org/10.1111/gbb.12117

Vermunt, J.K., Magidson, J., 2002. Latent Class Cluster Analysis, in: Applied Latent Class Analysis. https://doi.org/10.1186/1472-6947-9-47

Vieira, S., Pinaya, W.H.L., Mechelli, A., 2017. Neuroscience and Biobehavioral Reviews Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders : Methods and applications. Neurosci. Biobehav. Rev. 74, 58–75. https://doi.org/10.1016/j.neubiorev.2017.01.002

von Luxburg, U., Williamson BobWilliamson, R.C., Guyon, I., 2012. Clustering: Science or Art? JMLR Work. Conf. Proc. 27, 6579.

Wan, G., Kong, X., Sun, B., Yu, S., Tu, Y., Park, J., Lang, C., Koh, M., Wei, Z., Feng, Z., Lin, Y., Kong, J., 2018. Applying Eye Tracking to Identify Autism Spectrum Disorder

in Children. J. Autism Dev. Disord. 0, 0. https://doi.org/10.1007/s10803-018-3690-y

Wang, H., Chen, C., Fushing, H., 2012. Extracting Multiscale Pattern Information of fMRI Based Functional Brain Connectivity with Application on Classification of Autism Spectrum Disorders. PLoS One 7, 1–14. https://doi.org/10.1371/journal.pone.0045502

Wang, J., Wang, Q., Peng, J., Nie, D., Zhao, F., Kim, M., Zhang, H., Wee, C.Y., Wang, S., Shen, D., 2017. Multi-task diagnosis for autism spectrum disorders using multi-modality features: A multi-center study. Hum. Brain Mapp. 38, 3081–3097. https://doi.org/10.1002/hbm.23575

Wee, C.-Y., Wang, L., Shi, F., Yap, P.-T., Shen, D., 2014. Diagnosis of autism spectrum disorders using regional and interregional morphological features. Hum. Brain Mapp. 35, 3414–3430. https://doi.org/10.1002/hbm.22411

Weinberger, D.R., Goldberg, T.E., 2014. RDoCs redux. World Psychiatry 13, 36–38. https://doi.org/10.1002/wps.20096

Wolfers, T., Beckmann, C.F., Hoogman, M., Buitelaar, J.K., Franke, B., Marquand, A.F., 2019. Individual differences v. The average patient: Mapping the heterogeneity in ADHD using normative models. Psychol. Med. 1–10. https://doi.org/10.1017/S0033291719000084

Wolfers, T., Buitelaar, J.K., Beckmann, C.F., Franke, B., Marquand, A.F., 2015. From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. Neurosci. Biobehav. Rev. 57, 328–349. https://doi.org/10.1016/j.neubiorev.2015.08.001

Wolfers, T., Doan, N.T., Kaufmann, T., Alnæs, D., Moberget, T., Agartz, I., Buitelaar, J.K., Ueland, T., Melle, I., Franke, B., Andreassen, O.A., Beckmann, C.F., Westlye, L.T., Marquand, A.F., 2018. Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using Normative Models. JAMA Psychiatry 11, 1146–1155.

https://doi.org/10.1001/JAMAPSYCHIATRY.2018.2467

Wolfers, T., Llera, A., Onnink, A.M.H., Dammers, J., Hoogman, M., Zwiers, M.P., Buitelaar, J.K., Franke, B., Marquand, A.F., Beckmann, C.F., 2017. Refinement by integration: Aggregated effects of multimodal imaging markers on adult ADHD. J. Psychiatry Neurosci. 42, 386–394. https://doi.org/10.1503/jpn.160240

Xiao, X., Fang, H., Wu, J., Xiao, C.Y., Xiao, T., Qian, L., Liang, F.J., Xiao, Z., Chu, K.K., Ke, X., 2017. Diagnostic model generated by MRI-derived brain features in toddlers with autism spectrum disorder. Autism Res. 10, 620–630. https://doi.org/10.1002/aur.1711

Xu, D., Tian, Y., 2015. A Comprehensive Survey of Clustering Algorithms. Ann. Data Sci. 2, 165–193. https://doi.org/10.1007/s40745-015-0040-1

Xu, R., Wunsch, D.C., 2005. survey of clustering algorithms. IEEE Trans. Neural Networks 16, 645–678.

Yahata, N., Morimoto, J., Hashimoto, R., Lisi, G., Shibata, K., Kawakubo, Y., Kuwabara, H., Kuroda, M., Yamada, T., Megumi, F., Imamizu, H., Nanez, J.E., Takahashi, H., Okamoto, Y., Kasai, K., Kato, N., Sasaki, Y., Watanabe, T., Kawato, M., 2016. A small number of abnormal brain connections predicts adult autism spectrum disorder. Nat. Commun. 1–12. https://doi.org/10.1038/pj.2016.37

Zabihi, M., Oldehinkel, M., Wolfers, T., Frouin, V., Goyard, D., Loth, E., Charman, T., Tillmann, J., Banaschewski, T., Dumas, G., Holt, R., Baron-Cohen, S., Durston, S., Bölte, S., Murphy, D., Ecker, C., Buitelaar, J.K., Beckmann, C.F., Marquand, A.F., 2019. Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using Normative Models. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 1–12. https://doi.org/10.1016/j.bpsc.2018.11.013

Zhou, Y., Yu, F., Duong, T., 2014. Multiparametric MRI characterization and prediction in

autism spectrum disorder using graph theory and machine learning. PLoS One 9. https://doi.org/10.1371/journal.pone.0090405

# Tables

Table 1: Pattern classification studies of ASD

| Study | N | Age - years | Sex - male | Modality | Features | Classifier | Validation | Design | Accuracy # |
|---|---|---|---|---|---|---|---|---|---|
| (Ecker et al., 2010b) | TDC = 22; ASD = 22 | 27.00 +- 7.00 | 100% | Structural MRI | Voxel-based features | SVM | L2out-CV | cross-sectional | 86.00 % |
| (Ahmadlou et al., 2010) | TDC = 9; ASD = 8 | 10.8 +- n.s. | n.s. | EEG (rest) | All frequency bands | NN | 1/5-CV | cross-sectional | 90.00 % |
| (Jiao et al., 2010) | TDC = 16; ASD = 22 | 9.20 +- 2.10 | 84% | Structural MRI | Region-based features | multiple | 1/10-CV | cross-sectional | 87.00 % |
| (Ecker et al., 2010a) | TDC = 20; ASD = 20 | 33.00 +- 11.00 | 100% | Structural MRI | Region-based features | SVM | L1out-CV | cross-sectional | 85.00 % |
| (Uddin et al., 2011) | TDC = 24; ASD = 24 | 13.2 +- 0.60 | 91% | Structural MRI | Voxel-based features | SVM | 1/10-CV | cross-sectional | 90.00 % |
| (Anderson et al., 2011) | TDC = 40; ASD = 40 | 22.70 +- 7.40 | 100% | Functional MRI (rest) | Region-based features | not-specified | L1out-CV | cross-sectional | 79.00 % |
| (Ingalhalikar et al., 2011) | TDC = 45; ASD = 30 | 10.50 +- 2.50 | 75% | Diffusion MRI | Region-based features | SVM | L1out-CV | cross-sectional | 80.00 % |
| (Calderoni et al., 2012) | TDC = 38; ASD = 38 | 4.40 +- 1.50 | 0% | Structural MRI | Voxel-based features | SVM | L2out-CV | cross-sectional | 80.00 AUC |
| (Ahmadlou et al., 2012) | TDC = 9; ASD = 9 | 10.80 +- n.s. | n.s. | EEG (rest) | All frequency bands | NN | n.s. | cross-sectional | 95.50 % |
| (Duffy and Als, 2012) | TDC = 554; ASD = 430 | n.s. | 88% | EEG (rest) | Coherence measures | LDC | n.s. | cross-sectional | 87.20 % |
| (Murdaugh et al., 2012) | TDC = 14; ASD = 13 | 21.40 +- 3.90 | 100% | Functional MRI (rest) | Region-based features | LRC | L1out-CV | cross-sectional | 96.00 % |
| (Wang et al., 2012) | TDC = 29; ASD = 29 | n.s. | 83% | Functional MRI (rest) | Region-based features | LRC | L1out-CV | cross-sectional | 82.80 % |
| (Uddin et al., 2013) | TDC = 20; ASD = 20 | 9.90 +- 1.50 | 80% | Functional MRI (rest) | Network-based features | LRC | n.s.-CV | cross-sectional | 78.00 % |
| (Lim et al., 2013) | ADHD = 29; ASD = 29 | 14.90 +- 1.86 | n.s. | Structural MRI | Voxel-based features | GPC | L1out-CV | cross-sectional | 89.30 % |
| (Deshpande et al., 2013) | TDC = 15; ASD = 15 | 21.10 +- 0.90 | n.s. | Functional MRI (task) | Region-based features | SVM | 1/10-CV | cross-sectional | 95.90 % |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (Ingalhalikar et al., 2014) | TDC = 42; ASD = 57 | 10.40 +- 2.50 | n.s. | MEG (task)/ Diffusion MRI | All frequency bands/ Region-based features | ensemble | 1/5-CV | cross-sectional | 83.30 % |
| (Eldridge et al., 2014) | TDC = 30; ASD = 19 | 8.46 +- 1.30 | 15% | EEG (task) | Event-related potentials | SVM/ LRC/ NBC | L1out-CV | cross-sectional | 79.00 % |
| (Sabuncu and Konukoglu, 2014) | TDC = 325; ASD = 325 | 17.80 +- 7.40 | 88% | Structural MRI | Region-based features | SVM | 1/5-CV | cross-sectional | 60.00 % |
| (Wee et al., 2014) | TDC = 59; ASD = 58 | 10.80 +- 4.00 | 76% | Structural MRI | Region-based features | SVM | 1/2-CV | cross-sectional | 96.30 % |
| (Segovia et al., 2014) | TDC = 40; ASD = 52; ASD-sibs = 40 | 14.40 +- 1.70 | 67% | Structural MRI | Voxel-based features | SVM | n.s.-CV | cross-sectional | 80.00 % |
| (Just et al., 2014) | TDC = 17; ASD = 17 | 25.60 +- 6.70 | 94% | Functional MRI (task) | Voxel-based features | NBC | L1out-CV | cross-sectional | 97.00 % |
| (Zhou et al., 2014) | TDC = 153; ASD = 127 | 13.50 +- 6.00 | 86% | Structural MRI/ Functional MRI (rest) | Graph-based features | RFC | multiple-CV | cross-sectional | 70.00 % |
| (Gori et al., 2015) | TDC = 20; ASD = 21 | 4.10 +- 0.80 | n.s. | Structural MRI | Voxel/Region-based features | SVM | L2out-CV | cross-sectional | 74.00 AUC |
| (Chen et al., 2015) | TDC = 126; ASD = 126 | 14.80 +- 1.60 | 85% | Functional MRI (rest) | Region-based features | SVM/ RFC | TsetV | cross-sectional | 91.00 % |
| (Plitt et al., 2015) | TDC = 59; ASD = 59 | 17.70 +- 2.70 | 100% | Functional MRI (rest) | Region-based features | LRC/ SVM | L1out-CV | cross-sectional | 95.19 % |
| (Iidaka, 2015) | TDC = 328; ASD = 312 | 13.20 +- 3.10 | 84% | Functional MRI (rest) | Region-based features | NN | L1out-CV | cross-sectional | 90.00 % |
| (Crippa et al., 2015) | TDC = 15; ASD = 15 | 3.50 +- 7.70 | 80% | Behavior | Motor task | SVM | L1out-CV | cross-sectional | 96.70 % |
| (Libero et al., 2015) | TDC = 18; ASD = 19 | 27.10 +- 1.30 | 78% | Structural MRI/ Diffusion MRI/ MRS | Voxel/Region/ Concentration-based features | NN | L1out-CV | cross-sectional | 91.90 % |
| (Liu et al., 2016) | TDC = 29; TDC-IQ = 29; ASD = 29 | [4.0-11.0]* | 86% | Behavior | Face-perception task | SVM | L1out-CV | cross-sectional | 88.51 % |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (Ghiassian et al., 2016) | TDC = 458; ASD = 430 | 17.30 +- 8.40 | 89% | Structural MRI/ Functional MRI (rest) | - | - | 1/5-CV | cross-sectional | 75.00 % |
| (Chen et al., 2016) | TDC = 128; ASD = 112 | 14.80 +- 1.70 | 85% | Functional MRI (rest) | Network-based features | SVM | L1out-CV | cross-sectional | 79.17 % |
| (Chanel et al., 2016) | TDC = 14; ASD = 15 | 28.60 +- 1.70 | 86% | Functional MRI (task) | Voxel-based features | SVM | L1out-CV | cross-sectional | 92.30 % |
| (Yahata et al., 2016) | TDC = 107; ASD = 74 | 31.40 +- 8.50 | 82% | Functional MRI (rest) | Network-based features | LRC | L1out-CV | cross-sectional | 85.00% - CV; 75% - TsetV |
| (Emerson et al., 2017) | HR = 48; ASD = 11 | 2.00 +- 0.0 | 69% | Functional MRI (rest) | Network-based features | SVM | L1out-CV | longitudinal | 98.00% |
| (Hazlett et al., 2017) | HR = 145; ASD = 34 | 0.50 +- 0.00; 1 +- 0.00 | 63% | Structural MRI | Region-based features/ Demographics | NN | 1/10-CV | longitudinal | 94.00% |
| (Shen et al., 2017) | HR = 174; ASD = 47 | 0.50 +- 0.00 | 62% | Structural MRI | Cerebrospinal fluid | NN | 1/25-CV | longitudinal | 69.00% |
| (Guo et al., 2017) | TDC = 55; ASD = 55 | 12.70 +- 2.40 | 76% | Functional MRI (rest) | - | NN | 1/4-CV | cross-sectional | 86.36% |
| (Xiao et al., 2017) | DDC = 39; ASD = 46 | 2.25 +- 0.30 | 88% | Structural MRI | Region-based features | SVM/ RFC/ NBC | 1/3-CV | cross-sectional | 75.60% |
| (Li et al., 2017) | TDC = 16; ASD = 14 | 32.70 +- 7.69 | n.s. | Behavior | Motor task | SVM | L1out-CV | cross-sectional | 66.67% |
| (Sadeghi et al., 2017) | TDC = 31; ASD = 29 | 20.00 +- 6.16 | n.s. | Functional MRI (rest) | Region-based features | SVM/ NN | 1/5-CV | cross-sectional | 92.00% |
| (Grossi et al., 2017) | TDC = 10; ASD = 15 | 10.40 | 86% | EEG (rest) | - | multiple | L1out-CV | cross-sectional | 92.80% |
| (Jahedi et al., 2017) | TDC = 126; ASD = 126 | 17.31 +- 6.00 | 86% | Functional MRI (rest) | Region-based features | - | TsetV | cross-sectional | 71.00% |
| (Wang et al., 2017) | NYU: TDC = 58; ASD = 54; Stanford: TDC = 20; ASD = 20; UM_1: TDC = 31; ASD = 34; Yale: | site-specific | site-specific | Structural MRI | Region-based features | ensemble | - | cross-sectional | NYU: 76.51%; Stanford: 68.26%; UM-I: 68.40%; Yale: 67.04% |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | TDC = 22; ASD = 20 | | | | | | | | |
| (Subbaraju et al., 2017) | TDC = 530; ASD = 505 | site-specific | site-specific | Functional MRI (rest) | Region-based features | SVM | - | cross-sectional | 77.30% |
| (Abraham et al., 2017) | TDC =530; ASD = 505 | site-specific | site-specific | Functional MRI (rest) | Region-based features | SVM | L1SITEout-CV | cross-sectional | 67.30% |
| (Heinsfeld et al., 2018) | TDC = 530; ASD = 505 | site-specific | site-specific | Functional MRI (rest) | Region-based features | NN | L1SITEout-CV | cross-sectional | 70.00% |
| (Wan et al., 2018) | TDC = 37; ASD = 37 | 4.6 +- 0.7 | 89% | Behavior | Eye tracking | SVM | 1/5-CV | cross-sectional | 85.10% |
| (Sen et al., 2018) | TDC = 530; ASD = 506 | site-specific | site-specific | Functional MRI (rest) | Network-based features | SVM | 1/5-CV | cross-sectional | 62.25% |
| (Soussia and Rekik, 2018) | TDC = 186 ASD = 155 | 16.9 | 90% | Functional MRI (rest) | Low order morphological networks | Multiple | Multiple-CV | cross-sectional | 61.69% |
| (Simões et al., 2018) | TDC = 17 ASD = 17 | 16.4 +- 0.6 | 100% | EEG (task) | Visual stimulation task | SVM | 1/5-CV | cross-sectional | 81.00% |
| (Tariq et al., 2018) | TDC = 74 ASD = 119 | - | - | Behavior | Video watching features | Multiple | 1/10-CV | cross-sectional | 89,00% |
| (Jun et al., 2019) | TDC = 171 ASD = 121 | 14 +- 5.8 | 84% | Functional MRI (rest) | Network-based features | Multiple | 1/10-CV | cross-sectional | 75.86% |
| (Ghafouri-Fard et al., 2019) | TDC = 455 ASD = 487 | 10 +- 0.53 | 83% | Genetics | Single nucleotide polymorphisms | NN | 1/10-CV | cross-sectional | 73,67% |
| (Cheng et al., 2019) | TDC = 22 ASD = 25 | 9.3 +- 1.4 | 72% | EEG (rest) | - | SVM | Bootstrap | cross-sectional | 92.70% |
| (Payabvash et al., 2019) | TDC = 33 ASD = 14 | 8-12* | 100% | Diffusion MRI | Connectome edge density | Multiple | - | cross-sectional | 75.3% |
| (Parikh et al., 2019) | TDC = 430 ASD = 421 | 16.8 +- 7.7 | 88% | Behavior | - | Multiple | 1/25-CV | cross-sectional | 62.00% |
| (Kazemeinjad and Sotero, 2019) | TDC & ASD = 816 | - | - | Function MRI | Region-based features | SVM | 1/10-CV | cross-sectional | 95,00% |

Note: TDC = typically developing controls. ASD = Autism spectrum disorder. ASD-sibs = siblings of individuals with Autism spectrum disorder. age-years: we report the age in years and its standard deviation. sex-male: the percentage of males with male-sex. *age range instead of mean and standard deviation are reported. # the reported accuracy.

Table 2: Stratification studies of ASD

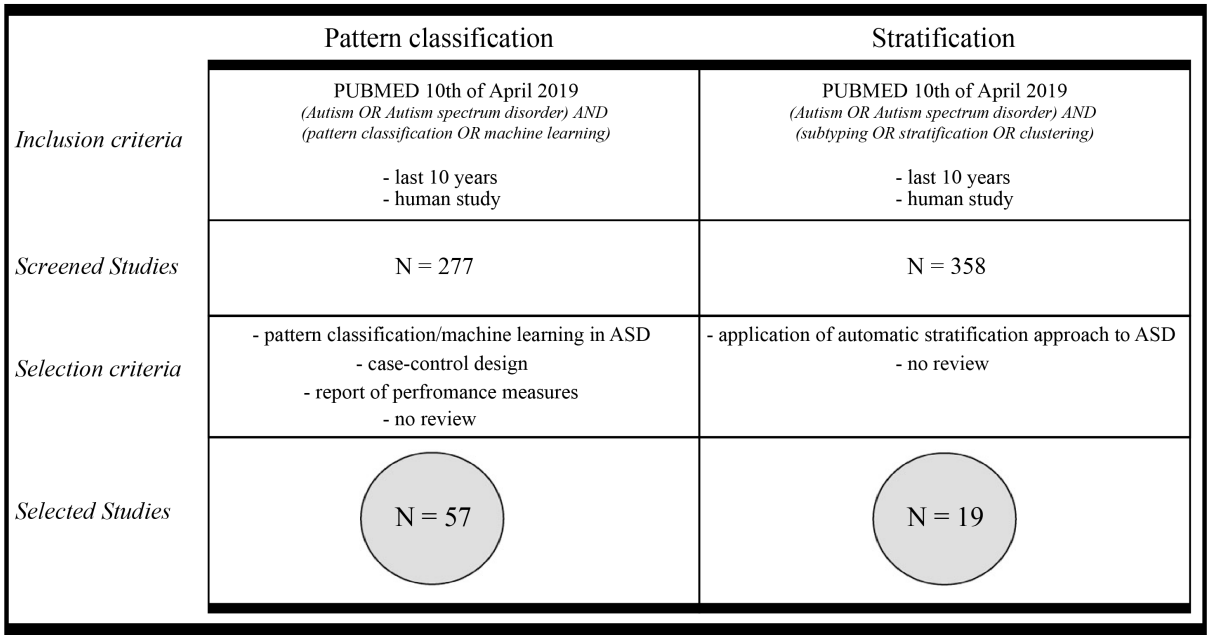
| study | N | Features | Stratification algorithm | Number of clusters/ components | Cluster descriptions | External validation |
|---|---|---|---|---|---|---|
| (Munson et al., 2008) | 245 ASD | IQ scores | Latent class cluster analysis and taxonometric analysis | 4 | low IQ; low verbal IQ, medium nonverbal; medium IQ; high IQ | symptom scores |
| (Bitsika et al., 2008) | 53 ASD | cognition, adaptive behaviors, | Ward's method | 3 | communication; social skills; adaptive behavior | cognition; symptom scores |
| )(Rapin et al., 2009) | 62 ASD | expressive phonology, comprehension, | hierarchical clustering/ Ward's method | 4 | low phenology and comprehension; low phenology and normal comprehension; normal phenology and low comprehension; normal phenology and comprehension | none |
| (Hu and Steinberg, 2009) | 1954 ASD | symptom scores | K means/ hierarchical clustering/ PCA; | 4 | severer language deficits; mild; savant skills; | none |
| (Lane et al., 2010) | 54 ASD | sensory processing | unclear | 3 | taste and smell sensitivity; movement related behavior; movement sensitivity; under responsive/seeks sensations; auditory filtering; low energy/weak; visual/auditory sensitivity; | symptom scores |
| (Sacco et al., 2012) | 245 ASD | demographic, clinical, case history, physiologic variables | K means | 4 | immune, circadian, non-sensory; circadian, sensory; stereotypic behaviors; mixed | none |
| (Fountain et al., 2012) | 6795 ASD | symptoms | Latent class growth analysis | 6 | high functioning; bloomers (substantial improvement); medium-high functioning; medium functioning; low-medium functioning; low functioning | demographics; autism risk factors |
| (Georgiades et al., 2013) | 391 ASD | symptom scores | Factor mixture modeling | 3 | social communication (-), repetitive behaviors (+); social communication (+), repetitive behaviors (-); social communication (-), repetitive behaviors (-); | demographics; cognitive measures |
| (Doshi-Velez et al., 2014) | 4927 ASD | electronic medical records | Ward's method | 4 | seizures; multisystem disorders; auditory disorders and infections; psychiatric disorders; not otherwise specified; | none |
| (Veatch et al., 2014) | 1261 ASD; 2563 (replication = 2563 ASD) | symptoms, demographics, somatic variables | Ward's method | 2 | severe, less severe | genomic data |
| (Kim et al., 2016) | 100 ASD-timepoint1; 100 ASD-timepoint2 | symptom scores | hierarchical clustering/ Ward's method | 4 | symptom severity; nonverbal and verbal skills; adaptive functioning; | none |
| (Ausderau et al., 2016) | 1307 ASD-timepoint1; 884 ASD-timepoint2 | sensory processing | latent profile transition analysis | 4 | mild; sensitive-distressed; attenuated-preoccupied; extreme-mixed; | adaptive behavior |

| | | | | | | |
|---|---|---|---|---|---|---|
| (Cholemkery et al., 2016) | 463 ASD | symptom scores | hierarchical clustering/ Ward's method | 3 | Impairments are social interaction; impairments in communication and language; restricted, repetitive and stereotyped behaviors | none |
| (Lombardo et al., 2016) | 378 ASD (replication = 123) | cognition | hierarchical clustering | 5 | mentalizing task; read complex emotions and mental states from the eyes; tapping ability; | none |
| (Hong et al., 2017) | 107 ASD | brain structure | hierarchical clustering | 3 | cortical thickness, intensity contrast; geodesic distance - decrease; geodesic distance - increase | symptom scores |
| (Feczko et al., 2018) | 47 ASD | brain function | community detection | 3 | stop accuracy task; BK span; Facial affect RT | symptom scores; adaptive behavior |
| (Easson et al., 2018) | 145 ASD | brain function | K means | 2 | connectivity pattern | symptom scores; cognitive measures |
| (Tomchek et al., 2018) | 400 ASD | core development; sensory features | latent cluster analysis | 4 | degree and quality of sensory information; age; differential presentation of developmental skills | none |
| (Duffy and Als, 2019) | 430 ASD | Electrophysiology | hierarchical clustering | 2 | Coherence measure of EEG signal | Demographics |

Note: ASD = Autism spectrum disorder.

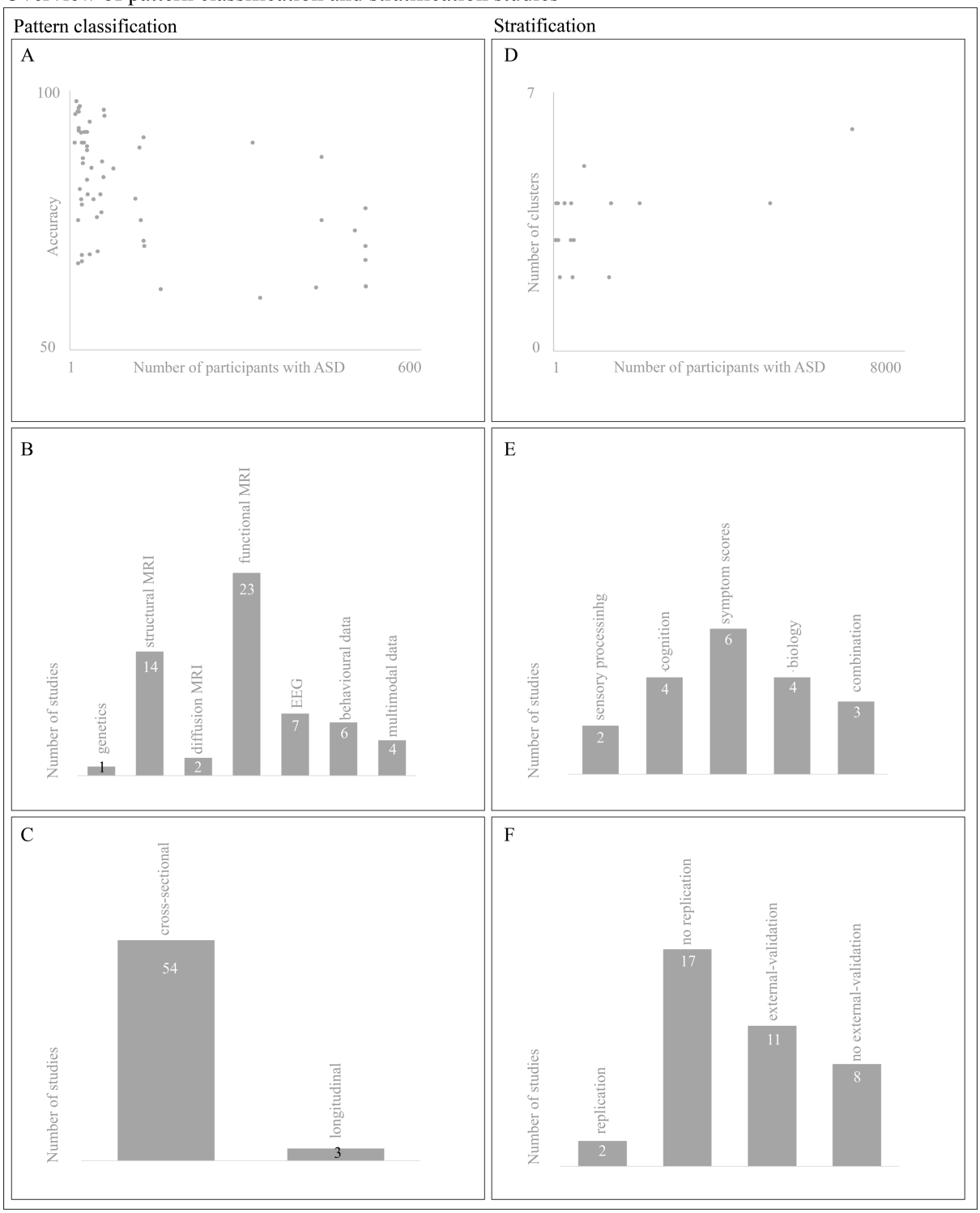
## Figures

**Figure 1**
Search strategy and method of selection




The selection process of the studies, inclusion criteria, the number of studies screened and selected, the search term and the date of the search are depicted.

**Figure 2**
Overview of pattern classification and stratification studies




A: The classification accuracy for Autism Spectrum Disorder (ASD) on the y-axis and the number of patients in the study on the x-axis. We observe a trend towards decreasing accuracy with increasing sample size. B: The number of studies that base their predictions on different data modalities. C: The number of studies that made cross-sectional versus longitudinal predictions. D: The number of identified clusters for Autism Spectrum Disorder (ASD) on the y-axis and the number of patients in the study on the x-axis. We observe a trend towards increasing number of clusters with increasing sample size. E: The number of studies that base

their stratifications on different data modalities. F: The number of studies replicate and validate their cluster solutions.