



# Applying multidimensional approaches to disentangle autism spectrum disorder heterogeneity

Angeline Mihailov

## ► To cite this version:

Angeline Mihailov. Applying multidimensional approaches to disentangle autism spectrum disorder heterogeneity. Psychiatrics and mental health. Université Paris-Saclay, 2021. English. NNT : 2021UPAST074 . tel-03510309

**HAL Id: tel-03510309**

**<https://theses.hal.science/tel-03510309>**

Submitted on 4 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Applying Multidimensional Approaches to Disentangle Autism Spectrum Disorder Heterogeneity

*Approches multidimensionnelles pour la caractérisation de l'hétérogénéité des troubles du spectre autistique*

### Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 575 - Physique et ingénierie: électrons, photons, sciences du vivant (EOBE)

Discipline: Imagerie et Physique Médicale

Spécialité de doctorat: Neuroimagerie Psychiatrique

Unité de recherche : Université Paris-Saclay, CEA, BAOBAB

Référent : Faculté des Sciences d'Orsay

Thèse présentée et soutenue à Paris-Saclay  
le 29/06/2021, par

**Angeline MIHAILOV**

#### Composition du Jury

**Monica ZILBOVICIUS**

*Directrice de recherche INSERM, Hôpital  
Necker, Université Paris Descartes*

Présidente

**Christine DERUELLE**

*Directrice de recherche CNRS, Institut de  
Neurosciences de la Timone*

Rapporteur

**Hervé LEMAÎTRE**

*Ingénieur de recherche CNRS, Université de  
Bordeaux*

Rapporteur

**Lisa EYLER**

*Professor, University of California San Diego*

Examinatrice

#### Direction de la thèse

**Vincent FROUIN**

*Directeur de recherche, CEA, Université Paris-  
Saclay*

Directeur

**Josselin HOUENOU**

*PUPH, Hôpital Universitaire Henri Mondor,  
Université Paris-Est Créteil*

Co-Directeur

**Cathy PHILIPPE**

*Chercheur, CEA, Neurospin, Université Paris-  
Saclay*

Encadrant, Examinatrice



## Thesis Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by social and communication deficits as well as restricted and repetitive behaviors. It has a prevalence rate of around 1 in 100 children, and a gender ratio of 4:1, males to females. A significant challenge in the understanding of ASD lies in its heterogeneity, with up to 70% of patients reporting an additional psychiatric, medical or genetic condition. Patients also present differences in neuroimaging, genetic and immune factors, all of which greatly complicate the understanding of ASD etiology. This widespread heterogeneity has caused difficulties in biomarker isolation, possibly due to widely used case-control experimental designs that combine ASD patients varying in behavioral, genetic and/or clinical profiles into one group. Though a diagnosis is dependent on behavioral presentations (namely social and communication deficits and repetitive behaviors), a high level of variation witnessed in several other biological and clinical factors has contributed greatly to the complexity of ASD symptomatology and etiology, making it difficult to develop proper therapies. Targeting only patients with a ‘pure’ diagnosis to participate in studies excludes those representing different ends of the functioning spectrum, ultimately impeding research. Such diagnostic ‘boxes’ can separate individuals that express similar traits, such as is the case for a patient that just made the cut-off versus one that just missed it. These two patients may in fact share symptoms and underlying physiological mechanisms that are more similar than two patients *within* the diagnostic threshold. Such practices may therefore confound information on the gradual emergence of psychopathology along development and on the study of prodromal risk factors.

One way to disentangle variability in ASD studies involves the use of dimensional approaches, which focus on the type and degree of several symptoms. Currently, the only way



to label an autistic individual is through diagnostic assessments, but often these assessments do not take into account the spectrum of accompanying genotypes and phenotypes, which makes it difficult to characterize patients on the *biological level*. The Research Domain Criteria (RDoC) has recently gained interest as an approach to solve this issue of characterization. RDoC is a framework for the investigation of psychiatric disorders, proposing the integration of several levels of information in order to explore dimensions spanning from normal to abnormal human behavior. Applying dimensional approaches, however, does not fully solve the issue of inconsistent results in autism research. This further necessitates the characterization of distinct subgroups of ASD. To this end, combining dimensional approaches with subgrouping strategies proves most relevant in solving the issue of inconsistent results across autistic literature.

### *Objectives*

The global aim of this work is to *better characterize autistic patients*, which is vital in the advancement of therapeutic strategies. We will aim to answer: *can dimensionally refined autistic subgroups provide us with more information on the etiology and underlying biological mechanisms of patients versus studying autistic patients as a whole?* Our proposal is two-fold, such that to successfully handle and disentangle the variability present in autistic patients it is crucial to, 1) apply dimensional approaches in order to recognize the disorder on a continual level and to incorporate several different types of data; and 2) to subgroup patients according to intra-group similarities in order to generate refined subpopulations. By applying these methods in ASD research, we can significantly improve the biological understanding of the disorder such that autistic patients can be better treated, whether preventatively or post-disorder. Furthermore, autistic studies have traditionally utilized cohorts with limited sizes, however this is slowly shifting towards increasingly available large multimodal cohorts.

Larger cohorts are necessary to perform the proposed approaches, as made possible by several big data initiatives including the Healthy Brain Network Cohort and the EU-AIMS Longitudinal European Autism Project Cohort, which will be used in this thesis work.

This work provides evidence of autistic and autistic-like subtypes, and confirms the necessity of applying dimensional approaches and subgrouping strategies in order to extract meaningful traits in autistic patients. We illustrate the importance of studying autism outside the realm of its diagnostic status by incorporating several levels of information including behavioral, neuroimaging, clinical, genetic and immunological data. We hope these approaches can eventually reduce heterogeneity within the disorder and pave the way to better understanding ASD etiologies and developmental pathways that will ultimately lead to the development of improved therapies and interventions.



# Acknowledgements

First of all, I would like to thank the members of the jury, Christine Deruelle, Hervé Lemaître, Monica Zilbovicius and Lisa Eyler, for agreeing to review my thesis manuscript. It is an honour.

My deepest gratitude to my supervisors, Vincent Frouin, Josselin Houenou and Cathy Philippe for their unwavering support and mentoring. I was lucky to be granted such kind, patient and insightful supervisors that truly made my entire PhD experience worthwhile. I learned a lot from you and it was a great pleasure to have worked with you these last 4 years.

Additionally, I would like to thank my colleagues. Antoine, for your immense help on several aspects of my work, and Lisa, for your constant support, kindness and your help on my scripts. Also, thank you to Lisa, Antoine, Thomas, Hélène, and Sam for all the fun coffee breaks that helped make the days more bearable!

I would like to exceptionally thank all the participants from the HBN and EU-AIMS cohorts, since this work would not have been possible without them and the time they invested in participating.

To my dear friend Steven, for your constant support and belief in me, you cannot imagine how much this motivated me throughout the years. And to my dear Izzy, I cannot thank you enough for your motivation and always being there to lighten the mood with our silly jokes.

To my mom who came all the way from Canada for several months to help out at home, there is no way I would have completed this work without you. And to my dad who has always supported me in anything I do, and who can flip around even the most stressful of situations.

Most importantly, I would like to thank Galliano for his constant support, both emotionally and scientifically. I am definitely lucky to have such a supportive husband.

Lastly, I would like to thank my darling daughter Amelia who was born half-way through this PhD, and who therefore granted me an experience that was truly worthwhile. You gave me real purpose and perspective.

This PhD work is dedicated to Ludmill Mihailov and Kiriaki Siampanis.

# Abbreviations

ADHD	Attention Deficit/Hyperactivity Disorder
ADI-R	Autism Diagnostic Interview-Revised
ADOS	Autism Diagnostic Observation Schedule
ARI	Affective Reactivity Index
ASD	Autism Spectrum Disorder
BIC	Bayesian Information Criterion
CBCL	Child Behavior Checklist
DAWBA	Development and Well-Being Assessment
DMN	Default Mode Network
DSM	Diagnostic and Statistical Manual of Mental Disorders
FSIQ	Full Scale Intelligence Quotient
GWAS	Genome-Wide Association Study
HBN	Healthy Brain Network
HLA	Human Leukocyte Antigen
ICD	International Classification of Diseases
IGI	Local Gyrfication Index
MDD	Major Depression Disorder
MNS	Mirror Neuron System
MRI	Magnetic Resonance Imaging
OCD	Obsessive Compulsive Disorder
OR	Odds Ratio
PCA	Principal Component Analysis
PDD-NOS	Pervasive Developmental Disorder Not Otherwise Specified
PGS	Polygenic Risk Score
RBS	Repetitive Behavior Scale
RDoC	Research Domain Criteria
ROI	Region of Interest
SCARED	Screen for Child Anxiety Related Disorders
SD	Standard Deviation
SDQ	Strength and Difficulties Questionnaire
SRS	Social Responsiveness Scale
TD	Typically Developing (Controls)
VWA	Vertex-Wise Analysis
WAIS	Wechsler Adult Intelligence Scale
WISC	Wechsler Intelligence Scale for Children

# Table of Contents

<b>Thesis Abstract</b> .....	ii
<b>Acknowledgements</b> .....	vi
<b>Abbreviations</b> .....	vii
<b>Table of Contents</b> .....	viii
<b>PART I: Introduction</b> .....	1
Chapter 1: Autism Spectrum Disorders (ASD).....	2
1.1.1. Evolution of Autism Spectrum Disorders.....	2
1.1.1.A. History.....	2
1.1.1.B. Evolution of Diagnostic Criteria.....	3
1.1.1.C. Diagnostic Assessments.....	5
1.1.1.C.i. Autism Diagnostic Interview - Revised (ADI-R).....	6
1.1.1.C.ii. Autism Diagnostic Observation Schedule - Generic (ADOS).....	6
1.1.1.C.iii. Diagnostic Interview for Social and Communication Disorders (DISCO).....	6
1.1.1.C.iv. Childhood Autism Rating Scale – Second Edition (CARS-2).....	7
1.1.1.C.v. Developmental, Dimensional, and Diagnostic Interview (3Di).....	7
1.1.2. Epidemiology.....	7
1.1.2.A. Prevalence.....	7
1.1.2.B. Sex Differences.....	8
1.1.2.C. Risk Factors.....	9
1.1.2.C.i. Prenatal Conditions.....	10
1.1.2.C.ii. Sociodemographic Factors.....	11
1.1.2.C.iii. Genetic Factors.....	11
Chapter 2: Heterogeneity in ASD.....	13
1.2.1. Clinical Outcomes in ASD Patients.....	13
1.2.1.A. Psychiatric Comorbidities.....	13
1.2.1.B. Cognitive Profiles.....	15
1.2.1.C. Medical Comorbidities.....	17
1.2.2. Genetics in ASD Patients.....	18
1.2.2.A. Twin Studies and Familial Heritability.....	19
1.2.2.B. Associated Monogenetic Disorders.....	20
1.2.2.C. Gene Variants.....	21
1.2.3. Immune Function in ASD.....	22
1.2.3.A. Inflammatory Prenatal Environment.....	23

1.2.3.B. The Immune System in ASD Throughout Life.....	25
1.2.3.C. The Human Leukocyte Antigen in ASD.....	27
Chapter 3: Neuroimaging Heterogeneity in ASD.....	29
1.3.1. Brain mapping in ASD.....	29
1.3.1.A. Structural MRI Neuroimaging.....	29
1.3.1.B. Functional MRI Neuroimaging.....	34
1.3.1.C. Other Neuroimaging Modalities.....	36
1.3.2. Social Cortical Paradigms in ASD.....	37
1.3.2.A. Mirror Neuron System.....	38
1.3.2.B. Theory of Mind.....	39
1.3.2.C. Default Mode Network.....	41
Chapter 4: Dimensional Approaches in Autism Studies.....	43
1.4.1. The Research Domain Criteria (RDoC).....	43
1.4.2. The Social Responsiveness Scale (SRS) as a Dimensional Construct.....	45
1.4.3. Studies Applying Dimensional Approaches in ASD.....	47
Chapter 5: Subgrouping in ASD.....	49
1.5.1. The Necessity to Subgroup.....	49
1.5.2. Studies Adopting Subgrouping Strategies in ASD.....	51
<b>PART II: Thesis Objectives.....</b>	<b>53</b>
<b>PART III: Materials and Methods Background.....</b>	<b>56</b>
Chapter 1: Cohorts.....	58
3.1.1. The Healthy Brain Network Cohort.....	58
3.1.2. The EU-AIMS Longitudinal European Autism Project Cohort.....	60
Chapter 2: Unsupervised Learning Techniques.....	61
3.2.1. Principal Component Analysis.....	62
3.2.3. Clustering Methods.....	63
Chapter 3: Neuroimaging Methods.....	67
3.3.1. Magnetic Resonance Imaging.....	67
3.3.2. Main Structural Imaging Approaches.....	69
3.3.3. FreeSurfer Tool.....	72
3.3.4. ComBat Tool: Harmonization of Multi-Site Imaging Data.....	75
Chapter 4: Genetic Analysis.....	76
3.4.1. Imputation of Genomic Data.....	76
3.4.2. PyHLA: Allele and Amino Acid Association Analysis.....	77

<b>PART IV: Studies Conducted in this thesis</b> .....	80
4.1. Study 1: <i>Cortical signatures in behaviorally clustered autistic traits subgroups: a population-based study</i> .....	81
4.2. Study 2: <i>In search of consistent behavioral and neuroimaging biomarkers in three samples with different autistic criteria</i> .....	102
4.3. Study 3: <i>HLA alleles: a stratifying factor for cortical signatures in Autism Spectrum Disorders</i> .....	145
<b>Conclusion</b> .....	161
<b>Résumé du Projet en Français</b> .....	166
<b>Bibliography</b> .....	173
<b>Abstracts</b> .....	198



# Part I: Introduction

# Chapter 1: Autism Spectrum Disorders (ASD)

## 1.1.1. Evolution of Autism Spectrum Disorders

### *1.1.1.A. History*

In 1943, child psychiatrist and physician Leo Kanner, born in present-day Ukraine, offered the first ever description of Autism, initially labelled ‘*autistic disturbances of affective contact*’<sup>1</sup>. He reported 8 boys and 3 girls that expressed the inability to relate to others, as well as disturbances in communication and apprehension following changes to non-social environments<sup>1</sup>. He reported that most of the children did not speak, and for those that did, they exhibited very unusual language such as echolalia, pronoun reversal and difficulties with socially-related language. Kanner also observed several atypical behaviors including repetitive motor movements involving arm flapping or body rocking. Kanner offered the example of one boy who preferred being alone, never cried to be with his mother, and attached extremely literal and inflexible meanings to words. For a significant amount of time following these initial reports, autism was thought to be an early onset form of schizophrenia. However, despite several similarities between the two disorders, autism was eventually classified as its own psychiatric disorder with a potentially strong genetic component. Kanner ultimately defined an autistic child as having the following behavioral characteristics: 1) failing to develop relationships before the age of 30 months; 2) deficits in normal language development; 3) ritualistic behavior and insistence on ‘sameness’. He addressed these impairments using the term ‘autism’.

Shortly after Kanner's reports, Austrian pediatrician Hans Asperger published work discussing a cohort of children who exhibited major issues in social interaction and motor function, despite preserved verbal skills<sup>2</sup>. He even used the term 'autism' to describe the collection of symptoms, unknowing of its use by Kanner. In general, Asperger described a cohort similar to that of Kanner's, however they were older and functioned at a higher level. He also described such patients as exhibiting restricted interests interfering with the acquisition of other skills. Due to the Second World War, publication of Asperger's work did not reach the English world until 1981 when it was reviewed by Lorna Wing<sup>3</sup>. Though Asperger's account of the disorder shared similarities to those of Kanner's, his work still heavily influenced the development of the Autism Spectrum Disorders classification.

#### *1.1.1.B. Evolution of Diagnostic Criteria*

Following Kanner and Asperger's initial reports, several studies validated the presence of such disabilities in children, leading to the first operational definition of 'Infantile Autism', published in 1980 in the Diagnostic and Statistical Manual of Mental Disorders (DSM-III), describing it as a *pervasive developmental disorder* and officially separating it from childhood schizophrenia<sup>4</sup>. Infantile Autism was to be classified in the first 30 months of life, and followed a three-fold criteria: 1) showing a lack of interest in others; 2) severe communication deficits; and 3) unnatural responses to the environment. This DSM-III definition was revised in 1987 dropping the first 30 months of life requirement and adding the mild diagnosis of pervasive developmental disorder-not otherwise specified (PDD-NOS)<sup>5</sup>. Though the term 'spectrum' was not yet incorporated into the title, this change already signified that autistic symptomatology was heterogeneous and thus made up of a spectrum of conditions.

In 1994, the DSM-III was updated to the DSM-IV and first recognized autism as a spectrum of disorders. The DSM-IV listed 5 conditions which included autism and PDD-NOS, and further added Asperger's Syndrome to the mild end of the spectrum, childhood disintegrative disorder (CDD) to the most severe end of the spectrum, and Rett's Syndrome<sup>6</sup>. However, by the year 2013, the DSM-IV was further updated to the DSM-5, where it adopted the classification of a spectrum labelled 'Autism Spectrum Disorders' (ASD) characterized by two criteria: 1) reciprocal social communication and interactions impairments, and 2) restricted and repetitive behaviors (removing CDD and Rett's Syndrome from the spectrum) (Table 1.1.1). The DSM-5 also offers guidelines on levels of severity for each symptom<sup>7</sup> (Table 1.1.2). This change from the DSM-IV was warranted by a shift of interest in autism genetics, as it was believed that finding associated genes, and eventually treatments, for this highly inheritable condition would prove more plausible for one disorder with associated traits as compared to five<sup>5,7</sup>.

Criteria	Description
<b>A. Persistent deficits in social communication and social interaction across multiple contexts, as manifested by the following, currently or by history.</b>	
	1. Deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions.
	2. Deficits in nonverbal communicative behaviors used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication.
	3. Deficits in developing, maintaining, and understanding relationships, ranging, for example, from difficulties adjusting behavior to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers
<b>B. Restricted, repetitive patterns of behavior, interests, or activities, as manifested by at least two of the following, currently or by history</b>	
	1. Stereotyped or repetitive motor movements, use of objects, or speech (e.g., simple motor stereotypies, lining up toys or flipping objects, echolalia, idiosyncratic phrases).
	2. Insistence on sameness, inflexible adherence to routines, or ritualized patterns or verbal nonverbal behavior (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat food every day).
	3. Highly restricted, fixated interests that are abnormal in intensity or focus (e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interest).
	4. Hyper- or hypo-reactivity to sensory input or unusual interests in sensory aspects of the environment (e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement).
<b>C. Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities or may be masked by learned strategies in later life).</b>	
<b>D. Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning.</b>	

**Table 1.1.1.** DSM-5 criteria for diagnosing ASD.

Severity Level	Social Communication	Restricted and Repetitive Behaviors
Level 3 "Requiring very substantial support"	Severe deficits in verbal and nonverbal social communication skills cause severe impairments in functioning, very limited initiation of social interactions, and minimal response to social overtures from others. For example, a person with few words of intelligible speech who rarely initiates interaction and, when he or she does, makes unusual approaches to meet needs only and responds to only very direct social approaches	Inflexibility of behavior, extreme difficulty coping with change, or other restricted/repetitive behaviors markedly interfere with functioning in all spheres. Great distress/difficulty changing focus or action.
Level 2 "Requiring substantial support"	Marked deficits in verbal and nonverbal social communication skills; social impairments apparent even with supports in place; limited initiation of social interactions; and reduced or abnormal responses to social overtures from others. For example, a person who speaks simple sentences, whose interaction is limited to narrow special interests, and how has markedly odd nonverbal communication.	Inflexibility of behavior, difficulty coping with change, or other restricted/repetitive behaviors appear frequently enough to be obvious to the casual observer and interfere with functioning in a variety of contexts. Distress and/or difficulty changing focus or action.
Level 1 "Requiring support"	Without supports in place, deficits in social communication cause noticeable impairments. Difficulty initiating social interactions, and clear examples of atypical or unsuccessful response to social overtures of others. May appear to have decreased interest in social interactions. For example, a person who is able to speak in full sentences and engages in communication but whose to- and-fro conversation with others fails, and whose attempts to make friends are odd and typically unsuccessful.	Inflexibility of behavior causes significant interference with functioning in one or more contexts. Difficulty switching between activities. Problems of organization and planning hamper independence.

**Table 1.1.2.** DSM-5 severity criteria for ASD.

The DSM-5 is not the only manual developed for the diagnosis of ASD. The International Classification of Diseases (ICD, currently in its 11<sup>th</sup> edition) is also widely used around the world and groups together several conditions into one ‘Autism Spectrum Disorder’, not including Rett’s Syndrome, much like the DSM-5 (<https://icd.who.int/browse11/l-m/en>). Though these two manuals are widely used, they exhibit some differences. The ICD-11, for example, is less culturally specific, distinguishes autism with and without intellectual disability, and acknowledges that older patients and women can sometimes camouflage their symptoms.

#### *1.1.1.C. Diagnostic Assessments*

Various tests for diagnosing and characterizing ASD rely on parent/caregiver reviews, self-assessments, or observational valuation. Reasons for this variability generally depend on the reason for procurement such as whether this information will contribute to a research study or to a clinical diagnosis. Most assessments are generally developed for early use in life since autism is a disorder with early symptoms, however many can be adapted for use throughout

life<sup>8</sup>. This section will review some of the most common diagnostic assessments used to measure ASD.

#### *1.1.1.C.i. Autism Diagnostic Interview - Revised (ADI-R)*

The Autism Diagnostic Interview – Revised (ADI-R) is considered to be a ‘gold standard’ semi-structured assessment of ASD where parents or caregivers report information on an individual assumed of having the disorder. This assessment is appropriate for children and adults above the age of 18 months and provides a diagnostic algorithm compatible with both the DSM-5 and ICD-11 criteria by assessing behavior across reciprocal social interaction, communication, and restricted and repetitive behaviors<sup>9</sup>.

#### *1.1.1.C.ii. Autism Diagnostic Observation Schedule - Generic (ADOS)*

Along with the ADI-R, the Autism Diagnostic Observation Schedule - Generic (ADOS) is also considered a ‘gold standard’ semi-structured assessment of ASD and evaluates social communication, interaction and play in an observational manner. It contains a series of modules adapted to the verbal capacity of the child or adult, which are each composed of activities allowing the examiner to determine the presence of behaviors that are in line with an ASD diagnosis<sup>10</sup>.

#### *1.1.1.C.iii. Diagnostic Interview for Social and Communication Disorders (DISCO)*

The Diagnostic Interview for Social and Communication Disorders (DISCO) is a semi-structured interview created for use with a suspected autistic patient by an assessor who is well

acquainted with the subject. It is meant for use in patients from infancy through to old age. This questionnaire adopts a dimensional approach to allow for the delineation of behavioral patterns that have developed over the course of many years<sup>11</sup>.

#### *1.1.1.C.iv. Childhood Autism Rating Scale – Second Edition (CARS-2)*

The Childhood Autism Rating Scale – Second Edition (CARS-2), was developed to distinguish between children with moderate or severe ASD, as well as between those with ASD and those with other cognitive disorders including disorders involving developmental delay. It is a measure often used in research studies and to support a clinical diagnosis in unison with other assessments, and can be administered by either parents/caregivers, teachers or clinicians<sup>12</sup>.

#### *1.1.1.C.v. Developmental, Dimensional, and Diagnostic Interview (3Di)*

The Developmental, Dimensional, and Diagnostic Interview (3Di) is a computerized assessment taken by the parents of suspected autistic patients, ranging from childhood to adulthood, to measure the intensity of symptoms across the autistic spectrum as well as offering information about potential comorbidities<sup>13</sup>.

### 1.1.2. Epidemiology

#### *1.1.2.A. Prevalence*

Initial studies of autistic prevalence took place in the 1970s indicating a rate of 4 cases per 10 000 children, implying a rare occurrence<sup>14–16</sup>. Prevalence rates have since increased to as many

as 1 in 100 children<sup>17,18</sup>. Indeed, such increases may be linked to environmental effects<sup>19</sup>, however more likely explanations for such a sharp increase in diagnoses include changes in diagnostic information and criteria, increased awareness and recognition, improved diagnostic methodologies, and an average overall younger age of diagnosis<sup>16,20,21</sup>.

### *1.1.2.B. Sex Differences*

Gender differences in ASDs are highly unbalanced resulting in a generally accepted ratio of 4:1, males to females<sup>22–24</sup>, with several studies reporting ratios varying from 2:1 to 7:1, males to females<sup>25–28</sup>. Explanations for such an imbalances have suggested a potential male bias resulting in females being historically under-recognised, or alternatively, that females harbour a protective effect making them less likely to develop autism. Some theories suggest females' higher drive to empathize and to socially conform, as well as greater chances of female ASD symptoms being presented through other conditions such as anorexia or borderline personality disorders<sup>29</sup>. It has also been suggested that ASD is an expression of the '*extreme male brain*', which is a theory that emphasizes the better ability of females to empathize, while males have a stronger ability to systemize (i.e. to analyze or construct rule-based systems)<sup>29–31</sup>.

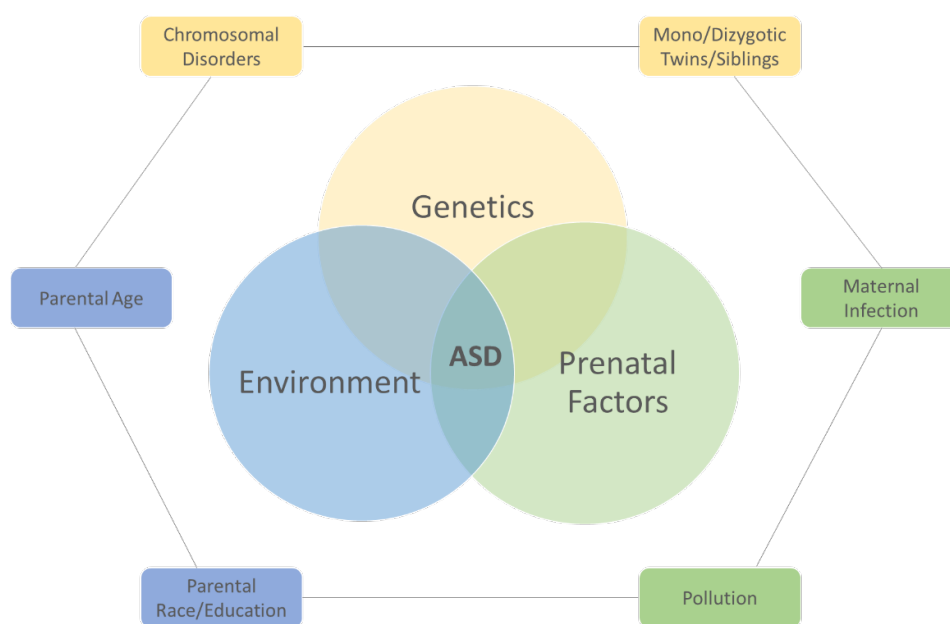
Biologically speaking, in relation to the extreme male brain theory, the fetal testosterone theory suggests that higher levels of testosterone in the amniotic fluid of mothers causes offspring having the improved ability to analyse and understand complex systems and patterns, while at the same time diminishing empathetic traits. This implies that these higher levels of testosterone push towards 'male' traits, as suggested by the extreme male brain theory, which in turn suggests that autistic brains show an exaggeration of features typically observed in males<sup>31</sup>.



Such features include a larger brain volume, over-connectivity within hemispheres, larger amygdalae, and decreased inter-hemispheric connectivity as a result of a diminished corpus callosum<sup>32,33</sup>. Moreover, studies have shown that children with an underdeveloped corpus callosum were more likely to receive a diagnosis of autism<sup>34</sup>. Lastly, further enforcing the testosterone theory, higher levels of testosterone in both males and females have reportedly been associated with autistic behavioral traits, such as avoidance of eye contact<sup>29,35</sup>. Finally, it is possible that if a bias does not exist, that classical diagnostic tools (such as the ADI-R or ADOS) are simply unequipped to detect the more subtle expression of autistic symptoms in females<sup>29</sup>.

### 1.1.2.C. Risk Factors

Considering that ASDs are regarded as a spectrum, this constitutes a multivariate profile that can be attributed to a wide variety of risks including prenatal conditions, sociodemographic backgrounds and genetic factors (Figure 1.1.1).



**Figure 1.1.1.** Summary of main risk factors in ASD development.

### *1.1.2.C.i. Prenatal Conditions*

The prenatal environment is a particularly significant factor in the risk of autism development in children. Specifically, maternal drug intake, illness, immune response, stress, and exposure to pollution have all been reported to affect autistic development in offspring. Maternal substance abuse, as well as the intake of alcohol or drugs for mental illness treatments such as selective serotonin uptake inhibitors and sodium valproate, have all been associated with an increased risk for ASD<sup>36–39</sup>. However, it is also possible that the notion of drug intake, be it strictly for mental illness or in an abusive manner, already indicates mental susceptibility of the mother and thus implies a genetic component to the development of a psychiatric disorder in the offspring. Maternal illnesses such as hypertension, obesity, asthma, diabetes and autoimmune disorders have also been linked to potential autistic symptoms in children<sup>40–43</sup>. As air pollution has been recognized by the World Health Organization (WHO) as one of the biggest health threats of our time, studies have not only shown that pregnant women living in highly polluted environments have a greater risk of having autistic children, but also that, postnatally, infants and children have a greater chance of developing autism throughout life<sup>44,45</sup>.

Maternal immune activation is one way in which the gestational environment can affect a child's neurodevelopmental outcome, and is thus a key area of research in ASD risk etiology. The maternal immune activation (MIA) model involves infecting pregnant rodents and observing effects in their offspring. This model posits that infected mothers cause neurodevelopmental and behavioral changes in offspring associated with psychiatric disorders such as schizophrenia and autism<sup>46–48</sup>. One study by *Shi et al., 2003*, showed how pregnant mice infected with human influenza subsequently gave birth to pups exhibiting abnormal social

interaction<sup>49</sup>. Additionally, it is suggested that seasonality is a significant factor in the risk development of autism due to variations in viral and/or other infections<sup>50</sup>.

#### *1.1.2.C.ii. Sociodemographic Factors*

Several studies have reported associations between sociodemographic and socioeconomic factors and the development of autism. Other than gender, which was previously discussed, factors that have been shown to associate with ASD risk include parental age, race, income and education. An American study conducted on consensus information in 2011 reported that white non-Hispanic mothers were more likely to have a child diagnosed with ASD<sup>51–53</sup>. This study, in line with many other studies, also suggested that mothers and fathers over the age of 34 were more likely to have children with ASD, as well as an increased risk of ASD in children whose mothers have a higher level of education<sup>53,54</sup>. Another study by *Bhasin and Schendel et al.*, 2007, has also suggested that higher maternal age, often accompanied by higher maternal education and median income, was linked with an increased risk for ASD development in children<sup>55</sup>. A possible explanation for this correlation is that families with higher incomes and education are more likely to be aware of potential developmental issues their children are having. They are also more likely to have the means to invest (time and resources) in proper assessments for their children.

#### *1.1.2.C.iii. Genetic Factors*

Though autistic etiology is currently poorly understood, several studies have suggested the involvement of a genetic component. Studies have shown that within monozygotic twins, if one

is diagnosed with ASD, the chances of the other having it ranges from 36%-95%<sup>56</sup>. The fact that this rate is not 100% suggests that non-genetic factors are probably at play in the predisposition of ASD development. In dizygotic twins however, if one sibling has ASD, the chances of the other having it plummets significantly to 0-30%<sup>57,58</sup>. Finally, in non-twin siblings, if one individual is diagnosed with ASD, his or her siblings only have a 4% chance of also developing the disorder<sup>59</sup>. Patients having a genetic or chromosomal condition such as Fragile X Syndrome tend to express higher levels of autistic symptoms<sup>60,61</sup>. Taken together, these studies provide evidence favoring a strong genetic component in ASD. Further details in genetic mechanisms and associated genes will be discussed in section 1.2.2.

## Chapter 2: Heterogeneity in ASD

A significant challenge in the characterization and understanding of ASD lies in its heterogeneity, with up to 70% of patients reporting an additional psychiatric, medical or neurological condition<sup>62–64</sup>. Presence and intensity of ASD symptoms also considerably vary between patients. Although a diagnosis is dependent on behavioral presentations (namely social and communication deficits and repetitive behaviors), a high level of variation witnessed in several other biological and clinical factors has contributed greatly to the complexity of ASD symptomatology and etiology, thus making it difficult to develop proper therapies. In this section, we will address this variability by discussing clinical, genetic and immunological presentations in ASDs.

### 1.2.1. Clinical Outcomes in ASD Patients

#### *1.2.1.A. Psychiatric Comorbidities*

Psychiatric comorbidities are well documented in ASD due to their common and consistent occurrence, leading the DSM-5 to recognize simultaneous psychiatric diagnoses alongside ASD. Widely discussed psychiatric comorbidities include Attention-Deficit Hyperactivity Disorder (ADHD), Anxiety, Depression, Bipolar Disorder, Obsessive Compulsive Disorder (OCD), Oppositional Defiant Disorder (ODD), Intellectual Disability and Childhood-onset Schizophrenia<sup>62,65–69</sup>. Behaviors often reported in ASD (aside from those defined in the

diagnosis criteria) are often believed to reflect these disorders and include attention problems, hyperactivity, anxiety, depression, irritability and aggression<sup>70-72</sup>.

One of the most common psychiatric comorbidity in ASD patients is ADHD, with a prevalence rate ranging from 29% - 55%. Both disorders also report higher instances of males over females and high impairment in daily life<sup>73</sup>. There are several biological, neuropsychological and behavioral overlapping mechanisms and factors between the two disorders. Though less consistent, neuroimaging studies have shown similarities between the two disorders in the form of disruptions in both resting and active brain networks as well as frontal and cerebellar alterations<sup>74-79</sup>. Commonly affected neuropsychological components in both ASD and ADHD include alterations in executive function and Theory of Mind (ToM)<sup>80,81</sup>. Aberrations in common biological mechanisms have also been reported between the disorders including in pathways and genes linked to GABA and glutamate levels<sup>82,83</sup>. Importantly, though several parallels have been reported, significantly more unique and mutually exclusive signatures exist, in all symptomatic aspects, which separate the two disorders and confirm their individual psychiatric status<sup>84</sup>.

Up to 80% of patients with ASD also present at least one type of anxiety disorder, with separation anxiety being the most common, followed by generalized anxiety and social phobia<sup>62,85</sup>. Having anxiety as an ASD patient typically exacerbates symptoms resulting in greater psychosocial dysfunction since patients exhibit more severe social avoidance, sleep issues, and family and peer interaction deficits<sup>86</sup>. Patients on the less extreme end of the

spectrum, as in Asperger's Syndrome, tend to exhibit higher levels of anxiety because of their awareness due to higher cognitive function<sup>87</sup>.

Another frequently reported psychiatric condition includes depression (with a prevalence rate between 10% and 50%), which is regularly mis- or under- diagnosed in ASD patients since depressive symptoms are often masked by ASD traits<sup>62,69</sup>. Obsessive Compulsive Disorder (OCD) is another common disorder in ASD patients with a rate of up to 37%, however it is difficult to differentiate since many features such as repetitive behaviors and intrusive thoughts are present in both disorders<sup>85,88</sup>. Bipolar disorder is frequently reported in patients and typically emerges during adolescence, with up to 30% of Bipolar patients carrying an ASD diagnosis<sup>68,89</sup>. Both Oppositional Defiant Disorder (ODD) and autism are characterized by a behavioral inhibition imbalance causing them to sometimes be co-diagnosed, with about one quarter of autistic patients also meeting ODD criteria<sup>90</sup>. Lastly, though schizophrenia is a disorder diagnosed later on in life, young autistic patients can develop Childhood-onset Schizophrenia (COS), which is a rare and severe form of schizophrenia occurring before the age of 13. Studies have reported the co-diagnosis of COS and ASD in roughly 30% of children who require a combination of aggressive pharmaco- and psychological therapy<sup>67</sup>.

#### *1.2.1.B. Cognitive Profiles*

Social cognition comprises the means of processing implicit and explicit information in order to attain understanding of others and self. This involves storing and applying information on facial and bodily expressions in order to deduce the identity, actions, and emotional status of another being and to elicit a suitable behavioral response<sup>91</sup>. ASD patients often exhibit deficits

somewhere along this interaction and consequently suffer greatly from social dysfunction. Thus, the most pronounced cognitive deficits in autistic patients stem from social cognition and perception and often involve atypical social interactions and communication. Furthermore, many supplementary cognitive deficits that are part of the complex cognitive process of socialization, and which have been found disrupted in ASD patients, include memory, attention, language processing, inhibition, motivation and emotional functioning<sup>70,90,92-94</sup>. Atypical language processing and development has manifested in difficulties surrounding phonology, grammar, syntax and semantic understanding. This can result in delayed and/or deviant comprehension, such as exceedingly literal interpretations of regular semantics<sup>95</sup>. Furthermore, several models have investigated socially-related dysfunction involving eye gaze studies, which have shown abnormal gaze and decreased eye contact in patients, as well as disruptions in face processing<sup>96,97</sup>. It has also been proven that patients do not perform well on non-verbal communication, imitation, and affective empathy, sympathy and mentalizing (which involves understanding the mental states of both self and others), as shown by deficits in theory of mind<sup>98,99</sup>. Specifically, theory of mind (ToM) describes the capacity of individuals to understand themselves and others by correctly characterizing self and others' mental states, and is said to be one of the cornerstones of effective social interaction<sup>100</sup>.

Further forms of cognitive decline have also been observed in autistic children in executive function and information processing domains. Executive function is defined as the mental processes underlying goal-directed behavior, language, cognitive flexibility, inhibition, working memory, and planning, and its dysfunction is implicated across development in both the social and repetitive behaviors criteria of ASD symptomatology<sup>101</sup>. Information processing is another domain in which autistic individuals show dysfunction, namely in the preferential



balance between bottom-up (local) and top-down (global) processing. In healthy controls, a preference for top-down processing is typically seen, while in ASD patients the opposite is observed where patients display superior low-level sensory-perceptual processing<sup>102</sup>. This could partially explain abnormal responsivity, in the form of hypo- or hyper-reactivity to one's environment as typically exhibited by patients<sup>103</sup>.

#### *1.2.1.C. Medical Comorbidities*

Though not included in the diagnostic criteria, several well-reported medical comorbidities also exist in autistic patients. The most common medical concerns described in patients with autism include epilepsy, sleep difficulties, gastrointestinal disorders, immune disorders, genetic disorders and neurological disorders such as tic and other motor disorders<sup>104–107</sup>. The presence of epilepsy has been widely investigated and proven in up to 39% of ASD patients<sup>108</sup>. Furthermore, the relationship between sleep and epilepsy is generally believed to be bidirectional, which could potentially explain the presence of both issues in autistic patients<sup>109,110</sup>. Many autistic patients also show alterations in gastrointestinal physiology, including increased intestinal permeability, higher levels of gut infections, and microbiota dysbiosis. This could be associated with the fact that the gut microbiome can communicate with the brain via the gut-brain axis by using signalling and immune-mediating molecules. Therefore, disequilibrium on one side of this axis can result in disruptions observed in both systems<sup>111,112</sup>. There is a tendency towards autoimmune disorders such as arthritis and thyroiditis in ASD patients and their families, probably linked to immune system dysregulation in patients<sup>48</sup>. Lastly, common neurological disorders reported in ASD patients include tic disorders such as Tourette Syndrome, which is typically worse during childhood and is characterized by voluntary tics that rid the patient of the urge to execute a certain movement or

action<sup>113</sup>. Several genetic disorders such as Fragile X Syndrome, Rett Syndrome, tuberous sclerosis and Down Syndrome are likewise associated with ASD, which will further be discussed in the following genetics section. Table 1.2.1 summarizes the most commonly reported comorbidities.

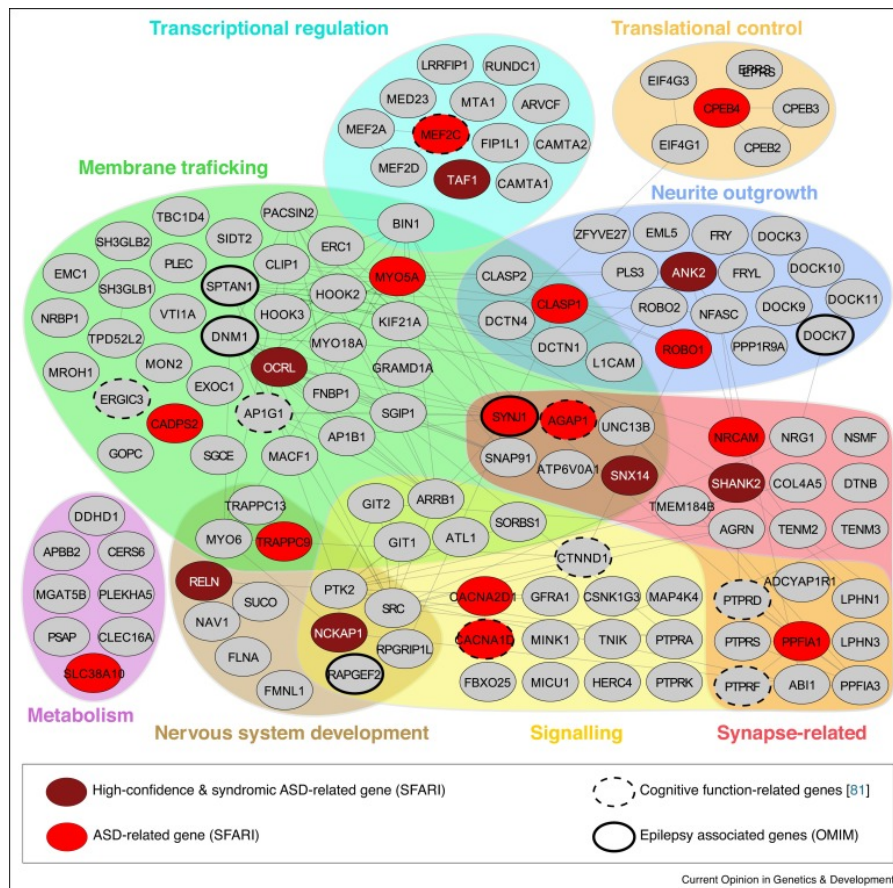
Psychiatric Comorbidities
Attention-Deficit Hyperactivity Disorder
Anxiety
Depression
Bipolar Disorder
Obsessive Compulsive Disorder
Oppositional Defiant Disorder (ODD)
Intellectual Disability
Childhood-onset Schizophrenia
Medical and Other Comorbidities
Epilepsy
Gastrointestinal Disorders
Immune Disorders
Sleep Disorders
Neurological and Tic Disorders
Chromosomal Disorders
Fragile-X Syndrome
Rett Syndrome
Tuberous Sclerosis
Down's Syndrome

**Table 1.2.1.** Summary of the most commonly observed comorbidities in ASD.

### 1.2.2. Genetics in ASD Patients

Since the delineation of autistic disorders, it has been clear that this syndrome relies on a genetic component. In general, genes reported in ASD are typically involved at some point along the molecular pathways of several cellular functions such as RNA processing and splicing, signalling, chromatin remodelling, synaptic plasticity, synaptic transmission, transcriptional regulation, translational control and nervous system development (Figure 1.2.1)<sup>114–118</sup>. Though

it has been difficult to isolate specific genetic causes, several studies to date have proven the existence of a strong genetic element in ASDs.



**Figure 1.2.1.** A neuronal microexon network associated with nervous system development and ASD, indicating genes specifically linked with autism (image from *Gonatopoulos-Pournatzis et al., 2020*).

### 1.2.2.A. Twin Studies and Familial Heritability

Twin studies have been pertinent in the understanding of genetic contributions to ASD development. Twin studies taking place between the 1970s and early 1990s first brought to light and revolutionized the understanding of a strong heritability in autistic disorders, and jump-started the search for genetic factors in ASDs<sup>119,120</sup>. This also eliminated the widely believed (at that time) theory of ‘*refrigerator mothers*’, which attributes autistic traits and diagnoses to distant maternal approaches<sup>121</sup>. Several studies since then have shown that the presence of ASD

in one monozygotic twin predicts its development in the second twin with an up to 95% chance (with an average of 80%)<sup>56</sup>. However, in dizygotic twins and non-twin siblings, the chances of both developing ASDs are 30% and 4%, respectively<sup>58,59</sup>. These differences in rates between different kinds of siblings highlight strong genetic underpinnings, but also the presence of epigenetic factors. Though heritability is an important area of study in ASD genetics, it has been proposed that families with multiple autistic children fall into one of two categories: 1) a majority low-risk group where the presence of ASD in families is primarily explained by a *de novo* variant; and 2) a minority high-risk group where the presence of ASD in families is explained by inherited variants, with a dominant transmission for males and a protective factor found in females<sup>122,123</sup>. Furthermore in multiplex families, ASD recurrence has been found in almost 50% of later-born males, while only showing a 20% recurrence in later-born females<sup>122,124</sup>.

#### *1.2.2.B. Associated Monogenetic Disorders*

Early studies in genetic causation of ASD implicated the presence of genetic syndromes, which account for 10% of ASD cases, including, but not limited to, Fragile X Syndrome, Rett Syndrome, and tuberous sclerosis<sup>125–127</sup>. These associated genetic syndromes are well-defined in the study of ASD due to their monogenic nature and thus relatively straightforward genetic etiology. Fragile X Syndrome studies have reported that up to 50% of males with the syndrome have ASD<sup>128</sup>. Patients with Fragile X Syndrome show dysregulation in Fragile Mental Retardation 1 protein (FMRP) production, resulting in deficient mRNA trafficking and synaptic plasticity, which is regulated by the Fragile Mental Retardation 1 gene (*FMRI*). Furthermore, *FMRI* regulates neuroligin, neurexin, and SHANK proteins, which are mutated in ASD<sup>129</sup>. Rett Syndrome is an X-linked disorder typically affecting females and characterized by extreme

neurodevelopmental delay and seizures. It is estimated that up to 40% of Rett Syndrome patients are co-diagnosed with ASD<sup>128</sup>. It is caused by mutations in the methyl-CpG binding protein 2 gene (*MeCP2*), which is heavily abundant in neurons of the mature nervous system and has functions in the silencing of unnecessary and/or harmful genes. Lastly, up to 60% of tuberous sclerosis patients are diagnosed with ASD. Tuberous sclerosis is a syndrome caused by dysfunction with the *TSC1* or *TSC2* gene, which functions in controlling dendritic proliferation<sup>128</sup>.

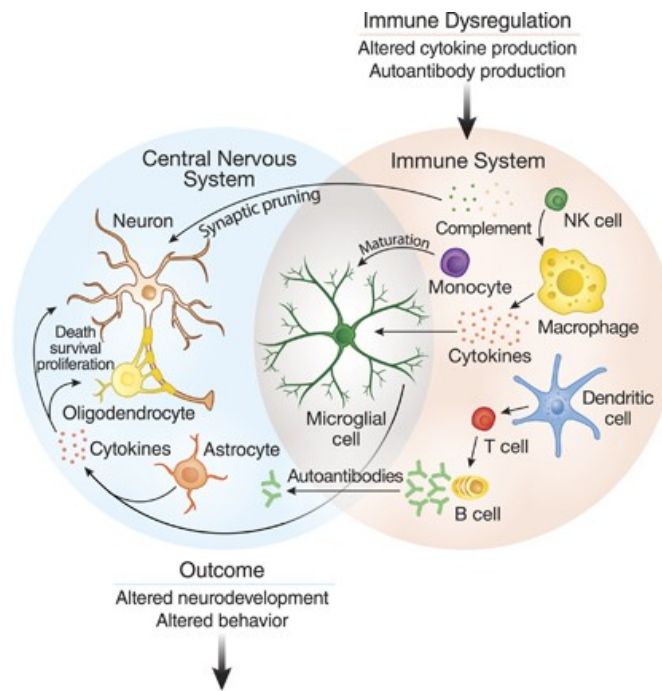
#### *1.2.2.C. Gene Variants*

Studies have also revealed that ASD patients present genetic factors not linked to specific syndromes with over 1000 identified genes<sup>130,131</sup>. Studies conducted by the Autism Genetic Resource Exchange (AGRE) and the Simon Simplex Collection reported rates of rare *de novo* copy-number variants (CNV) in ASD patients between 5%-10%<sup>132,133</sup>. Of note, however, is that in almost two-thirds of cases where ASD-associated CNVs were identified in patients, affected siblings did not share these same CNVs<sup>134</sup>. This genetic heterogeneity extends to the fact that not only many genes have been identified in ASD (such as *NRXN1*, *SHANK3* and *PTEN*) indicating locus heterogeneity, but also that most genetic variants have a high degree of pleiotropy (which means that they affect more than one phenotype)<sup>135</sup>. Considering the diversity of CNVs and number of genes, this enforces the widespread heterogeneity observed in autistic patients, even at the genetic level. Nevertheless, it is necessary to determine the genetic leaders underpinning autistic presentations.

Rare and common variants are also of interest in the study of autism genetics, and may sometimes even be presented together. Rare variants are frequently identified in autism and manifest themselves in the form of Mendelian genetic syndromes, chromosomal abnormalities, rare copy number variations, and *de novo* and transmitted point mutations, each having an occurrence rate between 5% and 10%<sup>136–138</sup>. Since several of these rare variants can be clinically identified, genetic screening is recommended upon receiving a diagnosis<sup>139</sup>. Concerning common variants, though several single nucleotide polymorphisms (SNPs) have been identified, not one has a sufficient enough effect to be considered individually causal<sup>130</sup>. Usually, common genetic polymorphisms together exert an additive ASD risk<sup>135</sup>.

### 1.2.3. Immune Function in ASD

The interaction between the immune and nervous systems has long been proven to be responsible for various imbalances in the brain due to a complex interaction among several cell types from both systems (Figure 1.2.2). In many psychiatric disorders, autism included, the immune system has affected several neurodevelopmental processes related to neurogenesis, synaptogenesis and synaptic pruning<sup>140</sup>. This section will explore the link between the immune system and autism.



**Figure 1.2.2.** Dysregulated interactions between cell types of the nervous and immune systems may accumulate and cause an unbalanced neuronal and synaptic architecture as well as unbalanced molecular functions. This could lead to several psychiatric disorders (*image from Meltzer and Van de Water, 2016*).

### 1.2.3.A. Inflammatory Prenatal Environment

The gestational environment could affect a fetus' development in several ways including adverse effects due to maternal immune activation, even in the absence of an actual infection. In particular, maternal immune system dysregulation and/or activation poses a risk in autism development in offspring via the creation of an inflammatory environment marked by an increased production of maternal cytokines that have the ability to target the placenta as well as enter the fetal compartment. This is evidenced by various reports indicating a link between bacterial or viral outbreaks at the time of pregnancy and increased rates of autism in later born children. Such examples include the 1964 rubella outbreak in the United States that witnessed a marked increase in autism prevalence in the children of mothers infected during that time, as well as correlations between influenza infections and/or fever and risks of ASD

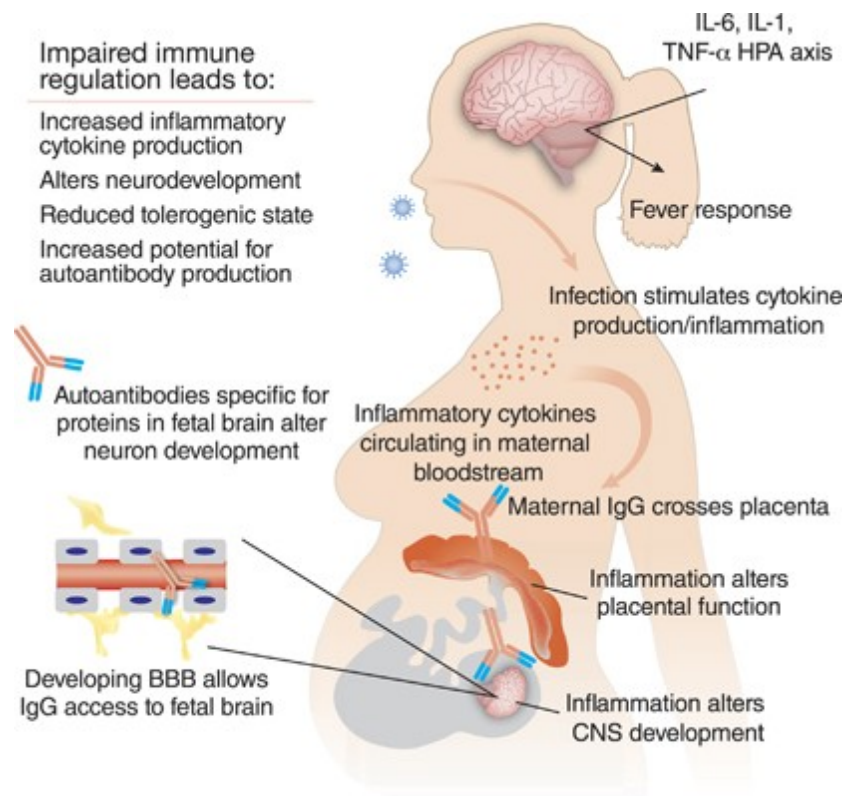
development<sup>141–143</sup>. This evidence collectively suggests that the type of infection may not necessarily be the culprit behind later ASD risks, but rather that the maternal immune activation itself remains responsible for the gestational dysregulations leading to increased incidences.

The most substantial evidence thus far indicating that maternal infection poses an increased risk for autism development comes from the maternal activation model, which was established based on a study where researchers infected rodent models (using influenza and *Escherichia coli* strains) and subsequently observed significant neurodevelopmental changes in pup offspring<sup>46,47</sup>. This model showed that offspring presented antisocial and repetitive behaviors upon maternal immune activation, with a higher incidence in males<sup>144,145</sup>. Concerning mechanisms of action upon maternal immune activation, it is suggested that the disruption of normal cytokine levels plays a role due to their pro-inflammatory functions. Cytokines are immunomodulatory cell-signalling proteins that act as endocrine messengers and are produced by two types of immune cells upon activation, namely innate and adaptive<sup>146</sup>. Studies have shown that the injection of pro-inflammatory cytokines in pregnant mice causes their pups to display autistic behaviors<sup>147</sup>.

It has also been suggested that maternal autoantibody activity can increase ASD-risk in children. Studies have reported that mothers with higher levels of anti-brain antibodies circulating in their plasma were four times more likely to have children with autism<sup>148,149</sup>. Animal studies have also supported this claim as evidenced by several studies that involved injecting pregnant mothers with the serum of mothers with autistic offspring, and subsequently



observing pro-autistic behaviors in offspring<sup>150,151</sup>. To collectively visualize all discussed effects on maternal activation, refer to Figure 1.2.3.



**Figure 1.2.3.** Activation of the maternal immune response can cause an increased risk for autism. Such activation can be attributed to a combination, or independent effects, of autoantibodies, viral/bacterial infections, and an increased level of immune molecules such as cytokines (image from Meltzer and Van de Water, 2016)<sup>48</sup>.

### 1.2.3.B. The Immune System in ASD Throughout Life

Evidence of immune dysfunction is also vastly reported in autistic patients themselves, however progress in this aspect of ASD, as in all others, is obstructed by inherent heterogeneity. Nonetheless, few consistent observations remain. Many studies have reported that children with autism present higher levels of autoantibodies in their serum (compared to controls or their non-affected siblings), which are linked to atypical development and are distinct from those measured in the serum of their mothers during pregnancy<sup>152–155</sup>. A study by Piras *et al.*, 2014,

showed that autoantibodies in children with autism correlated with disorder severity, motor impairment, IQ and social interaction<sup>156</sup>. They also reported that while a relationship exists between maternal and child autoantibodies, both maintain their own unique characteristics.

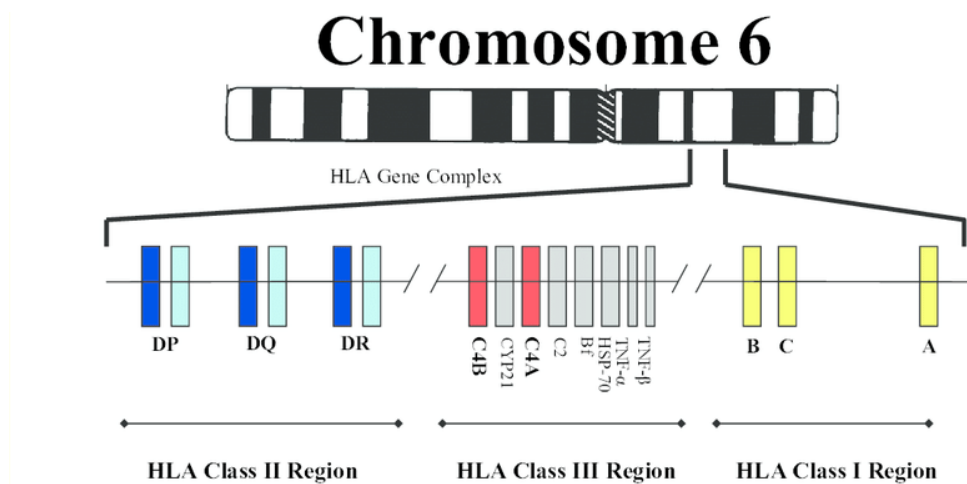
Specific immunogenetic factors have also been reported in autistic populations. Studies have found an association between autism and SNPs located in the promotor of the *MET* oncogene, which codes for a pleiotropic receptor that functions as a negative immune regulator and is also critical for neuronal migration in the development of the cerebellum<sup>157</sup>. Upon genetic ontology analysis, it has been shown that the autistic transcriptome is heavily enriched for immune response and neuronal activity-dependent genes<sup>158</sup>. Furthermore, one study examined gene expression in autistic and control brains and found that genes upregulated in patients held inflammatory response and immunomodulatory functions<sup>114</sup>. It is important to understand that genetic elements do not act alone, warranting a deeper understanding of their impact on ongoing networks contributing to neuroimmune dysfunction in ASD patients.

General immune system dysregulation has been reported as imbalances in immune components such as helper CD4<sup>+</sup> and cytotoxic CD8<sup>+</sup> T cells, natural killer cells, and immunoglobulins such as IgG and IgM in autistic patients<sup>159–161</sup>. Furthermore, same as reported in the serum of pregnant mothers, autistic patients themselves also exhibit distorted levels of cytokines and chemokines, including IL-6, IL-8, TFN, and IFN- $\gamma$ , linked to altered states of inflammation<sup>162,163</sup>. Several investigations have described pro-inflammatory states in autistic patients, including a constantly stimulated neuroinflammatory environment with activated microglia and astroglia, in the brains of autistic patients<sup>162,164</sup>. Furthermore, increased levels of

these pro-inflammatory cytokines have also been linked to increased severity in autistic behaviors<sup>165</sup>.

### *1.2.3.C. The Human Leukocyte Antigen in ASD*

The Human Leukocyte Antigen (HLA) genes are a highly polymorphic group of genes located on the Major Histocompatibility Complex (MHC) on the short arm of chromosome 6, which express HLA proteins responsible for regulating the immune system (Figure 1.2.4). These genes have numerous alleles for the purpose of serving several sub-functions involved in fine-tuning the adaptive immune system and its response, and aids in essential physiological processes involved in brain development and homeostasis<sup>166,167</sup>. HLA-A, -B, and -C antigens are encoded by HLA class I genes and help govern the cellular immune response; while HLA-DRB1, -DQB1 and -DPB1 are encoded by the HLA class II gene cluster and offer crucial functions in the mediation of the humoral immune response. The polymorphism in MHC manifests in numerous alleles for these antigens. Since the immune response has been linked to changes in neurodevelopment, this genetic region has proven to be one of the most significant genetic risk loci for several severe psychiatric diseases, ASD included<sup>168</sup>. The co-existence of inflammation and autoimmunity reported in autistic patients indicate the potential presence of an essential susceptibility marker located in the highly polymorphic HLA super-locus<sup>169</sup>.



**Figure 1.2.4.** Image showing the HLA complex located on chromosome 6 (*Image from Westover et al., 2011*)<sup>170</sup>.

Since the 1980s, several studies have reported a connection between ASD diagnoses and HLA alleles<sup>171,172</sup>. In particular, class II HLA allele differences have been extensively mapped in autistic patients, with a particular focus on DRB1 and DQB1 HLA genes where DRB1\*11 and DQB1\*04 have been associated with ASD-risk<sup>173–175</sup>. Class I HLA genes have also been studied in ASD individuals, namely HLA-A and HLA-C, with a higher incidence in patients and their mothers<sup>176,177</sup>. Though researchers have established an important relationship between HLA genes and autism neurodevelopment, due again to the heterogeneous presentation of the disorder, consistent and stable results are yet to be achieved.

## Chapter 3: Neuroimaging Heterogeneity in ASD

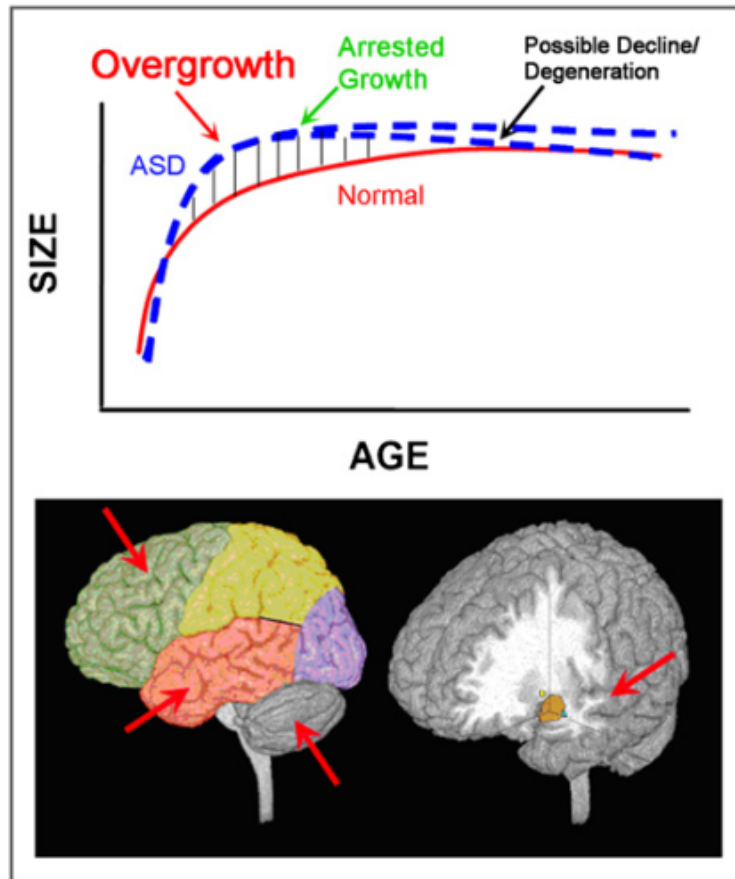
Widespread heterogeneity is often observed in the profiles of autistic patients, potentially explaining the frequently inconsistent results in the literature. Unfortunately, this does not spare neuroimaging markers. One source of this variability in the literature may be attributed to widely used case-control experimental designs, which combine autistic individuals varying in behavioral, genetic and/or clinical profiles into one group. Regions of the brain that have been typically implicated in neuroimaging studies of autistic patients are said to be part of the “social brain” network, and often include: temporal regions such as the fusiform face area and posterior superior temporal sulcus, frontal regions including the inferior and superior frontal gyrus, the insula, and the amygdala. The following sections will discuss important results observed in the neuroimaging markers of autistic groups, as well as neuroimaging paradigms pertinent in the brains of ASD patients.

### 1.3.1. Brain Mapping in ASD

#### *1.3.1.A. Structural MRI Neuroimaging*

Autistic patients demonstrate age-specific anatomical abnormalities, which is an important factor to take into account during the development of any experimental design and in the interpretation of research. One of the most well-reported and consistent structural result in ASD patients is an abnormal early brain development, specifically a larger brain volume overgrowth<sup>178,179</sup>. Autistic patients typically exhibit this early overgrowth in volume until six

to eight years old on average, for both males and females, when their brain size intersects with that of typically developing controls. Subsequently, an accelerated decline in size is observed in patients until adolescence, with possible degeneration continuing on until late middle age (Figure 1.3.1)<sup>180,181</sup>. Furthermore, this atypical neurodevelopmental trajectory is not consistent across brain regions, with frontal and temporal lobes more affected than parietal and occipital lobes, suggesting a temporally regional effect operating in the posterior-anterior direction<sup>182</sup>. Physiological processes underlying this observation in brain volume remain unclear, and have been attributed to fluctuating dynamics between cortical thickness and surface area<sup>183</sup>. Specifically, these two cortical features present distinct aspects in their development (i.e. having different progenitor cells) and structure (i.e. thickness is attributed to the number of neuronal cortical columns, while surface area is attributed to the number of neuronal radial columns)<sup>184</sup>. This abnormal development in the grey matter reflects the structural connectivity of the white matter. For example, parallel to an early brain overgrowth in grey matter volume, studies have also observed abnormal white matter tract organization in patients between the ages of six and twenty-four months who were diagnosed with an ASD<sup>185</sup>. In general, with the exception of accelerated age-related decline in brain volume (as well as its two components: thickness and surface area), many cortical alterations reported throughout the literature are not always reliable in autistic patients<sup>186</sup>.



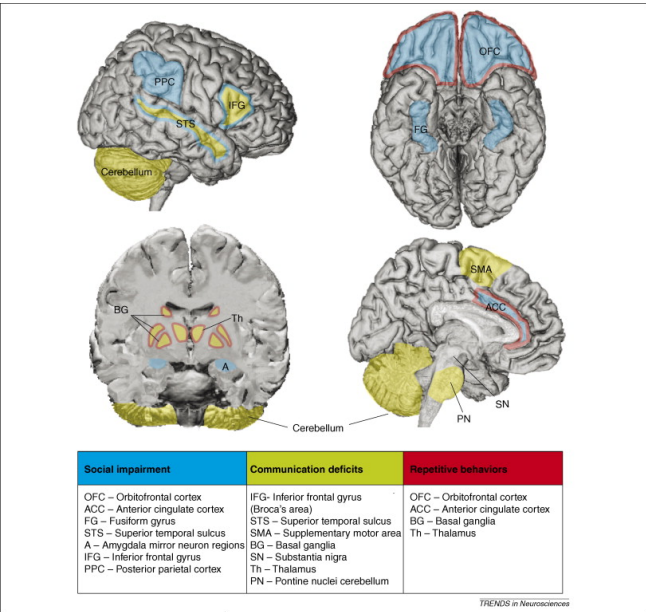
**Figure 1.3.1.** A plot showing the general progression of brain volume growth for autistic patients (blue) and controls (red), indicating a typical overgrowth early on in life, followed by an accelerated decline later in life for patients. Red arrows point to the sites of regional overgrowth in autistic patients, namely the frontal lobe, the frontal parts of the temporal lobe, cerebellum and the amygdala (*figure taken from Courchesne et al., 2007*)<sup>181</sup>.

Across autistic patients, structural abnormalities in specific areas have been reliably reported within regions such as the cerebellum, anterior and posterior cingulate, subcortical amygdala, hippocampus, basal ganglia, the insula and global frontotemporal and frontoparietal regions<sup>187</sup>. These regions are associated with behaviors that are often reported dysfunctional in autistic patients, such as, and most importantly, social communication deficits (the temporal lobe, the insula), as well as language abnormalities (Broca's and Wernicke's area), socioemotional processing (amygdala), and repetitive behaviors (the frontal striatal system and basal ganglia)<sup>188–192</sup>. Furthermore, autistic patients also display structural changes in brain regions linked to external behaviors that are not part of the autistic diagnostic criteria, but that are nonetheless present in the behavioral profile. A highly affected region in ASD patients is the

cerebellum, which mainly holds functions in motor control and cognitive processing (in executive functioning, visuospatial, and memory domains) and is frequently diminished in volume in patients<sup>193</sup>. The amygdala is a highly affected region in patients due to its functions in aggression, fear, emotion, and social interaction, and is typically increased in size in autistic patients<sup>187,194,195</sup>. Additionally, it has been reported that white matter integrity is diminished between the amygdala and neocortical regions in ASD patients<sup>196</sup>. The frontal lobe has also received considerable attention in the field of autism due to its implication in executive functions, social behavior and communication, with abnormalities reported in cortical thickness and volume<sup>180,197</sup>. Another important region, the fusiform gyrus, is involved in face recognition and therefore social functioning, thus its implication in ASD does not come as a surprise<sup>198</sup>. Studies have reported both increased and asymmetrical volumes (between hemispheres) in the fusiform gyri of autistic patients, as well as increases in cortical thickness<sup>199–201</sup>. An interesting feature that frequently abnormal in patients is the degree of curvature, more specifically, the gyrification. Increases in gyrification have been reported in autistic patients in the frontal lobes of children and adolescents<sup>202,203</sup>. Other changes in the cortical gyrification of patients includes increases in temporal regions and decreases in the supramarginal gyrus<sup>203–205</sup>. In particular, results involving changes to gyrification are interesting since this feature is believed to be developmentally determined<sup>206</sup>. Lastly, one of the most reportedly altered regions in the brain of autistic patients is the temporal lobe since it encompasses the superior temporal sulcus, the fusiform gyrus, and temporal gyri, all of which are linked to autistic-like behaviors<sup>207–210</sup>. The superior temporal sulcus is strongly implicated in language and social cognition, with several studies reporting decreases in volume and several functional alterations, which will be discussed in the next section<sup>209,210</sup>. The discussed regions are also altered in structural connectivity, with reports of decreased fractional anisotropy between areas responsible for social cognition, specifically between the bilateral temporal lobes and amygdala, and regions



adjacent to the superior temporal sulcus<sup>211</sup>. Weaker connectivity integrity was also found in the superior longitudinal fasciculus, which is located between the superior temporal sulcus and the inferior frontal gyrus in autistic patients<sup>212</sup>.



**Figure 1.3.2.** A diagram indicating many of the regions affected in ASD (figure taken from Amaral et al., 2008)<sup>213</sup>.

Indeed there are several well-reported cortical abnormalities in patients diagnosed with ASD (Figure 1.3.2), however evidence has suggested that these abnormalities may not always be disorder-specific. An example illustrating this concept involves amygdala alterations in autistic patients. Differences in amygdalae morphology have been reported, in general, in disorders exhibiting similar behaviors (i.e. in emotion and social interactions) including anxiety and Oppositional Defiant Disorder or Conduct Disorder<sup>214,215</sup>. This is suggestive of an observed alteration being linked to a behavior versus a disorder, and alludes to the necessity of investigating dimensional subconstructs within disorders as to better disentangle existing heterogeneity.

### *1.3.1.B. Functional MRI Neuroimaging*

Structurally altered regions observed in ASD patients often suggest network-level abnormalities, therefore as expected, many of these regions also present functional irregularities in patients. Functional MRI (fMRI) studies are often categorized into two types: 1) resting state, which is a baseline measure that observes the brain's functionality at rest; and 2) task-based, which requires the participant to conduct a task while observing their functional cortical response. In general, reduced brain activity is considered a main functional hallmark of several cognitive disorders, including ASD and is typically seen in task-based fMRI<sup>216</sup>. However, other studies have reported increased brain activity, which is also seen in resting-state fMRI.

In the context of functional studies that are conducted at rest, the Default Mode Network (DMN) is an important brain network since it is active when an individual is in wakeful rest, such as mind-wandering. This network comprises portions of the frontal lobe, the posterior midline (including the cingulate and precuneus) and the inferior parietal lobule, and is often perturbed in autistic patients<sup>217,218</sup>. It has been suggested that abnormal characteristics of the DMN could be linked to atypical information integration, as well as inflexibility in processing and responding to social stimuli in autistic patients, which contributes to the further exacerbation of symptoms<sup>219</sup>. Studies have also shown differences in activity outside the DMN in regions involving social processing in patients with ASD. Specifically, reduced functional connectivity has been reported in ASD patients versus controls between the amygdala and the ventromedial prefrontal cortex and, separately, between the amygdala and the inferior occipital gyrus<sup>220,221</sup>. Reduced functional connectivity has also been reported in autistic patients versus controls in the superior parietal and visuospatial areas<sup>222</sup>. Another study by *Tyszka et al., 2014*, reported decreased connectivity in temporal and frontal regions without any global abnormalities<sup>223</sup>.

Alternatively, a small number of studies have also reported hyper-connectivity, particularly in autistic patients with severe social dysfunction<sup>224</sup>.

Task-based fMRI studies have likewise shown aberrant social networks in ASD patients. These tasks often employ stimuli meant to invoke social networks in functions such as facial processing, for example. A recent study investigated implicit and explicit facial expressions in the form of gender and emotional recognition, respectively, in autistic patients compared to controls and found decreased activity in numerous social regions of the brains when processing emotional facial expressions (specifically in the amygdala, fusiform gyrus and superior temporal sulcus)<sup>225</sup>. Another study measured ASD patients' brain activity in response to biological motion before and after being asked to conduct a 'social skills' training course spanning 5 weeks. What they found was a significant positive correlation between activity in the superior temporal sulcus and mentalizing scores post training<sup>226</sup>. Additionally, studies have reported reduced activity in the middle temporal gyrus, fusiform gyrus, amygdala, medial prefrontal cortex, and inferior frontal gyrus of autistic patients in response to dynamic versus static facial expressions<sup>225,227</sup>. These same studies observed diminished bi-directional connectivity between the primary visual cortex – middle temporal gyrus – inferior frontal gyrus circuit in autistic patients compared to controls<sup>227</sup>. Furthermore, in a study by *Lynn et al., 2018*, ASD patients and controls were asked to encode and recognize images of faces and cars, which showed diminished functional connectivity in autistic patients between the fusiform gyrus and other regions (including the frontal and primary visual cortices) when looking at faces versus cars<sup>228</sup>. Another commonly studied task is the social rewards task, which often causes increased activity in the anterior cingulate gyrus and left frontal gyrus and decreased activation in the nucleus accumbens in autistic patients versus controls<sup>229</sup>.

### *1.3.1.C. Other Neuroimaging Modalities*

A significant amount of neuroimaging research in ASD uses MRI data, however other modalities such as electroencephalography (EEG), magnetoencephalography (MEG) and positron emission tomography (PET) are also vastly used and can offer additional insight into cortical pathologies. Data from these modalities can complement and fill gaps observed in fMRI and structural MRI (sMRI) autistic research by providing high temporal resolution features, offering complementary perspectives in the documentation of cortical (dys)functions in ASD patients.

EEG is a non-invasive functional method that records the electrophysiological activity of the brain by measuring the integration of neuronal ion current oscillations<sup>230</sup>. Several studies have shown EEG irregularities in the brains of ASD patients, suggesting a potential diagnostic validity in using EEG to characterize patients<sup>231–233</sup>. EEG studies are pertinent in the field of ASD due to their utility in investigating epilepsy and sleep disorders, two conditions that are extremely common in autistic patients with up to 46% of patients being diagnosed with epilepsy<sup>234</sup>. This prevalence in epilepsy in ASD patients suggests common physiological mechanisms between the two disorders, namely in alterations in GABA-ergic neurotransmission that is often reported in autistic patients and is linked to seizure susceptibility<sup>235</sup>. Approximately two-thirds of ASD patients also exhibit chronic sleep disorders, including insomnia, often contributing to cognitive disabilities<sup>236,237</sup>. Sleep issues are measured via questionnaires, however many recent studies are shifting towards the use of EEG and MEG techniques (including animal studies)<sup>238,239</sup>.

MEG is a non-invasive functional neuroimaging method elucidating high-resolution temporal and spatial neural activity. Though not currently used for diagnostic purposes, MEG techniques have similarly been conducted on autistic patients in the name of research in order to better comprehend paradigms used to understand cortical function, since many studies have shown domain and resting-state differences in autistic patients using MEG<sup>240</sup>. In particular, several studies have employed MEG to investigate language acquisition, cerebral laterality, regional connectivity and auditory hypersensitivity<sup>241–243</sup>.

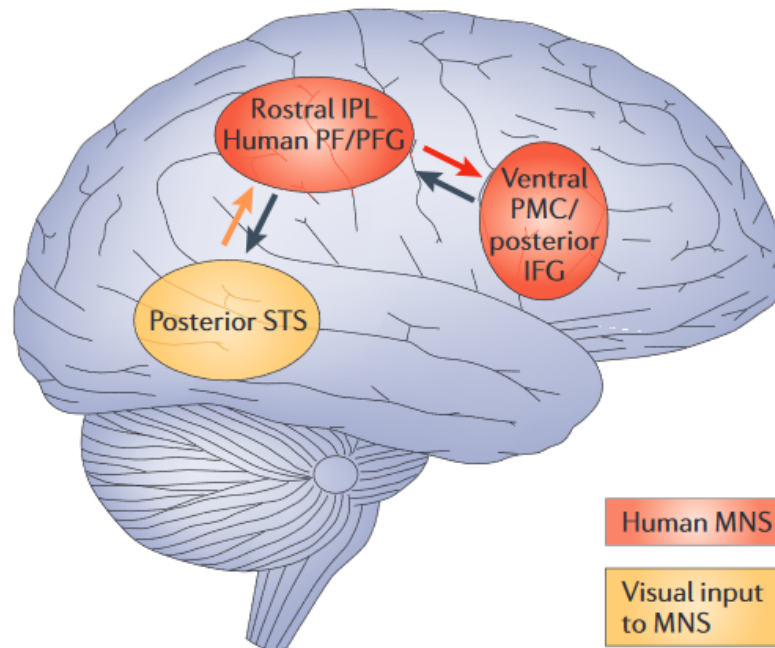
PET is another functional imaging method using injected radioactive substances to measure changes in metabolic brain processes through deviations in blood flow and regional chemical composition<sup>244</sup>. In particular, PET studies in autism have been conducted to better understand how underlying biochemical mechanisms of neuropeptides such as oxytocin and serotonin effect in the brain<sup>245</sup>. PET studies have also been particularly useful in studying synaptic density in neuropsychiatric disorders, including autism, through the development of PET ligands capable of targeting synaptic vesicles, which offers a molecular perspective to cortical abnormalities in the brain<sup>246</sup>.

### 1.3.2. Social Cortical Paradigms in ASD

Cortical and functional alterations in the brains of autistic patients are not only expressed regionally, but also at the network-level. Subsequent sections will describe some of the abnormal cortical systems reported in ASD (some already briefly discussed) that are often linked to behavioral symptoms typically witnessed in patients.

### *1.3.2.A. Mirror Neuron System*

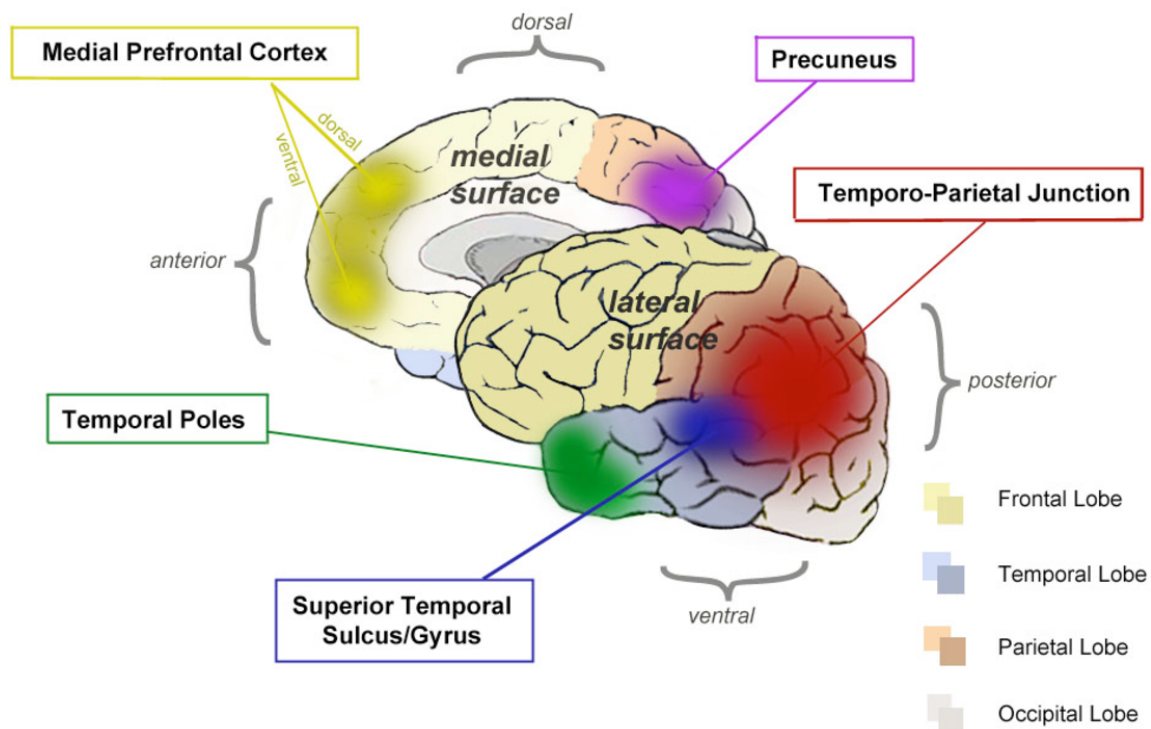
The Mirror Neuron System (MNS) was first discovered in monkeys as a group of neurons that fire during the performance of a goal-directed action and during the observation of others performing that same action<sup>247</sup>. This second function of the MNS has inspired studies in humans involving imitation and its applicability to social cognition in the understanding of others' intentions and emotions<sup>248</sup>. In humans, the MNS network is traced between the inferior frontal gyrus and inferior parietal lobule, with visual input originating in the superior temporal sulcus (Figure 1.3.3)<sup>249,250</sup>. Due to this network's implication in social cognition, it has been an important framework of study in autism research with evidence from structural and functional neuroimaging studies showing dysfunctions and cortical alterations in the MNS of ASD patients. Autistic patients in particular have shown decreased activity in the MNS during social tasks as well as correlations between symptom severity and MNS functionality, thus offering the hypothesis that MNS dysfunction may be considered a core neuroanatomical deficit in autistic patients<sup>248</sup>. Several neuroimaging studies have reported changes in or along the MNS neural network including disordered functional connectivity between implicated regions, as well as decreases in grey matter in MNS areas of the brain<sup>209,251</sup>. Studies have likewise shown reduced activity during imitation and social mirroring tasks, contributing to the overall large volume of literature suggesting the MNS as an effective biomarker in autism<sup>252,253</sup>.



**Figure 1.3.3.** The main regions involved in the Mirror Neuron System, including the superior temporal sulcus, the inferior parietal area, and the inferior frontal gyrus (image taken from Iacoboni et al., 2006)<sup>254</sup>.

### 1.3.2.B. Theory of Mind

Theory of mind (ToM) is synonymous to mentalizing and is defined as the capacity of individuals to understand themselves and others by correctly identifying self and others' mental states, which is a vital component in successful social interactions<sup>100</sup>. Rather than taking on an emotional state, ToM describes objective knowledge about others' mental states, thus adopting more of a socio-cognitive process versus a socio-affective process (which involves emotions such as empathy)<sup>255,256</sup>. Brain regions comprising the ToM network include the ventral temporoparietal junction, the superior temporal sulcus, the temporal poles, the medial prefrontal cortex and the precuneus/posterior cingulate regions (Figure 1.3.4)<sup>257,258</sup>.



**Figure 1.3.4.** Main regions involved in the Theory of Mind (ToM) network, including the temporoparietal junction, the super temporal sulcus, the temporal poles, the prefrontal cortex and the precuneus (image taken from Bowman and Wellman, 2014)<sup>258</sup>.

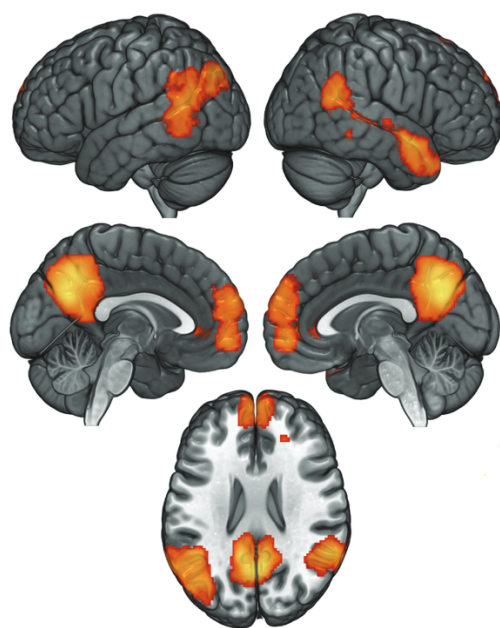
Deficits in the theory of mind network are one of the most notable biomarkers in autistic disorders. Alterations in this network often lead social, communication and behavioral impairments<sup>259</sup>. Autistic patients generally exhibit decreased functional connectivity in medial prefrontal, temporoparietal, posterior cingulate, motor and sensorimotor regions during mentalizing as compared to controls, which are regions implicated in ToM and also further enforce the general anterior-posterior pattern of functional and structural differences found in autistic patients<sup>260–263</sup>. EEG and MEG studies have also provided some evidence in dysfunctionality of the ToM network in autistic patients showing that during ToM tasks, greater stimulation was reported in the temporoparietal junction of subjects with autistic traits compared to controls<sup>186</sup>. Additionally, though more indirectly, structural studies have further offered insight into the study of ToM in the sense that regions typically involved in this network have repeatedly shown structural abnormalities in ASD patients when compared to typically developing individuals<sup>187,208,209,211</sup>.



### 1.3.2.C. Default Mode Network

This network has already been briefly discussed in section 1.3.1.B, however to briefly reiterate, the Default Mode Network (DMN) activates upon wakeful rest such as mind-wandering, and comprises portions of the frontal lobe, the posterior midline (including the posterior cingulate and precuneus), and the inferior parietal lobule and posterior temporal areas, which are all often perturbed in autistic patients (Figure 1.3.5)<sup>217,218</sup>. Dysfunction of the DMN affects information integration, as well as inflexibility in processing and responding to social stimuli in autistic patients, causing further symptom aggravation<sup>219</sup>. The DMN encompasses many of the regions involved in the ToM network, as well as in social tasks, suggesting a high overlap between these systems. This has lead scientists to suggest a direct link between physiological and psychological ‘baselines’, proposing social cognition as the default mode of thought<sup>264,265</sup>. It is widely accepted that the DMN significantly contributes to social cognition, or *vice versa*, however mechanistic insights are yet to be further developed. Majority of DMN-related studies in ASD patients have described deficits in the previously mentioned regions. Specifically, functional connectively resting-state studies have consistently reported over-connectivity between DMN regions in ASD children, as well as under-connectivity in adolescent and adult patients. This suggests a developmental heterogeneity not normally seen in typically developing peers<sup>220,266,267</sup>. Even within the DMN itself, differences have been reported in the name of hyper-connectivity between the posterior cingulate cortex and temporal and frontal regions, and hypo-connectivity between remaining regions<sup>219,266</sup>. Generally, the literature has converged on two major connectivity observations in the DMN in ASD patients: 1) increased within-DMN connectivity, particularly between the posterior cingulate and the medial prefrontal; and 2) reduced connectivity between DMN nodes and outside brain regions, both of which acquire an even more complex dynamic with age. Task-based studies have also offered insight into the understanding of the DMN in autistic patients, particularly revolving around socially related

tasks. Studies using self- and/or other- referential tasks generally involve the prefrontal cortex as well as aberrant patterns between the posterior cingulate and frontal regions during self-representation tasks in patients with ASD compared to controls<sup>267,268</sup>. Additionally, ToM tasks typically show reduced function of the temporoparietal junctions and prefrontal cortex in autistic adults<sup>269</sup>. Lastly, as predicted, structural abnormalities also exist in the DMN in patients with autism in several modalities including thickness, gyrification and volume. For example, increased cortical thickness has been reported in the posterior cingulate and the prefrontal cortex in all ages of autistic participants, as well as reduced grey matter volume in the right temporoparietal junction<sup>270–272</sup>. Gyrification studies have shown reductions in the brains of male autistic patients, but not females, in the medial prefrontal cortex when compared to controls<sup>273</sup>. Structural neurodevelopmental studies have further reported atypical age-related development in the DMN nodes throughout life in ASD patients, specifically, accelerated thinning in the bilateral posterior cingulate, and slowed volume reduction in the medial prefrontal cortex and temporoparietal junction<sup>274,275</sup>.



**Figure 1.3.5.** Main regions implicated in the Default Mode Network, including the frontal lobe, the posterior cingulate, the precuneus, the inferior parietal lobule and posterior temporal areas (*image taken from Kaplan et al., 2017*)<sup>276</sup>.

## Chapter 4: Dimensional Approaches in Autism Studies

As numerous mentioned throughout this thesis, autism studies have seen an extensive array of inconsistent results in almost all data modalities. This widespread heterogeneity points to several underlying physiological and etiological differences that are evident in the disorder. One way to try and disentangle this variability involves the use of dimensional approaches, which focus on the type and degree of several symptoms rather than placing individuals into one diagnostic category. This operates well in the study of diverse disorders such as autism and psychiatric conditions in general. Indeed, the only way to label an autistic individual is through diagnostic assessments, but often these assessments do not take into account the spectrum of genotypes, behaviors, and physiological phenotypes that accompany an autistic individual, making it difficult to characterize patients on the biological level. This section will consider current insights in the field of autism dimensionality.

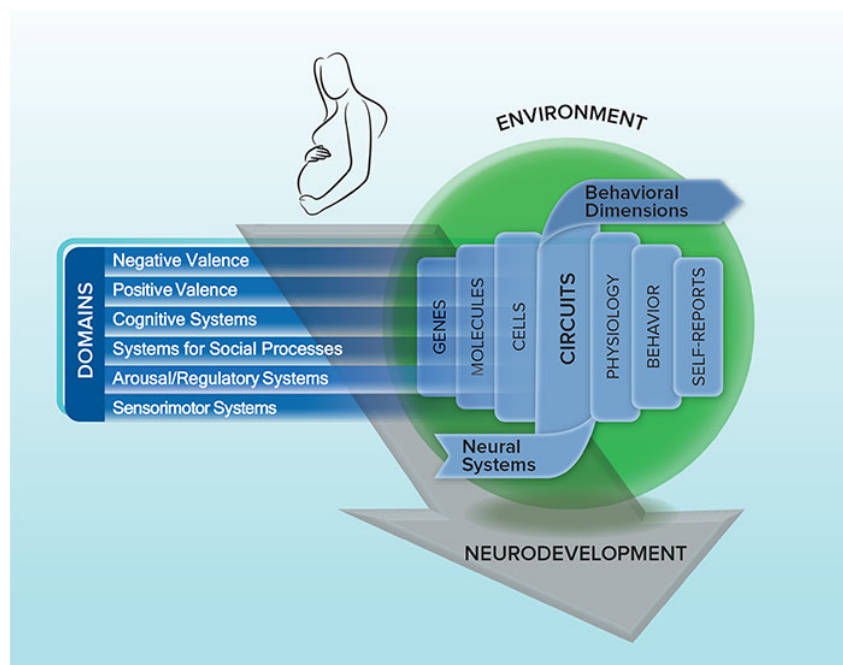
### 1.4.1. The Research Domain Criteria (RDoC)

The Research Domain Criteria (RDoC) is a framework for the investigation of psychiatric disorders that proposes the integration of several levels of information (including genetic, clinical assessments, biological, etc.), in order to explore dimensions spanning from normal to abnormal human behavior. Over the years, diagnostic categories have failed to capture underlying pathophysiological mechanisms of psychiatric disorders, and thus rarely successfully explain the biological aspects of such disorders. Historically, many psychiatric

disorders once thought to align in clinical presentations, have actually shown widespread negative results. Mental illnesses are unlike diseases such as cancer, which can be quantified and diagnosed with the help of laboratory tests (for the most part), therefore they are typically characterized by behaviors that are often shared with other psychiatric disorders. Moreover, these behaviors are challenging to quantify. Therefore, due to the qualitative nature of assessing mental illnesses, as witnessed in the variability of patients qualifying for the same symptom-based diagnosis, basing research on diagnostic categories can give rise to vast heterogeneity. Additionally, the existence of various comorbidities in psychiatric disorders further complicates the study of certain disorders in isolation. Therefore, targeting only patients with ‘pure’ diagnoses to participate in studies excludes those representing different ends of the functioning spectrum, ultimately hindering research. Diagnostic boxes can separate individuals that are in fact close in behavior, for instance a patient that just made the cut-off versus one that just missed it. These two patients may in fact share symptoms and underlying physiological mechanisms that are more similar than two patients *within* the diagnostic threshold. Such practices may therefore confound true results in developmental psychopathology and on the study of prodromal risk factors<sup>277</sup>.

The strategy behind the RDoC framework is implemented as a matrix of elements focusing on six major domains of human functioning, which are further divided into constructs measured using several units of analysis such as genetic, physiological, behavioral and clinical data points (Figure 1.4.1). This framework was created in 2009 with the idea that it would be filled and expanded as findings progress, eventually providing information on cognitive and biological processes underlying mental illnesses and paving the way to the development of better tools, diagnostic systems and treatments. While psychiatric disorders have until now been based on limited clinically-based classifications, this paradigm hopes to shift focus by incorporating

pathophysiological data to eventually help target better treatments and clinical decision making for patients with mental illnesses<sup>277</sup>.

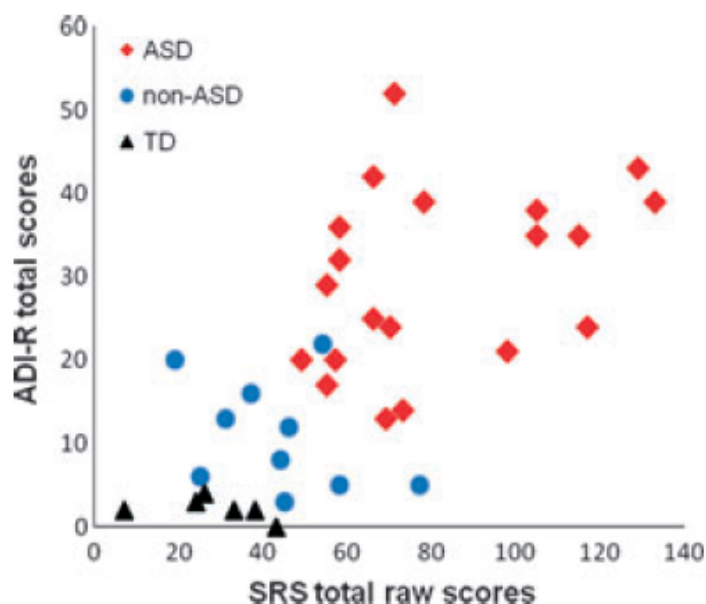


**Figure 1.4.1.** A schema of the RDoC framework showing the various levels of data constructs and dimensions needed in order to fully comprehend the spectrum between normal and abnormal human psychology (source: <https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/>).

#### 1.4.2. The Social Responsiveness Scale (SRS) as a Dimensional Construct

The Social Responsiveness Scale (SRS) is a qualitative measure assessing the continuum of clinical autistic traits in both children and adults. It consists of 65 items divided into five subscales representing different domains of social deficits including social awareness, social cognition, social motivation, social communication and autistic mannerisms (restricted interests and repetitive behaviors)<sup>278</sup>. Though not a diagnostic assessment, the SRS has proven to be a valid measure of autistic traits and is highly correlated with an ASD diagnosis. It is also often administered during clinical interviews to help provide a comprehensive understanding of ASD. The SRS exhibits high inter-rater and cross-cultural reliability, and correlates greatly with

ADOS and ADI-R diagnostic assessments for autism from the DSM-5, making it a robust measure in the dimensional study of ASD behaviors (Figure 1.4.2)<sup>278–281</sup>.



**Figure 1.4.2.** Plot showing the significant concordance between the SRS scale and diagnostic assessments for autism (*image taken from Kamio et al., 2013*). Non-ASD refers to patients with other psychiatric disorders.

As mentioned in the previous section, it is important to avoid excluding subjects falling short of the autistic diagnosis cut-off, as these subjects may in fact have similar symptoms to diagnosed patients and therefore be of great use in the search for biomarkers and therapies. Using a continuous trait versus a categorical one (i.e. the SRS versus an ASD diagnosis) also avoids selection and environmental biases that often accompany diagnosed individuals. It has been suggested that autistic traits run along a continuum extending into the general population, therefore a shift toward dimensional studies in this heterogeneous disorder has been steadily increasing. Furthermore, autistic traits assessed by the SRS exhibited a continuous distribution in the general child population (in a study conducted on 22 529 individuals) indicating no evidence of a natural gap differentiating diagnosed vs. non-diagnosed patients<sup>279</sup>. This however does not fully solve the issue of inconsistent results in autism research. Since etiological, developmental and biological heterogeneity is nevertheless present, it is still enormously

challenging to understand the disorder. This necessitates the characterization of distinct subgroups of ASD, which will be discussed in the next section. To this end, achieving such subgroups proves most relevant in combination with a dimensional approach, which promotes stratification according to behavioral and biological features as suggested by the RDoC. The SRS remains a strong candidate to use in such experimental designs when studying autism due to its continuity within ASD and in the general population, making it a good candidate to enter into the framework of RDoC. Furthermore, the SRS has already been employed in understanding autism in a wide range of publications varying from behavioral, to genetic, to neuroimaging studies<sup>278–280,282–287</sup>. As previously mentioned, the RDoC aims to integrate various constructs, phenotypes, and genotypes in order to explore a spectrum ranging from normal to abnormal human behavior, and as such, becomes conceivable in ASD when combining the study of the SRS scale with other variables.

#### 1.4.3. Studies Applying Dimensional Approaches in ASD

Indeed, the issue of heterogeneity in autism has shifted focus towards better characterizing the disorder, resulting in the interest of using autistic traits as proxies in studying ASD. The SRS has been used in several neuroimaging studies to test correlations between cortical morphology and symptom severity, showing that more symptomatic individuals tended to have a thinner cortices and decreased gyrification in temporal regions<sup>282,287,288</sup>. Furthermore, functional studies have also used the SRS to deduce information in the field of autism. Namely, several studies have reported altered activity in the cingulate region by comparing brain activity against the SRS and the Autism Spectrum Screening Questionnaire (ASSQ)<sup>289–291</sup>. These studies have

concluded pertinent information in the study of autism by conducting investigations centralized around autistic traits in the field of neuroimaging.

Autistic traits, as quantified by the SRS, have also been investigated and used in fields outside of neuroimaging. For example, vitamin D deficiency as well as higher concentrations of phthalate metabolites (which can be found in pollution and plastics) during pregnancy has lead to offspring with higher autistic traits as measured by the SRS<sup>292,293</sup>. The SRS scale has also been used to show that patients with higher symptomatology are more inclined to suffer from sleep disturbances and epilepsy, which are two well-documented factors experienced by autistic patients<sup>294,295</sup>. Interestingly, dimensional approaches have further been applied in genetic studies related to autism, including one study characterizing an increased incidence of autistic traits (as measured by the SRS) in a pool of patients diagnosed with neurofibromatosis type 1. Authors concluded that the neurofibromatosis type 1 gene is likely a quantitative trait locus of ASD<sup>296</sup>. All these studies prove that focusing on the dimensional aspect of autistic traits can perhaps deliver comprehensive and less restrictive evidence in the development and understanding of ASD.

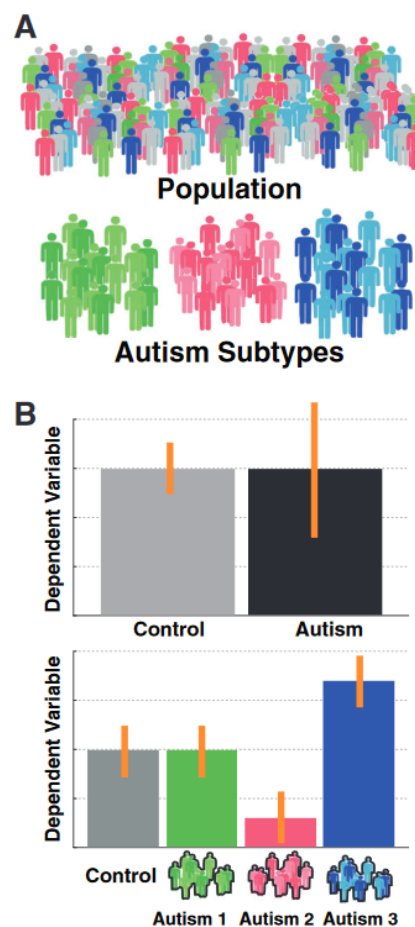


## Chapter 5: Subgrouping in ASD

### 1.5.1. The Necessity to Subgroup

In addition to using dimensional approaches to better understand autism, another concept extremely important in disentangling ASD heterogeneity is subgrouping. The inability to establish consistent empirical biomarkers due to general case-control designs in the literature has greatly impeded the understanding of autistic etiology and the development of proper treatment and therapies (Figure 1.4.3). The issue with autism classification is that the DSM-5 bases a diagnosis on three major symptoms: social deficits, social communication issues and restricted and repetitive behaviors. However, autistic patients actually manifest a wider variety of abnormal behaviors extending outside these bounds and into differences in biological outcomes. Thus, regardless of how often the DSM changes its definition of autistic disorders, using a categorical label smooths over tons of information that can be useful in advancing autism research. One example is that you can have an individual, *A*, who presents severe social deficits and negligible levels of repetitive behaviors as well as high levels of aggression and anxiety, is epileptic and experiences sleep disturbances. On the other hand, individual *B* presents minimal social deficits and high levels of repetitive behaviors as well as attention problems and gastrointestinal issues, but presents no epilepsy, sleep disturbances, aggression or anxiety. Since both of these patients are diagnosed with autism, could you administer them the same treatment or therapy and expect an attenuation of symptoms? Probably not. At the very least clinicians could probably suggest cognitive behavioral therapy for patient *A* to deal with his severe social deficits (as well as additional medications to treat the epilepsy and sleep

disorder), and serotonin receptor inhibitors for patient *B* to help alleviate the repetitive behavior impulses<sup>297,298</sup>. Thus, a clinician would be more likely to treat the individual symptoms, *dimensionally*, versus giving a *one size fits all* autistic medication. Such a targeted treatment does not seem to exist yet, and even if it did, it would probably not be useful for all patients along the autistic spectrum. Furthermore, not only do a variable combination of symptoms exist in ASD, but each symptom also fluctuates in severity. This could be one of the reasons as to why case-control paradigms in the literature have failed to yield consistent and reproducible results (of course with the exception of a few ground truths). Therefore, this thesis work highlights the necessity to advance autistic research within the realm of dimensional models as well as the need to further characterize common subgroups. Applying these approaches may offer hope in disentangling and understanding the heterogeneity present in ASD.



**Figure 1.4.3.** Conducting case-control experimental designs in the study of autism can smooth over important information about subjects. Isolating autistic subgroups provides detailed information about the disorder, which is more likely to lead to the development of better targeted therapies (image taken from Lombardo et al., 2019)<sup>299</sup>.

### 1.5.2. Studies Adopting Subgrouping Strategies in ASD

The variability of autism has perplexed researchers and clinicians alike, causing several of them to begin subgrouping the disorder. Work has been conducted on fractionating autistic patients according to cognitive profiles, clinical phenotypes, developmental patterns, genetic factors, and/or biological data. For example, a study by *Ellegood et al., 2015*, used mouse models to summarize the heterogeneity typically observed in neuroanatomical signatures of autistic patients by clustering 26 different autistic mouse models, which separated into three distinct subgroups each with unique cortical patterns<sup>300</sup>. The authors advocated the importance of subgrouping in autism research to better understand the disorder with supposition that each subgroup's homogeneous profile reflects a common underlying pathophysiology. Another study by *Stevens et al., 2019*, applied Gaussian Mixture Models with Hierarchical Clustering to identify behavioral subgroups of autism and subsequently examined treatment responses. Authors looked at academic, cognitive, play, motor, executive, language and social variables and isolated 16 subgroups including some subgroups that were similar in variables affected, but different in severity<sup>301</sup>. Authors also found that children within clusters responded more similarly to proposed treatments compared to a group comprised of all children diagnosed with ASD. What made this study particularly interesting was the incorporation of a treatment option for patients within clusters showing that treatment options work better when decreasing some of the heterogeneity, which is in line with the idea that homogenization of underlying mechanisms permits a better treatment response. It is also worth noting that the treatments administered in this study were 'cluster-specific', i.e. they were tailored to treat symptoms that were deficit within each subgroup, indicating the use of a dimensionally-oriented approach. Lastly, a study conducted by our group in *Mihailov et al., 2020*, revealed specific neuroimaging signatures associated to behaviorally clustered autistic traits subgroups<sup>302</sup>. Particularly, these signatures only appeared upon clustering, and disappeared in a case-control comparison, thus

again enforcing the necessity to subtype heterogeneous patients. Furthermore, some of the results in this study were also observed using another cohort where authors were able to generate many of the same behavioral subgroups, as well as a similar neuroanatomical alteration in one of the subgroups. Specifically, both studies generated an autistic traits subgroup high in ADHD-like deficits with cortical alterations in the motor area of the brain (*Mihailov et al., In Preparation*). Although replication between both studies was not impeccable, the fact that we were able to observe a stable cortical phenotype associated with a certain behavioral profile in an autistic traits subgroup already indicates the importance of limiting heterogeneity in search of consistent biomarkers in the field of ASD.

## Part II: Thesis Objectives

As thoroughly described in the preceding sections, and throughout this manuscript thus far, autism is an enormously heterogeneous disorder differing in etiology, biological traits, behaviors, clinical presentations, medical and psychiatric comorbidities. This established heterogeneity has hardly been overcome by the research conducted in the field of autism over the past couple of decades, most of which was centered around a case-control paradigm typically leading to the introduction of rampant variability. It seems the time has come either to retire this approach, or to introduce new and improved modes of analysis. The global objective of this work is to *better characterize autistic patients*, which is vital in the advancement of appropriate treatments and therapies. Our proposal to achieve this is two-fold, such that to successfully handle the variability present in autistic patients and disentangle the complex etiology of the disorder, it is crucial to: 1) apply dimensional approaches in order to recognize the disorder on a continual level and to include several different types of data, extending out from the core criteria necessary for an autistic diagnosis; and 2) subgroup patients according to intra-group similarities in order to potentially identify common traits and shared underlying pathophysiological mechanisms. By applying these approaches to ASD research, we hope to improve the biological understanding of the disorder in order to help enhance the development of therapeutic interventions.

The way this thesis work was able to achieve its goals was through the help of several specialists from a vast network of fields. This multidisciplinary environment permitted expertise insight from methodologists to improve experimental design, from computer scientists to expedite analysis protocols, from statisticians to ensure good statistical practices, and from psychiatrists to interpret and discuss theoretical questions. Such a rich environment laid the groundwork for a fruitful PhD with access to several experts and many resources. Another important factor in completing the objectives of this thesis work was access to large multidimensional cohorts

including neuroimaging data, clinical assessments, behavioral questionnaires, genetic information and even some biological data. Such richly phenotyped cohorts allow for the application of dimensional approaches due to their variety of data types, and for the stratification of autistic traits due to their large number of subjects. The next chapter will discuss the data used and the methodologies applied in order to conduct our objective of dimensionally-oriented stratification and overall characterization of autistic, and autistic-like subjects.

# Part III: Materials and Method

## Background



This part will aim to clarify and explain the theory behind the materials and methods used to conduct this thesis work in order to provide a baseline understanding of the methodologies presented in the studies.

# Chapter 1: Cohorts

Owing to the concept of machine learning, several research initiatives began curating richly phenotyped multidimensional data in order to comprehensively understand psychiatric disorders from several angles. Two cohorts were used to conduct this thesis work: The Healthy Brain Network Cohort, and the EU-AIMS Longitudinal European Autism Project Cohort, which will be discussed in the coming sections.

## 3.1.1. The Healthy Brain Network Cohort

The Healthy Brain Network Cohort (HBN) is a general population cohort that was launched by the Child Mind Institute as an initiative to help fast-track biomarker identification and to cause researchers to re-think the practices surrounding labels from clinically defined nosology, such as is from the DSM or ICD<sup>303</sup>. This cohort was established due to the necessity of studying the developing brain as well as factors underlying the emergence of psychiatric disorders throughout life, since it is reported that up to 75% of psychiatric disorders begin before the age of 24<sup>304</sup>. They also question the relevance of diagnostic boundaries and emphasize the need for transdiagnostic subtyping using rich dimensions from various types of data in order to address the heterogeneity existing in developmental psychopathology. In order to help researchers achieve these goals, the HBN cohort was created with the goal of obtaining 10 000 subjects (currently over 3 500 subjects have been registered) between the ages of 5 to 21 residing within the New York City area. The biobank includes data in the following domains: behavior,

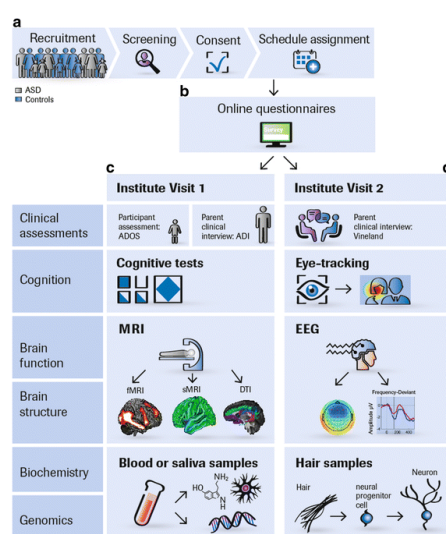
cognition, clinical presentations, brain imaging (including structural and functional MRI, connectivity, and EEG), genetic data, voice samples, and actigraphy. Among these data, there are measures also accounting for psychiatric status, learning assessments, familial structure, and environmental and lifestyle variables (Figure 3.1.1). Patients were recruited via community advertisements and word of mouth from clinician offices, which appealed to parents who held concerns about the mental health of their children or adolescents. This resulted in many of the subjects in the cohort being at-risk, or clinically diagnosed with a psychiatric disorder. This data is freely available to institutions interested in developing a project answering questions revolving around the neurodevelopment of psychiatric outcomes, making it a suitable cohort for applying dimensional approaches using symptomatic subjects.

General Information	Behavioral Measures
Demographics CMI Symptom Checker Edinburgh Handedness Inventory Intake Interview Physical Activity Questionnaire for Older Children (PAQ-C) (8-14) Physical Activity Questionnaire for Adolescents (PAQ-A) (14-19) Barratt Simplified Measure of Social Status Financial Support Questionnaire Medical History Questionnaire – Family Pregnancy and Birth Questionnaire Parent Mental Health Screen (PMHS)	Child Behavior Checklist (CBCL) Youth Self Report (YSR) (11-18) Adult Self Report (ASR) (18+) Screen for Child Anxiety Related Disorders (SCARED) – Parent Report & Self Report (8-18) Mood & Feelings Questionnaire (MFQ) – Parent Report & Self Report (8+) Affective Reactivity Index – (ARI-S) Self Report Columbia Suicide Severity Rating Scale (C-SSRS) – Self Report (7+) Extended Strengths and Weaknesses Assessment of Normal Behavior (E-SWAN) (5-17) Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale (SWAN) (6+) Conners ADHD Rating Scales Self Report Short Form (Conners) (8+) Repetitive Behavior Scale (RBS) Dishion Teacher (Dishion_Teacher)
Physical Measures	Autism Spectrum Screening Questionnaire (ASSQ) Social Communication Questionnaire (SCQ) Social Responsiveness Scale-2 (SRS-2) Strengths and Difficulties Questionnaire The Columbia Impairment Scale (CIS) Parent and self report Social Aptitudes Scale (SAS) WHO Disability Assessment Schedule (WHODAS) Parent and Self-Report Food Frequency Questionnaire (FFQ) (5-17) Positive and Negative Affect Scale (PANAS) Inventory of Callous-Unemotional Traits - Parent and Self Report General Self Efficacy (GSE) GRIT Scale (GRIT) Positive Behavior Scale (PBS) Child Flourishing Scale (CFS) Gilliam Autism Rating Scale-3 (GARS-3)
FITNESSGRAM (Pushups, Curl-ups, Trunk-Lift, Sit and Reach, Grip Strength) Cardiovascular/Endurance Fitness Test Vitals (Heart Rate, Blood pressure) Measurements (Height/weight, Waist circumference, Bio-impedance) Blood Draw Saliva and Hair Samples Baby Tooth Collection Urine Sample (Toxicology screen, Pregnancy test: 11+) Ishihara Color Vision Test Electroencephalography (EEG) Eye Tracking Magnetic Resonance Imaging (MRI) Peterson Puberty Scale (6-17) Sleep Disturbance Scale for Children (SDSC) (6-15) McMaster Pediatric Migraine Questionnaire (Migraine) Traumatic Brain Injury (TBI) Grooved Pegboard (Pegboard)	Family Structure, Stress and Trauma
Cognition and Language Tasks	Family History-Research Diagnostic Criteria (FH-RDC) Parental Stress Index IV (PSI-IV) Alabama Parenting Questionnaire – Self Report (APQ) (6-18) Alabama Parenting Questionnaire – Parent Report (APQ) (6-18) Children's Perception of Interparental Conflict (CPI-C) (8-18) Distress Tolerance Index – Parental Self Report Children's Coping Strategies Checklist – Revised (CCSC) (8-18) Negative Life Events Scale (NLES) – Parent & Self Report (8-18) Adverse Childhood Experiences Scale (ACES) (18+)
NIH Toolbox Tasks: Flanker, Card Sort and Processing Speed Temporal Discounting Task Adaptive Cognitive Evaluation (ACE) (5-12) Wechsler Intelligence Scale for Children-V (WISC-V) (6-17) Wechsler Adult Intelligence Scale-IV (WAIS-IV): (17+) Wechsler Abbreviated Scale of Intelligence-II (WASI): (17+) Wechsler Individual Achievement Test – II (WIAT) Differential Ability Scales – II (DAS) (5 or IQ below 70) Clinical Evaluation of Language Fundamentals – 5th Edition (CELF-5) Goldman-Fristoe Test of Articulation – II (GFTA) Comprehensive Test of Phonological Processing – II (CTOPP) Test of Word Reading Efficiency (TOWRE) (6+) Expressive Vocabulary Test (EVT) (when indicated) Peabody Picture Vocabulary Test (PPVT) (when indicated)	Substance Use and Addiction Measures
Diagnostic Assessments	Fagerstrom Test for Nicotine Dependence (FTND) (18+) Alcohol Use Disorders Identification Test (AUDIT) (11+) Modified Fagerstrom Tolerance Questionnaire- Adolescents (FTQA) (13-17) European School Survey Project on Alcohol & Other Drugs (ESPAD) (10+) Internet Addiction Test (IAT) Internet Usage Questionnaire (IUQ) - Parent & Self Report Parent-Child Internet Addiction Test (PCIAT) Yale Food Addiction Scale (YFAS) and YFAS-Child
Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) Child and Adolescent Psychiatric Assessment Schedule (Cha-PAS) (when indicated) Vineland Adaptive Behavior Scale – Parent/Caregiver Rating Form (when indicated) Yale Global Tic Severity Scale (YGTSS) (when indicated, 6+) Yale-Brown Obsessive Compulsive Scale (Y-BOCS) (when indicated, 18+) Children's Yale-Brown Obsessive Compulsive Scale (when indicated, 6-18)	Longitudinal Follow Up Measures
	Youth Services Survey (YSS) & Services Assessment for Children and Adolescents (SACA) Follow Up: CBCL Follow Up: Columbia Impairment Scale Parent and Self Report Follow Up: WHODAS Parent and Self Report

**Figure 3.1.1.** The full list of clinical, behavioral, and environmental assessments acquired in the HBN protocol (*image taken from Alexander et al., 2017*)<sup>303</sup>.

### 3.1.2. The EU-AIMS Longitudinal European Autism Project Cohort

The EU-AIMS Longitudinal European Autism Project Cohort was created for the purpose of studying ASD heterogeneity<sup>305</sup>. This cohort was initiated due to concerns that autism research was failing to produce seriously consistent results, which in turn has been slowing the development of proper treatments and therapies. As previously mentioned in the current manuscript, investigators similarly stress on issues with case-control paradigms and believe that a shift towards investigating intra-group variability may prove more insightful. For this reason, the EU-AIMS cohort was created and is the largest European multi-centre, multidimensional observational study appealing to researchers aspiring to achieve accurate ASD biomarker stratifications. Currently, the cohort includes 764 richly phenotyped individuals that are either controls (typically developed) or have been diagnosed with autism. Using an accelerated longitudinal design, participants are systematically characterized on several dimensions including behavior, clinical outcomes, neurocognition, environmental and familial structure, brain structure and function, and genetics (Figure 3.1.2). This cohort was created to fulfill the purpose of confirming, rejecting and/or refining current understandings of autism, with hopes of identifying meaningful autism subgroups to further help distinguish underlying mechanisms and etiologies, thus making it a perfect cohort for the purpose of this thesis work.



**Figure 3.1.2.** Schema illustrating the EU-AIMS cohort protocol and data types (image taken from Loth et al., 2017)<sup>305</sup>.

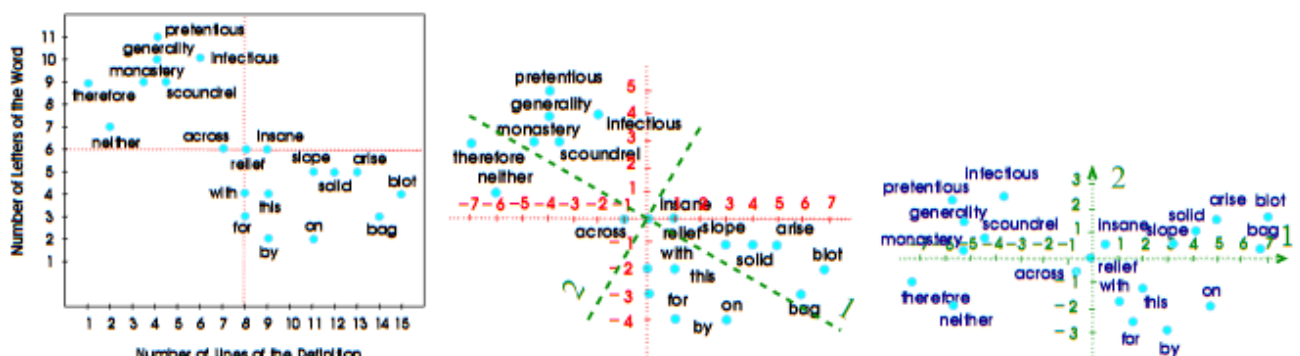
## Chapter 2: Unsupervised Learning Techniques

Unsupervised learning uses algorithms to find patterns in chaotic and unlabelled data in order to build a representation<sup>306</sup>. These types of algorithms are particularly useful in autonomously exploring large amounts of multidimensional data (or points in a multidimensional space) to delineate factors responsible for variability in a population. Unsupervised learning is different from supervised learning which is trained using a labelled dataset to fit a function of many covariates that can be used to make a prediction (an example of this is linear regression). The use of unsupervised techniques allows researchers to recognize similarities or dissimilarities between subjects without any labelled dataset, which is particularly useful in dimensional studies investigating heterogeneous psychiatric disorders. It allows the researcher to reveal underlying patterns using variables of interest without the presence of an a priori hypothesis, also preventing preceding biases. Additionally, such learning methods can aid in clustering patients by finding intra-similar subgroups throughout the data<sup>307</sup>. These unsupervised techniques are often derived from dimension reduction approaches, which remove factors that do not contribute to variability in the model, such as a principle component analysis<sup>308</sup>. In any case, this technique is extremely useful in psychiatric studies based on large multivariate cohorts.

### 3.2.1. Principle Component Analysis

A Principle Component Analysis (PCA) is a multivariate statistical technique using underlying mathematical principles to reduce dimensionality in a dataset with several possibly correlated variables. This is a common first step in analysing large datasets and entails transforming these several variables into fewer variables (obtained as linear combinations of the initial variables) that concisely explain the variability in the dataset, called principle components (PCs)<sup>309</sup>. The examination of the dataset in the ‘reduced’ space offers insight into observable patterns and/or trends in the data. The geometric process of obtaining the first two PCs can be summed up in three simple steps, which are summarized in Figure 3.2.1<sup>310</sup>:

- 1) First, the variables are centered and then plotted against each other.
- 2) Next, the main direction of the cloud of points is delineated, denoted as the first principle component, to obtain the minimum of the sum of squared distances from each of the points to the first component. The second component is added orthogonally to the first one using the minimum sum of squared distances.
- 3) Lastly, once the first two components have been established, the figure is rotated so that the first component lays horizontally, and the second vertically, thus acting as the new  $x$ - and  $y$ -axes.

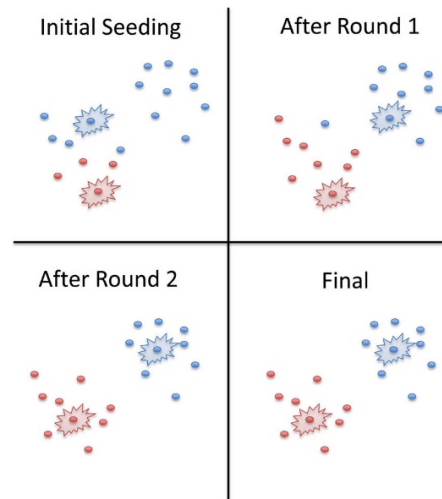


**Figure 3.2.1.** A basic summary to the geometric processes involved in the dimensionality reduction of a principle component analysis (image taken from Abdi and Williams, 2010)<sup>310</sup>.

### 3.2.3. Clustering Methods

Perhaps one of the most widely used unsupervised data mining techniques in neuropsychiatric research and beyond is clustering. The main purpose of clustering is to divide a population or data points into subgroups with maximum intra-cluster similarity and inter-cluster dissimilarity, segregating subgroups with similar traits<sup>307</sup>.

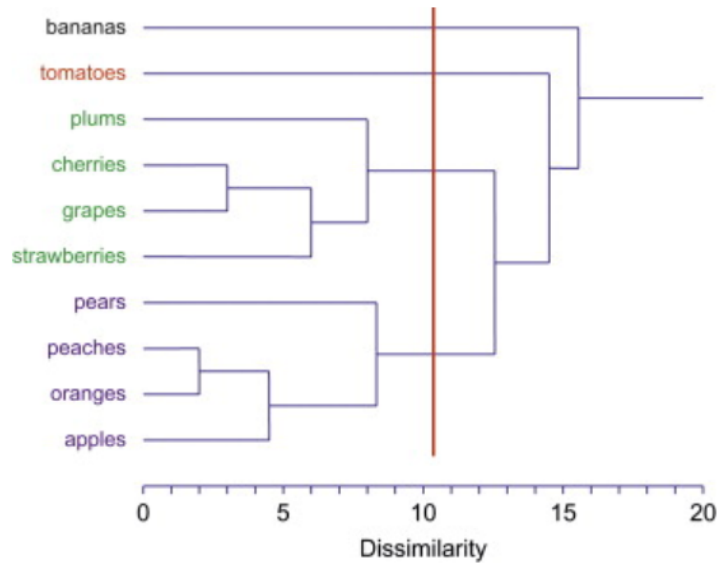
One of the most commonly used clustering techniques is the  $k$ -means clustering, which is an iterative algorithm aiming to find the local maxima in each iteration based on a specified number of clusters<sup>311,312</sup>. After specifying the desired number of  $k$  clusters, this algorithm randomly assigns each data point to a cluster after which the centroid is calculated using the mean value of each assimilated cluster. Each point is then *re-assigned* to the closest centroid, taking into account all surrounding data points per centroid. The final step in this method is to continuously re-compute the centroid means until all data points are assigned to the best possible centroid and no more improvements are possible (Figure 3.2.2)<sup>311</sup>. An alternative to  $k$ -means clustering is  $k$ -medoids, which computes the centroid medoids (analogous to linear medians, i.e. the most central point in a cloud of data points) instead of computing centroid means<sup>313</sup>.  $K$ -medoids can offer greater interpretability and is less noisy and sensitive to outliers as compared to  $k$ -means since the proposed cluster centroids are actual data points, versus a computed mean.



**Figure 3.2.2.** A simplified visual summary of the iterative steps taken to conduct a  $k$ -means clustering analysis, using the example of a 2 cluster solution (image taken from Page et al., 2014)<sup>314</sup>.

Hierarchical clustering is another technique for the isolation of subgroups. This method develops a hierarchy of clusters using a bottom-up algorithm that begins by assigning each data point its own cluster, and then successively merges pairs of the most similar clusters together until all clusters have been agglomerated into one single overarching cluster containing all the data points<sup>315</sup>. Normally, after assigning each data point its own cluster, a distance metric measuring the distance between two clusters is selected and used as a guide to iteratively combine two clusters at a time based on its lowest metric value. Then, as mentioned, this process is repeated until the root of the tree generates one large cluster containing all the data. The resulting clusters are normally represented in a dendrogram due to the hierarchical nature of this method (Figure 3.2.3). This technique does not require the preselection of a cluster number, but equivalently a ‘cutting point’ must be selected that decides where to split the tree into several branches<sup>315</sup>. A major difference between  $k$ -means and hierarchical clustering techniques is that hierarchical clustering cannot handle big data well, while  $k$ -means is less computationally intensive. Furthermore, hierarchical clusterings may be more reproducible than  $k$ -mean ones since the  $k$ -means method begins with an arbitrary choice of clusters, which may lead to differing results upon repeated runs<sup>311,315</sup>.

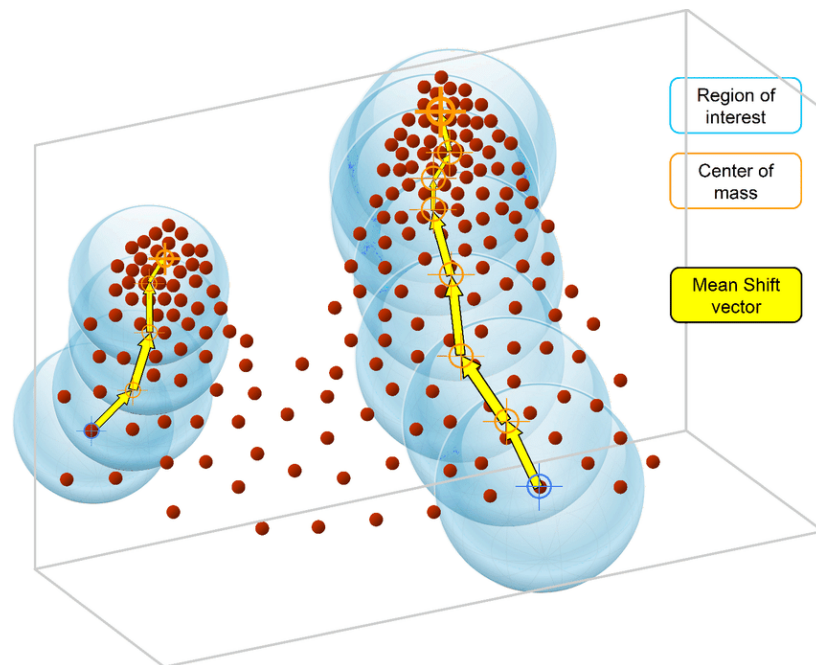




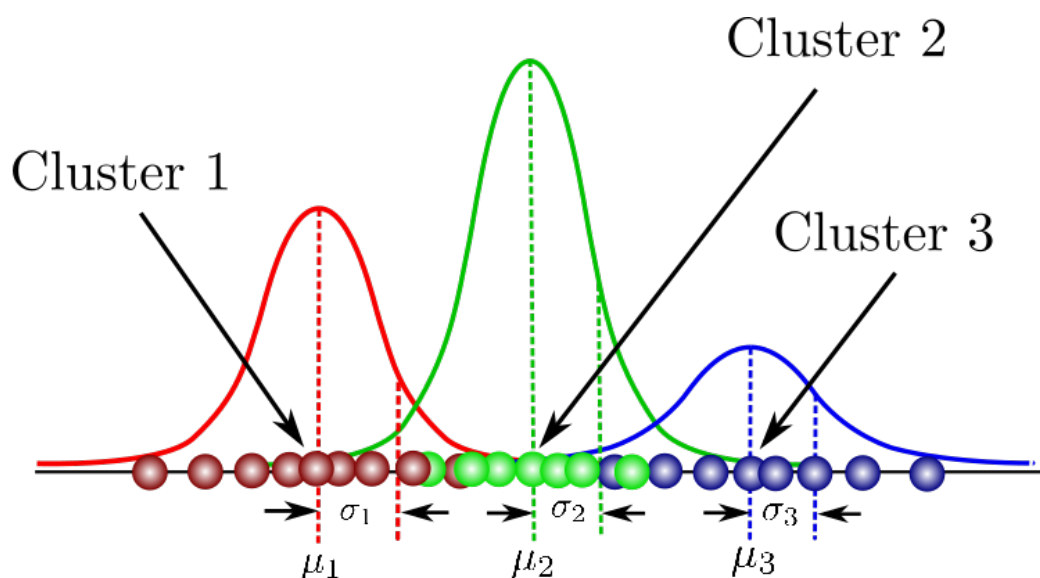
**Figure 3.2.3.** An example of a dendrogram typically generated using a hierarchical clustering method (*image taken from Tullis and Albert, 2013*)<sup>316</sup>.

Other less popular but still widely used clustering methods include mean-shift clustering and Gaussian Mixture Models. Mean-shift clustering is a centroid-based algorithm that attempts to find dense areas of data points<sup>317</sup>. It uses a ‘sliding window’ technique that moves throughout the data space and continuously updates center points as the average of points within the current ‘sliding window’ (Figure 3.2.4). This technique involves a post-processing stage that filters each window in order to exclude near-duplicates, finalizing a set of centroid points and their resultant subgroups. A benefit to using mean-shift clustering as opposed to  $k$ -means is that it is not necessary to choose a  $k$  number of clusters beforehand as this technique discovers this alone<sup>318</sup>. Additionally, Gaussian Mixture Models (GMM) are also interesting in the use of partitioning data since they extend the  $k$ -means algorithm in terms of the shape of clusters (i.e. perhaps the patterns do not always manifest in a ‘cloud-like’ manner with the same spatial extension)<sup>312,319</sup>. GMM use two parameters instead of one to interpret the shape of clusters, namely the mean and the standard deviation, and calculates the Gaussian distribution of each cluster (Figure 3.2.5). Subsequently, according to this distribution, the probability of belonging to a cluster is computed for each data point. This is iterated until near-complete convergence is

achieved. Since this process uses probabilities, each data point can have more than one ‘percentage chance’ of belonging to a cluster (i.e. a 60% chance of belonging to cluster A, and a 40% chance of belonging to cluster B).



**Figure 3.2.4.** An illustration of the mean-shift clustering algorithm showing its implementation of the ‘sliding window’ technique (image taken from Chen et al., 2018)<sup>318</sup>.

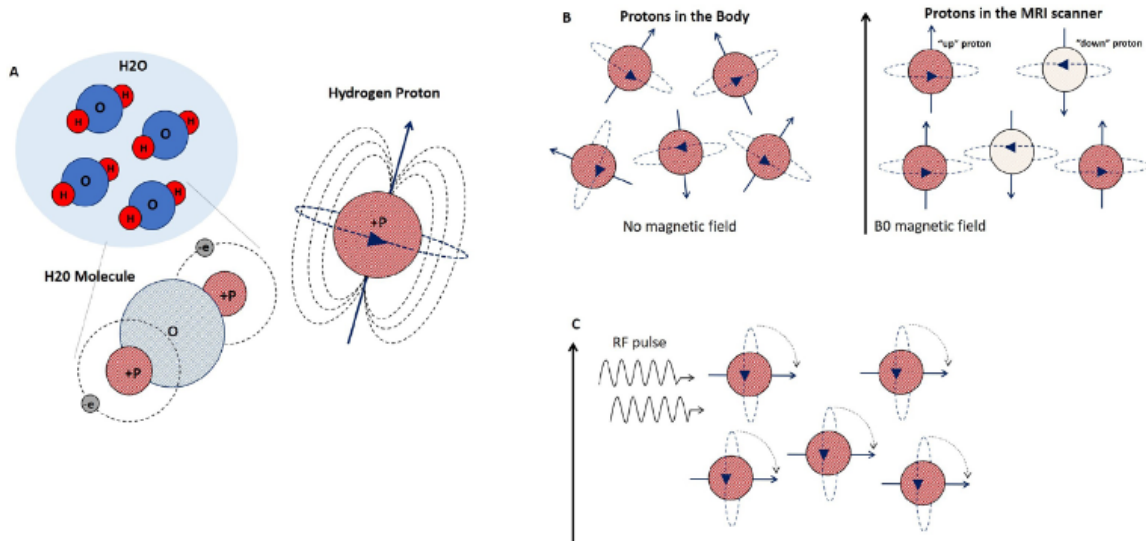


**Figure 3.2.5.** A schema outlining the method behind GMM clustering.  $\sigma$  represents the standard deviation and  $\mu$  represents the mean within each group (image taken from <https://towardsdatascience.com/gaussian-mixture-models-explained>).

## Chapter 3: Neuroimaging Methods

### 3.3.1. Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is an extremely important technique used in the acquisition of brain and general body images. Once the body is placed in an MRI machine, the energy emitted by hydrogen atoms within the alternating magnetic field results in MRI data in the form of internal images<sup>320</sup>. Specifically, MRIs employ powerful magnets that force the body's protons to align with their fields to which a radiofrequency is pulsed through the patient causing these protons to spin out of equilibrium. Once the radiofrequency is halted, the protons realign with the magnetic field, but in the process release energy that is detected by the MRI (Figure 3.3.1)<sup>321</sup>. It is this data that is eventually transformed into the resultant image. The protons releasing this energy essentially come from the hydrogen atoms of the water molecule. The specific weighting of the acquisition process enables the acquisition of images related to the local concentration of water, or to the local environment of the water molecules (respectively T1 weighting or T2 weighting), yielding either structural or functional images. A great advantage of this technique is that it is safe (even for sensitive populations such as pregnant mothers since it does not use ionizing radiation), and that it allows researchers to non-invasively obtain *in-vivo* information about organs in the body. MRI also comes with a few disadvantages such that it is extremely expensive to maintain and that it is sensitive to motion, which can be a problem for troubled patients that cannot remain still in tight spaces for long<sup>320</sup>.



**Figure 3.3.1.** A basic illustration showing the physical technique underlying MRI data acquisition (image taken from <https://frontiersin.org/article/10.3389/frym.2019.00023>).

In neuroimaging research, MRI allows for the study of structural and functional aspects of the brain. Structural brain MRI data is visualized by detecting contrasts between structures or tissues since they vary in intensity thus allowing us to directly distinguish, for example, grey and white matter, as well as subcortical structures, cerebrospinal fluid and blood vessels<sup>320,321</sup>. Functional neuroimaging visualization on the other hand, relies on the blood oxygenation level (BOLD) contrasts in the brain tissue, which occur when an area is active due to metabolic activity causing it to alter oxygen levels. Specifically, the regional oxygenation state is altered from paramagnetic deoxyhemoglobin to diamagnetic oxyhemoglobin causing a change in relaxivity rates, which modifies the local environment of the water molecules, and ultimately allows the MRI to visualize this BOLD contrast<sup>322</sup>.

In the study of structural neuroimaging, it is generally necessary to preprocess MR images prior to analysis<sup>323–326</sup>. The general preprocessing steps are described below:

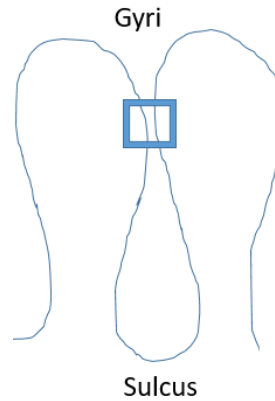
- A) *Bias Correction*: When a subject is placed inside the MRI magnetic field, variations in image intensities in the form of inhomogeneities may occur that need to be corrected otherwise the image may remain corrupted throughout the entire preprocessing stream.
- B) *Skull Stripping and Segmentation*: Skull stripping involves the removal of non-brain tissues from the image, while segmentation partitions remaining brain tissues such as grey matter, white matter, and cerebrospinal fluid.
- C) *Normalization*: This step is particularly important in group analyses as it spatially normalizes all subjects to a common space in order to limit inter-subject variability, and is necessary before running any statistical analyses. This process is normally executed in one of two ways: either by averaging all subject scans into one to produce a group representation, or by normalizing all subjects scans onto a predetermined template.
- D) *Smoothing*: This optional step is recommended in order to increase signal-to-noise ratio by further removing small anatomical differences between subjects. This is typically generated by averaging signal intensities of targeted data points between neighbouring points, ultimately resulting in seemingly smoother images.
- E) *Parcellation*: This step is often done in research studies to obtain average and robust measures in different brain regions. It is a process that partitions the brain into labelled parcels, or neuroanatomical ‘regions of interest’ according to a predetermined atlas.

### 3.3.2. Main Structural Imaging Approaches

There are three main approaches adopted in the literature for studying neuroanatomical structures of the brain. The first investigates the surface and is called vertex-wise analysis,

involving the recreation of the cortical surface into a mesh and aligning the vertices of this mesh to a common standard. It enables the visualization of the brain by computing local measurements along thousands of points on a 3D cortical representation<sup>327</sup>. The second approach is volumetric, called voxel-based morphometry, and involves obtaining features of the brain at the level of voxels instead of vertices<sup>328</sup>. Lastly, ‘regions of interest’ (ROI) analyses involve obtaining the average voxel or vertex value within a pre-defined brain region and consists of studying the brain from a more global perspective<sup>326</sup>.

Vertex-wise analysis (VWA) uses the vertex as a unit of analysis in the brain, which can be defined as a point on the surface of a 3D reconstructed cortex, extracted from MRI data<sup>327</sup>. What makes the VWA method exceptional is that its normalization procedure allows for analyzing a subject’s brain in its native space, which widely diminishes distortion. Registration between vertices is done on a point-by-point basis aligning individual curvature patterns. This allows for a high level of precision during statistical analyses between groups since morphometric measurements are computed in the subject’s native space at each corresponding vertex<sup>329</sup>. Another major advantage of this method is that it takes into account specific surface morphology, which is advantageous in the event that two regions are close in terms of Euclidean distance, but rather far in terms of surface distance (as is often illustrated by two sulcal edges; Figure 3.3.2). A popular neuroimaging program using this type of analysis is FreeSurfer (<http://surfer.nmr.mgh.harvard.edu/>)<sup>327</sup>.



**Figure 3.3.2.** Sometimes two regions that are close in terms of Euclidean distance may both be included in the same voxel (indicated by the blue box), even though they are in fact far from one another in surface morphology.

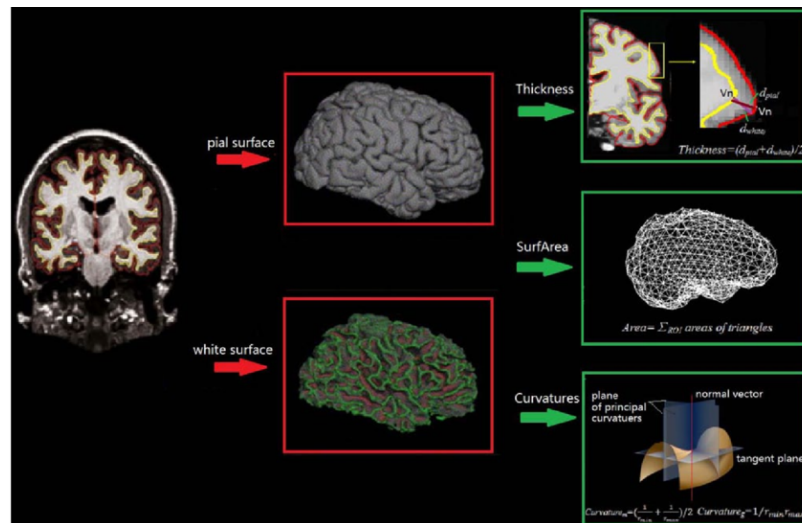
Voxel-based morphology (VBM) is an unbiased and completely automated process using MRI data to compare structural brain differences between groups of subjects using voxels, which are cube-shaped units of analysis that can be considered analogous to a pixel<sup>328</sup>. VBM is a method that is exceptionally useful in volumetric studies, particularly of the grey and white matter. As is also seen in VWA, VBM detects morphological differences between groups by adopting a whole-brain analysis technique, which is typically seen in exploratory analyses. Unlike VWA, VBM removes inter-individual differences by running a nonlinear spatial normalization technique that registers averaged images to a predefined template. This process is nonlinear in order to permit the warping of corresponding features of an individual image to a template before statistical group comparisons. A neuroimaging program that conducts this type of analysis is Statistical Parametric Mapping (SPM).

Both VWA and VBM methods are extremely important in the quantification and analysis of cortical morphology, however for this thesis work, VWA was chosen. Due to the nature and heterogeneity of the data used, it seemed more suitable to use a technique that performs a normalization in the native space of subjects. This means that the use of a template is not always

necessary, thus minimizing distortions. Also, conducting analyses from a native space is highly useful in such a multidimensional cohort where patients often express large cortical variabilities. Furthermore, using a VWA method allows for the study of gyrification and surface area, compared to VBM that focuses on volumes.

### 3.3.3. FreeSurfer Tool

The neuroimaging analysis in this thesis work was conducted using the FreeSurfer program, which is able to conduct all of the preprocessing elements described above (<http://surfer.nmr.mgh.harvard.edu/>). FreeSurfer is a semi-automated program used to visualize and analyze structural subcortical and cortical features such as cortical thickness, surface area, gyrification and volume, on the individual or group level. As mentioned above, this program runs a vertex-wise analysis on cortical thickness, surface area, and gyrification across a 3D reconstructed brain on over 300 000 vertices dispersed throughout the cortical surface. Cortical thickness is computed as the shortest distance between the pial and white matter surface, per vertex (Figure 3.3.3). Surface area is computed as the sum of the area of all triangles surrounding a vertex (Figure 3.3.3).

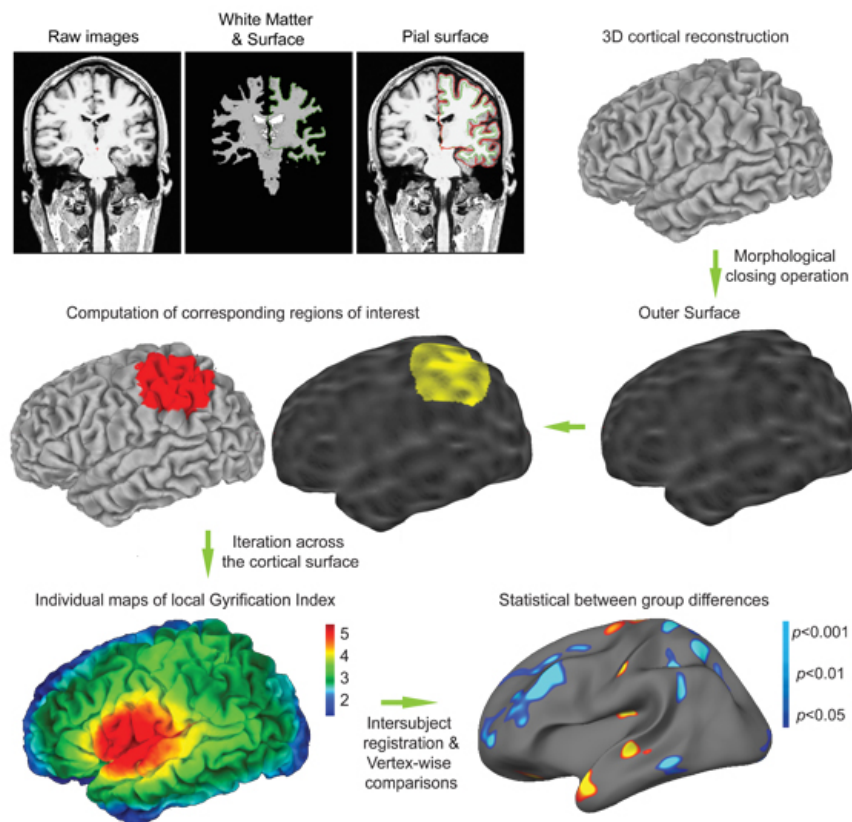


**Figure 3.3.3.** A schematic showing how FreeSurfer computes features of interest including surface area, thickness and curvature (image taken from Wang et al., 2014)<sup>330</sup>.



The local gyrification index (LGI), measuring gyral complexity, is slightly more complex and is calculated at each vertex as follows (Figure 3.3.4)<sup>331</sup>:

- 1) An outer surface is reconstructed to surround the entire pial surface in the form of a convex envelop (or simply put, as if the brain has been tightly wrapped in a bag).
- 2) Overlapping circular regions of interest are centered around each vertex along this outer enveloped surface.
- 3) Each overlapping circular region of interest has a corresponding circular region along the pial surface.
- 4) The LGI is computed as the ratio between each outer and inner circular region of interest along the entire brain.



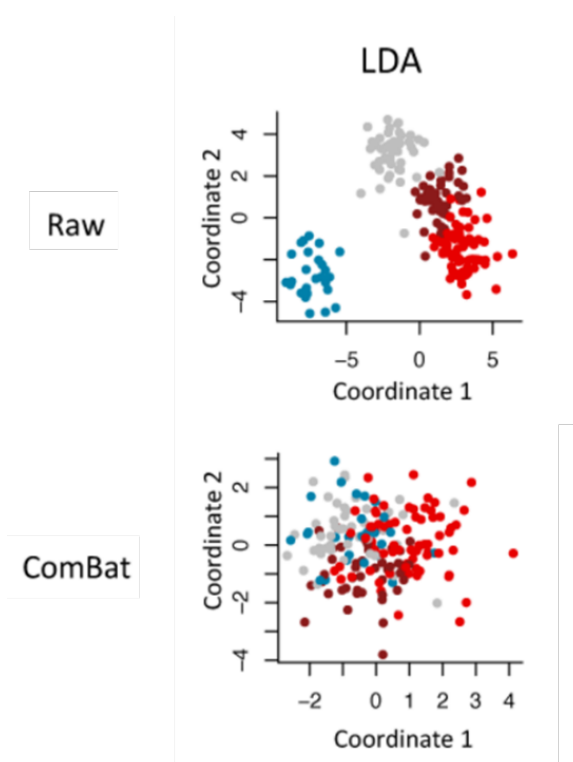
**Figure 3.3.4.** An overview of the process used in computing the local gyrification index (*image taken from Schaer et al., 2008*)<sup>331</sup>.

FreeSurfer is also capable of producing parcellated volumetric cortical and subcortical data, since it is not advised to conduct volumetric analysis using a VWA. This is done by automatically assigning a neuroanatomical label to each ROI according to a reference atlas. In FreeSurfer, there are two main ROI atlases that are used: the Desikan-Killiany atlas, comprised of 34 ROIs per hemisphere (68 total)<sup>332</sup>; and the Destrieux atlas comprised of 74 slightly smaller ROIs per hemisphere (148 total)<sup>333</sup>. FreeSurfer also provides data in subcortical volumes via an automated subcortical segmentation<sup>334</sup>.

One of the main reasons that FreeSurfer is used is for its functions in group analysis, which can be done via the command line or using the built-in user interface application called Query, Design, Estimate, Contrast (QDEC). Once the data has been preprocessed, the way in which FreeSurfer conducts its group analysis is by performing a general linear model (GLM) at each vertex according to the feature of interest. Linear modelling is typically used to describe data as a linear combination of explanatory factors in addition to noise. The quality of the description is determined by how well the data fits the linear model. The notation used by FreeSurfer is  $y = X*beta + Y*gamma$ , where  $y$  denotes the vector data (i.e. thickness or surface area for each subject per each vertex),  $X$  represents the design matrix covariates (such as diagnosis),  $Y$  represents the covariates of non-interest (such as age or gender), and  $beta$  denotes the unknown parameter estimates vector. The nullity of the **beta** parameter can be tested within QDEC, which reports results at the cluster level, i.e., at the level of a group of vertices. Since a huge number of tests are achieved at the cluster level, a specific non-parametric internal procedure is used to correct for multiple testing.

### 3.3.4. ComBat Tool: Harmonization of Multi-Site Imaging Data

With growing interest in big data, more and more cohorts are now built by including multiple sites for imaging acquisition. Technical variability exists across acquisition centres, which often introduces scanner effects from differences in scanner parameters, manufacturers and field strengths. This is common in the neuroimaging measurements extracted from modalities including MRI, fMRI and DTI. In order to address this issue and to maximize statistical power, data harmonization techniques like ComBat offer post-processing removal of unwanted site and/or scanner effects that may hinder results and cause specious findings (Figure 3.3.5)<sup>335,336</sup>. This method operates by adjusting the mean value and variance of feature measures across sites. ComBat, initially a popular batch-effect correction tool widely used in genomics studies, has proven successful in removing inter-site biological variability, even in cohorts containing few subjects. Empirical Bayes are used to improve the estimation of additive and multiplicative site effect parameters.



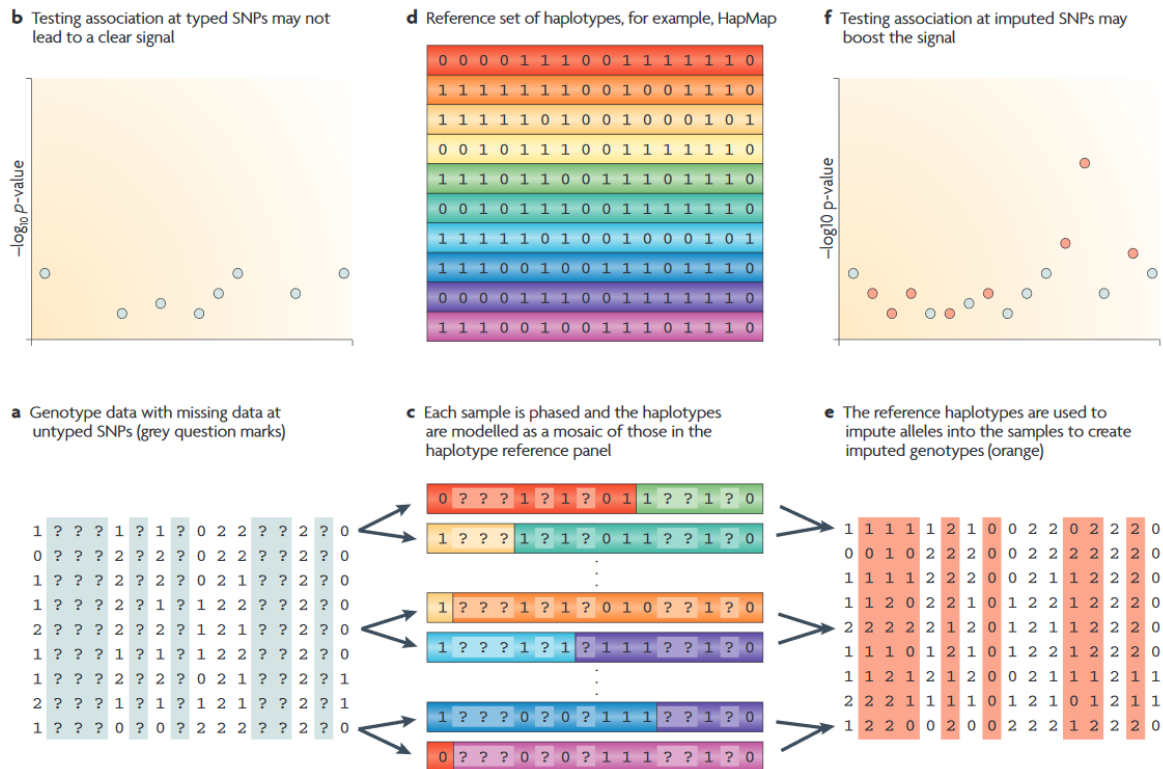
**Figure 3.3.5.** The plot above shows a typical Linear Discriminant Analysis (LDA, a dimension reduction technique) run on a dataset to show present variabilities due to site effects. The plot below shows the same dataset that has been homogenized by the removal of site effects using ComBat (image taken from Fortin et al., 2018)<sup>336</sup>.

## Chapter 4: Genetic Analysis

### 3.4.1. Imputation of Genomic Data

Genotype information is obtained via two major methods, next-generation sequencing and single nucleotide polymorphism (SNP) arrays. However, when it comes to obtaining whole-genome measurements to perform genome-wide association studies (GWAS) to detect loci, it often requires larger sample sizes, larger sequencing depths for whole-genome sequencing, and denser SNP arrays for microarray-based genotyping<sup>337,338</sup>. For cost effective reasons, compromises need to be reached and generally GWAS are carried out with measurements on limited sets of common SNPs (~1 million), which are provided by microarray-based genotyping. Nevertheless, the limited set of common variants chosen by the microarray manufacturer can be advantageously augmented to a border catalogue. This issue is solved using the process of imputation, which estimates missing genotypic data from haplotype or genotype reference panels via statistical inference, taking advantage of similar linkage disequilibrium patterns between reference and study cohorts<sup>338</sup>. Specifically, genetic imputation allows for the accurate evaluation of genetic markers that are not directly assayed, as it attempts to reveal commonalities between underlying haplotypes of study participants and the same haplotypes in a reference set in order to ‘uncover’ missing alleles. An example of a widely used reference set is the HapMap and 1000 Genomes Project panels<sup>339</sup>. The theoretical groundwork of genetic linkage studies and haplotype mapping approaches is formed on the idea that a modest set of genetic variants from a group of subjects can provide further data on other genetic variants in these same individuals<sup>338</sup>. This framework applies only if imputed samples refer to

individuals belonging to populations sufficiently represented in the reference panels. A summary of the full imputation process can be seen in Figure 3.4.1. This process of imputation is widely used in the analysis of genome-wide association studies (GWAS) since it increases SNP density (thus reducing the distance between SNPs) and can significantly boost power and fine-map associations<sup>337</sup>. By applying these enhancements, researchers increase their chances of identifying causal variants (such as for a disease) as well as conducting multi-platform meta-analyses<sup>340</sup>. Once genetic data are imputed, further analyses can be conducted on the new complete data including zygosity tests, allele association analyses and/or amino acid association analyses.



**Figure 3.4.1.** A visual summary of the genetic imputation process (*image taken from Marchini et al., 2010*)<sup>337</sup>.

### 3.4.2. PyHLA: Allele and Amino Acid Association Analysis

Allele and amino acid association analyses are typically conducted to test the association between them and susceptibility to a certain disease. The HLA loci are known to be one of the

most polymorphic and gene dense regions of the entire human genome with several associations between HLA variants and psychiatric and immune disorders. HLA nomenclature maintains the following pattern: *gene locus - antigen specificity (allele group) - specific allele*. The asterisk, ‘\*’, indicates that typing was performed using a molecular method and the colon, ‘:’, acts as a field separator. For example, for DQB1\*06:03, ‘DQB1\*06’ refers to the group of alleles encoding for that antigen, while the ‘:03’ refers to a specific allele encoding for a unique protein<sup>341</sup>. Numerous tools have been created to analyze HLA types, however most have been found inadequate in running association analyses of HLA types that have been realized from GWAS and next-generation sequencing (NGS) data. This inadequacy stems from the fact that many of these tools were designed for multi-allelic associations, and thus do not scale well with large and incessantly increasing amounts of reported HLA types<sup>342</sup>. The tailor-made PyHLA package (implemented in Python) was therefore created to solve this challenge, and mainly functions in allele and amino acid associations, however it can also run homo- or heterozygosity tests and interaction tests<sup>342</sup>.

PyHLA can be used to detect HLA alleles linked to diseases of interest by running an allele association analysis, with the possibility of applying different methods depending on the experimental design. First, at the most basic level, this program can run a Pearson’s chi-squared or a Fisher’s exact test on a 2 x 2 contingency table, which considers minor and major allele counts for a single locus in cases and controls by comparing one allele against all remaining alleles congregated together. This means that if an HLA gene (such as HLA-A or HLA-DRB1, for example) has  $n$  common alleles, then  $n$  tests are run where one allele is compared against the other  $n-1$  alleles grouped together. By running Pearson’s chi-squared or Fisher’s exact tests however, it is not possible to take into account covarying factors or to investigate continuous traits. To overcome this, PyHLA is also capable of running logistic and linear regression

analyses, both of which can incorporate multiple binary and/or continuous confounding variables. In order to examine amino acid associations with a particular disease, just as for the allele associations, PyHLA employs Pearson's chi-squared and/or Fisher's exact tests on a contingency table, as well as the possibility to run logistic and linear regression for more profound experimental designs. Amino acid and allele sequences are obtained from the IMGT/HLA database, which is a catalogue of highly curated HLA sequences<sup>343</sup>. Finally, all analyses have the possibility of being adjusted for multiple comparisons using Bonferroni or false discovery rate (FDR) corrections. All resources, tutorials and scripts are freely available at <https://github.com/felixfan/PyHLA>.

## Part IV: Studies Conducted in This Thesis



## 4.1. Study 1

### Cortical signatures in behaviorally clustered autistic traits subgroups: a population-based study

This study was published in June 2020 and can be accessed at: <https://doi.org/10.1038/s41398-020-00894-3>

***Cortical signatures in behaviorally clustered autistic traits subgroups: a population-based study***

*Angeline Mihailov, Cathy Philippe, Arnaud Gloaguen, Antoine Grigis, Charles Laidi, Camille Pigué, Josselin Houenou and Vincent Frouin. Translational Psychiatry, 2020.*

#### Study Summary

Extensive heterogeneity in Autism Spectrum Disorders (ASD) has hindered the characterization of consistent biomarkers, which has led to widespread inconsistent results in the literature. Thus, in order to better characterize ASD, homogenization *within* the disorder is necessary. To achieve this, we applied a dimensional approach, as suggested by the Research Domain Criteria (RDoC) Initiative, to conduct a study performing a clustering analysis on subjects differing in autistic traits. Isolating homogenized subtypes could provide insight into underlying biological mechanisms and an overall better understanding of ASD. The current study aims to delineate behavioral differences among ASD-like subjects and to subsequently further observe unique cortical signatures of identified subgroups. This will provide us with biological explanations for behavioral subtypes found within one of the core subconstructs of

ASD: social communication deficits, which is believed to be distributed along a continuum extending into the general population.

A total of 1093 participants from the population-based “Healthy Brain Network” Cohort (Child Mind Institute in the New York City area, USA) were selected based on score availability in behaviors relevant to ASD, aged 6 to 18 and FSIQ  $\geq 70$ . This consists of individuals at-risk for developing psychiatric conditions as well as typically developing participants. Core autistic traits are often concomitant with behavioral traits varying in type and degree and can often indicate the presence of a comorbid psychiatric disorder, thus further complicating the characterization of ASD. To this end, all participants were submitted to an unsupervised clustering analysis on behavioral dimensions that uncovered three subgroups with clinically significant levels of social deficits (serving as our impaired ASD traits subgroups): (1) high in emotionally dysfunctional traits, (2) high in ADHD-like traits, and (3) high in anxiety and depressive symptoms. Further cortical analyses were conducted to determine if elicited behavioral profiles were linked to unique structural brain signatures, which included 527 subjects with good quality structural MRI T1 data. Site effects on cortical features were adjusted using the ComBat method, which offers post-processing data harmonization to remove inter-site biological variability. Neuroimaging analyses compared cortical thickness, gyrification and surface area, and were controlled for age, gender, and FSIQ, and corrected for multiple comparisons. Upon investigating cortical phenotypes, after combining all three subgroups into one heterogeneous ‘ASD-like’ group and comparing it to controls, no significant differences were detected. However, unique cortical surface signatures were observed for each individual subgroup in comparison to controls.

These observations provide evidence of ASD traits subtypes, and confirm the necessity of applying dimensional approaches in order to extract meaningful differences. We hope these approaches can eventually reduce heterogeneity within the disorder and pave the way to better understanding ASD etiologies and developmental pathways that will eventually lead to the development of improved therapies and interventions.

## Materials and Methods

### *Part I: Clinical Profiles*

#### ***HBN Cohort and Participants***

The Healthy Brain Network (HBN) Cohort initiative within the Child Mind Institute began in 2015 with the goal of collecting brain imaging, cognitive/behavioral, and genetic data for 10 000 children and adolescents (5-21 years old) to investigate the heterogeneity behind neuropsychiatric and neurocognitive development<sup>303</sup>. It comprises a population of individuals at-risk for developing psychiatric disorders and typically developing participants. Subjects were recruited through flyer dissemination and subsequently assessed on clinical questionnaires at three sites in New York City, USA: Staten Island, Mobile Van, Midtown.

In the status of the HBN cohort, consensus diagnostics are not available for most of the subjects enrolled, however this does not preclude the possibility to carry out our dimensional study since subjects were not selected based on an ASD diagnosis, but rather on the presence of behavioral constructs relevant in the field of ASD. There were 1800 subjects participating at the time of this study, of which 1093 remained based on available overlap in behavioral scores assessing social deficits, hyperactivity, anxiety, irritability, depression, aggression and attention

problems, and having a full-scale Intelligence Quotient (FSIQ)  $\geq 70$ . We selected these 7 behaviors due to their presence in comorbid psychiatric disorders commonly reported in ASD patients, and therefore their frequent emergence along the autistic behavioral spectrum, implicating them in the understanding of ASD behavioral neuropathology<sup>62,85,89,344,345</sup>. Full-scale IQ was measured using the Wechsler Adult Intelligence Scale (WAIS-III, for those over 16) or the Wechsler Intelligence Scale for Children (WISC-III). Written informed consent was obtained from legal guardians and from participants themselves. This cohort study initiative was approved by the Chesapeake Institutional Review Board.

### ***Behavioral Assessments***

One of the most prominent dimensions in ASD patients is social impairment. Here, we used data from the widely used 65-item parent Social Responsiveness Scale (SRS) as a quantitative measure of clinical autistic traits, making it the central variable of interest in our study and in fact the score from which we separate out our autistic traits individuals. The SRS has been proven as a valid measure of autistic traits and thus has been used as a measure of autistic traits (for the purpose of understanding ASD) in several behavioral, genetic, and neuroimaging studies<sup>278–280,282–287</sup>. Though not a diagnostic tool, the SRS exhibits high inter-rater and cross-cultural reliability, and correlates greatly with the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview—Revised (ADI-R) diagnostic assessments for ASD from the DSM-5, making it a robust measure to use in the dimensional study of ASD behaviors<sup>278–281</sup>. Similarly, for the remaining behaviors we did not use diagnostic assessments but rather scales measuring behavioral trait severity. Hyperactivity levels were determined using the hyperactivity subscale within the Strengths and Difficulties Questionnaire (SDQ)<sup>346</sup>; anxiety was measured using the total score from the Screen for Child Anxiety Related Disorders Parent-Report (SCARED-P)<sup>347</sup>; irritability was defined using the total score of the Affective

Reactivity Index Parent-Report (ARI-P)<sup>348</sup>; and lastly, levels of depression, aggression, and attention problems were determined using subscales of the same names within the Child Behavioral Checklist (CBCL)<sup>349</sup>.

### ***Unsupervised Clustering Analysis (K-means)***

We conducted a *k*-means analysis on scaled z-scores of the previously mentioned 7 behaviors. This extracted subgroups varying in SRS and other accompanying behavioral characteristics. Briefly, *k*-means is an algorithm identifying mean cluster centroids, which serves to partition a sample into *k* subgroups<sup>350</sup>. A substantial challenge in such analyses lies in determining the number of clusters, which is a user-defined parameter. To address this problem, the chosen number of clusters *k* was determined using a Bayesian Information Criterion (BIC) distribution (Figure 4.1.1)<sup>351</sup>.

Mean behavioral scores, FSIQ and age were compared between subgroups using non-parametric 2-sided Mann-Whitney *U* tests, while gender differences were determined using a chi-square test. Python version 2.7 and R 3.4.0 were used on a Linux platform to perform all analyses in this study. Python packages used include Pandas (version 0.19.2), SciPy (version 1.1.0), and Matplotlib (version 1.5.1).

## ***Part II: Neuroimaging Analysis of Cortical Surface Features***

### ***Structural MRI Acquisition and Processing***

MRI acquisition took place at three different sites: mobile 1.5T Siemens Avanto in Staten Island, 3T Siemens Tim Trio at Rutgers University Brain Imaging Center, and 3T Siemens

Prisma at the CitiGroup Cornell Brain Imaging Center (acquisition protocols are extensively described in Alexander et al., 2017).

T1-weighted images were processed using the FreeSurfer software version 6.0.0 (<https://surfer.nmr.mgh.harvard.edu/>). For more information on precise methods of image analysis and the construction of anatomical information for each individual done by this software, refer to <sup>327,352</sup>. Briefly, the FreeSurfer analysis stream includes intensity normalization, skull stripping, and segmentation of gray (pial) and white matter surfaces<sup>327</sup>. Subsequent tessellation, as well as various topology corrections and inflation, leads to 3D meshes of cortical surfaces in different resolutions. Our work is based on a tessellation with ~160 000 vertices per hemisphere and used the FreeSurfer *fsaverage* template. We focused on three morphological measures of which the processing stream created vertex-wise maps for analysis: cortical thickness (CT), surface area (SA), and gyrification (IGI). The local gyrification index is measured as the ratio between buried and visible cortex<sup>331</sup>. All images were manually inspected in-house, in addition to using the Euler number as a metric of quality by retaining images at a threshold of -217, as specified in *Rosen et al., 2018*<sup>353</sup>.

### ***Elimination of Site Effects on Cortical Features Using ComBat***

A harmonization process was performed to account for the multiple acquisition sites. Features extracted from structural MR images are prone to technical variability across acquisition centers such as differences in scanning parameters, scanner manufacturers and field strengths. In order to remove cortical feature bias and variability caused by the unwanted site effects, the ComBat technique was applied to harmonize feature data along our three acquisition centres. This method adjusts the mean value and variance of feature measures across sites<sup>354</sup>.

## ***Statistical Analysis***

Vertex-wise statistical analyses were conducted using the command-line group analysis stream in FreeSurfer. Cortical surfaces for each participant were first registered to a study-specific template, then smoothed using a full-width-at-half maximum (FWHM) kernel of 10 mm for CT and SA, and 5 for IGI. A general linear model was fit at each vertex  $i$  to compare the three morphological measures between groups, using gender as a categorical covariate, and age and FSIQ as continuous covariates (site effects were already accounted for at the vertex level), and including the residual error:

$$y_i = \beta_0 + \beta_1 \text{Group} + \beta_2 \text{Sex} + \beta_3 \text{Age} + \beta_4 \text{FSIQ} + \varepsilon_i$$

We performed a cluster-level analysis using a cluster-forming threshold of  $p = 0.01$ . We report clusters with cluster-wise  $p$ -value ( $cwp$ ) of  $cwp < 0.05$ . These  $p$ -values were corrected for multiple comparisons using the *mri\_glmfit-sim* precomputed MonteCarlo simulation.

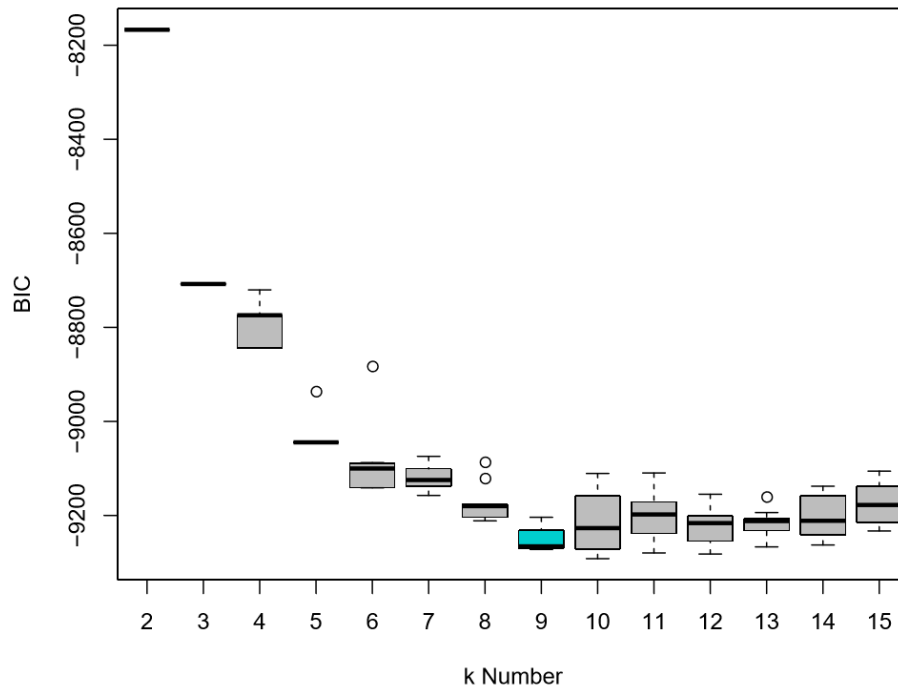
## **Results**

### ***Part I: Clinical Profiles***

#### ***Data-Driven Behavioral Subgroups in HBN Cohort***

Based on the computed BIC value distribution, a  $k$  value of 9 was retained as our supervised partitioning for this study (Figure 4.1.1). Thus, upon running the clustering analysis, we obtained 9 subgroups with various behavioral profiles (Table 4.1.1).

### Bayesian Information Criterion for Optimal Cluster Selection



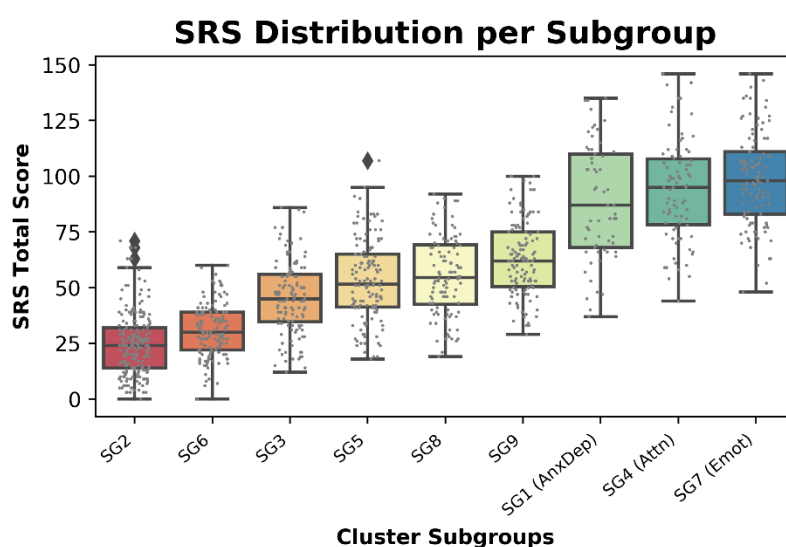
**Figure 4.1.1.** This curve suggests the optimal choice for  $k=9$ . We calculated Bayesian Information Criterion (BIC) distributions corresponding to clusters ranging from  $k=2$  to  $k=15$ , in order to gain confidence in our cluster selection. Moreover, this procedure was repeated with 10 different randomly chosen starting number points for each cluster number to observe and ensure validity.

Subgroups	<i>n</i> (Total = 1093)	SRS (s.d.): Social Deficits	SCARED-p (s.d.): Anxiety	ARI-p (s.d.): Reactivity	SDQ- Hyperactivity (s.d.)	CBCL-AB (s.d.): Aggression	CBCL-AP (s.d.): Attention Problems	CBCL-WD (s.d.): Depression	Mean Age (s.d.)	Gender Ratio	Mean FSIQ (s.d.)
<b>Subgroup 1 (AnxDep)</b>	61	<b>88.7* (26.0)</b>	<b>40.0* (13.0)</b>	6.1 (3.0)	5.9 (1.9)	12.6 (3.8)	<b>10.8* (3.4)</b>	<b>8.4* (3.1)</b>	12.7 (2.7)	35:26	96.9 (14.2)
Subgroup 2	221	24.4 (13.8)	6.7 (5.1)	0.8 (1.4)	2.0 (1.5)	1.9 (2.3)	2.1 (2.1)	0.8 (1.2)	11.2 (2.9)	114:107	101.9 (15.2)
Subgroup 3	124	45.6 (16.2)	22.8 (10.2)	2.8 (2.1)	3.1 (1.7)	5.4 (3.6)	4.9 (3.0)	4.5 (2.5)	12.3 (3.1)	59:65	100.1 (15.6)
<b>Subgroup 4 (Attn)</b>	82	<b>93.5* (22.5)</b>	20.0 (9.4)	2.2 (1.9)	<b>7.5* (1.9)</b>	7.4 (4.0)	<b>12.6* (2.7)</b>	5.5 (2.4)	11.7 (3.1)	57:25	93.3 (15.3)
Subgroup 5	130	53.3 (18.3)	11.6 (7.7)	7.5 (1.9)	6.0 (1.9)	14.0 (5.0)	8.7 (2.8)	2.6 (2.0)	10.6 (2.9)	90:40	99.5 (14.5)
Subgroup 6	141	30.6 (11.5)	5.9 (5.0)	1.1 (1.5)	6.2 (1.6)	3.9 (3.1)	8.0 (2.6)	0.9 (1.0)	10.8 (2.9)	99:42	98.7 (15.1)
<b>Subgroup 7 (Emot)</b>	107	<b>98.3* (20.7)</b>	24.7 (11.8)	<b>8.7* (2.3)</b>	<b>8.4* (1.5)</b>	<b>20.1* (5.0)</b>	<b>13.3* (3.4)</b>	5.9 (3.0)	10.7 (2.8)	71:36	95.7 (15.4)
Subgroup 8	104	55.5 (17.9)	14.3 (9.6)	3.0 (2.0)	<b>8.6* (1.4)</b>	10.1 (4.6)	<b>12.9* (2.3)</b>	1.7 (1.5)	10.1 (2.3)	78:26	98.3 (15.0)
Subgroup 9	123	63.3 (16.9)	14.0 (7.4)	1.7 (1.7)	5.9 (1.4)	4.4 (2.9)	6.9 (2.1)	2.0 (1.6)	10.6 (2.8)	77:46	96.1 (15.6)

**Table 4.1.1.** Mean behavioral scores and demographic data for each subgroup, including our 3 high autistic traits subgroups of interest (highlighted). Behavioral scores with a clinically high threshold were bolded and denoted with a '\*' superscript (*s.d* = *Standard Deviation*). \*Values surpassing high clinical levels of each score. Ranges for each score: SRS: 0 to 123+; SCARED-p: 0 to 82; ARI-p: 0 to 12; SDQ-Hyperactivity: 0 to 10; CBCL-AB: 0 to 40+; CBCL-AP: 0 to 22+; CBCL-WD: 0 to 17+. Age ranges: Subgroup 1 (AnxDep): 6.9 to 17.1; Subgroup 2: 6.0 to 17.7; Subgroup 3: 5.8 to 17.9; Subgroup 4 (Attn): 6.6 to 17.7; Subgroup 5: 5.8 to 17.6; Subgroup 6: 6.0 to 17.7; Subgroup 7 (Emot): 6.1 to 17.3; Subgroup 8: 6.2 to 16.4; Subgroup 9: 6.1 to 17.6.



The average SRS levels were used to decide which subgroups represented high autistic traits participants and which were controls. From these 9 subgroups, 3 expressed high levels of SRS, representing our socially impaired ‘high autistic traits’ subgroups. The SRS levels of these 3 subgroups fall within the ‘*severe*’ or at least upper ‘*moderate*’ classification of the SRS scale (an SRS value above ~80), thus indicating a high level of social impairment, providing us with greater confidence that subjects within these subgroups have ‘autistic-like’ traits (Figure 4.1.2).

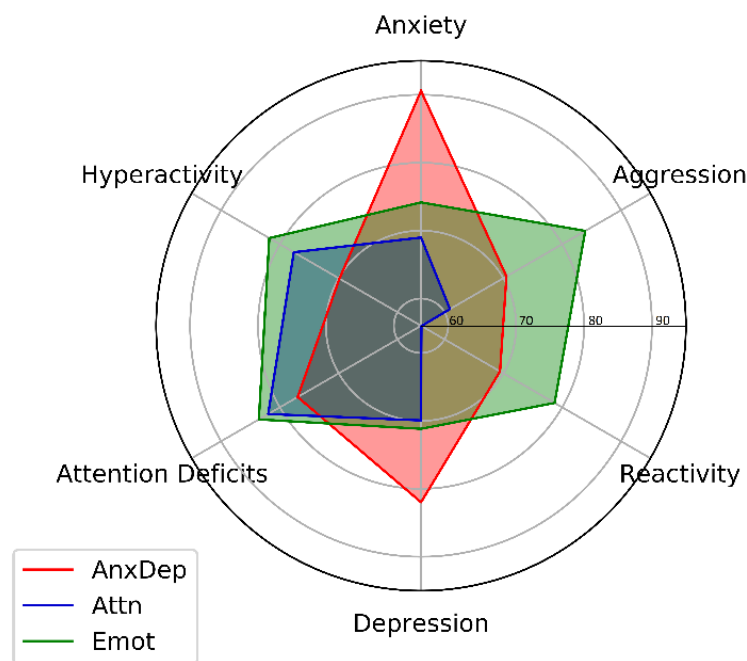


**Figure 4.1.2.** As seen in the above, 3 subgroups (from the 9 generated in the cluster analysis) showed high levels of SRS and were thus chosen as our high autistic traits subgroups for the current study. They were chosen based on raw values falling within the ‘*severe*’ or upper ‘*moderate*’ classification of the SRS (above ~80), thus indicating a truly high level of social impairment and providing us with greater confidence that subjects within these subgroups are indeed representative of having high autistic traits. Each boxplot indicates the median and ranges of raw SRS scores within each subgroup. Diamonds represent outlier values.

Additionally, mean SRS values in these 3 subgroups are comparable to the average SRS level of ~86 reported in diagnosed ASD patients<sup>279,355,356</sup>. Regarding the behavioral compositions of our 3 high autistic traits subgroups, one subgroup showed high levels of reactivity, aggression, and ADHD-like symptoms (hyperactivity and attention issues),  $n = 107$  (described as emotional dysregulation- *Emot*); the second maintained normal levels in all behavioral scores except for attention problems and hyperactivity,  $n = 82$  (described as attention problems- *Attn*); and the third showed high levels of anxiety and depression,  $n = 61$  (described as anxiety depression-

*AnxDep*)(Figure 4.1.3). Clinically high levels were determined for each behavioral measure according to the literature <sup>278,347,348,357–359</sup>. Though the remaining 6 subgroups contained subjects with SRS values ranging from low to high, each of these subgroups maintained an overall low SRS mean and were thus combined as our control group ( $n = 843$ ) with the aim of creating a representative general population without autistic traits subjects. Studies often barely obtain additional behavioral information on their controls other than a ‘non-diagnosis’ or ‘low SRS’. By combining these remaining 6 subgroups into one control group, we smooth out several behavioral heterogeneities and yield a control group composed of a wide behavioral spectrum, while still maintaining low mean levels of SRS (our target variable of interest to be contrasted in subsequent analyses).

## High Autistic Traits



**Figure 4.1.3.** A clustering analysis yielded 9 subgroups varying in behavioral composition. From these, 3 exhibited high SRS levels. The first subgroup (*Emot*), coloured in green, had strong emotional dysregulation (aggression and reactivity) with ADHD-like symptoms. The second subgroup (*Attn*), coloured in blue, showed ADHD-like tendencies. Lastly, the third subgroup (*AnxDep*), coloured in red, exhibited high levels of anxiety and depression, as well as attention deficits. This plot was built upon normalized scores that were converted to a scale of 1 to 100 (as indicated by each encircling grey line) for simplification.

Upon comparing the 3 socially impaired subgroups to one another, several significant differences in behavioral scores were found (Table 4.1.2A, ‘*Comparisons Between Subgroups*’). We also compared each autistic traits subgroup to controls and observed significant differences in all behavioral scores, except for reactivity in the *Attn* subgroup. Lastly, we decided to combine all 3 high autistic traits subgroups into one ‘autistic-like’ group (combined high SRS, *hSRS*) to compare against controls, which yielded highly significant differences in every behavioral score (Table 4.1.2A, ‘*Comparisons to Controls*’).

	SRS: Social Deficits	SCARED-p: Anxiety	ARI-p: Reactivity	SDQ: Hyperactivity	CBCL-AB: Aggression	CBCL-AP: Attention Problems	CBCL-WD: Depression	Age	Gender (K <sup>1</sup> )	FSIQ
<b>A. Behavioral Data Cohort (n = 1093)</b>										
Comparisons Between Subgroups										
Emot vs. Attn	0.07	<b>4.21x10<sup>-3</sup>*</b>	<b>7.30x10<sup>-3</sup>*</b>	<b>5.43x10<sup>-4</sup>*</b>	<b>1.85x10<sup>-3</sup>*</b>	0.053	0.15	<b>0.014 *</b>	0.32	0.13
Emot vs. AnxDep	<b>0.01*</b>	<b>1.77x10<sup>-11</sup>*</b>	<b>3.52x10<sup>-8</sup>*</b>	<b>4.58x10<sup>-14</sup>*</b>	<b>3.80x10<sup>-17</sup>*</b>	<b>2.03x10<sup>-5</sup>*</b>	<b>1.43x10<sup>-5</sup>*</b>	<b>6.80x10<sup>-4</sup>*</b>	0.76	0.25
Attn vs. AnxDep	0.16	<b>2.72x10<sup>-16</sup>*</b>	<b>4.72x10<sup>-14</sup>*</b>	<b>1.41x10<sup>-6</sup>*</b>	<b>8.13x10<sup>-12</sup>*</b>	<b>1.70x10<sup>-5</sup>*</b>	<b>5.92x10<sup>-5</sup>*</b>	<b>0.02 *</b>	0.19	0.056
Comparisons to Controls										
Emot	<b>3.71x10<sup>-16</sup>*</b>	<b>1.88x10<sup>-36</sup>*</b>	<b>3.33x10<sup>-36</sup>*</b>	<b>8.07x10<sup>-35</sup>*</b>	<b>1.26x10<sup>-34</sup>*</b>	<b>1.13x10<sup>-36</sup>*</b>	<b>3.17x10<sup>-36</sup>*</b>	0.15	0.37	<b>8.94x10<sup>-3</sup>*</b>
Attn	<b>3.75x10<sup>-11</sup>*</b>	<b>3.57x10<sup>-14</sup>*</b>	0.37	<b>3.48x10<sup>-17</sup>*</b>	<b>1.15x10<sup>-4</sup>*</b>	<b>2.49x10<sup>-36</sup>*</b>	<b>5.38x10<sup>-36</sup>*</b>	<b>0.03 *</b>	0.18	<b>2.01x10<sup>-4</sup>*</b>
AnxDep	<b>4.92x10<sup>-23</sup>*</b>	<b>8.92x10<sup>-34</sup>*</b>	<b>8.47x10<sup>-17</sup>*</b>	<b>7.51x10<sup>-4</sup>*</b>	<b>7.58x10<sup>-18</sup>*</b>	<b>9.30x10<sup>-16</sup>*</b>	<b>2.68x10<sup>-34</sup>*</b>	<b>2.69x10<sup>-5</sup>*</b>	0.64	0.13
<i>hSRS</i> (Emot + Attn + AnxDep)	<b>1.68x10<sup>-102</sup>*</b>	<b>1.19x10<sup>-36</sup>*</b>	<b>2.44x10<sup>-38</sup>*</b>	<b>5.40x10<sup>-40</sup>*</b>	<b>1.91x10<sup>-3</sup>*</b>	<b>2.15x10<sup>-47</sup>*</b>	<b>4.18x10<sup>-36</sup>*</b>	<b>7.14x10<sup>-3</sup>*</b>	0.3	<b>7.06x10<sup>-5</sup>*</b>
<b>B. Neuroimaging Cohort (n = 527; Subset of Behavioral Data Cohort)</b>										
Comparisons Between Subgroups										
Emot vs. Attn	0.09	0.49	<b>1.81x10<sup>-14</sup>*</b>	<b>4.01x10<sup>-3</sup>*</b>	<b>5.95x10<sup>-15</sup>*</b>	0.35	0.24	<b>4.49x10<sup>-3</sup>*</b>	0.76	0.21
Emot vs. AnxDep	<b>0.02 *</b>	<b>3.27x10<sup>-7</sup>*</b>	<b>2.78x10<sup>-5</sup>*</b>	<b>2.96x10<sup>-7</sup>*</b>	<b>1.08x10<sup>-10</sup>*</b>	<b>4.61x10<sup>-2</sup>*</b>	<b>3.93x10<sup>-1</sup>*</b>	<b>4.36x10<sup>-1</sup>*</b>	0.99	0.29
Attn vs. AnxDep	0.12	<b>1.88x10<sup>-7</sup>*</b>	<b>2.38x10<sup>-7</sup>*</b>	<b>4.14x10<sup>-9</sup>*</b>	<b>8.12x10<sup>-6</sup>*</b>	0.08	<b>5.83x10<sup>-4</sup>*</b>	0.06	0.66	0.11
Comparisons to Controls										
Emot	<b>2.41x10<sup>-26</sup>*</b>	<b>3.03x10<sup>-3</sup>*</b>	<b>1.93x10<sup>-23</sup>*</b>	<b>8.93x10<sup>-17</sup>*</b>	<b>2.05x10<sup>-26</sup>*</b>	<b>3.19x10<sup>-18</sup>*</b>	<b>1.23x10<sup>-15</sup>*</b>	0.09	0.31	<b>0.94 *</b>
Attn	<b>9.22x10<sup>-23</sup>*</b>	<b>1.25x10<sup>-8</sup>*</b>	0.31	<b>8.18x10<sup>-9</sup>*</b>	<b>1.15x10<sup>-3</sup>*</b>	<b>2.19x10<sup>-15</sup>*</b>	<b>8.80x10<sup>-17</sup>*</b>	<b>4.93x10<sup>-3</sup>*</b>	0.12	<b>1.85x10<sup>-1</sup>*</b>
AnxDep	<b>4.32x10<sup>-13</sup>*</b>	<b>5.01x10<sup>-17</sup>*</b>	<b>1.84x10<sup>-9</sup>*</b>	<b>2.35x10<sup>-3</sup>*</b>	<b>3.01x10<sup>-12</sup>*</b>	<b>1.94x10<sup>-9</sup>*</b>	<b>2.14x10<sup>-17</sup>*</b>	<b>5.34x10<sup>-1</sup>*</b>	0.61	0.16
<i>hSRS</i> (Emot + Attn + AnxDep)	<b>2.33x10<sup>-47</sup>*</b>	<b>3.10x10<sup>-25</sup>*</b>	<b>4.53x10<sup>-19</sup>*</b>	<b>1.75x10<sup>-20</sup>*</b>	<b>6.96x10<sup>-23</sup>*</b>	<b>1.09x10<sup>-19</sup>*</b>	<b>1.15x10<sup>-38</sup>*</b>	<b>8.25x10<sup>-3</sup>*</b>	0.07	<b>1.15x10<sup>-1</sup>*</b>

**Table 4.1.2.** Demographic and behavioral score comparisons, presented as p-values, between autistic traits subgroups, and between each autistic traits subgroup and controls. The top section shows these comparisons in the behavioral cohort (‘*A. Behavioral Data Cohort (n = 1093)*’), with autistic traits subgroups (i.e. *Emot*, *Attn*, *AnxDep*) compared to each other (under ‘*Comparisons Between Subgroups*’), each autistic traits subgroup compared to controls, and finally a combination of all three autistic traits subgroups (*hSRS*) compared to controls (under ‘*Comparisons to Controls*’). The bottom section (*B. Neuroimaging Cohort (n = 527; Subset of Behavioral Data Cohort)*) presents the same comparisons in the neuroimaging cohort, which is composed of a subset of subjects from the behavioral cohort with usable structural MRI data. Significant p-values are bolded with a ‘\*’ superscript.

With respect to demographic information, there were no significant differences in gender and FSIQ between all subgroups. However, there were reported differences in age, though age ranges were similar (mean age = 10.8, SD = 3.4)(Table 4.1.2A, ‘*Comparisons Between Subgroups*’). Upon comparing each subgroup to controls, we again found no differences in gender. We did however find differences in age between all subgroups and controls, except for *Emot*. Although age differences were present, age ranges were again similar. FSIQ differed between all subgroups and controls, except for *AnxDep*, which is to be expected since autistic traits are generally accompanied by differences in FSIQ. Lastly, the comparison between the *hSRS* group and controls yielded significant differences in age and FSIQ, but not gender (Table

4.1.2A, ‘*Comparisons to Controls*’). Due to these differences, we deemed it important to control for FSIQ, age and gender in the subsequent neuroimaging analysis.

## Part II: Neuroimaging Analysis of Cortical Surface Features

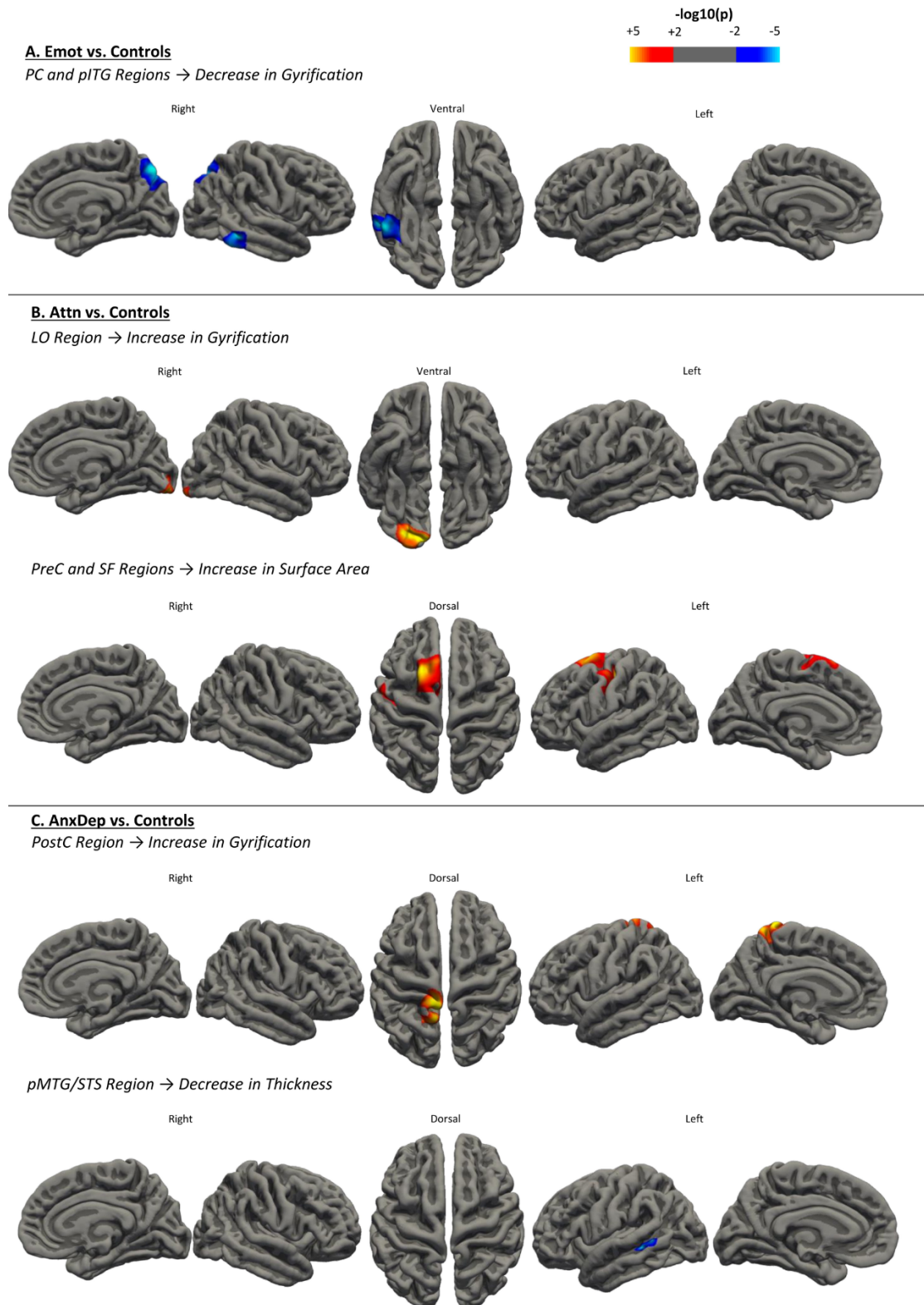
### ***Morphological Comparisons***

After removing subjects that have not undergone MRI acquisition and/or did not pass the T1 image quality check, as well as those removed during the outlier detection step, we obtained a sample of:  $n = 47$  in the “*Emot*” group,  $n = 39$  in the “*Attn*” group,  $n = 31$  in the “*AnxDep*” group, and  $n = 410$  controls, producing a total of 527 subjects with available T1 data participating in the study (Table 4.1.3). Behavioral score and demographic information comparisons were nearly identical to the behavioral cohort (Table 4.1.2B, ‘*Neuroimaging Cohort*’). To delineate the interest and significance of subtyping in an autistic traits population, we first combined all 3 subgroups into one large group (*hSRS*,  $n = 117$ ) and compared cortical thickness, local gyrification and surface area against controls. Indeed, this comparison did not yield significant differences in any of the measured surface features. We then compared the same surface features between each of our 3 subgroups against controls. After correction for multiple comparisons, the *Emot* subgroup exhibited decreases in gyrification in the right hemisphere in two separate clusters, one spanning the precuneus (including parts of the superiorparietal area)( $p < 0.01$ , Cohen’s  $d = 0.51$ ), and another in the temporal lobe (including the posterior inferior temporal gyrus and the middle temporal)( $p < 0.01$ , Cohen’s  $d = 0.48$ )(denoted as PC and pITG)(Figure 4.1.4A). The *Attn* subgroup displayed elevated local gyrification peaking in the lateraloccipital area of the right hemisphere (denoted as LO)( $p < 0.01$ , Cohen’s  $d = 0.41$ ). Additionally, the *Attn* subgroup also exhibited two separate clusters in the left hemisphere showing increases in surface area in the precentral cortex (along the central

sulcus)( $p < 0.01$ , Cohen's  $d = 0.61$ ), and superiorfrontal regions ( $p < 0.01$ , Cohen's  $d = 0.58$ )(denoted as PreC and SF)(Figure 4.1.4B). Lastly, the *AnxDep* subgroup showed increases in gyrification spanning the left postcentral and precuneus regions (PostC)( $p < 0.01$ , Cohen's  $d = 0.33$ ), and decreases in thickness in the left posterior middle temporal gyrus lining the superior temporal sulcus (pMTG/STS)( $p < 0.01$ ; Cohen's  $d = 0.55$ )(Figure 4.1.4C)(Table 4.1.4).

Subgroups	Subgroup Characteristics	<i>n</i>	SRS (s.d.): Social Deficits	SCARED-p (s.d.): Anxiety	ARI-p (s.d.): Reactivity	SDQ-Hyperactivity (s.d.)	CBCL-AB (s.d.): Aggression	CBCL-AP (s.d.): Attention Problems	CBCL-WD (s.d.): Depression	Mean Age (s.d.)	Gender Ratio	Mean FSIQ (s.d.)
Emot	High attention deficits, hyperactivity, aggression and reactivity	47	<b>95.2* (19.4)</b>	21.4 (10.6)	<b>8.6* (2.3)</b>	<b>8.2* (1.4)</b>	<b>20.8* (4.8)</b>	<b>13.0* (3.4)</b>	6.0 (3.3)	10.8 (3.4)	30:17	96.7 (17.2)
Attn	High attention deficits and hyperactivity	39	<b>89.2* (23.9)</b>	21.1 (9.2)	2.3 (2.1)	<b>7.2* (1.9)</b>	7.4 (4.1)	<b>12.5* (2.8)</b>	6.2 (2.5)	12.5 (2.5)	27:12	93.3 (13.0)
AnxDep	High anxiety and depression, as well as attention deficits	31	<b>82.2* (26.9)</b>	<b>37.6* (12.9)</b>	6.0 (2.9)	5.9 (2.0)	13.0 (3.0)	<b>11.3* (3.5)</b>	<b>8.7* (3.4)</b>	13.4 (2.4)	19:12	98.2 (15.3)
hSRS (Combination of previous 3)	High autistic traits	117	<b>89.7* (23.5)</b>	25.6 (13.0)	5.9 (3.7)	<b>7.2* (2.0)</b>	14.3 (7.1)	<b>12.4* (3.3)</b>	6.8 (3.2)	12.1 (3.0)	76:41	96.0 (15.4)
Controls	General population with low or absent autistic traits	410	40.3 (21.4)	12.0 (9.3)	2.6 (2.9)	4.5 (2.8)	5.6 (5.5)	6.1 (4.4)	2.1 (2.3)	11.3 (3.0)	225:185	101.0 (15.4)

**Table 4.1.3.** Demographic information for subgroups from the neuroimaging cohort. Mean and standard deviation information is included for autistic traits subgroups on age, FSIQ, and behavioral scores, as well as gender ratio. The same information is also included for the combined *hSRS* group and the control group. (*s.d* = *Standard Deviation*). Clinically high values were bolded and denoted with a ‘\*’ superscript. Ranges for each score: SRS: 0 to 123+; SCARED-p: 0 to 82; ARI-p: 0 to 12; SDQ-Hyperactivity: 0 to 10; CBCL-AB: 0 to 40+; CBCL-AP: 0 to 22+; CBCL-WD: 0 to 17+.



**Figure 4.1.4:** Surface feature comparisons between each subgroup and controls. A) The *Emot* subgroup yielded decreases in gyrification in the right precuneus and temporal regions ( $cwp = 0.0004$  and  $0.005$ , respectively). B) The *Attn* subgroup exhibited increases in gyrification in the left lateraloccipital region ( $cwp = 0.002$ ), and increases in surface area in the left precentral and superiorfrontal regions ( $cwp = 0.02$  and  $0.02$ , respectively). C) The *AnxDep* subgroup showed increases in gyrification in the left postcentral area ( $cwp = 0.02$ ), and decreases in thickness in the left middle temporal gyrus/superior temporal sulcus ( $cwp = 0.04$ ). Colours represent the  $-\log_{10}(p\text{-value})$ , with red(+) indicating an increase and blue(-) indicating a decrease compared to controls in affected morphological features.

Subgroups	Feature	Hemisphere	Cluster Name	All areas included in cluster	Peak region	Peak MNI co-ordinates			Size of region (mm <sup>2</sup> )	Cluster-wise p-value (cwp)
						X	Y	Z		
Emot	↓ Gyrfication	Right	<b>PC</b> (Precuneus)	Precuneus, Superiorparietal	Precuneus	14	-71.4	40.9	1860.87	<b>0.0004</b>
	↓ Gyrfication	Right	<b>pITG</b> (Posterior Inferiortemporal)	Inferiortemporal, Middletemporal	Middletemporal	59	-42.4	-10.4	1237.64	<b>0.005</b>
Attn	↑ Gyrfication	Right	<b>LO</b> (Lateraloccipital)	Lateraloccipital, Lingual	Lateraloccipital	26.5	-87.2	-10.7	1395.61	<b>0.002</b>
	↑ Surface Area	Left	<b>PreC</b> (Precentral)	Precentral, Central Sulcus	Precentral	-32	-18.1	44.2	960.47	<b>0.02</b>
	↑ Surface Area	Left	<b>SF</b> (Superiorfrontal)	Superiorfrontal	Superiorfrontal	-18.1	10.5	60.8	985.47	<b>0.02</b>
Anxdep	↑ Gyrfication	Left	<b>PostC</b> (Postcentral)	Postcentral, Central Sulcus	Postcentral	-8.2	-38.8	74.1	969.46	<b>0.02</b>
	↓ Thickness	Left	<b>pMTG/STS</b> (posterior Middletemporal/Superior Temporal Sulcus)	Middletemporal, Superior Temporal Sulcus	Middletemporal	-49.1	-39	-5.6	399.23	<b>0.04</b>

**Table 4.1.4.** Neuroimaging cluster coordinates Information. The type of morphological feature (↑ indicating an increase, and ↓ indicating a decrease), hemisphere, and affected brain regions are indicated above for each autistic traits subgroup. Additional information on MNI coordinate data, region size, peak region and cluster-wise p-value for each result within each autistic traits subgroup are also reported.

## Discussion

Autism Spectrum Disorder studies have unceasingly demonstrated heterogeneity, warranting a shift in focus towards initially characterizing these differences before subsequent analysis, and steering away from case-control studies. To this end, a dimensional approach proves most relevant. To the best of our knowledge, this is the first study using an unsupervised clustering analysis on a population-based cohort to investigate how autistic traits cluster with other behavioral dimensions into subgroups, with subsequent isolation of subgroup cortical signatures. Recent evidence advocates that autistic traits fall along a continuum within the general population, which was why this study was not limited to diagnosed individuals, but rather focused on autistic traits as absolute constructs in order to avoid potential selection or environmental biases often accompanying diagnosed patients. We obtained 3 autistic traits subgroups in our unsupervised clustering analysis with the following behavioral profiles: 1) high aggression, reactivity, and ADHD-like traits (*Emot*), 2) high in attention deficits and hyperactivity (*Attn*), and 3) high in anxiety and depression, as well as attention deficits (*AnxDep*). Furthermore, upon neuroanatomical investigation, we found that comparing each subgroup to controls uncovers unique cortical signatures. Namely, after correcting for multiple

comparisons, the *Emot* subgroup showed decreased gyrification in precuneus and posterior inferior temporal regions (PC and pITG); the *Attn* subgroup displayed increases in gyrification in the lateraloccipital (LO) and increases in surface area in the precentral and superiorfrontal regions (SF); and lastly, the *AnxDep* subgroup exhibited an increase in gyrification in the postcentral cortex (PostC), as well as a decrease in thickness in the posterior middle temporal and superior temporal sulcus area (pMTG/STS). Most interestingly, we found that by comparing the structural brain features of one heterogeneous autistic traits group (composed by combining all 3 subgroups) to controls, we were unable to uncover any cortical signatures. Simply comparing behaviorally diverse ASD-like cases to controls proved far too rudimentary to yield consistent features.

Though several case-control studies have reported neuroanatomical differences in ASD populations, these studies have remained grossly inconsistent, possibly due to ASD heterogeneity. Here, we showed that by running a direct comparison between subjects having high versus low/absent autistic traits, no cortical differences were reported. In a study by Haar et al., authors compared cortical thickness differences between ASD and controls and ran both univariate and multivariate comparisons<sup>270</sup>. Results were strikingly weak and were attributed to the considerable heterogeneity of the ASD population. The authors ultimately suggested that previously reported neuroanatomical differences between cases and controls held low clinical significance, and advocated the necessity of subdividing ASD groups by genetic, clinical and/or behavioral traits in the identification of unique neuroanatomical abnormalities<sup>270</sup>. Further studies in animal research have also encouraged subtyping in ASD, namely a study by Ellegood et al., which ran a clustering analysis on ASD neuroanatomy in a cohort comprising several varieties of mouse models, and subsequently observed resulting clusters' corresponding gene and behavior patterns<sup>300</sup>. The mentioned studies, along with several others, promote a shift



towards subtyping ASD and autistic traits populations in order to better understand and treat the disorder.

Compared to our high vs. low/absent autistic traits contrast that yielded no results, by isolating behaviorally refined autistic traits subgroups we observed cortical signatures despite having lower statistical power than the combined sample. Decreased gyrification detected in the right PC and pITG region in the *Emot* subgroup is consistent with studies in ASD<sup>218,360</sup>. In general, the precuneus is highly implicated in the default mode network (DMN) as well as in visuospatial processing, empathy and memory, while the temporal lobe correlates to memory, audition, theory of mind and visual processes<sup>361–364</sup>. Considering that this group bears high in aggression, studies have also reported a general decrease in gyrification in aggressive patients<sup>365,366</sup>, as well as decreased functional connectivity between the precuneus and other brain regions in patients exhibiting higher aggression traits, possibly due to its role in the DMN and empathy<sup>257,367</sup>. Additionally, this subgroup exhibited high ADHD-like symptoms which have also shown links to precuneus regions of the brain<sup>368,369</sup>. In our second subgroup, *Attn*, we observed increases in surface area in the precentral (primary motor) cortex, which is involved in voluntary motor control<sup>370,371</sup>, and the superior frontal gyrus, which is part of the motor control network and also harbors functions in attention, working memory, executive functioning and in the default mode network<sup>369,372,373</sup>. A study has even suggested that early motor impairments are predictors of future social communication delays, further indicating the importance of understanding this region in relation to ASD risk<sup>374</sup>. Specifically, within ASD, atypical motor functioning has been measured in patients from infancy until well into adulthood<sup>375,376</sup>. Seeing as how we observed extensive structural alterations throughout the motor control network in the *Attn* subgroup, this warrants further investigation into the relationship between ADHD-like traits and motor control in ASD and autistic traits populations. The *Attn* subgroup also presented decreases in

gyrification in the lateraloccipital region, which is heavily implicated in visual perception, and specifically in face recognition, which greatly influences social communication<sup>377–379</sup>. The last subgroup, *AnxDep*, exhibited increases in gyrification in the postcentral region (primary somatosensory cortex), which functions as the main sensory receptive area of the brain<sup>380,381</sup>. In ASD, atypical sensory reception, more specifically over-responsivity to tactile sensory inputs, is a very prevalent symptom<sup>382–384</sup>. This suggests that autistic individuals could easily be overwhelmed, perhaps forging a link to the development of anxious and depressive behaviors, as observed in this subgroup. The *AnxDep* subgroup additionally showed a decrease in thickness in the pMTG/STS region, which has been greatly implicated in language and social aspects, and thus an extremely important cortical region of interest in ASD behavioral studies<sup>385–387</sup>. Therefore, taking into account dimensional constructs of behavior in ASD can better prepare subgroups for the identification of biological mechanisms. Further investigation is warranted into the relationship between affected regions and corresponding subgroup behaviors in the context of ASD since these regions have been consistently reported within the ASD literature.

The behaviors observed in our subgroups enforce the fact that ASD is highly concurrent with several psychiatric conditions in up to 80-95% of patients<sup>62,89</sup>. Reported comorbid disorders include ADHD, depression disorders, anxiety disorders, obsessive compulsive disorder (OCD), and conduct disorders<sup>62,85,89,102,345</sup>. This high degree of comorbidity (based on diagnostic information) corresponds to our dimensional results, which describe these associations in an even more descriptive and spectral manner using behavioral constructs. Having access to a multidimensional cohort containing assessments of behaviors reported in ASD allowed us to explore how core autistic traits inherently distribute with other symptoms in a dimensionally continuous population. By running a data-driven clustering analysis on a population-based

cohort, we isolated 3 main autistic traits subgroups. The *AnxDep* subgroup is composed of subjects high in anxiety, depression and attention deficits. This is in line with findings reported in the literature where anxiety and depression appear to be some of the most common psychiatric comorbidities in ASD patients<sup>62,345</sup>. The *Attn* subgroup could represent an isolated population consisting purely of ADHD and autistic traits in an otherwise behaviorally muted subclass. This may perhaps become the optimal subgroup for studying the overlap between ADHD and ASD. Lastly, the *Emot* subgroup has ADHD-like traits in combination with emotional regulation abnormalities as evidenced by high degrees of aggression and reactivity. The *Emot* subgroup suggests a third combination of behavioral traits showing that the aggressive behaviors often observed in autistic traits participants can in effect co-occur with ADHD-like traits. The diverse behavioral profile of each subgroup highlights the importance of combining independent behaviors into one multivariate analysis to observe how they distribute. For example, as mentioned previously, both the *Attn* and *Emot* subgroups show high levels of attention problems and hyperactivity, and remain relatively close in anxiety and depression levels. The *Emot* subgroup, however, exhibits exceptionally high levels of aggression and reactivity, a factor that separates one ADHD-like autistic traits subgroup into two (i.e. *Attn* and *Emot*), thus increasing behavioral homogeneity and the likelihood of extracting biological features from cortical images.

Notably, the unsupervised clustering analysis yielded high autistic traits subgroups with gender ratios (averaging 2:1, male to female) differing from those usually reported ASD populations (averaging 4:1, male to female)<sup>23</sup>. However, this gender disequilibrium is not entirely surprising as this difference can be explained by the fact that studies sampling from the general population often show a lower ratio (3:1)<sup>388</sup>, and that overall variability may play a role. Moreover, several studies have reported ratios ranging from 2:1 to 7:1<sup>25–28</sup>, indicating a heterogeneity that warrants

further exploration as well as a diversity in gender ratio that depends on how cohorts are built. Most importantly, within the current investigation gender differences were controlled for in the neuroimaging analysis.

This study has several potential limitations. Firstly, though the SRS included a repetitive behavior subscale, it would have been interesting to include an independent repetitive behavior component within the clustering analysis. Concerning the unsupervised clustering, inherent limitations include the somewhat arbitrary determination of the number of clusters, and difficulties to reproduce the same partitioning in another dataset. Also, it is challenging to account for covariates in unsupervised clustering analyses. Alternative approaches could also have been applied on this dataset that would prove interesting in future studies including clustering based on SRS subscales (with subsequent study of behavioral and morphological traits), or clustering on a broader range of scales (not only pertaining to behaviors central in ASD) with subsequent isolation of subgroups high in SRS. Additionally, the present study used a general population-based cohort, and not one tailored for ASD studies, thus warranting the careful isolation of behavioral variables relevant to our objective. Within the morphometric results, it is possible we did not observe further thickness differences due to the wide age range of our cohort. Thickness changes more with age and environment and may thus present larger heterogeneities than does gyrification (which is typically developed in-utero and shortly after birth), leading us to observe greater gyrification alterations within our results<sup>389,390</sup>. Also, the average age of subjects in the current study (~11.4 years old) could indicate that our cortical results are consequences of differential child development, a hypothesis however that can only be confirmed using a longitudinal, prospective design. Clinical diversity in autistic traits may be further explained by other modalities, thus next steps would involve considering genetic, volumetric, diffusion, and functional differences between the acquired subgroups.

In conclusion, we showed that subtypes of autistic traits yield refined signatures and therefore stress the importance of stratification using a dimensional approach. Several studies, including the current one, have demonstrated the difficulty in yielding significant biological features in case-control comparisons, leading to large-scale inconsistencies within ASD literature. Since several of the behavioral associations and affected cortical regions discussed in this study have similarly been implicated in ASD studies, our findings maintain the growing assumption that outcomes in autistic traits are related to variations observed in ASD patients. By uncovering better-defined subtypes of ASD, studies can finally begin to truly understand the underlying genetic, biological and behavioral mechanisms of this syndrome.

## 4.2. Study 2

### In search of consistent behavioral and neuroimaging biomarkers in three samples with different autistic criteria

The manuscript for this study is currently in preparation to be published.

***In search of consistent behavioral and neuroimaging biomarkers in three samples with different autistic criteria***

*Angeline Mihailov, Cathy Philippe, Antoine Grigis, Josselin Houenou and Vincent Frouin. In Preparation.*

#### Study Summary

To date, there is no known or officially approved treatment for ASD other than few translational medicines targeting the alleviation of certain symptoms. This may in part be largely due to the lack of proper biological, genetic and behavioral characterization of this heterogeneous disorder. The experimental paradigm used to investigate ASD, which was based on case-control studies until recent time, is recognized as inadequate. Therefore, in order to better understand ASD, experimental designs tend to stratify individuals in the case group, with the addition of dimensional approaches, in order to divide individuals into homogenized subgroups. Since case-control comparisons have thus far generated poorly reproducible results (with the exception of a few examples, of course), isolating autistic subtypes may help elucidate underlying mechanisms in these subgroups, with hopes of eventually accelerating the discovery of therapeutic interventions<sup>391,392</sup>.

The objective of the current study is two-fold. First, we will cluster autistic and autistic-like samples according to semi-dimensional or dimensional criteria, respectively, to observe if any behavioral subgroups are consistently generated across these different samples. Second, we will investigate subgroups' neuroimaging and polygenic scores (PGS) to see if we observe similar traits across samples, and more importantly, to understand what these traits represent. The experimental design from *Mihailov et al., 2020* will be used within each sample to conduct the current study.

Using the European EU-AIMS LEAP cohort, participants were selected based on availability in behavioral assessments and diagnostic status, as well as having an FSIQ>70. Specifically, two samples entitled '*Dimensional*' and '*Dimensional + Repetitive*' will follow a dimensional approach and were created based on availability in assessments measuring social deficits, anxiety, depression, behavioral misconduct, attention problems and hyperactivity, and additionally, the availability of a repetitive behaviors assessment for the *Dimensional + Repetitive* sample only. These dimensional analyses will be blind to an ASD diagnosis and thus represent a general population sample differing in psychiatric traits. A third semi-dimensional sample, entitled the '*ASD*' sample, will be created based on availabilities in all previously mentioned assessments (except for social deficits) as well as a clinical diagnosis of ASD. A total of 256 subjects will participate in the first two analyses adopting the dimensional approach (*Dimensional* and *Dimensional + Repetitive* samples, differing only in the presence of repetitive behaviors), and 255 subjects (159 autistic patients + 96 controls) will be included in the third semi-dimensional analysis (the *ASD* sample). Since autistic and autistic-like traits are often concomitant with behavioral traits associated to a comorbid psychiatric disorder (which occur in up to 95% of autistic patients), we decided to study these behaviors dimensionally in order

to elucidate refined information regarding behavioral symptomatology in ASD<sup>62,89</sup>. To this end, as inspired by *Mihailov et al., 2020*, each of these three samples will be independently submitted to an unsupervised  $k$ -means clustering analysis based on the previously described behaviors. Subsequently, generated clusters will be scrutinized on neuroimaging and PGS data to determine if any unique traits are associated with certain subgroups, and furthermore, to observe if any consistencies remain across all three samples.

For the *Dimensional* and *Dimensional + Repetitive* analyses, both samples were clustered into 5 subgroups (4 autistic traits + one low autistic traits, serving as our controls), while the semi-dimensional analyses (*ASD* sample) clustered into 4 subgroups (with controls externally added for subsequent analysis since only ASD subjects were submitted to the clustering). The most consistent behavioral subgroups across all three samples included an autistic traits subgroup that was asymptomatic in all behaviors (except for social deficits, of course), *ASYMPTO*; a symptomatic subgroup, *SYMPTO*, which was high in all behaviors (except for depression); and lastly an autistic traits subgroup high in ADHD traits, *ADHD-like*, which exhibited high attention deficits and hyperactivity. Furthermore, good quality T1 MRI data was available for 85% of subjects allowing us to investigate cortical thickness, surface area and gyrification, while controlling for age, gender, site and FSIQ, and correcting for multiple comparisons. Cortical analyses revealed the same result for all *ADHD-like* subgroups across samples, namely a decrease in surface area in the right precentral region ( $p < 0.005$ ). This result in particular is similar to one generated in *Mihailov et al., 2020* showing an alteration in the precentral area in the *Attn* subgroup (an *ADHD-like* subgroup). Indeed these neuroimaging results were not the only differences observed, however they were the most consistent across samples. Moreover, upon combining all subgroups into one in order to run a case-control comparison per sample, results were extremely limited. Lastly, we also observed associations between the ADHD PGS



and the *ADHD-like* subgroup. Upon running case-control comparisons to look at differences in PGS, we observed a significant association between the autistic/autistic-like groups and the Major Depressive Disorder and Empathy PGS.

Current results prove that by subtyping ASD and ASD-like subjects, we are more likely to uncover refined patterns in behavioral, neuroimaging and genetic data. Results also confirm the interest of applying dimensional subtyping approaches in order to extract meaningful differences. We hope these approaches can eventually reduce heterogeneity within the disorder and pave the way to better understanding ASD etiologies and developmental pathways that will eventually lead to the development of improved therapies and interventions.

## Materials and Methods

### *Part I: Clinical Profiles*

#### ***EU-AIMS LEAP Cohort and Participants***

The EU-AIMS Longitudinal European Autism Project (LEAP) Cohort is one of the tools built by the EU-AIMS project devoted to a better understanding, diagnosis and ultimately care of ASD persons. In line with the recent consensus regarding the complex and intermingled etiology of the disease, not only is there a demand for the search of diagnostic biomarkers, but also prognostic and stratification biomarkers. The latter are actively searched for since they are of primary interest in the general research on autism. The unique data of this multicentre European-wide initiative includes longitudinal multidisciplinary observations in ASD patients in order to tackle developmental research questions, disentangle heterogeneity and establish links between biological and clinical profiles. This neuroimaging cohort is comprised of brain

imaging, cognitive/behavioral and genetic data for 764 participants including patients with an autistic condition and typically developed controls, between the ages of 6 and 30 years old having an FSIQ between 50 and 148. Six centres contributed to the curation of data: Institute of Psychiatry, King's College London, United Kingdom; Autism Research Centre, University of Cambridge, United Kingdom; Radboud University Medical Centre, Nijmegen, the Netherlands; University Medical Centre Utrecht, the Netherlands; Central Institute of Mental Health, Mannheim, Germany; and University Campus Bio-Medico, Rome, Italy. Autism diagnoses were confirmed using the Diagnostic and Statistical Manual (DSM-IV or DSM-5), or the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10). This study was approved by local ethical committees in each participating center, and written consent was provided by all participants and their legal guardians (for those younger than 18 years old). FSIQ was assessed in all participants using the Wechsler Abbreviated Scales of Intelligence. Further details on experimental design, clinical assessments and other data can be found in *Loth et al., 2017*<sup>305</sup>.

In the current study, after filtering for participants with an FSIQ > 70, we were left with 681 participants including autistic patients and controls. Experimental analyses were run on three samples based on different criteria. Table 4.2.1, Figure 4.2.1 and Figure 4.2.2 all aim to visually describe the experimental design:

Sample 1: This group is entitled the '**Dimensional**' cohort and will undergo a dimensional analysis (blind to an ASD diagnoses) with participants from along the entire spectrum (including patients with ASD and controls) selected based on availability of behavioral scores assessing social deficits, anxiety, depression, behavioral misconduct, attention problems and hyperactivity. This resulted in 256

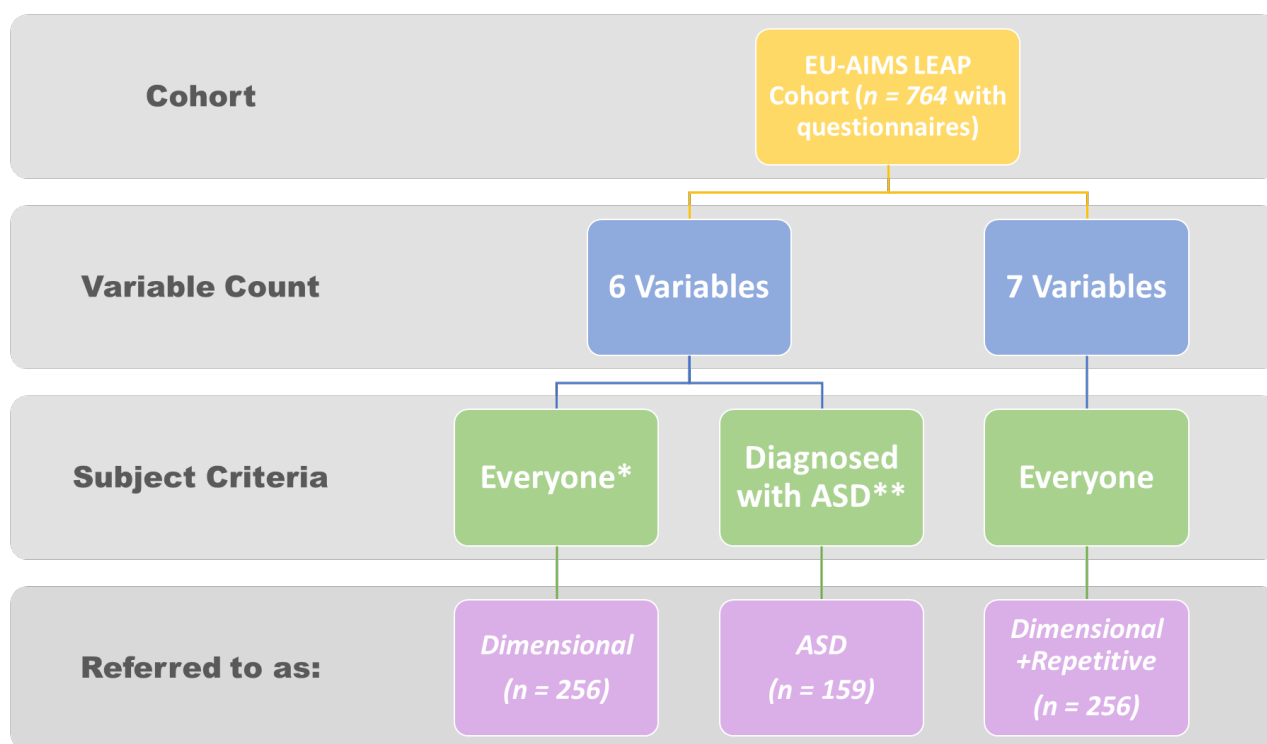
participants. The purpose of this first analysis is to study the replicability, to the best of our ability, of the behavioral-based clustering in *Mihailov et al., 2020*<sup>302</sup>.

Sample 2: This sample is entitled the ‘***Dimensional + Repetitive***’ cohort and will undergo a dimensional analysis (blind to an ASD diagnoses) with participants selected based on availability of behavioral scores assessing social deficits, anxiety, depression, behavioral misconduct, attention problems, hyperactivity *and* repetitive behaviors, resulting in 256 participants. The purpose of this second analysis is to add the important repetitive and restricted behaviors scale since this was not sufficiently available in the HBN cohort used in *Mihailov et al., 2020*<sup>302</sup>.

Sample 3: This cohort is entitled the ‘***ASD***’ sample and only includes participants diagnosed with ASD. Participants were chosen based on the presence of a behavioral assessment in scores assessing anxiety, depression, behavioral misconduct, attention problems, hyperactivity and repetitive behaviors. This did not included the SRS, which measures social deficits, since these subjects are already characterized as autistic. These criteria resulted in 159 participants. Contrast analyses will further include true controls ( $n = 96$ ), which will bring this sample to 255 subjects.

	Dimensional	Dimensional + Repetitive	ASD
<b>Purpose</b>	To apply a dimensional approach to extract behavioral subgroups from a sample (blind to ASD diagnoses) as seen in <i>Mihailov et al., 2020</i> by using the same behavioral variables	To apply a dimensional approach to extract behavioral subgroups from a sample (blind to ASD diagnoses) by using the same behavioral variables as in <i>Mihailov et al., 2020</i> , with the addition of repetitive behaviors	To apply a dimensional approach on an autistic cohort using same behavioral variables as in <i>Mihailov et al., 2020</i> , with the addition of repetitive behaviors
<b>Included Behaviors</b>	Social Deficits Anxiety Depression Attention Problems Hyperactivity Behavioral Misconduct	Social Deficits Anxiety Depression Attention Problems Hyperactivity Behavioral Misconduct Repetitive Behaviors	Anxiety Depression Attention Problems Hyperactivity Behavioral Misconduct Repetitive Behaviors

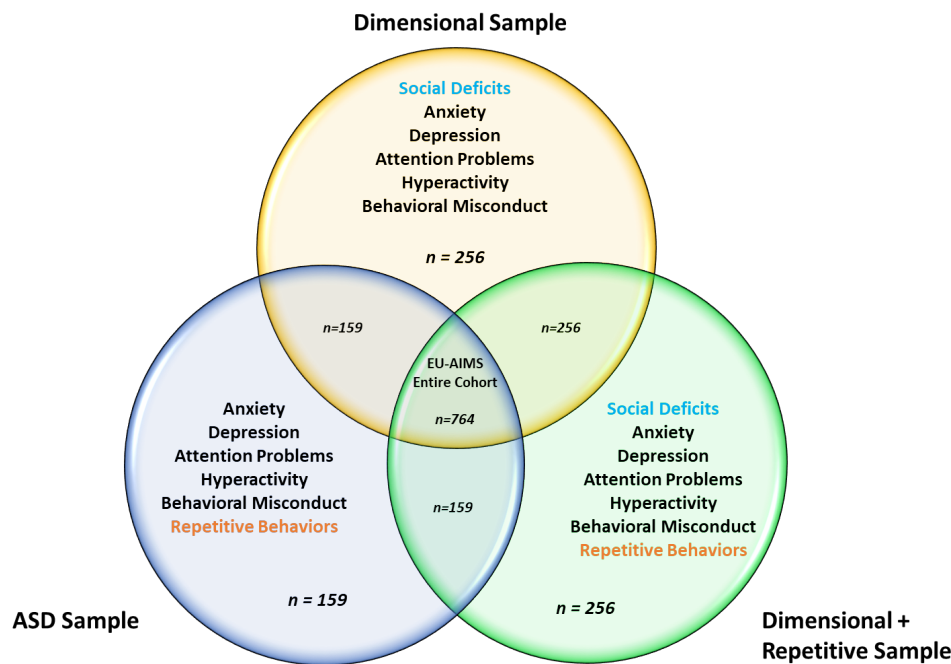
**Table 4.2.1.** Visual summary of the reasoning behind each sample criteria, as well as the included behaviors.



\*Repetitive behaviors measure not included.

\*\*Social Deficits (SRS) measure not included.

**Figure 4.2.1.** Flow chart summarizing the paths taken to build each of the three samples.



**Figure 4.2.2.** Venn diagram summarizing the behavioral scores investigated in each sample, as well as overlapping subjects. Of note, social deficits were not studied in the *ASD* sample, and repetitive behaviors were not studied in the *Dimensional* Sample.

We chose to conduct the same experiment in these three samples that differ in dimensional (*Dimensional* and *Dimensional + Repetitive*) and categorical (*ASD*) criteria in order to observe if consistencies remain across behavioral, neuroimaging and genetic factors.

### ***Behavioral Assessments***

One of the most prominent dimensions in ASD patients are social impairments. Here, we used data from the widely used 65-item parent Social Responsiveness Scale (SRS) as a quantitative measure of clinical autistic traits, making it a central variable in our study and the score from which we isolate our autistic traits individuals in our dimensional analyses (samples 1 and 2)<sup>278</sup>. The SRS is a widely used measure of autistic traits and has been used for the purpose of understanding autism in several behavioral, genetic, and neuroimaging studies<sup>278–280,282–287</sup>. Though not a diagnostic tool, the SRS exhibits high inter-rater and cross-cultural reliability,

and correlates greatly with the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview—Revised (ADI-R) diagnostic assessments for ASD from the DSM-5, making it a robust measure to use in the dimensional study of ASD behaviors (Figure 1.4.2)<sup>278–281</sup>. For the remaining behaviors, since they are incorporated dimensionally, we did not use diagnostic assessments and instead used scales measuring behavioral trait severity. We measured attention problems and hyperactivity using the subscales ‘Inattention’ and ‘Hyperactivity/Impulsivity’ from the ADHD Rating Scale<sup>393</sup>; repetitive behaviors were determined from the Repetitive Behaviors Scale (RBS)<sup>394</sup>; and finally anxiety, depression, and behavioral misconduct were established using the subscales ‘Anxiety’, ‘Depression’ and ‘Behavioral Disorder’ from the Development and Well-Being Assessment (DAWBA)<sup>395</sup>. These variables were chosen as the best possible behaviors matching those used in the study by *Mihailov et al., 2020* (Table 4.2.2).

	HBN	EU-AIMS
Social Deficits	SRS	SRS
Anxiety	SCARED	DAWBA
Depression	CBCL	DAWBA
Attention Problems	CBCL	ADHD Rating Scale
Hyperactivity	SDQ	ADHD Rating Scale
Aggression	CBCL	DAWBA
Reactivity	ARI	
Repetitive Behaviors	N/A	RBS

**Table 4.2.2.** Clinical assessments in the current study were chosen based on the best possible similarity to those used in *Mihailov et al., 2020*.

### ***Unsupervised Clustering Analysis (K-means)***

A *k*-means analysis was conducted on scaled z-scores of the previously mentioned behaviors, per sample. This yielded subgroups differing in behaviors but high in autistic traits (as measured

by the SRS) in the dimensional analyses (samples 1 and 2), and autistic subgroups with varying behavioral profiles in the semi-dimensional analysis (sample 3). Briefly,  $k$ -means is an algorithm identifying mean cluster centroids, which serves to partition a sample into  $k$  subgroups<sup>311,312</sup>. A substantial challenge in such analyses lies in determining the number of clusters, which is a user-defined parameter. To address this problem, the chosen number of clusters  $k$  was determined using a Bayesian Information Criterion (BIC) distribution<sup>351</sup>.

Mean behavioral scores, FSIQ and age were compared between subgroups within each of the three samples using non-parametric 2-sided Mann-Whitney  $U$  tests, while gender differences were determined using a chi-square test. Python 2.7 and R 3.4.0 were used to perform all analyses in this study. Python packages used include Pandas (version 0.19.2), SciPy (version 1.1.0), and Matplotlib (version 1.5.1), which were used for data handling, statistical analysis and visualization, respectively.

## Part II: Neuroimaging Analysis of Cortical Surface Features

### ***Structural MRI Acquisition and Processing***

MRI acquisition (all 3T scanners) took place at six different sites previously mentioned above: GE Medical Systems Discover MR 750 at King's College London, United Kingdom; Siemens Verio at the Autism Research Centre, University of Cambridge, United Kingdom; Siemens Skyra at Radboud University Medical Centre, Nijmegen, the Netherlands; Philips Medical Systems Achieva/Ingenia CX at University Medical Centre Utrecht, the Netherlands; Siemens TimTrio at Central Institute of Mental Health, Mannheim, Germany; and GE Medical Systems

Sigma HDxTt at University Campus Bio-Medico, Rome, Italy. For further scanning parameters please refer to Table 4.2.3.

Site	Manufacturer	Model	Software Version	Acquisition sequence	Coverage	Slices	Thickness [mm]	Resolution [mm <sup>3</sup> ]	TR [s]	TE [ms]	FA [°]	FOV
Cambridge	Siemens	Verio	Syngo MR B17	Tf3d1_ns	256*256	176	1.2	1.1*1.1*1.2	2.3	2.95	9	270
London	GE	Discovery	LX MR	SAG ADNI	256*256	196	1.2	1.1*1.1*1.2	7.31	3.02	11	270
	Medical systems	mr750	DV23.1_V02_1317.c	GO ACC SPGR								
Mannheim	Siemens	TimTrio	Syngo MR B17	MPRAGE ADNI	256*256	176	1.2	1.1*1.1*1.2	2.3	2.93	9	270
Nijmegen	Siemens	Skyra	Syngo MRD13	Tf3d1_16ns	256*256	176	1.2	1.1*1.1*1.2	2.3	2.93	9	270
Rome	GE	Signa HDxt	24/LXMR	SAG ADNI	256*256	172	1.2	1.1*1.1*1.2	5.96	1.76	11	270
	Medical systems		HD16.0_V02_1131.a	GO ACC SPGR								
Utrecht	Philips	Achieva			256*256	170	1.2	1.1*1.1*1.2	6.76	3.1	9	270
	Medical Systems	/Ingenia CX	3.2.3, 3.2.3.1	ADNI GO 2								

TR: repetition time; TE: echo time; FA: flip angle; FOV: field of view.

**Table 4.2.3.** Detailed acquisition parameters from each of the six sites participating in the current study.

T1-weighted images were processed using the FreeSurfer software version 6.0.0 (<https://surfer.nmr.mgh.harvard.edu/>). For more information on precise methods of image analysis and the construction of anatomical information for each individual done by this software, refer to *Dale et al., 1999* and *Fischl et al., 2002*<sup>327,352</sup>. Briefly, the FreeSurfer analysis stream includes intensity normalization, skull stripping, and segmentation of gray (pial) and white matter surfaces<sup>327</sup>. Subsequent tessellation, as well as various topology corrections and inflation, leads to 3D meshes of cortical surfaces in different resolutions. Our work is based on a tessellation with ~160 000 vertices per hemisphere and used the FreeSurfer *fsaverage* template. We focused on three morphological measures of which the processing stream created vertex-wise maps for analysis: cortical thickness, surface area, and gyrification using the local gyrification index (LGI) which is measured as the ratio between buried and visible cortex<sup>331</sup>. All images were rated by three independent raters, manually edited, and finally re-assessed by the same raters. Manual edits were performed by applying changes to the grey or white matter, or



by cloning (performed if a large number of voxels were missing due to erroneous removal during skull-stripping or other editing procedures).

### ***Statistical Analysis***

Vertex-wise statistical analyses were conducted using the command-line group analysis stream in FreeSurfer. Cortical surfaces for each participant were first registered to fsaverage, and then smoothed using a full-width-at-half maximum (FWHM) kernel of 10 mm for cortical thickness and surface area, and 5 for IGI. A general linear model was fit at each vertex  $i$  to compare the three morphological measures between contrasts, using gender and site as categorical covariates, and age and FSIQ as continuous covariates and including the residual error:

$$y_i = \beta_0 + \beta_1 \text{Group} + \beta_2 \text{Sex} + \beta_3 \text{Age} + \beta_4 \text{FSIQ} + \varepsilon_i$$

We performed a cluster-level analysis using a cluster-forming threshold of  $p < 0.005$ . Clusters with cluster-wise p-value ( $cwp$ ) of  $cwp < 0.05$  are reported. These p-values were corrected for multiple comparisons using the *mri\_glmfit-sim* precomputed MonteCarlo simulation.

### **Part III: Polygenic Scores Analysis**

#### ***Polygenic Scores***

In order to determine underlying genetic profiles, Polygenic (Risk) Scores (PGS) were considered for each individual in all resultant subgroups in each of the three samples. These PGS were pre-computed by the EU-AIMS Consortium using the tool PRSice-2<sup>396</sup>. Genome-Wide Association Study summary statistics were used as a reference for each PGS. For the linkage disequilibrium based SNP pruning, only SNPs with a Minor Allele Frequency (MAF)  $> 1\%$  and an  $R^2 < 0.1$  in windows of 500kb were used. PGS were adjusted for ancestry using

principle components 1 through 4 available from the EU-AIMS consortium. The PGS corresponding to four traits of interest were chosen for this analysis based on their association with the behaviors studied, namely the PGS for Attention-Deficit and Hyperactivity Disorder (ADHD), Autism, Empathy, and Major Depressive Disorder (MDD).

### ***Statistical Analysis***

Each subgroup within each sample was compared against their corresponding controls using a logistic regression controlling for age and sex. An overall categorical contrast (i.e. combining all subgroups into one high SRS or ASD group to compare against controls) was also run for each sample using the same model. All resulting p-values were Bonferroni corrected for the four tests.

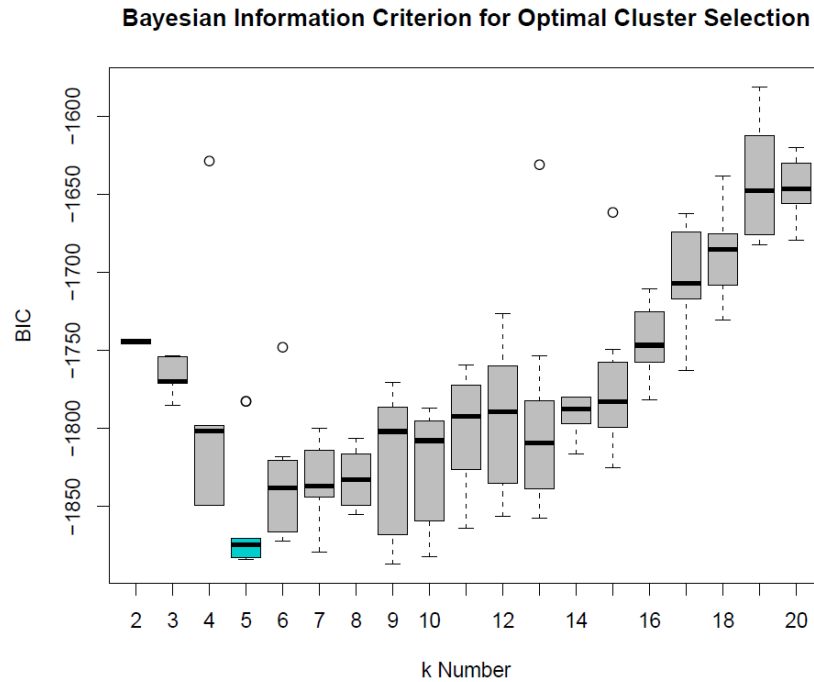
## **Results**

### ***Part I: Clinical Profiles***

#### ***Data-Driven Behavioral Subgroups in the EU-AIMS Cohort per Sample***

##### ***Sample 1: Dimensional***

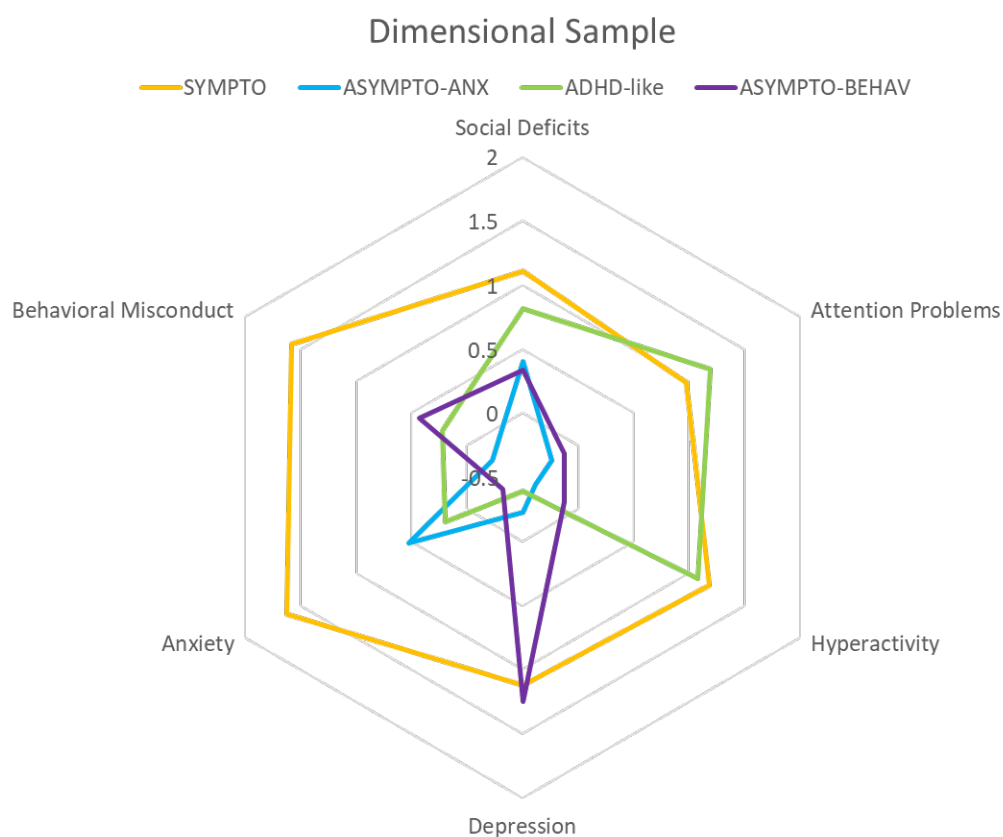
Based on the computed BIC value distribution, a  $k$  value of five was retained as our supervised partitioning for the *Dimensional* sample (Figure 4.2.3).



**Figure 4.2.3.** The BIC value distribution suggested a  $k$  value of five to be used in the clustering analysis for the *Dimensional* sample.

Thus, upon running the clustering analysis, we obtained five subgroups with various behavioral profiles. The average SRS levels were used to decide which subgroups represented high autistic traits participants and which were controls. From these 5 subgroups, 4 expressed high levels of SRS, thus acting as our socially impaired ‘high autistic traits’ subgroups, and one expressed low levels of SRS acting as our control group. Mean SRS levels falling within the ‘*severe*’ or at least upper ‘*moderate*’ classification of the SRS scale (an SRS value above ~80), indicates a high level of social impairment, providing us with greater confidence that subjects within the autistic traits subgroups are indeed ‘autistic-like’. Additionally, mean SRS values in the four autistic traits subgroups are comparable to the average SRS level of ~86 typically reported in diagnosed ASD patients<sup>279,355,356</sup>. Regarding the behavioral compositions of our four high autistic traits subgroups, one subgroup was asymptomatic (except for social deficits) but displayed elevated levels of anxiety,  $n = 38$  (described as *ASYMPTO-ANX*); the second was also asymptomatic (except for social deficits) but displayed elevated levels of behavioral

misconduct,  $n = 27$  (referred to as *ASYMPTO-BEHAV*); the third showed high levels of ADHD-like behaviors including attention deficits and hyperactivity,  $n = 51$  (described as *ADHD-like*); and the fourth subgroup was symptomatic in all behaviors, except for depression,  $n = 35$  (referred to as *SYMPTO*). Cut-off levels were determined for each behavioral measure according to the literature<sup>397–400</sup>. The last subgroup, which was composed of low levels of SRS,  $n = 105$ , will represent a general population without autistic traits and will serve as our control group contrast in subsequent analyses (Figure 4.2.4 and Table 4.2.4).



**Figure 4.2.4.** A radar plot summarizing the behavioral profiles of each autistic traits subgroup generated in the *Dimensional* sample clustering analysis. Numbers along the plot indicate z-scores. To note, it may seem like the depression scale is clinically significant, however this is due to a highly skewed distribution and a small range. Therefore, due to the nature of how z-scores are calculated, the resulting value was sensitive to smaller changes. However, as seen in Table 4.2.4, indeed the average depression levels are low.

Subgroup	n (Total = 256)	SRS: <i>Social Deficits</i>	ADHD-RS: <i>Inattention</i>	ADHD-RS: <i>Hyperactivity</i>	DAWBA: <i>Depression</i>	DAWBA: <i>Anxiety</i>	DAWBA: <i>Behavioral Misconduct</i>	Mean Age	Gender Ratio	Mean FSIQ
<b>ASYMPTO-ANX</b> (Subgroup 1)	38	<b>84.2*</b>	2.6	1.2	1.4	<u>2.7</u>	1.9	13.6	26:12	100.9
<b>SYMPTO</b> (Subgroup 2)	35	<b>114.5*</b>	<b>6.6*</b>	<b>5.6*</b>	1.7	<b>4.1*</b>	<b>4.0*</b>	12.1	22:13	103.2
<b>ADHD-like</b> (Subgroup 3)	51	<b>102.0*</b>	<b>7.3*</b>	<b>5.3*</b>	0.2	2.2	2.4	11.9	38:13	97.4
<b>ASYMPTO-BEHAV</b> (Subgroup 4)	27	<b>81.4*</b>	3.0	1.9	1.8	1.5	<u>2.7</u>	12.9	22:5	105.9
<b>Controls</b> (Subgroup 5)	105	24.0	0.8	0.2	0.2	0.9	1.4	13.3	67:38	106.6

**Table 4.2.4.** Behavioral score values in each subgroup generated in the *Dimensional* sample. Stars (\*) indicate values surpassing high clinical levels of each score. Though statistically insignificant, underlined values indicate elevated levels of the score in order to distinguish between both *ASYMPTO* subgroups. Ranges for each score: **SRS**: 0 to 123+; **ADHD-RS-Inattention**: 0 to 9; **ADHD-RS-Hyperactivity**: 0 to 9; **DAWBA-Depression**: 0 to 5; **DAWBA-Anxiety**: 0 to 5; **DAWBA-Behavioral Misconduct**: 0 to 5.

Upon comparing the four autistic traits subgroups to one another, several significant differences in behavioral scores were found (Table 4.2.5A, ‘*Comparisons Between Subgroups*’). We also compared each autistic traits subgroup to controls (the fifth generated low-SRS subgroup), and observed significant differences in all behavioral scores, except for depression in the *ADHD-like* and *ASYMPTO-ANX* subgroups. Lastly, we decided to combine all three high autistic traits subgroups into one ‘autistic-like’ group (combined high SRS, *Dimensional-hSRS*) to compare against controls, which yielded highly significant differences in every behavioral score (Table 4.2.5A, ‘*Comparisons to Controls*’).

With respect to demographic information, there were no significant differences in gender and FSIQ between all subgroups, with the exception of a difference in FSIQ between the *ADHD-like* and *ASYMPTO-BEHAV* subgroups. There were reported differences in age, though age ranges were similar (11.9 – 13.6 mean age range)(Table 4.2.5A, ‘*Comparisons Between Subgroups*’). Upon comparing each subgroup to controls, we again found no differences in

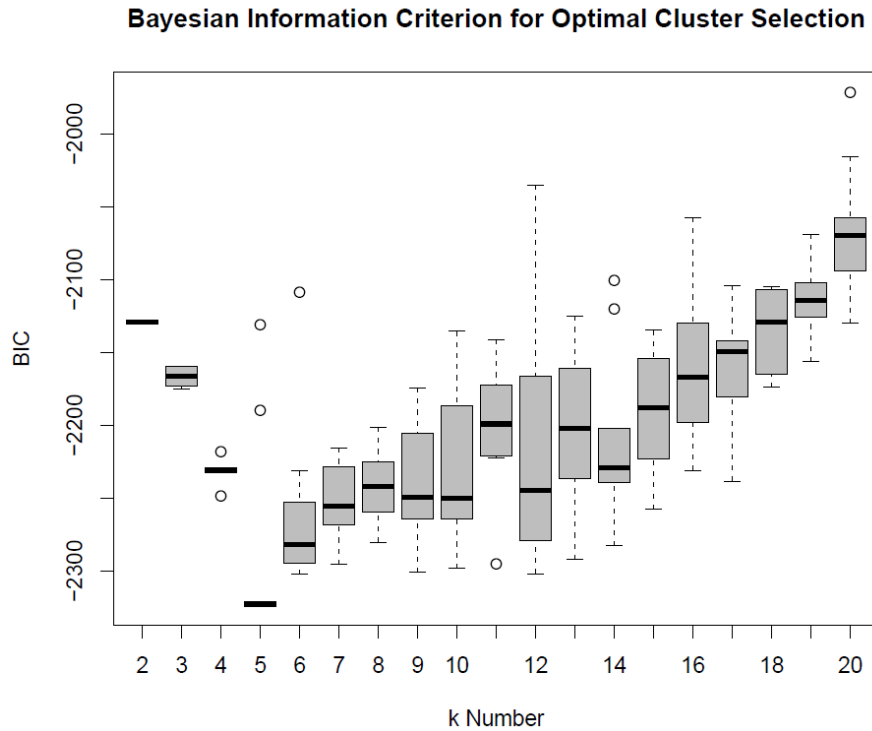
gender. We did find differences in age, however the ranges were again similar (11.9 – 13.3 mean age range). FSIQ only differed between the *ADHD-like* subgroup and controls. Lastly, the comparison between the *Dimensional-hSRS* group and controls yielded significant differences in FSIQ, but not age or gender (Table 4.2.5A, ‘*Comparisons to Controls*’). Due to presented differences, we deemed it important to control for FSIQ, age and gender in the subsequent neuroimaging analysis.

	SRS: Social Deficits	ADHD-RS: Inattention	ADHD-RS: Hyperactivity	DAWBA: Depression	DAWBA: Anxiety	DAWBA: Behavioral Misconduct	Age	Gender (X)	FSIQ
<b>A. Behavioral Cohort (n = 256)</b>									
<b>Comparisons Between Subgroups</b>									
ASYMPTO-ANX vs. SYMPTO	0.00000732*	2.772E-10*	2.142E-15*	4.386E-06*	8.22E-09*	7.14E-15*	0.0251*	0.7998	0.5250
ASYMPTO-ANX vs. ADHD-like	0.0054*	4.518E-19*	9.24E-14*	0.6918	0.339	0.0384*	0.0067*	0.6937	0.3422
ASYMPTO-ANX vs. ASYMPTO-BEHAV	0.999	0.999	0.3522	1.896E-09*	0.0000738*	0.0078*	0.4021	0.3711	0.2182
SYMPTO vs. ADHD-like	0.1062	0.492	0.999	2.568E-09*	1.578E-13*	1.44E-11*	0.7650	0.3592	0.1108
SYMPTO vs. ASYMPTO-BEHAV	0.00000678*	2.802E-07*	6.9E-09*	0.999	6.3E-19*	1.866E-06*	0.2573	0.1869	0.4685
ADHD-like vs. ASYMPTO-BEHAV	0.003*	2.586E-14*	1.164E-07*	4.992E-14*	0.0258*	0.999	0.1399	0.6798	0.0398*
<b>Comparisons to Controls</b>									
ASYMPTO-ANX vs. Controls	1.374E-32*	5.358E-06*	1.656E-08*	0.174	3.534E-25*	0.0006*	0.6040	0.7548	0.0503
SYMPTO vs. Controls	3.81E-50*	6.42E-33*	2.916E-49*	1.434E-16*	9.48E-59*	8.88E-42*	0.0490*	0.9192	0.2326
ADHD-like vs. Controls	4.806E-51*	4.68E-51*	3.786E-44*	0.999	1.41E-17*	2.826E-14*	0.0094*	0.2483	0.0009*
SYMPTO vs. ASYMPTO-BEHAV	1.212E-26*	1.716E-06*	6.6E-14*	4.932E-23*	0.0000684*	5.31E-14*	0.6299	0.1292	0.8276
hSRS vs. Controls	5.508E-67*	8.04E-32*	3.144E-28*	3.174E-08*	3.75E-28*	3.534E-21*	0.0617	0.2425	0.0068*
<b>B. Neuroimaging Cohort – Subset of Behavioral Cohort (n = 219)</b>									
<b>Comparisons Between Subgroups</b>									
ASYMPTO-ANX vs. SYMPTO	0.0000174*	1.062E-08*	3.522E-14*	0.0000597*	1.848E-06*	1.92E-11*	0.0179*	0.8621	0.2514
ASYMPTO-ANX vs. ADHD-like	0.0726	1.752E-13*	5.508E-13*	0.999	0.405	0.096	0.0296*	0.7288	0.8022
ASYMPTO-ANX vs. ASYMPTO-BEHAV	0.999	0.999	0.1932	5.424E-09*	0.000261*	0.0384*	0.6359	0.5855	0.2598
SYMPTO vs. ADHD-like	0.0534	0.999	0.999	0.00000636*	3.198E-10*	1.008E-08*	0.7557	0.6907	0.1736
SYMPTO vs. ASYMPTO-BEHAV	0.0001218*	0.00001602*	2.118E-07*	0.999	6.84E-14*	0.0000834*	0.0778	0.5591	0.8929
ADHD-like vs. ASYMPTO-BEHAV	0.0000174*	1.062E-08*	3.522E-14*	0.0000597*	1.848E-06*	1.92E-11*	0.1278	0.9846	0.1883
<b>Comparisons to Controls</b>									
ASYMPTO-ANX vs. Controls	1.326E-29*	0.00003858*	5.646E-06*	0.1224	1.62E-24*	0.000453*	0.2329	0.8995	0.0676
SYMPTO vs. Controls	2.43E-44*	1.062E-29*	1.302E-46*	4.002E-15*	1.248E-50*	1.434E-38*	0.1498	0.9636	0.7586
ADHD-like vs. Controls	3.108E-39*	1.14E-40*	8.22E-43*	0.999	6.9E-17*	7.56E-14*	0.2163	0.3860	0.0311*
ASYMPTO-BEHAV vs. Controls	1.608E-25*	4.986E-07*	3.738E-12*	6.9E-24*	0.0005832*	4.428E-13*	0.5940	0.3203	0.9102
hSRS vs. Controls	1.752E-57*	2.052E-27*	5.67E-24*	7.68E-09*	3.912E-27*	1.98E-19*	0.6909	0.3406	0.0666

**Table 4.2.5:** Demographic and behavioral score comparisons in the *Dimensional* sample, presented as p-values, between autistic traits subgroups and between each autistic traits subgroup and controls. The top section shows these comparisons in the behavioral cohort (‘A. Behavioral Cohort (n=256)’), with autistic traits subgroups compared to each other (under ‘Comparisons Between Subgroups’) and with each autistic traits subgroup compared to controls (under ‘Comparisons to Controls’). *hSRS* denotes a group that was created by combining all autistic traits subgroups in order to compare against controls. The bottom section (B. Neuroimaging Cohort (n= 219; Subset of Behavioral Cohort)) presents the same comparisons in the neuroimaging cohort, which is composed of a subset of subjects from the behavioral cohort with usable structural MRI data. P-values were adjusted using the Bonferroni method. Significant p-values are bolded with a ‘\*’ superscript.

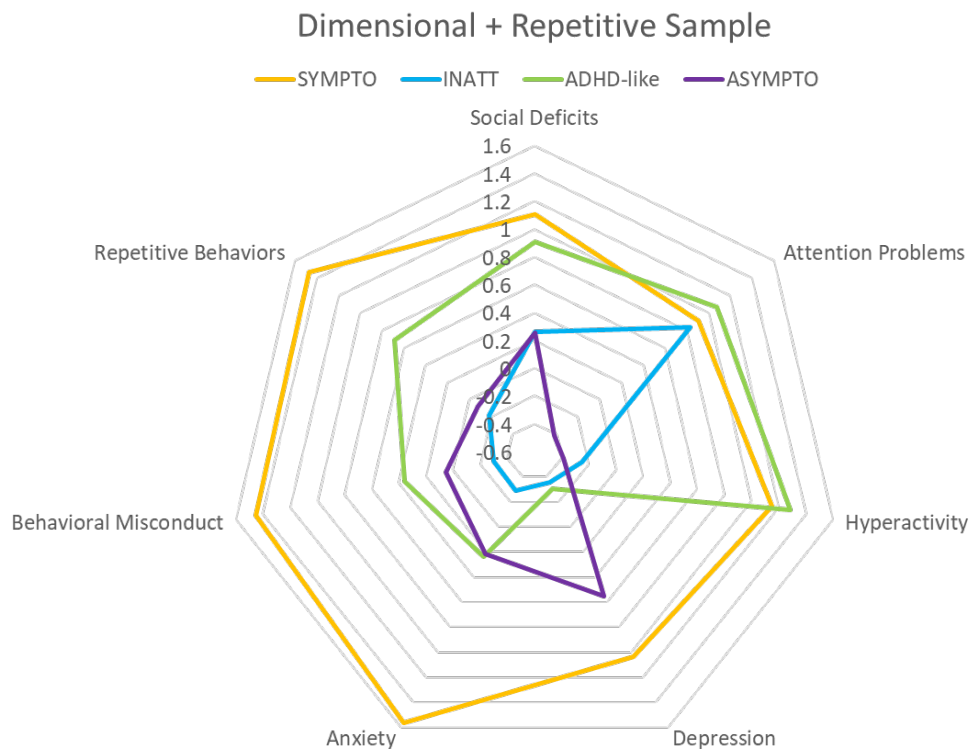
## Sample 2: Dimensional + Repetitive

Based on the computed BIC value distribution, a *k* value of five was retained as our supervised partitioning for the *Dimensional + Repetitive* sample (Figure 4.2.5).



**Figure 4.2.5.** The BIC value distribution suggested a  $k$  value of five to be used in the clustering analysis for the *Dimensional + Repetitive* sample.

Thus, upon running the clustering analysis, we obtained five subgroups with various behavioral profiles. Same as for the *Dimensional* sample, average SRS levels were used to decide which subgroups represented high autistic traits participants and which were controls. From these five subgroups, four expressed high levels of SRS, thus acting as our socially impaired ‘high autistic traits’ subgroups. Of these four high autistic traits subgroups, in addition to high social deficits, one subgroup presented low levels in all other behaviors,  $n = 46$  (referred to as *ASYMPTO*); the second showed high levels of attention problems,  $n = 32$  (described as *INATT*); the third subgroup exhibited significant ADHD-like behaviors (inattention and hyperactivity),  $n = 41$  (described as *ADHD-like*); and the last autistic-like subgroup showed high levels in all targeted behaviors, except for depression,  $n = 40$  (referred to as *SYMPTO*). Again, as for the *Dimensional* sample, there was a fifth low-SRS subgroup,  $n = 97$ , which will serve as a control contrast in subsequent neuroimaging analyses (Figure 4.2.6 and Table 4.2.6).



**Figure 4.2.6.** A radar plot summarizing the behavioral profiles of each autistic traits subgroup generated in the *Dimensional + Repetitive* sample clustering analysis. Numbers along the plot indicate z-scores. To note, it may seem like the depression scale is clinically significant, however this is due to a highly skewed distribution and a small range. Therefore, due to the nature of how z-scores are calculated, the resulting value was sensitive to smaller changes. However, as seen in Table 4.2.6, indeed the average depression levels are low.

Subgroup	n (Total = 256)	SRS: Social Deficits	ADHD-RS: Inattention	ADHD-RS: Hyperactivity	DAWBA: Depression	DAWBA: Anxiety	DAWBA: Behavioral Misconduct	RBS: Repetitive Behaviors	Mean Age	Gender Ratio	Mean FSIQ
<b>SYMPTO</b> (Subgroup 1)	40	<b>114.2*</b>	<b>6.4*</b>	<b>5.5*</b>	1.6	<b>4.1*</b>	<b>3.8*</b>	<b>28.8*</b>	12.3	26:14	104.3
<b>INATT</b> (Subgroup 2)	32	<b>78.0*</b>	<b>6.2*</b>	1.6	0.3	1.6	1.8	8.8	13.9	24:8	98.8
<b>ADHD-like</b> (Subgroup 3)	41	<b>105.7*</b>	<b>7.0*</b>	<b>5.9*</b>	0.3	2.3	2.6	19.3	11.6	31:10	98.1
<b>Controls</b> (Subgroup 4)	97	22.1	0.4	0.2	0.2	0.9	1.4	1.3	13.1	62:35	107.3
<b>ASYMPTO</b> (Subgroup 5)	46	<b>77.9*</b>	2.0	1.2	1.1	2.2	2.2	10.0	13.1	32:14	102.2

**Table 4.2.6.** Behavioral score values in each subgroup generated in the *Dimensional + Repetitive* sample. Stars (\*) indicate values surpassing high clinical levels of each score. Ranges for each score: **SRS**: 0 to 123+; **ADHD-RS-Inattention**: 0 to 9; **ADHD-RS-Hyperactivity**: 0 to 9; **DAWBA-Depression**: 0 to 5; **DAWBA-Anxiety**: 0 to 5; **DAWBA-Behavioral Misconduct**: 0 to 5; **RBS-Repetitive Behaviors**: 0 to 100.



Upon comparing the four socially impaired subgroups to one another, several significant differences in behavioral scores were found (Table 4.2.7A, '*Comparisons Between Subgroups*'). We also compared each autistic traits subgroup to controls (the fifth generated low-SRS subgroup) and observed significant differences in all behavioral scores, except for depression in the *ADHD-like* and *INATT* subgroups. Lastly, we decided to combine all three high autistic traits subgroups into one 'autistic-like' group (combined high SRS, *Dimensional + Repetitive-hSRS*) to compare against controls, which yielded highly significant differences in every behavioral score (Table 4.2.7A, '*Comparisons to Controls*').

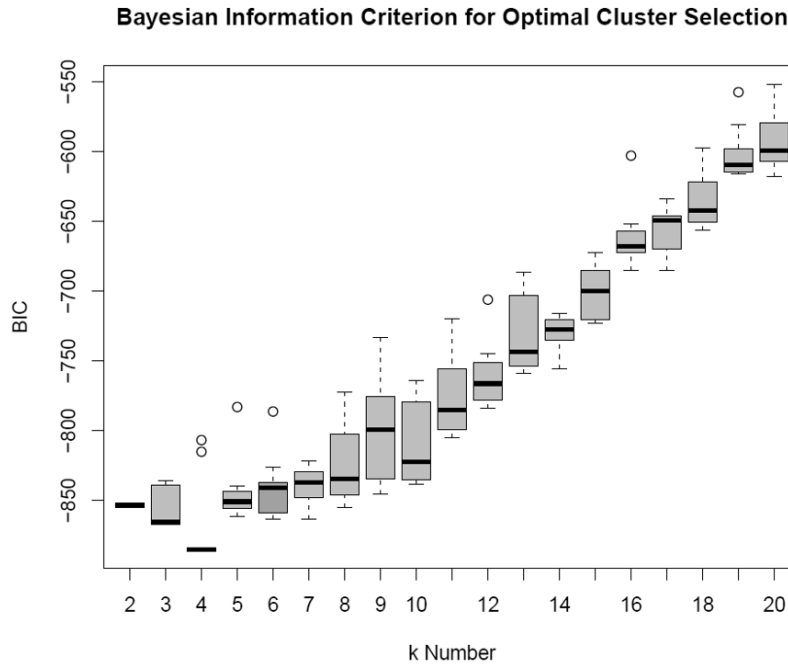
With respect to demographic information, there were no significant differences in gender and FSIQ between all subgroups. There were reported differences in age, though age ranges were similar (11.6 – 13.9 mean age range)(Table 4.2.7A, '*Comparisons Between Subgroups*'). Upon comparing each subgroup to controls, we again found no differences in gender. We did find differences in age between the *ADHD-like* subgroup and controls, however the ranges were again similar (11.6 – 13.1 mean age range). FSIQ differed between all subgroups and controls, with the exception of the *SYMPTO* subgroup. Lastly, the comparison between the *Dimensional + Repetitive-hSRS* group and controls yielded significant differences in FSIQ, but not in age or gender (Table 4.2.7A, '*Comparisons to Controls*'). Due to presented differences, we deemed it important to control for FSIQ, age and gender in the subsequent neuroimaging analysis.

	SRS: Social Deficits	ADHD-RS: Inattention	ADHD-RS: Hyperactivity	DAWBA: Depression	DAWBA: Anxiety	DAWBA: Behavioral Misconduct	RBS: Repetitive Behaviors	Age	Gender (X)	FSIQ
A. Behavioral Cohort (n = 256)										
Comparisons Between Subgroups										
SYMPTO vs. INATT	6.384E-07*	0.999	7.42E-12*	0.00001162*	3.857E-18*	2.149E-13*	1.435E-10*	0.01048*	0.51061	0.11648
SYMPTO vs. ADHD-like	0.7525	0.999	0.999	1.253E-06*	2.765E-11*	2.422E-07*	0.0006636*	0.28404	0.42246	0.08521
SYMPTO vs. ASYMPTO	6.041E-10*	1.687E-13*	1.316E-17*	0.5418	3.066E-12*	1.652E-10*	1.47E-11*	0.18966	0.82591	0.56951
INATT vs. ADHD-like	0.000105*	0.5215	5.145E-13*	0.999	0.0217*	0.0014*	4.655E-09*	0.00066*	0.83081	0.84899
INATT vs. ASYMPTO	0.999	5.684E-13*	0.999	0.0004151*	0.0245*	0.3234	0.999	0.22269	0.78802	0.38827
ADHD-like vs. ASYMPTO	8.05E-07*	4.438E-19*	4.697E-19*	0.0001183*	0.999	0.5005	2.387E-07*	0.02130*	0.69698	0.28696
Comparisons to Controls										
SYMPTO vs. Controls	4.13E-53*	4.669E-44*	2.765E-47*	3.374E-15*	2.275E-52*	1.141E-36*	1.12E-40*	0.1368	0.9396	0.2529
INATT vs. Controls	5.32E-26*	3.577E-46*	3.332E-10*	0.999	0.0000196*	0.0056*	1.561E-17*	0.2237	0.3488	0.0042*
ADHD-like vs. Controls	1.673E-49*	4.942E-58*	3.703E-49*	0.8267	2.17E-16*	6.076E-16*	5.789E-48*	0.0088*	0.2542	0.0014*
ASYMPTO vs. Controls	4.907E-35*	3.122E-09*	1.05E-07*	2.611E-11*	1.078E-16*	6.636E-09*	5.656E-18*	0.9498	0.6340	0.0710*
hSRS vs. Controls	1.246E-61*	4.235E-39*	6.671E-24*	9.8E-08*	3.41E-24*	4.501E-19*	2.786E-28*	0.2233	0.2914	0.0020*
B. Neuroimaging Cohort – Subset of Behavioral Cohort (n = 219)										
Comparisons Between Subgroups										
SYMPTO vs. INATT	3.864E-06*	0.999	1.274E-10*	0.00006398*	2.233E-14*	1.96E-11*	1.351E-10*	0.0093*	0.6226	0.0477*
SYMPTO vs. ADHD-like	0.2436	0.999	0.999	0.0004984*	5.341E-07*	0.00004396*	0.0001547*	0.7278	0.8940	0.2045
SYMPTO vs. ASYMPTO	8.47E-08*	2.492E-10*	1.372E-14*	0.8365	8.96E-08*	1.582E-08*	1.148E-10*	0.1905	0.9958	0.2794
INATT vs. ADHD-like	0.0112*	0.5124	1.064E-12*	0.999	0.0175*	0.0004536*	0.00006566*	0.0072*	0.9908	0.6406
INATT vs. ASYMPTO	0.999	1.001E-10*	0.999	0.0014*	0.0063*	0.3899	0.999	0.2112	0.5796	0.5005
ADHD-like vs. ASYMPTO	0.0028*	9.94E-14*	6.195E-17*	0.007*	0.999	0.3745	0.0028*	0.1257	0.8629	0.8588
Comparisons to Controls										
SYMPTO vs. Controls	3.822E-47*	3.017E-40*	8.54E-44*	1.813E-14*	2.989E-44*	1.022E-34*	1.316E-43*	0.4003	0.9255	0.8312
INATT vs. Controls	7.07E-24*	3.997E-44*	4.879E-09*	0.999	0.0001428*	0.0063*	8.96E-18*	0.0733	0.3141	0.0086*
ADHD-like vs. Controls	8.26E-37*	5.992E-52*	5.089E-50*	0.371	9.17E-15*	7.98E-17*	5.824E-38*	0.2550	0.5786	0.0745
ASYMPTO vs. Controls	7.77E-33*	1.127E-08*	4.669E-06*	2.891E-11*	2.184E-17*	1.176E-08*	1.022E-16*	0.5563	0.8844	0.0883
hSRS vs. Controls	1.19E-52*	5.509E-35*	6.685E-20*	3.584E-08*	8.75E-23*	2.184E-17*	1.022E-25*	0.8387	0.3943	0.0218*

**Table 4.2.7:** Demographic and behavioral score comparisons in the *Dimensional + Repetitive* sample, presented as p-values, between autistic traits subgroups and between each autistic traits subgroup and controls. The top section shows these comparisons in the behavioral cohort ('A. Behavioral Cohort (n=256)'), with autistic traits subgroups compared to each other (under 'Comparisons Between Subgroups') and with each autistic traits subgroup compared to controls (under 'Comparisons to Controls'). *hSRS* denotes a group that was created by combining all autistic traits subgroups in order to compare against controls. The bottom section (B. Neuroimaging Cohort (n= 219; Subset of Behavioral Cohort)) presents the same comparisons in the neuroimaging cohort, which is composed of a subset of subjects from the behavioral cohort with usable structural MRI data. P-values were adjusted using the Bonferroni method. Significant p-values are bolded with a '\*' superscript.

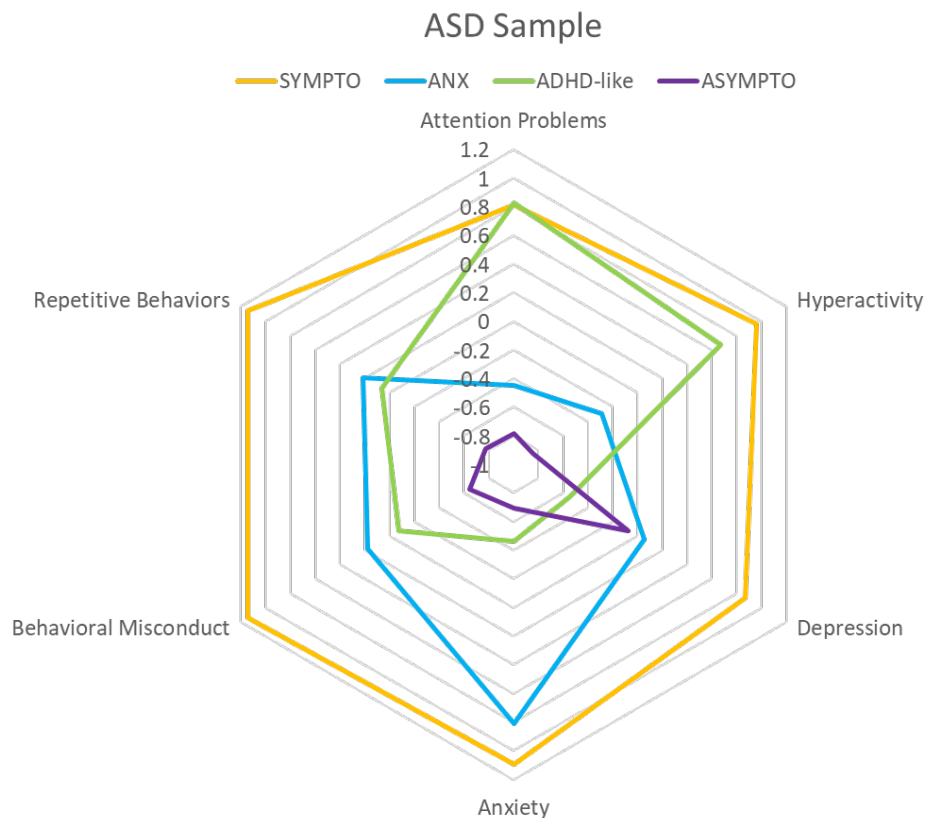
### Sample 3: ASD

Based on the computed BIC value distribution, a  $k$  value of four was retained for our supervised partitioning of the *ASD* sample (Figure 4.2.7).



**Figure 4.2.7.** The BIC value distribution suggested a  $k$  value of four to be used in the clustering analysis for the *ASD* sample.

Thus, upon running the clustering analysis, we obtained four subgroups of autistic patients with various behavioral profiles. The first subgroup expressed high levels of anxiety,  $n = 32$  (described as *ANX*); the second subgroup was asymptomatic in the studied behaviors,  $n = 55$  (described as *ASYMPTO*); the third subgroup exhibited high levels of ADHD-like behaviors,  $n = 42$  (referred to as *ADHD-like*); and the fourth was symptomatic in all behaviors, except for depression,  $n = 30$  (referred to as *SYMPTO*)(Figure 4.2.8 and Table 4.2.8). For this sample, true controls (according to the EU-AIMS consortium) were added as a control contrast in further neuroimaging and genetic analyses.



**Figure 4.2.8.** A radar plot summarizing the behavioral profiles of each subgroup generated in the *ASD* sample clustering analysis. Numbers along the plot indicate z-scores. To note, it may seem like the depression scale is clinically significant, however this is due to a highly skewed distribution and a small range. Therefore, due to the nature of how z-scores are calculated, the resulting value was sensitive to smaller changes. However, as seen in Table 4.2.8, indeed the average depression levels are low.

Subgroup	n (Total = 256)	ADHD-RS: Inattention	ADHD-RS: Hyperactivity	DAWBA: Depression	DAWBA: Anxiety	DAWBA: Behavioral Misconduct	RBS: Repetitive Behaviors	Mean Age	Gender Ratio	Mean FSIQ
<b>ANX</b> (Subgroup 1)	32	3.6	2.6	0.8	<b>3.6*</b>	2.8	19.4	12.7	19:13	101.4
<b>SYMPTO</b> (Subgroup 2)	30	<b>7.4*</b>	<b>6.2*</b>	1.7	<b>4.0*</b>	<b>3.9*</b>	<b>30.3*</b>	12.3	22:8	106.3
<b>ADHD-like</b> (Subgroup 3)	42	<b>7.5*</b>	<b>5.4*</b>	0.2	1.9	2.5	17.6	11.6	32:10	98.8
<b>ASYMPTO</b> (Subgroup 4)	55	2.5	0.9	0.7	1.6	1.9	7.7	13.6	39:16	100.5
<b>True Controls</b>	96	0.9	0.4	0.3	1.0	1.5	1.2	13.2	62:34	107.0

**Table 4.2.8.** Behavioral score values in each subgroup generated in the *ASD* sample. Stars (\*) indicate values surpassing high clinical levels of each score. Ranges for each score: **ADHD-RS-Inattention**: 0 to 9; **ADHD-RS-Hyperactivity**: 0 to 9; **DAWBA-Depression**: 0 to 5; **DAWBA-Anxiety**: 0 to 5; **DAWBA-Behavioral Misconduct**: 0 to 5; **RBS-Repetitive Behaviors**: 0 to 100.

Finally, as was done for the previous two samples, all *ASD* subgroups were compared to one another on behavioral scores, isolating several significant differences (Table 4.2.9A, '*Comparisons Between Subgroups*'). Each *ASD* subgroup was also compared to controls (typically developing, as defined by the EU-AIMS cohort), with every subgroup exhibiting significant differences in all behaviors except for depression in the *ADHD-like* subgroup. Again, we combined all four subgroups to compare against controls in a case-control fashion, which yielded significant differences in every behavioral score (Table 4.2.9A, '*Comparisons to Controls*').

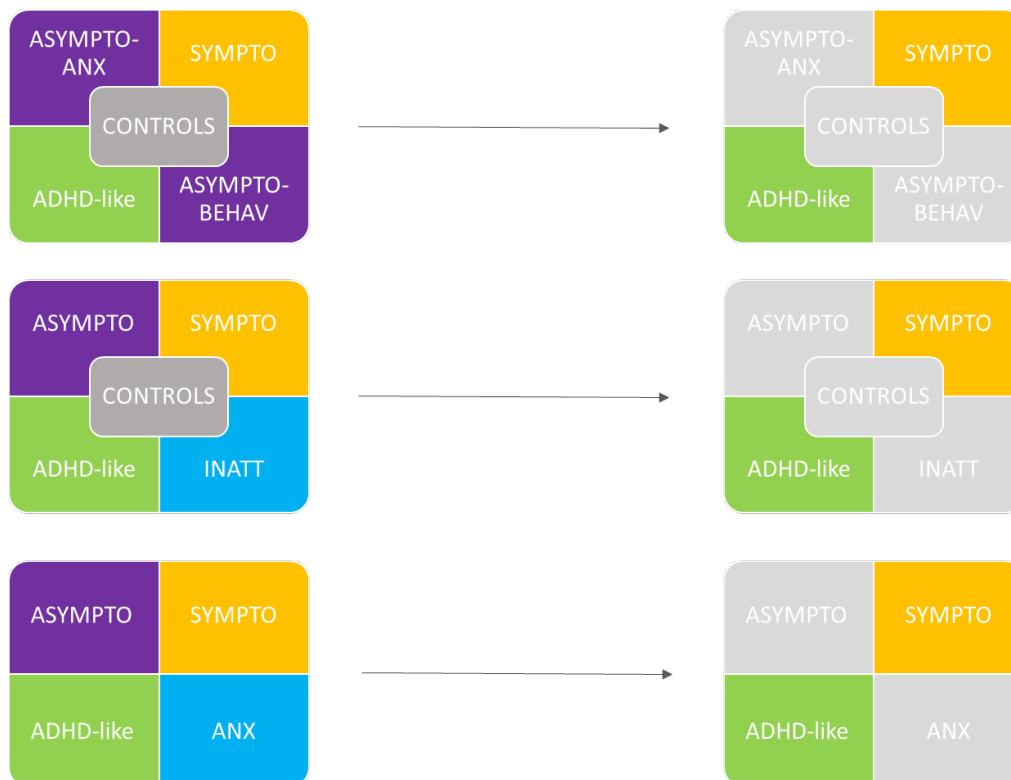
Within the demographic information for the between subgroup comparisons, we found no differences in gender. Additionally, we found a difference in age only between the *ADHD-like* and *ASYMPTO* subgroups, and a difference in FSIQ only between the *SYMPTO* and *ADHD-like* subgroups (Table 4.2.9A, '*Comparisons Between Subgroups*'). Upon comparing each subgroup to controls, we found differences in all behaviors in all subgroups, with the exception of depression in the *ADHD-like* subgroup and hyperactivity in the *ASYMPTO* subgroup. There were no differences in gender, and age only differed between the *ADHD-like* subgroup and controls. Also, there were differences in FSIQ between the *ADHD-like* and *ASYMPTO* versus controls. Lastly, the comparison between the autistic group and controls yielded significant differences FSIQ, but not age or gender (Table 4.2.9A, '*Comparisons to Controls*'). Due to these differences, we deemed it important to control for FSIQ, age and gender in the subsequent neuroimaging analysis.

	ADHD-RS: Inattention	ADHD-RS: Hyperactivity	DAWBA: Depression	DAWBA: Anxiety	DAWBA: Behavioral Misconduct	RBS: Repetitive Behaviors	Age	Gender (X)	FSIQ
<b>A. Behavioral Cohort (n = 255)</b>									
<b>Comparisons Between Subgroups</b>									
<i>ANX vs. SYMPTO</i>	2.082E-09*	1.404E-09*	0.0294*	0.7356	0.0002214*	0.0036*	0.5489	0.3724	0.2346
<i>ANX vs. ADHD-like</i>	4.182E-12*	0.00002178*	0.00036*	3.726E-11*	0.999	0.999	0.0818	0.1954	0.5268
<i>ANX vs. ASYMPTO</i>	0.357	0.000096*	0.999	4.356E-16*	0.00002712*	2.64E-09*	0.1890	0.3872	0.8192
<i>SYMPTO vs. ADHD-like</i>	0.999	0.8034	5.286E-08*	2.928E-14*	1.704E-07*	0.0000078*	0.2725	0.9874	0.0409*
<i>SYMPTO vs. ASYMPTO</i>	3.75E-15*	2.856E-25*	0.0012*	1.62E-19*	3.396E-16*	1.302E-17*	0.0522	0.9882	0.1207
<i>ADHD-like vs. ASYMPTO</i>	2.16E-19*	2.406E-17*	0.0114*	0.423	0.0012*	8.04E-12*	0.0009*	0.7259	0.6429
<b>Comparisons to Controls</b>									
<i>ANX vs. Controls</i>	5.064E-09*	2.04E-10*	0.00005904*	9.3E-33*	6.42E-15*	3.228E-31*	0.4267	0.7508	0.0690
<i>SYMPTO vs. Controls</i>	5.214E-35*	1.398E-39*	4.566E-13*	3.204E-37*	9.9E-33*	6.84E-41*	0.1352	0.5057	0.7843
<i>ADHD-like vs. Controls</i>	4.488E-43*	7.56E-30*	0.999	3.984E-09*	2.304E-12*	1.056E-42*	0.0026*	0.2511	0.0029*
<i>ASYMPTO vs. Controls</i>	0.0000246*	0.0726	0.0018*	0.00002232*	0.006*	7.26E-19*	0.4248	0.5385	0.0128*
<i>ASD vs. Controls</i>	1.896E-25*	3.24E-18*	0.00004752*	1.014E-20*	3.894E-16*	1.122E-28*	0.1438	0.4040	0.0051*
<b>B. Neuroimaging Cohort – Subset of Behavioral Cohort (n = 218)</b>									
<b>Comparisons Between Subgroups</b>									
<i>ANX vs. SYMPTO</i>	3.414E-07*	2.214E-08*	0.0222*	0.999	0.0003084*	0.0558	0.3434	0.6798	0.4168
<i>ANX vs. ADHD-like</i>	2.142E-09*	0.0002874*	0.006*	4.266E-08*	0.999	0.6192	0.1335	0.5594	0.7355
<i>ANX vs. ASYMPTO</i>	0.999	0.0216*	0.999	6.06E-12*	0.0002646*	1.158E-07*	0.3274	0.6095	0.5672
<i>SYMPTO vs. ADHD-like</i>	0.999	0.5934	1.884E-06*	6.72E-10*	5.604E-06*	0.0000189*	0.5983	0.8578	0.1544
<i>SYMPTO vs. ASYMPTO</i>	8.52E-12*	1.056E-21*	0.0012*	3.504E-14*	1.776E-15*	1.656E-15*	0.0415*	0.8254	0.1132
<i>ADHD-like vs. ASYMPTO</i>	7.68E-15*	2.958E-13*	0.0852	0.6384	0.000195*	6.42E-08*	0.0063*	0.9563	0.7834
<b>Comparisons to Controls</b>									
<i>ANX vs. Controls</i>	0.00000678*	0.00001236*	0.000399*	1.596E-27*	2.298E-12*	3.162E-28*	0.8544	0.9543	0.3153
<i>SYMPTO vs. Controls</i>	1.56E-29*	9.9E-36*	4.188E-14*	3.24E-31*	4.248E-33*	1.716E-41*	0.3433	0.7116	0.9173
<i>ADHD-like vs. Controls</i>	2.112E-35*	2.496E-24*	0.999	5.718E-08*	2.46E-13*	7.86E-38*	0.0980	0.5533	0.0888
<i>ASYMPTO vs. Controls</i>	0.0000666*	0.2988	0.003*	0.00005556*	0.0438*	4.902E-17*	0.1126	0.6010	0.0246*
<i>ASD vs. Controls</i>	2.568E-21*	2.292E-14*	0.00002778*	4.548E-19*	2.184E-14*	3.642E-26*	0.9546	0.5374	0.0482*

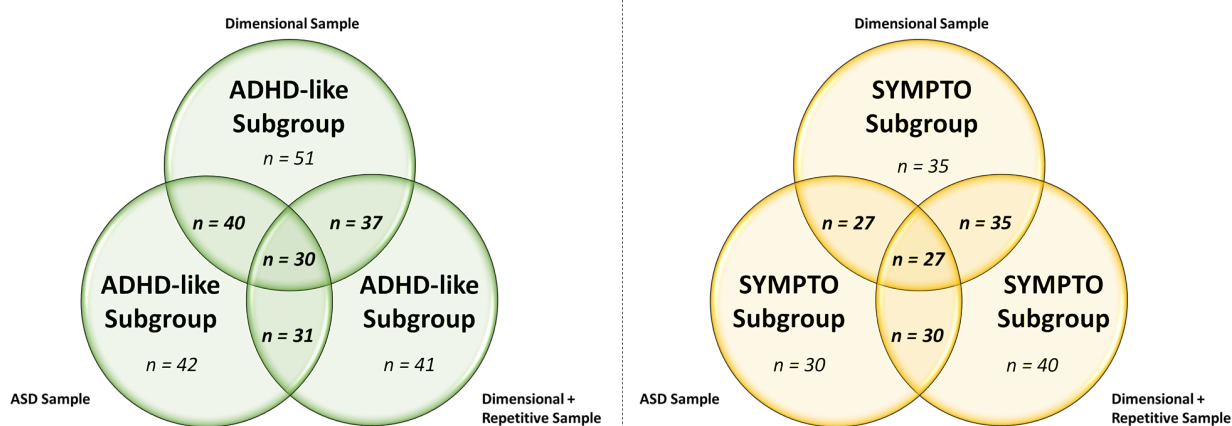
**Table 4.2.9:** Demographic and behavioral score comparisons in the *ASD* sample, presented as p-values, between autistic subgroups and between each autistic subgroup and controls. The top section shows these comparisons in the behavioral cohort ('A. Behavioral Cohort (n= 159 ASD + 96 Controls = 255)'), with autistic subgroups compared to each other (under 'Comparisons Between Subgroups') and with each autistic subgroup compared to controls (under 'Comparisons to Controls'). *hSRS* denotes a group that was created by combining all autistic subgroups in order to compare against controls. The bottom section (B. Neuroimaging Cohort (n= 218; Subset of Behavioral Cohort)) presents the same comparisons in the neuroimaging cohort, which is composed of a subset of subjects from the behavioral cohort with usable structural MRI data. P-values were adjusted using the Bonferroni method. Significant p-values are bolded with a '\*' superscript.

### *Commonalities in the Behavioral Subgroups Between the Three Samples*

As can be seen above, there were some repetitions among the generated subgroups. The most consistent subgroups were the *ADHD-like* and *SYMPTO* subgroups, which appeared in every clustering analysis indicating the presence of relatively stable dimensional profiles (Figure 4.2.9 and Figure 4.2.10). The *ASYMPTO* subgroup also appeared more than once in all three sample analyses, again indicating stability of this behavioral profile in autistic and autistic-like individuals. Lastly, Table 4.2.10 summarizes the proportion of diagnosed ASD patients per subgroup per sample.



**Figure 4.2.9.** The clustering analysis within each of the three samples collectively generated two subgroups with consistent profiles: the *SYMPTO* and *ADHD-like* subgroups. Also the *ASYMPTO* subgroup was consistently generated.



**Figure 4.2.10.** A Venn diagram indicating the number of overlapping subjects between samples, and overall, for the *ADHD-like* and *SYMPTO* subgroups.

Sample	Clusters	<i>n</i>	<i>n</i> Diagnosed with ASD (proportion of subjects)
Dimensional ( <i>n</i> = 256)	ASYMPTO-ANX	38	37 ( <b>97.3%</b> )
	SYMPTO	35	35 ( <b>100%</b> )
	ADHD-like	51	48 ( <b>94.1%</b> )
	ASYMPTO-BEHAV	27	24 ( <b>88.8%</b> )
	Controls (low social deficits)	105	16 ( <b>15.2%</b> )
Dimensional + Repetitive ( <i>n</i> = 256)	INATT	32	26 ( <b>81.2%</b> )
	SYMPTO	40	40 ( <b>100%</b> )
	ADHD-like	41	39 ( <b>95.1%</b> )
	ASYMPTO	46	43 ( <b>93.5%</b> )
	Controls (low social deficits)	97	12 ( <b>12.4%</b> )
ASD ( <i>n</i> = 255, including 159 ASD + 96 Controls)	ANX	32	32 ( <b>100%</b> )
	SYMPTO	30	30 ( <b>100%</b> )
	ADHD-like	42	42 ( <b>100%</b> )
	ASYMPTO	55	55 ( <b>100%</b> )
	True Controls	96	0 ( <b>0%</b> )

**Table 4.2.10.** A summary of diagnosed ASD patients per subgroup per sample.

## Part II: Neuroimaging Analysis of Cortical Surface Features

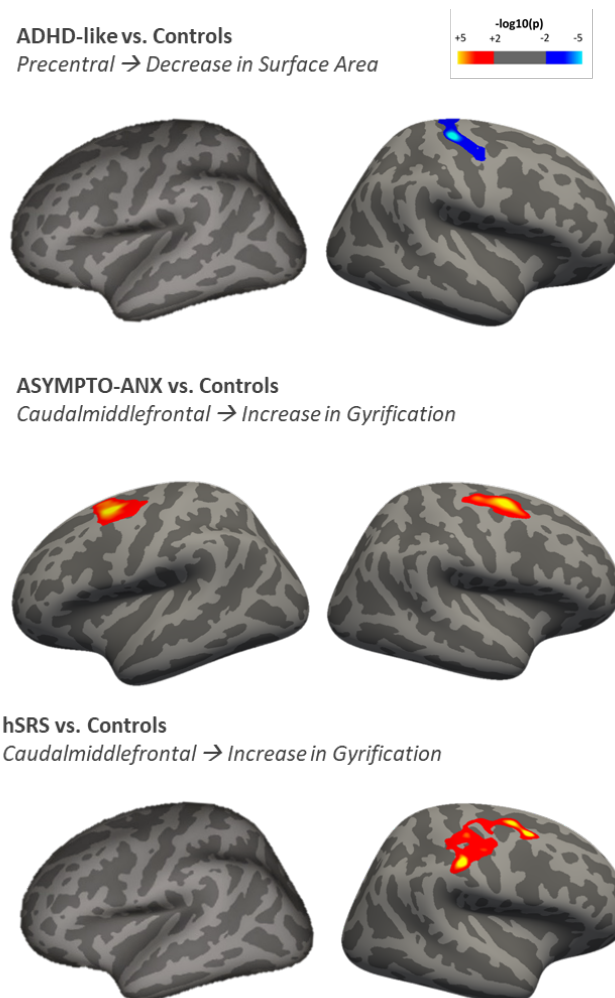
### ***Morphological Comparisons***

#### *Sample 1: Dimensional*

After removing subjects that have not undergone MRI acquisition and/or did not pass the T1 image quality check, we obtained subgroups of: *n* = 33 in the “ASYMPTO-ANX” group, *n* = 22 in the “ASYMPTO-BEHAV” group, *n* = 34 in the “ADHD-like” group, *n* = 26 in the “SYMPTO” group, and *n* = 104 controls, producing a total of 219 subjects with available T1 data participating in the neuroimaging analysis. Behavioral score and demographic information comparisons were nearly identical to the behavioral cohort (Table 4.2.5B, ‘*Neuroimaging Cohort*’). To delineate the interest and significance of subtyping in an autistic traits population, we first combined all four subgroups into one large group (*Dimensional-hSRS*, *n* = 115) and



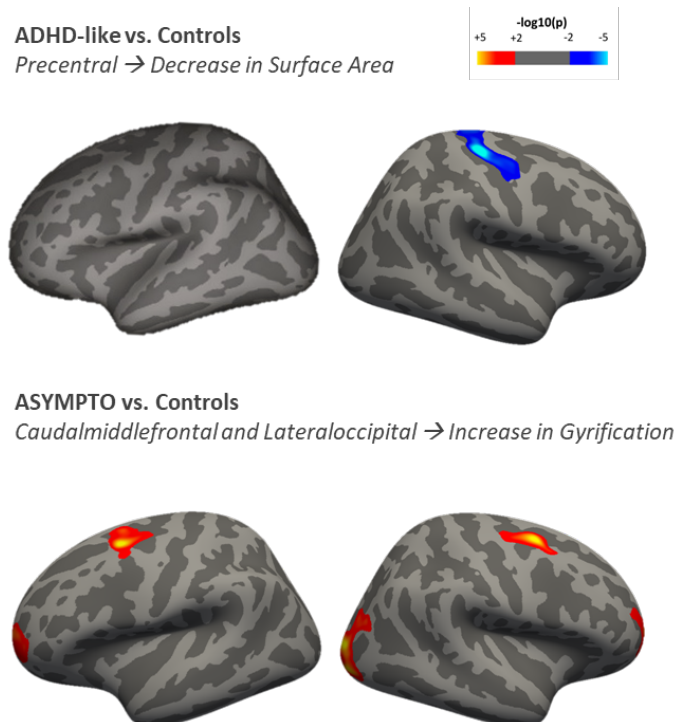
compared cortical thickness, local gyrification and surface area against respective controls ( $n = 104$ ). This comparison yielded a significant difference in increased gyrification peaking in the caudalmiddlefrontal region, but also comprising the supramarginal area ( $p < 0.005$ ), after correction for multiple comparisons. We then compared the same surface features between each of our four subgroups and controls. After correction for multiple comparisons, the *ASYMPTO-ANX* subgroup exhibited an increase in bilateral gyrification in the caudalmiddlefrontal ( $p < 0.005$ ), and the *ADHD-like* subgroup showed a decrease in the right hemisphere surface area in the precentral region ( $p < 0.005$ )(Figure 4.2.11).



**Figure 4.2.11.** Cortical brain regions that were significantly different between subgroups and their respective controls in the *Dimensional* sample. A difference between the *hSRS* group and controls is also reported in this sample. Colours represent the  $-\log_{10}(p\text{-value})$ , with red(+) indicating an increase and blue(-) indicating a decrease compared to controls in affected morphological features.

### *Sample 2: Dimensional + Repetitive*

After removing subjects that have not undergone MRI acquisition and/or did not pass the T1 image quality check in the *Dimensional + Repetitive* sample, we obtained subgroups of:  $n = 39$  in the “*ASYMPTO*” group,  $n = 29$  in the “*INATT*” group,  $n = 25$  in the “*ADHD-like*” group,  $n = 30$  in the “*SYMPTO*” group, and  $n = 96$  controls, producing a total of 219 subjects with available T1 data participating in the neuroimaging analysis. Behavioral score and demographic information comparisons were nearly identical to the behavioral cohort (Table 4.2.7B, ‘*Neuroimaging Cohort*’). To delineate the interest and significance of subtyping in an autistic traits population, we first combined all four subgroups into one large group (*Dimensional + Repetitive-hSRS*,  $n = 123$ ) and compared cortical thickness, local gyrification and surface area against respective controls ( $n = 96$ ). This comparison did not yield a significant result after correction for multiple comparisons. We then compared the same surface features between each of our four subgroups and controls. After correction for multiple comparisons, the *ASYMPTO* subgroup exhibited an increase in bilateral gyrification in the caudalmiddlefrontal, rostralmiddlefrontal and lateraloccipital ( $p < 0.005$ ). Also, we again observed a decrease in surface area in the *ADHD-like* subgroup in the right hemisphere precentral region ( $p < 0.005$ )(Figure 4.2.12).

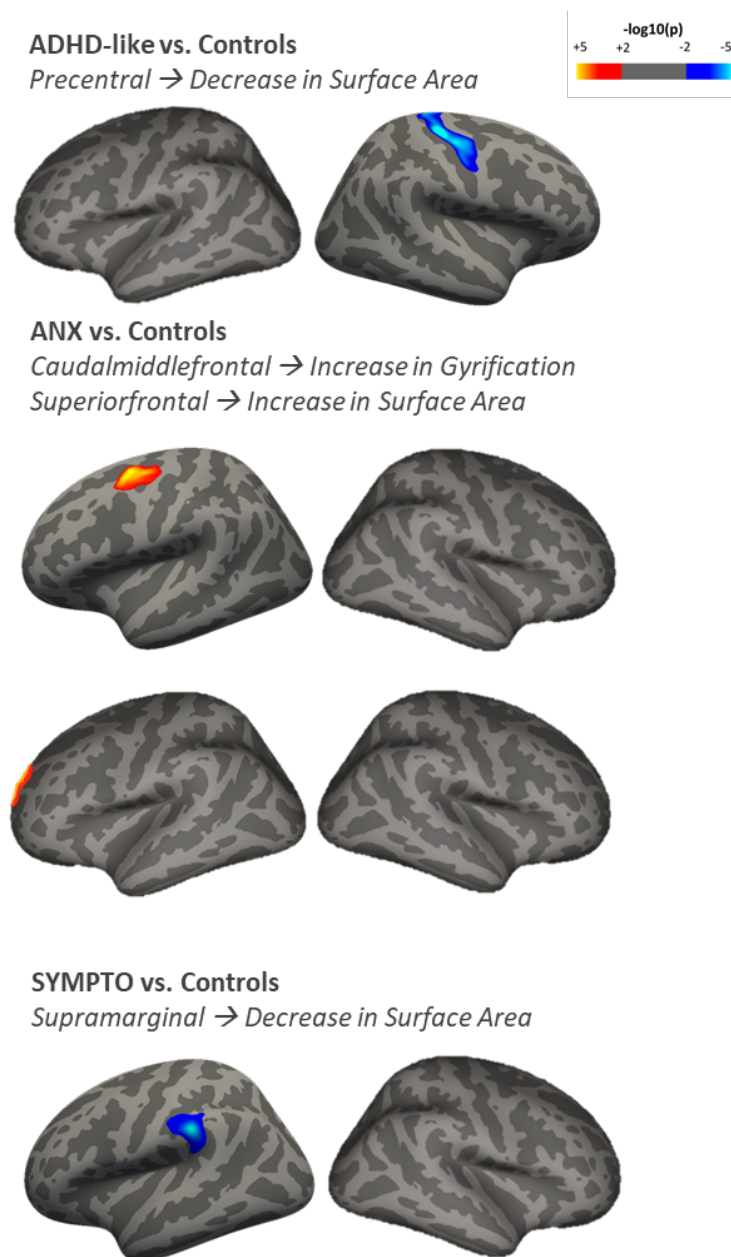


**Figure 4.2.12.** Cortical brain regions that were significantly different between subgroups and their respective controls in the *Dimensional + Repetitive* sample. Colours represent the  $-\log_{10}(p\text{-value})$ , with red(+) indicating an increase and blue(-) indicating a decrease compared to controls in affected morphological features.

### Sample 3: ASD

For the *ASD* sample, after removing subjects that have not undergone MRI acquisition and/or did not pass the T1 image quality check we obtained subgroups of:  $n = 23$  in the “*ANX*” group,  $n = 47$  in the “*ASYMPTO*” group,  $n = 29$  in the “*ADHD-like*” group,  $n = 24$  in the “*SYMPTO*” group, and  $n = 95$  controls (controls officially registered by the EU-AIMS consortium), producing a total of 218 subjects with available T1 data participating in this analysis. Behavioral score and demographic information comparisons were nearly identical to the behavioral cohort (Table 4.2.9B, ‘*Neuroimaging Cohort*’). To delineate the interest and significance of subtyping in an autistic population, we first combined all four subgroups into one large group (*ASD*,  $n = 123$ ) and compared cortical thickness, local gyrfication and surface area against controls ( $n = 95$ ). This comparison did not yield a significant result after correction for multiple comparisons. We then compared the same surface features between each of our four subgroups and controls.

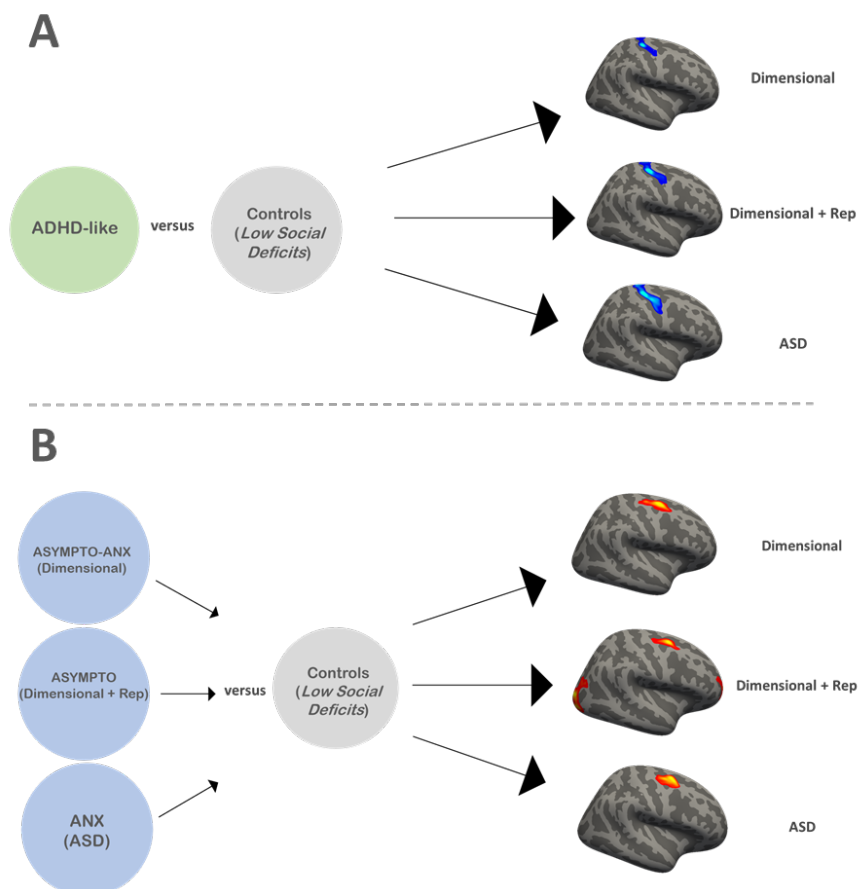
After correction for multiple comparisons, the *ANX* subgroup exhibited an increase in gyrification in the right caudalmiddlefrontal, and an increase in surface area in the left superiorfrontal ( $p < 0.005$ ). We also observed a decrease in surface area in the left supramarginal in the *SYMPTO* subgroup. Lastly, yet again, we observed a decrease in surface area in the *ADHD-like* subgroup in the right hemisphere precentral region ( $p < 0.005$ )(Figure 4.2.13).



**Figure 4.2.13.** Cortical brain regions that were significantly different between subgroups and their respective controls in the *ASD* sample. Colours represent the  $-\log_{10}(p\text{-value})$ , with red(+) indicating an increase and blue(-) indicating a decrease compared to controls in affected morphological features.

## Commonalities in the Neuroimaging Features Between the Three Samples

Just as we observed consistent behavioral subgroups between the three samples, we also saw coherent cortical features associated with these subgroups. The most stable result is a decrease in surface area in all of the *ADHD-like* subgroups across all samples (Figure 4.2.14A). Additionally, we saw an increase in gyrification in the right caudalmiddlefrontal in the *ASYMPTO-ANX* and *ANX* subgroups of the *Dimensional* and *ASD* samples, respectively, and in the *ASYMPTO* subgroup of the *Dimensional + Repetitive* sample (Figure 4.2.14B). Though anxiety was not statistically significant in the *ASYMPTO-ANX* and *ASYMPTO* subgroups, a major anxiety trend was present, therefore this common brain difference may potentially be attributed to anxiety levels. Overall, this indicates relatively consistent cortical biomarkers associated with both ADHD and anxiety traits in both autistic and autistic-like subjects.



**Figure 4.2.14.** A) A decrease in surface area in the right precentral region was observed in all *ADHD-like* subgroups across all samples. B) An increase in gyrification in the right caudalmiddlefrontal was seen in all subgroups having either a significantly high level of anxiety, or at least an elevated trend in anxiety.

### Part III: Polygenic Risk Scores

Combining all subgroups into one and comparing it against controls showed a significant association with the Empathy (lower) and Major Depressive Disorder (MDD) PGS in all three samples. Furthermore, each subgroup within each sample was compared against respective controls on all four PGS of interest.

#### *Sample 1: Dimensional*

In the *Dimensional* sample, the *ASYMPTO-ANX* and *ASYMPTO-BEHAV* subgroups were significantly linked to the MDD ( $p_{\text{adj}} = 0.016$ ) and a lower Empathy ( $p_{\text{adj}} = 0.0014$ ) PGS, respectively. Also, the ADHD PGS did not technically survive Bonferroni correction in the *ADHD-like* subgroup, however it was extremely close ( $p = 0.0128$ ; compared to the threshold of  $0.0125 = 0.05/4$ ).

#### *Sample 2: Dimensional + Repetitive*

In the *Dimensional + Repetitive* sample, the *SYMPTO* subgroup was associated with a lower Empathy PGS ( $p_{\text{adj}} = 0.022$ ), while the *ASYMPTO* subgroup was linked to the MDD ( $p_{\text{adj}} = 0.013$ ) and lower Empathy ( $p_{\text{adj}} = 0.012$ ) PGS. Again, for the *ADHD-like* subgroup, though the ADHD and MDD PGS p-values did not survive Bonferroni correction, they were close (both having  $p = 0.0135$ ; compared to the threshold of  $0.0125 = 0.05/4$ ).

### Sample 3: ASD

Lastly, in the *ASD* sample, the *ADHD-like* subgroup was significantly associated with the ADHD PGS ( $p_{\text{adj}} = 0.041$ ), while the *ASYMPTO* subgroup was significantly linked to the MDD ( $p_{\text{adj}} = 0.016$ ) and lower Empathy ( $p_{\text{adj}} = 0.044$ ) PGS.

A grand summary of all reported results is illustrated in Figure 4.2.15.

Sample	Clustering Results	Neuroimaging Results Against Controls	PGS Associations
<b>Dimensional</b>		<p>ADHD-like  <math>\Rightarrow</math> <math>\downarrow</math> Right <b>Precentral</b> Surface Area</p> <p>ASYMPTO-ANX  <math>\Rightarrow</math> <math>\uparrow</math> Bilateral <b>Caudal middle frontal</b> Gyrification</p> <p>hSRS  <math>\Rightarrow</math> <math>\uparrow</math> Right <b>Caudal middle frontal</b> Gyrification</p>	<p>ADHD-like <math>\Rightarrow</math> <b>ADHD</b> (trend)</p> <p>ASYMPTO-ANX <math>\Rightarrow</math> <b>MDD</b></p> <p>ASYMPTO-BEHAV <math>\Rightarrow</math> <b>Empathy</b></p> <p>Categorical <math>\Rightarrow</math> <b>Empathy, MDD</b></p>
<b>Dimensional + Repetitive</b>		<p>ADHD-like  <math>\Rightarrow</math> <math>\downarrow</math> Right <b>Precentral</b> Surface Area</p> <p>ASYMPTO  <math>\Rightarrow</math> <math>\uparrow</math> Bilateral <b>Caudal middle frontal</b> Gyrification  <math>\uparrow</math> Right <b>Lateral occipital</b> Gyrification</p>	<p>ADHD-like <math>\Rightarrow</math> <b>ADHD, MDD</b> (trend)</p> <p>ASYMPTO <math>\Rightarrow</math> <b>Empathy, MDD</b></p> <p>SYMPTO <math>\Rightarrow</math> <b>Empathy</b></p> <p>Categorical <math>\Rightarrow</math> <b>Empathy, MDD</b></p>
<b>ASD</b>		<p>ADHD-like  <math>\Rightarrow</math> <math>\downarrow</math> Right <b>Precentral</b> Surface Area</p> <p>ANX  <math>\Rightarrow</math> <math>\uparrow</math> Right <b>Caudal middle frontal</b> Gyrification  <math>\downarrow</math> Left <b>Superior frontal</b> Surface Area</p> <p>SYMPTO  <math>\Rightarrow</math> <math>\downarrow</math> Left <b>Supramarginal</b> Surface Area</p>	<p>ADHD-like <math>\Rightarrow</math> <b>ADHD</b></p> <p>ASYMPTO <math>\Rightarrow</math> <b>Empathy, MDD</b></p> <p>Categorical <math>\Rightarrow</math> <b>Empathy, MDD</b></p>

**Figure 4.2.15.** A visual summary of all results reported in the current study. Orange PGS scores indicate a lower score value.

## Discussion

Recent interests in autism characterization through stratification have shifted away from case-control studies due to excessive heterogeneity reported within the disorder, resulting in the use of dimensional approaches. Our group recently published an original study on the population-based HBN cohort in order to investigate how autistic traits cluster with other behavioral traits, and to further observe how these clusters express cortically in the brain. Based on this paper we

decided to replicate these objectives in three samples acquired from the EU-AIMS cohort. The reasoning behind running the same analyses in parallel samples was to observe if dimensional analyses in a ‘general population’ could generate similar phenotypes as a dimensional analysis conducted in a strictly ASD sample. Specifically, for the dimensional approach, we curated two samples: one depending on availability of the same behavioral variables (to the best of our ability) as used in *Mihailov et al., 2020* (entitled the *Dimensional* sample), and a second with these same behaviors plus the addition of repetitive behaviors, since this variable was insufficiently available in the HBN cohort (entitled the *Dimensional + Repetitive* sample). Next, since the EU-AIMS cohort was created specifically for studying autism, we decided to build one more sample based on availability in the same behaviors (including repetitive behaviors) in a group of ASD-only patients (entitled the *ASD* sample). All samples contained overlapping subjects, since they are part of the same cohort. Next, we looked at structural neuroimaging traits and polygenic scores (PGS) in all subgroups in each sample. Though several unique behavioral profiles were elicited in each sample following a clustering analysis, of greatest interest to us was to see if any profiles were consistently generated despite changes in sample criteria. Indeed there were, since in all three sample clusterings we observed subgroups that were symptomatic in all behaviors except depression (*SYMPTO*), subgroups that were asymptomatic in all behaviors except, of course, social deficits (*ASYMPTO*, *ASYMPTO-ANX* and *ASYMPTO-BEHAV*), and finally subgroups high in ADHD-like behaviors (*ADHD-like*). Furthermore, in the neuroimaging analysis, the only subgroup type that generated the same result across all samples was the *ADHD-like* subgroup, which displayed a significant decrease in surface area in the right precentral region. Additionally, we observed an increase in gyrification in the caudalmiddlefrontal in three subgroups: the *ASYMPTO-ANX* of the *Dimensional* sample, the *ASYMPTO* of the *Dimensional + Repetitive* sample, and the *ANX* subgroup in the *ASD* sample. Most interestingly, we noticed that by comparing brain features



of one heterogeneous autistic or autistic traits group (composed by combining all subgroups) to controls, limited results were obtained. The only time we observed a significant result was by comparing the *Dimensional-hSRS* group to controls, which resulted in increased gyrification peaking in the caudalmiddlefrontal region. Lastly, we compared PGS between each subgroup and their respective controls, and consistently found that the *ADHD-like* subgroups across all samples were either associated significantly, or almost significantly, with the PGS calculated for ADHD disorders. We suggest that our results advocate for a shift towards subtyping autistic patients in studies in order to better characterize the disorder since case-control comparisons are not refined enough to yield unique and reliable features.

Behaviors studied in our subgroups were chosen due to their occurrence in several psychiatric disorders comorbid with ASD<sup>62,89</sup>. Up to 95% of autistic patients have reported a psychiatric comorbidity including ADHD, anxiety disorders, mood disorders, OCD, and conduct disorders, which are defined by the behaviors reported in our subgroups<sup>62,85,345</sup>. Since these comorbidities are often reported based on diagnostic information, and thus not characterized very well in alignment with ASD research, the current study investigates them from a dimensional perspective due to the accessibility of a multidimensional cohort. We believe this will provide a more refined and spectral explanation of exactly how autistic traits distribute with other symptoms. In the current study, we performed such analyses in both a dimensional and an autistic sample in order to see if consistencies remained in order to confirm our dimensional methodology. We believe that studying autistic traits dimensionally is a superior method in the understanding of ASD, and that observing similar subgroups between dimensional and autistic samples confirms the validity of this approach and offers more comprehensive information. We observed three subgroups that were generated in each of the three autistic traits/autistic samples. The first, *ADHD-like*, had clinically significant levels of attention problems and hyperactivity,

which is in line with the *Attn* subgroup generated in *Mihailov et al., 2020*<sup>302</sup>. ADHD is profusely reported in ASD patients, therefore this subgroup could be representative of subjects exhibiting only these traits, with no overlap in other behaviors<sup>62,85</sup>. The second consistent subgroup we observed, the *SYMPTO* subgroup, had autistic or autistic-like patients with significantly high levels in all behaviors except depression. This subgroup is comparable to the *Emot* subgroup produced in *Mihailov et al., 2020*, since it also exhibits high levels of ADHD behaviors in addition to misconduct-related behaviors. The only difference is that the *SYMPTO* subgroup also exhibits significant anxiety levels while the *Emot* does not. It is worth noting however, that anxiety levels were elevated in the *Emot* subgroup. The *SYMPTO* subgroup is interesting as it seems to represent a subtype of autism that is symptomatic in all traits, thus probably representing patients with severe cases of autism. Since ADHD is one of the most widely reported comorbidities in ASD, it is interesting that the *SYMPTO* subgroup also exhibits ADHD traits which may perhaps represent another subtype of autistic patients with ADHD traits who additionally experience external psychiatric traits (as opposed to the *ADHD-like* subgroup that strictly expresses ADHD traits). The third consistent behavioral subgroup involved autistic or autistic-like individuals that were asymptomatic in all traits (except social deficits), which could represent a subtype of higher-functioning autistic patients.

An important trait to discuss in these results is depression. It is indeed strange that though it is widely reported in autistic patients, depression was the only behavioral trait not reaching clinical significance in any subgroups. This could potentially be explained by constant levels of depression in autistic patients that are not different enough to elicit subgroups, or that depression is unstable and therefore difficult to characterize in autistic patients. Alexithymia (difficulty to identify and express emotions) is common in subjects with ASD and may also explain the low observed levels of depression<sup>401</sup>. Alternatively, it could also indicate that

autistic patients with significantly high levels of depression are quite small in number, and that running an analysis on such a limited number of subjects does not yield enough power to isolate a subgroup with high levels of depression. Indeed, depression levels are significantly different between most of the subgroups, however these differences are due to variations in low levels of depression. In any case, reported subgroups have clarified at least part of the variability present in autistic and autistic-like samples, as evidenced by their differing behavioral profiles, which offers an increased likelihood of finding interesting traits such as cortical signatures and genetic differences.

In each sample, conducting a case-control comparison in cortical morphology looking at cortical thickness, gyrification and surface area, yielded none or extremely limited results. The only difference we found was in the *Dimensional* sample upon comparing *Dimensional-hSRS* versus controls where we found a peak increase in gyrification in the right caudalmiddlefrontal and part of the supramarginal ( $cwp < 0.005$ ), which are regions typically involved in emotional regulation, addiction behaviors and language processing<sup>402–404</sup>. This cortical result, however, was also found in the *Dimensional ASYMPTO-ANX* subgroup, bilaterally ( $cwp < 0.005$ ), therefore we suspect that the strength of this result could have resonated in the case-control comparison upon combination of subgroups. Interestingly, this caudalmiddlefrontal alteration was also a consistent result observed in all three samples. Specifically, we witnessed this increase in gyrification bilaterally in the *ASYMPTO-ANX* subgroup of the *Dimensional* sample, the *ASYMPTO* subgroup of the *Dimensional + Repetitive* sample, and only in the right hemisphere in the *ANX* subgroup of the *ASD* sample ( $cwp < 0.005$ ), when compared to corresponding controls. We suspect this result to be linked to anxiety levels since we observed it in the *ANX* subgroup, as well as in the *ASYMPTO* and *ASYMPTO-ANX* subgroups (though anxiety did not reach significance in these last two subgroups, it was certainly elevated).

Furthermore, this finding of increased gyrification in the superior frontal area has already been reported in patients with anxiety<sup>405</sup>. The second consistently reported neuroimaging result in this study was a decrease in surface area in the right hemisphere precentral gyrus ( $cwp < 0.005$ ) in all *ADHD-like* subgroups from all samples upon comparison to their corresponding controls. This cortical signature is particularly interesting since in *Mihailov et al., 2020*, we also found alterations in the precentral region in the *Attn* (having ADHD traits) subgroup, however it was in the opposite direction and hemisphere. Indeed, this discrepancy is not favorable, however at the very least it allows us to conclude that ASD and ASD-like subjects with ADHD-like traits tend to display alterations in the precentral region. The precentral region typically functions in motor control, with studies reporting links between ASD, ADHD and motricity, as well as alterations in precentral functioning in ASD-ADHD subgroups<sup>406–408</sup>. Seeing as how structural alterations were observed within the motor area in all *ADHD-like* subgroups, further investigation is warranted into the relationship between ADHD traits and motor control in autistic and/or autistic-like individuals. Neuroimaging results reported in this study prove that subgrouping autistic and autistic-like samples may ease the characterization of biological traits, especially since we observed consistent results across subgroups, across samples. Several of the current results have been sporadically reported in the literature, which demonstrates that by refining subgroups we can pinpoint sources of variability typically described in ASD studies.

Due to the availability of genetic data in the EU-AIMS cohort, relevant polygenic (risk) scores (PGS) were compared between subgroups and their respective controls, as well as in a case-control manner, across all three samples. Interestingly, the case-control comparisons (i.e. *hSRS/ASD* versus controls) showed a significant difference in the MDD and Empathy PGS in all samples. It is indeed puzzling as to why we did not observe an association with the ASD PGS, however the Empathy results are still quite interesting. Certainly, as explained thoroughly

throughout this manuscript, heterogeneity complicates autism characterization, which can be one reason as to why we do not observe any results with the ASD PGS. Additionally, the fact that the Empathy PGS separates well between autistic/autistic-like subjects and controls could indicate that this trait is well characterized genetically, further accentuating the necessity of using dimensional constructs over diagnostic categories to understand psychiatric diseases. The reasoning behind the association with the MDD PGS is probably more complex, since indeed out of all the behaviors included in the study, depression was the one construct that never reached a clinical threshold in any of the subgroups. Yet we observe it on a genetic level in case-control and subgroup-control comparisons. Most interestingly however, were the ADHD PGS associations that we observed in all *ADHD-like* subgroups. Specifically, in the *ASD* sample, we detected a significant association between the ADHD PGS and the *ADHD-like* subgroup, and furthermore, though the associations between the ADHD PGS and *ADHD-like* subgroups in the remaining two samples did not survive Bonferroni correction, they were still at the limit ( $p_{adj} = 0.051$  and  $p_{adj} = 0.054$ ). This suggests that from all PGS scores, ADHD traits are best characterized, perhaps indicating a solid genetic component. However, though we observed interesting results, these genetic characterizations are still quite variable in nature, indicating the presence of vast heterogeneity in the genetics of ASD patients and warranting further study into this field.

The purpose of this study was to investigate the presence of similar traits using samples constructed based on differing subject criteria. Though there were some consistencies within subgroups, the most consistent results by far were generated by the *ADHD-like* subgroups, from the beginning until the end (Figure 4.2.16). First, *ADHD-like* subgroups were generated from the clustering analysis in all samples with the same behavioral profile, next we observed decreases in surface area in the right precentral region, and finally, we found relevant

associations between these subgroups and the ADHD PGS. Furthermore, the *ADHD-like* subgroup and its neuroimaging results were the only observations that moderately mirrored what was reported in *Mihailov et al., 2020*. It seems that ADHD traits in autistic and autistic-like individuals show less biological and behavioral heterogeneity as compared to other constructs. Indeed, ASD heterogeneity will probably not be refined using a one size fits all partitioning method, but the current study was at least able to define a robust ADHD traits subgroup in autistic participants, which isolates a well-defined subtype. If one ASD subtype can already be well characterized, this could lead to a better understanding of underlying biological mechanisms and eventually jump-start the development of treatments and therapies to treat at least this portion of patients. The current study was able to successfully characterize a fragment of the heterogeneity present in ASD by taking into account several traits (behavioral, neuroimaging and genetic) and using diverse methods (a combination of clustering and dimensional approaches) while still producing consistent results. We therefore stress the importance of further investigating autistic patients with ADHD traits, with associated precentral region alterations, as this may represent a potentially substantial finding.

Population	Clustering Results	Neuroimaging Results Against Controls	PGS Associations
<b>Dimensional</b>		<p>ADHD-like  → ↓ Right Precentral Surface Area</p> <p>ASYMPTO-ANX  → ↑ Bilateral Caudalmiddlefrontal Gyrfication</p> <p>hSRS  → ↑ Right Caudalmiddlefrontal Gyrfication</p>	<p>ADHD-like → ADHD (trend)</p> <p>ASYMPTO-ANX → MDD</p> <p>ASYMPTO-BEHAV → Empathy</p> <p>Categorical → Empathy, MDD</p>
<b>Dimensional + Repetitive</b>		<p>ADHD-like  → ↓ Right Precentral Surface Area</p> <p>ASYMPTO  → ↑ Bilateral Caudalmiddlefrontal Gyrfication → ↑ Right Lateraloccipital Gyrfication</p>	<p>ADHD-like → ADHD, MDD (trend)</p> <p>ASYMPTO → Empathy, MDD</p> <p>SYMPTO → Empathy</p> <p>Categorical → Empathy, MDD</p>
<b>ASD</b>		<p>ADHD-like  → ↓ Right Precentral Surface Area</p> <p>ANX  → ↑ Right Caudalmiddlefrontal Gyrfication → ↑ Left Superiorfrontal Surface Area</p> <p>SYMPTO  → ↓ Left Supramarginal Surface Area</p>	<p>ADHD-like → ADHD</p> <p>ASYMPTO → Empathy, MDD</p> <p>Categorical → Empathy, MDD</p>

**Figure 4.2.16.** A visual summary of results in the *ADHD-like* subgroup across all samples. This subgroup generated the most consistent results in behavioral, cortical and PGS features.

This study has several limitations. Firstly, the determination of the number of clusters in the unsupervised  $k$ -means clustering approach is somewhat arbitrary, making it challenging to replicate the same partitioning in other datasets. However, we tried to address this challenge by implementing a BIC approach to systematically choose a value. Additionally, clustering analyses do not typically allow for the covarying of confounding variables. Also, after filtering for behavioral assessment and neuroimaging data availability, the study sample significantly decreased. Therefore, though we report interesting and consistent observations, future studies with larger subgroups are still necessary to confirm these results. Additionally, this study aimed to replicate the methodology of *Mihailov et al., 2020*, which it did not completely achieve. Though many of the behavioral clusterings were similar, or even the same (for example, the *Attn* and *ADHD-like*, and the *Emot* and *SYMPTO*), we did not witness many of the neuroimaging results, except for the precentral alteration. This however is not entirely surprising since there were major differences between the cohorts including that the HBN cohort is a population-based one, while the EU-AIMS cohort is clinical. Also, though behaviors studied were as similar as possible, the instruments used to quantify them were different between cohorts (with the exception of the SRS scale), which could have also introduced several variabilities and biases. Lastly, though we investigated interesting variables, this list is by no means exhaustive in the understanding of ASD biological traits. Diversity associated with autism and autistic traits needs to be further explained by other modalities, therefore next steps would involve considering more sophisticated genetic data, as well as other neuroimaging features including volumetric, diffusion, and functional differences.

In conclusion, we report a behaviorally and neuroanatomically relevant cluster in ASD, namely a subgroup of subjects that have ADHD traits and alterations in the precentral region (similar to what was observed in *Mihailov et al., 2020*). We achieved the two main objectives of this

project, which was first to apply a combination of dimensional and subgrouping approaches to disentangle heterogeneity in ASD, and second, to uncover stable features across samples built from different criteria. We prove that subgrouping offers refined results, while case-control comparisons provide limited information. Uncovering better-defined subtypes of autism leads to improved delineation of underlying genetic, biological and behavioral mechanisms, which is essential in the development of proper therapies.



### 4.3. Study 3

## HLA alleles: a stratifying factor for cortical signatures in Autism Spectrum Disorders

The manuscript for this study is currently in preparation to be published.

***HLA alleles: a stratifying factor for cortical signatures in Autism Spectrum Disorders***

*Angeline Mihailov, Cathy Philippe, Sigrid LeClerc, Laura Lombardi, Antoine Grigis, Ryad Tamouza, Marion Leboyer, Josselin Houenou and Vincent Frouin. In Preparation.*

### Study Presentation

Inflammation, infection and auto-immunity are linked to the development of Autism Spectrum Disorders (ASD), both in-utero and throughout life<sup>169</sup>. Specifically, it has been reported that ASD manifestation is not only linked to alterations in brain development and differences in behaviors, but also changes in immune processes including pro-inflammatory states, abnormal cell-mediated immunity, inflammatory-mediated gut dysbiosis and the presence of anti-brain autoantibodies<sup>155,409–411</sup>. Genome-wide association studies have shown that the Human Leukocyte Antigen (HLA) genes, which are highly polymorphic immune genes implicated in several sub-functions that fine-tune the immune system, are one of the most significant risk loci for several psychiatric disorders including ASD<sup>166,168</sup>. Furthermore, studies have reported that ASD is influenced by the genetic regulation of the HLA system, particularly by class I and class II genes including HLA-A, -B, -DRB1 and -DQB1<sup>173,412</sup>. In light of this information, and the

fact that HLA involvement in ASD has yet to be fully explored and understood, we decided to conduct an allele association study looking at HLA class I (A, B and C) and HLA class II (DRB1, DQA1, DQB1 and DPB1) genes in order to better understand potential allelic structure associated with ASD risk. Furthermore, should we isolate any alleles significantly associated with ASD, we will further investigate the effect of these alleles on cortical regions of interest (ROI).

A total of 531 participants from the EU-AIMS Longitudinal European Autism Project (LEAP) cohort were selected based on genetic data availability, European ancestry and having an FSIQ > 75. This included both subjects diagnosed with ASD ( $n = 300$ ) as well as (non-related) controls ( $n = 231$ ). First, we submitted participants to a 2-digit allele association analysis that compared allele frequencies between ASD patients and controls using a logistic regression model covarying for age, sex and the first four principal components to account for cryptic population stratification. Results showed a lower frequency of HLA alleles DQB1\*03 and DRB1\*11 in ASD patients versus controls. Furthermore, we observed a trend higher frequency of the HLA allele DQB1\*06 in ASD patients versus controls. Since these three alleles showed significant differences in frequencies in ASD versus controls, we decided to study their effect on the brains of patients: do they exert a specific pattern that is different from the commonly observed changes between the brains of autistic patients and controls? A total of 434 subjects (236 ASD patients and 198 controls) from the initial cohort had available good quality neuroimaging data. For the neuroimaging analysis, the presence/absence of each of the three alleles allowed the construction of two contrasts per each ROI of the Desikan FreeSurfer atlas: 1) ASD patients *with* the allele versus controls, and 2) ASD patients *without* the allele versus controls. These analyses compared cortical thickness, gyrification, surface area and subcortical structures, and were covaried for age, gender, FSIQ and site (and subcortical volumes in the

subcortical comparisons), with statistical results being Bonferroni adjusted. We found that ASD patients *without* the DQB1\*06 allele had unique decreases in surface area in the right bank of the superior temporal sulcus and the left pars triangularis. Furthermore, we found that ASD patients *with* the DRB1\*11 allele exhibited decreases in surface area in the right inferiorparietal region. Lastly, within the ASD group we also compared patients with the DQB1\*06 allele versus those without, and patients with the DRB1\*11 allele versus those without, on a battery of behavioral and cognitive assessments but found no differences.

In conclusion, we suggest that the DQB1\*06 and DRB1\*11 HLA alleles are associated with the presence of cortical anomalies in the ASD population. We further suggest that these HLA alleles may offer insight into the *stratification* of cortical morphology features, but not in behavioral features since no differences were detected, in ASD patients. Based on our results, we are convinced that heterogeneity in ASD, which is known to arise from multiple factors including immunity, is deeply rooted in the variability of the immune system itself, and furthermore, that these differences can be linked to cortical alterations.

## Materials and Methods

### *Part I: Allele Association Analysis*

#### ***EU-AIMS LEAP Cohort and Participants***

The EU-AIMS Longitudinal European Autism Project (LEAP) cohort was created to help understand autism by aiding in the stratification of biomarkers in patients. It is a multicentre European-wide initiative aiming to tackle developmental research questions based on longitudinal multidisciplinary observations in ASD patients, and disentangle heterogeneity and

establish links between biological and clinical profiles. This cohort is comprised of brain imaging, cognitive/behavioral and genetic data for 764 individuals including autistic patients and controls between the ages of 6 and 30 years old and an FSIQ between 50 and 148. Six centres contributed to the data curation: Institute of Psychiatry, King's College London, United Kingdom; Autism Research Centre, University of Cambridge, United Kingdom; Radboud University Medical Centre, Nijmegen, the Netherlands; University Medical Centre Utrecht, the Netherlands; Central Institute of Mental Health, Mannheim, Germany; and University Campus Bio-Medico, Rome, Italy. Further details on experimental design, clinical assessments and remaining data can be found in Loth et al., 2017<sup>305</sup>. Autism diagnoses were confirmed using the Diagnostic and Statistical Manual (DSM-IV or DSM-5), or the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10).

Criteria for inclusion in the present study was having available genetic data and an FSIQ > 75, which resulted in 663 participants. This study was approved by local ethical committees in each participating center, and written informed consent was provided by all participants and/or their legal guardians (for those younger than 18 years old). FSIQ was assessed in all participants using the Wechsler Abbreviated Scales of Intelligence.

### ***Genotyping, Imputation and Quality Check***

Genotyping was done using the Infinium OmniEXpress-24v1 BeadChip (<700 000 markers) from Illumina by the CNRGH (French National Centre for Human Genetics Research). Starting from non-imputed genotypes provided by Pasteur Institute, SNPs missing more than 5% of subjects, with a Minor Allele Frequency (MAF) lower than 5%, not in Hardy-Weinberg equilibrium ( $p < 10^{-6}$ ) were filtered out. Subjects with more than 5% of missing data, sex

discrepancy, non-European ancestry or cryptic relatedness were discarded. This quality check resulted in 531 participants comprising 300 individuals diagnosed with ASD and 231 typically developing individuals. In order to obtain HLA allele profiles for each subject, HLA allele at HLA-A, B, C, DPB1, DQA1, DQB1, and DRB1 were imputed using the HLA Genotype Imputation with Attribute Bagging (HIBAG) Bioconductor package in R<sup>413</sup>, with a filter on imputation probability of  $< 0.5$  discarded.

### ***Allele Association Analysis***

Once the data was prepared according to the previously described methods, it underwent an allele association analysis using the PyHLA package, which is specifically designed for association analyses of HLA gene alleles from genome-wide genotyping and next-generation sequencing data, and can also perform zygosity tests and interaction tests between HLA alleles and diseases<sup>342</sup>. A logistic regression model was run to test the association of HLA alleles typed with a 2-digit resolution while allowing for multiple covariates. Covariates included age, sex and the first four principle components that were computed to account for cryptic stratification of individuals in the cohort. Statistics underwent a Benjamini-Hochberg correction. Results were reported in terms of specific allele frequencies between ASD subjects and controls.

## **Part II: Neuroimaging Analysis**

### ***Structural MRI Acquisition and Processing***

MRI acquisition (all 3T scanners) took place at 6 different sites previously mentioned above: GE Medical Systems Discover MR 750 at King's College London, United Kingdom; Siemens Verio at the Autism Research Centre, University of Cambridge, United Kingdom; Siemens

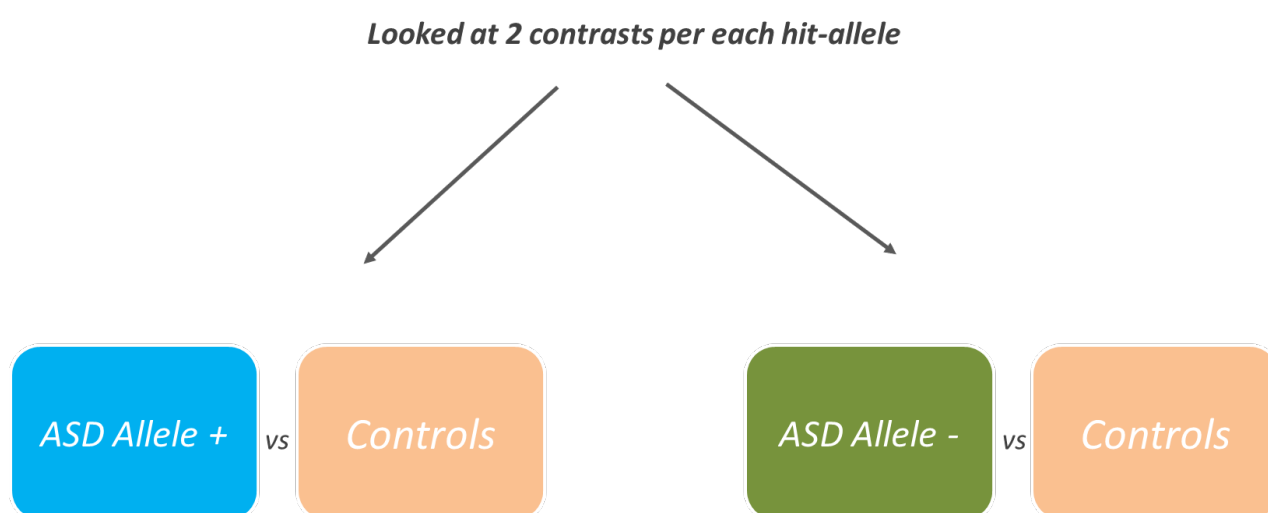
Skyra at Radboud University Medical Centre, Nijmegen, the Netherlands; Philips Medical Systems Achieva/Ingenia CX at University Medical Centre Utrecht, the Netherlands; Siemens TimTrio at Central Institute of Mental Health, Mannheim, Germany; and GE Medical Systems Sigma HDxTt at University Campus Bio-Medico, Rome, Italy.

Cortical and subcortical features extracted with the FreeSurfer software version 6.0.0 were considered in this study (<https://surfer.nmr.mgh.harvard.edu/>). The high resolution T1-weighted images were processed with recon-all. For more information on precise methods of image analysis and the construction of anatomical information for each individual done by this software, refer to<sup>327,352</sup>. Briefly, the FreeSurfer analysis stream includes intensity normalization, skull stripping, and segmentation of gray (pial) and white matter surfaces, followed by a surface tessellation resulting in a standardized triangular grey matter cortical mesh per hemisphere<sup>327</sup>. Individual brains were registered to a spherical atlas according to cortical folding patterns in order to geometrically match brains across subjects. Cortical thickness, surface area, gyrification and subcortical features are available for each ROI according to the Desikan-Killiany atlas from FreeSurfer<sup>332</sup>. This resulted in thirty-four ROIs per hemisphere for each cortical feature (thickness, surface area and gyrification), and seven ROIs per hemisphere for the subcortical features. After filtering for subjects with available ROI data, we were left with a total of 434 subjects for the neuroimaging analysis divided into 236 ASD patients and 198 controls.

### ***Statistical Analysis***

Based on the results from the allele association analysis, hit-alleles with different frequencies in ASD patients will be used to define subgroups of patients according to these alleles' status.

Specifically, if hit-allele  $X$  has a different frequency in ASD patients, then two neuroimaging contrasts will be run: ASD *with* allele  $X$  versus controls (denoted ASD/ $X^+$  vs controls), and ASD *without* allele  $X$  versus controls (denoted ASD/ $X^-$  vs controls) (Figure 4.3.1). Differences in thickness, surface area, gyrification and subcortical features were assessed in each contrast by running a linear model controlling for age, sex, FSIQ, and site (and total subcortical grey volume for the subcortical features) and looking at the affect of diagnosis. P-values were adjusted on a whole-brain level (for sixty-eight ROIs) using the Bonferroni method.



**Figure 4.3.1.** The two contrasts that will be conducted per each hit-allele from the allele association analysis. This will be run for cortical thickness, area, gyrification and subcortical features.

## Results

### *Part I: Allele Association Analysis*

We found a lower frequency of ASD patients with the DQB1\*03 allele ( $p < 0.019$ , frequency = 0.273, OR = 0.6), and again a lower frequency of ASD patients with the DRB1\*11 ( $p < 0.019$ , frequency = 0.053, OR = 0.4). Next, we observed a trend greater frequency of ASD patients with the DQB1\*06 allele ( $p < 0.055$ , frequency = 0.300, OR = 1.6). Even though after correction for multiple comparisons, the p-value just barely surpassed the nominal significance level of 0.05, we decided to still maintain this allele in the analysis since we deemed it an interesting

trend. Also of note, the DQB1\*06 allele was the only allele of the DQB1 gene whose p-value reached 0.05, even if slightly surpassed, as all other alleles had much larger p-values (Table 4.3.1).

Allele	A*_ASD	B*_ASD	A_CTRL	B_CTRL	Freq_Case	Freq_T D	OR	L95	U95	p-value: adj
DQB1*06	180	420	95	367	0.300	0.205	1.598	1.160	2.201	<b>0.055*</b>
DQB1*03	164	436	176	286	0.273	0.381	0.617	0.462	0.822	<b>0.019</b>
DRB1*11	32	586	55	407	0.053	0.119	0.431	0.263	0.704	<b>0.019</b>

**Table 4.3.1.** The subject counts, frequencies (*Freq<sub>x</sub>*), odds ratios (*OR*), confidence intervals (*L95 – U95*) and p-values of the three alleles generated from the allele association analysis. A = carrying the allele, B = not carrying the allele. \*This p-value surpassed the significance threshold after correction, however we still deemed an interesting trend and therefore kept it in the study.

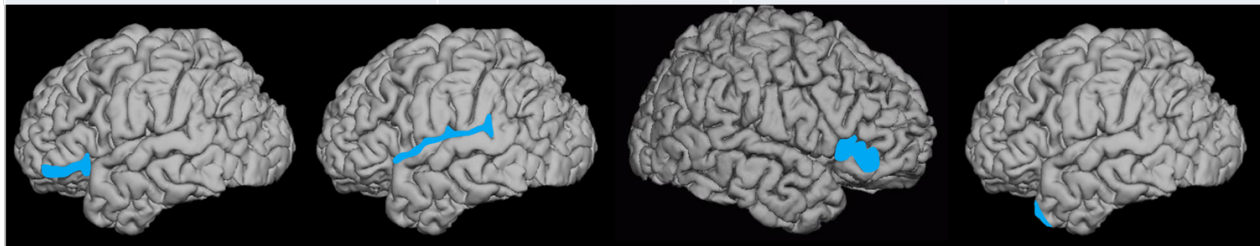
## Part II: Neuroimaging Analysis

### ***ASD versus Controls Contrast***

Before constructing the allele status contrast maps based on the presence/absence of the hit-HLA alleles (DQB1\*06, DQB1\*03 and DRB1\*11), a global comparison in all ROI features was run in ASD versus controls in order to rule out any subsequent results that may be due to a general case-control contrast. Statistically significant results comprised a decrease in surface area in the left pars orbitalis ( $p = 0.000127$ ;  $p_{adj} = 0.0086$ ), insula ( $p = 0.0000486$ ;  $p_{adj} = 0.0033$ ) and the right pars triangularis ( $p = 0.000112$ ;  $p_{adj} = 0.0076$ ) in autistic patients. We also observed a decrease in local gyrification in the left temporal pole ( $p = 0.000195$ ;  $p_{adj} = 0.013$ ) in autistic patients (Figure 4.3.2). FSIQ was significantly different between ASD patients and controls ( $p = 0.00365$ ), but not age or gender ( $p = 0.754$  and  $p = 0.143$ , respectively).



Characteristics		ASD	Controls	Adjusted p-value
Region	Feature			
Left pars orbitalis	Surface Area	743.05 ↓	766.73	0.0086
Left insula	Surface Area	2446.43 ↓	2533.92	0.0033
Right pars triangularis	Surface Area	1597.13 ↓	1672.70	0.0076
Left temporal pole	Gyrification	2.52 ↓	2.57	0.013
Demographic Information				
Age		17.51	17.69	0.754
Gender (M:F)		(168:68)	(127:71)	0.143
FSIQ		104.95	108.97	0.00365
n		236	198	



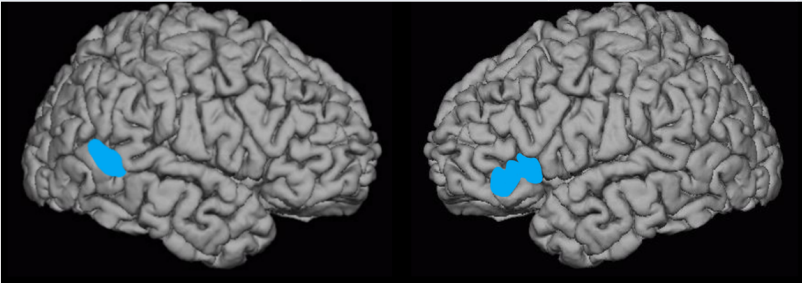
**Figure 4.3.2.** Significant cortical differences in an ASD-controls comparison. Surface area is measured in mm<sup>2</sup> and gyrification is a unitless ratio.

### *Allele Status Comparisons*

Upon looking at the effects of the alleles of interest (i.e. DQB1\*06, DQB1\*03 and DRB1\*11) on the cortical morphometry of ASD patients, we obtained results in two contrasts: ASD patients *without* the DQB1\*06 allele (ASD/DQB1\*06<sup>-</sup>) versus controls, and ASD *with* the DRB1\*11 allele (ASD/DRB1\*11<sup>+</sup>) versus controls. ASD/DQB1\*06<sup>-</sup> patients presented decreases in surface area in the right bank of the superior temporal sulcus ( $p = 0.000484$ ;  $p_{adj} = 0.032$ ) and the left pars triangularis ( $p = 0.000315$ ;  $p_{adj} = 0.021$ ). FSIQ also significantly differed between groups ( $p = 0.00836$ ), but not age or gender ( $p = 0.371$  and  $p = 0.187$ , respectively). This contrast also exhibited a decrease in surface area in the left insula ( $p = 0.000091$ ;  $p_{adj} = 0.0061$ ), the right pars triangularis ( $p = 0.000157$ ;  $p_{adj} = 0.011$ ), and a decrease in gyrification in the left temporal pole ( $p = 0.00000447$ ;  $p_{adj} = 0.00030$ ), however we omitted these results from this contrast since for this comparison we are interested in isolating results

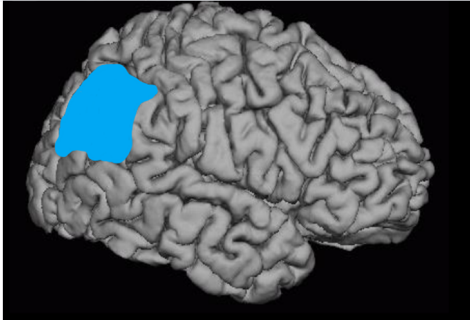
unique to this contrast (Figure 4.3.3). ASD/DQB1\*11<sup>+</sup> patients presented a decrease in surface area in the right inferior parietal region ( $p = 0.000351$ ;  $p_{adj} = 0.023$ ), and differed in age ( $p = 0.015$ ), but not FSIQ or gender ( $p = 0.097$  and  $p = 0.237$ , respectively)(Figure 4.3.4).

DQB1 Gene	Characteristics		ASD/DQB1*06 <sup>-</sup>	Controls	Adjusted p-value
	Region	Feature			
	Right bank of the superior temporal sulcus	Surface Area	939.25 ↓	987.17	0.028
	Left pars triangularis	Surface Area	1378.53 ↓	1453.83	0.026
	Demographic Information				
	Age		18.33	17.69	0.371
	Gender (M:F)		(81:31)	(126:70)	0.187
	FSIQ		104.58	109.05	0.00836
	n		112	196	



**Figure 4.3.3.** Cortical differences and demographic information for the contrast comparing ASD/DQB1\*06<sup>-</sup> versus controls. Surface area is measures in mm<sup>2</sup>.

DRB1 Gene	Characteristics		ASD/DRB1*11 <sup>+</sup>	Controls	Adjusted p-value
	Region	Feature			
	Right inferiorparietal	Surface Area	5134.36 ↓	5821.36	0.023
	Demographic Information				
	Age		20.72	17.66	0.015
	Gender (M:F)		(20:5)	(118:61)	0.237
	FSIQ		103.95	108.51	0.097
	n		25	179	



**Figure 4.3.4.** Cortical differences and demographic information for the contrast comparing ASD/DRB1\*11<sup>+</sup> versus controls. Surface area is measures in mm<sup>2</sup>.

## Comparisons in Clinical and Behavioral Assessments

Since we observed changes in the brain linked to the absence of the DQB1\*06 and the presence of the DRB1\*11 alleles, we decided to run post-hoc analyses comparing allele presence versus absence within ASD patients on clinical and behavioral assessments. This was done in order to determine if neuroanatomical changes could be explained by certain traits. We compared groups on several assessments available in the EU-AIMS cohort (Table 4.3.2) by running a linear model controlling for age, gender and FSIQ, and looking at group effects (correcting for multiple comparisons using the Bonferroni method). These contrasts generated no differences.

Level	IA/ OA	Domain/task	Time point	Schedule <sup>a</sup>			
				A	B	C	D
Clinical diagnosis							
Level 1	IA	Autism Diagnostic Interview-Revised (ADI-R) <sup>c</sup>	Base	p <sup>b</sup>	p <sup>b</sup>	p <sup>b</sup>	p <sup>b</sup>
Level 1	IA	Autism Diagnostic Observation Schedule (ADOS or ADOS-2) <sup>c</sup>	Base & FU	S <sup>b</sup>	S <sup>b</sup>	S <sup>b</sup>	S <sup>b</sup>
Dimensional measures of ASD symptoms							
Level 2	OA	Social Responsiveness Scale-2nd Edition (SRS-II) <sup>c</sup>	Base & FU	S & p <sup>b</sup>	S & P	P	P
Level 2	OA	Repetitive Behaviour Scale-Revised (RBS-R) <sup>c</sup>	Base & FU	p <sup>b</sup>	P	P	P
Level 2	OA	Short Sensory Profile (SSP) <sup>c</sup>	Base & FU	p <sup>b</sup>	P	P	P
Level 2	OA	Children's Social Behaviour Questionnaire (CSBQ) <sup>c</sup>	Base & FU	–	P	P	P
		Adults' Social Behaviour Questionnaire (ASBQ) <sup>c</sup>		S & p <sup>b</sup>	–	–	–
Level 2	OA	Autism Quotient (AQ), AQ-Adol, AQ-Child <sup>c</sup>	Base & FU	S & p <sup>b</sup>	P	P	P
Level 2	OA	Aberrant Behaviour Checklist	FU	p <sup>b</sup>	P	P	P
Level 2	OA	Adult Routines Inventory (ARI) or Childhood Routines Inventory-Revised (CIR-R)	FU	S	–	–	–
				–	P	P	P
Level 2	OA	Sensory Experiences Questionnaire—short version (SEQ 3.0)	FU	p <sup>b</sup>	P	P	P
Level 2	OA	Global Score of Change	FU	p <sup>b</sup>	p <sup>b</sup>	p <sup>b</sup>	p <sup>b</sup>
Comorbidities							
Level 2	OA	Development and Well-Being Assessment (DAWBA) <sup>c</sup>	Base	S & p <sup>b</sup>	S & P	P	P
Level 2	OA	Strengths and Difficulties Questionnaire (SDQ) <sup>c</sup>	Base & FU	S & p <sup>b</sup>	S & P	P	P
Level 2	OA	DSM-5 ADHD rating scale <sup>c</sup>	Base & FU	S & p <sup>b</sup>	P	P	P
Level 3	OA	Beck Anxiety Inventory (BAI)	Base & FU	S	S	P	P
Level 3	OA	Beck Depression Inventory (BDI)	Base & FU	S	S	P	P
Quality of life/adaptive behaviour							
Level 1	IA	Vineland Adaptive Behaviour Scale-2nd Ed (VABS-2)	Base & FU	p <sup>b</sup>	p <sup>b</sup>	P	P
Level 2	OA	Columbia Impairment Scale (CIS) <sup>c</sup>	Base & FU	S & p <sup>b</sup>	S & P	P	P
	OA	Child Health and Illness Profile (CHIP-CE) or World Health Organisation Quality of Life (WHOQOL-BREF)	Base & FU	S	–	–	–
Medical or psychiatric history							
Level 1	OA	NIH ACE Subject Medical History Questionnaire <sup>c</sup>	Base	S <sup>c</sup> or p <sup>b</sup>	P	P	P
Level 2	OA	NIH ACE Family History Form	Base	S <sup>c</sup> or p <sup>b</sup>	P	P	P
Level 2	OA	Medical Psychiatric History Perinatal Environmental Risk Questionnaire <sup>c</sup>	FU	p <sup>b</sup>	P	P	P
Level 2	OA	Brief Life Events Questionnaire, anchored in pregnancy	FU	p <sup>b</sup>	P	P	P
Level 2	OA	Children's Sleep Habits Questionnaire or Adult version (FU only)	Base & FU	– S	S & P –	P –	P –
Level 2	IA	Family Medical History Interview	FU	S <sup>c</sup> or p <sup>b</sup>	P	P	P
Cognitive and psychological profile							
Level 1	IA	WASI or WISC / WAIS (4 subtests) <sup>c</sup>	Base	S	S	S	S
Level 2	OA	HRS-MAT online adaptive IQ test	FU	S	S	S	S
Level 2	IA	Probabilistic reversal learning <sup>c</sup>	Base & FU	S	S	S	S
Level 2	IA	Spatial working memory <sup>c</sup>	Base & FU	S	S	S	S
Level 2	IA	UnSegmented block design task <sup>c</sup>	Base	S	S	S	S
Level 2	IA	Animated shapes narratives task <sup>c</sup>	Base & FU	S	S	S	S
Level 3	OA	Empathy Quotient	Base & FU	S	P	–	–
Level 3	OA	Systemising Quotient	Base & FU	S	P	–	–
Level 3	OA	Child EQ-SQ	Base & FU	–	–	P	–

**Table 4.3.2.** List of assessments available in the EU-AIMS cohort.

## Discussion

The association between HLA genes and autistic development, specifically their relationship to cortical features, is a field of study in need of further exploration. In particular, we suggest that the HLA region offers insight into ASD pathology since patients characteristically express immunological dysfunction<sup>166</sup>. To the best of our knowledge, this is the first study investigating the effect of relevant HLA alleles on cortical morphology in an autistic cohort. In particular, we found a decreased frequency of HLA alleles DQB1\*03 and DRB1\*11, and an increased frequency of HLA allele DQB1\*06, in autistic patients compared to controls. Since these alleles were different in ASD patients, this further inspired us to examine their effect on brain morphology in patients compared to controls. Indeed, we found a decrease in surface area in the right bank of the superior temporal sulcus (rSTS) and left pars triangularis (IPT) in autistic patients *without* the DQB1\*06 allele (ASD/DQB1\*06<sup>-</sup>), which are regions involved in several social cognition processes<sup>180,187</sup>. We also observed a decrease in surface area in the right inferiorparietal (rINP) in autistic patients *with* the DRB1\*11 allele (ASD/DRB1\*11<sup>+</sup>), which is a region heavily involved in emotional and sensory perception, and motor deficits<sup>190,385,414</sup>. Our study was motivated by the idea of further fractionating autistic patients according to HLA allele status in order to isolate refined subgroups presenting unique and explanatory biological traits. By doing this, we expect to reduce some of the heterogeneity present in autistic populations and to eventually better understand existent pathophysiology as we expect these traits to reflect underlying biological mechanisms.

HLA class II genes produce transmembrane glycoproteins that function in presenting peptides to CD4<sup>+</sup> lymphocytes<sup>415</sup>. They are highly polymorphic with diverse structural variations that are the foundation of the antigenic specificity<sup>416</sup>. Furthermore, they play a crucial role in autoimmune diseases, such as multiple sclerosis and type 1 diabetes<sup>417,418</sup>. The HLA gene

cluster is proven to be one of the most significant genetic risk loci for several psychiatric disorders, including ASD<sup>168</sup>. Therefore, it is not surprising that we observed significant differences in HLA alleles in an autistic population, namely decreased frequencies of the DQB1\*03 and DRB1\*11 alleles, and an increased frequency of the DQB1\*06 allele. In general, HLA alleles in autistic populations have been studied but highly variable, with most reported differences in the HLA-A, -DRB1, and -DQB1 alleles<sup>173,174,412,419,420</sup>. In particular, the DQB1\*06 HLA allele has reported variable frequencies in autistic populations, and is heavily linked to multiple sclerosis (MS) and narcolepsy<sup>173,412,421–423</sup>. Multiple studies have discussed the very common occurrence of sleep disorders in autistic patients, and though narcolepsy is not thoroughly studied in patients, the fact that the DQB1\*06 allele is linked to a sleep disorder could perhaps shed light on sleep issues in ASD patients<sup>86,237</sup>. *Posar et al., 2020* even specifically advocated researching the link between narcolepsy and autism, suggesting possible common mechanisms between the two disorders<sup>424</sup>. This allele has also been reported in reviews suggesting a possible common mechanism between microbiome immunity and the central nervous system in ASD and MS<sup>425–427</sup>. The DQB1\*03 HLA allele typically exhibits lower frequencies in ASD, and has also been linked to multiple sclerosis as well as an increased risk of neuropathic pain<sup>173,174,428,429</sup>. Lastly, reports have shown higher and lower frequencies of the DRB1\*11 HLA allele in autistic populations, as well as links to thyroid malfunctions<sup>173,174,412,430</sup>. Though HLA allele results remain heterogeneous, warranting further study into the delineation of consistent associations, outcomes associated with these three alleles can offer insight into autistic development.

In the neuroimaging analysis, we first conducted a case-control comparison in order to isolate cortical differences attributed to an ASD diagnosis alone. We found that ASD patients exhibited a decrease in surface area in the left insula and pars orbitalis, the right pars triangularis, and a

decrease in gyrification in the left temporal pole. The insula and temporal pole are typically involved in social cognition, empathy and awareness, and are thus pertinent brain features in the study of ASD<sup>190,385,431</sup>. The pars triangularis and par orbitalis are also heavily involved in social cognition and communication and are again important regions to understand in the context of ASD development. Though these results are in line with the literature, the literature itself is still quite variable. These results are well-reported but still not consistently seen in case-control comparisons, therefore subgrouping according to various constructs is a necessary feat in the understanding of autism. Upon isolating autistic groups based on allele status, we found neuroanatomical alterations in two contrasts: ASD/DQB1\*06<sup>-</sup> patients versus controls, and ASD/DRB1\*11<sup>+</sup> patients versus controls. Upon comparing ASD/DQB1\*06<sup>-</sup> against controls, in addition to observing decreases in surface area in the rSTS and IPT, we report decreased surface area in the left insula and right pars triangularis, and a decrease in gyrification in the left temporal pole. The last three alterations were also observed in the case-control comparison and may perhaps be attributed to an overall diagnosis effect. Furthermore, the rSTS may be the only truly unique result in the ASD/DQB1\*06<sup>-</sup> versus controls contrast since the IPT is close to the left pars orbitalis (which is another region elicited in the ASD-controls contrast) on the Desikan atlas. The rSTS is an extremely important region that is often altered in autistic patients, likely due to its functions in social cognition, thus, linking this area to the DQB1\*06 allele can be an interesting subtype in need of further exploration<sup>190,385</sup>. Next, upon comparing ASD/DRB1\*11<sup>+</sup> to controls, we observed a decreased surface area in the rINP, which is involved in motor performance and therefore often dysfunctional in autistic patients, even from a young age<sup>374–376</sup>. Most interestingly however, was that this contrast showed no additional differences linked to the global case-control comparison, suggesting a truly independent subgroup involving a unique, but equally pertinent, region that is altered in ASD patients. Collectively, the regions affected by these two alleles are all involved in the mirror neuron

system (MNS), namely the right parstriangularis (as seen in case versus controls, and ASD/DQB1\*06<sup>-</sup> versus controls contrasts), the rSTS (as seen in the ASD/DQB1\*06<sup>-</sup> versus controls) and the rINP (as seen in the ASD/DRB1\*11<sup>+</sup> versus controls contrasts)<sup>249,254</sup>. The MNS is important in autistic studies since it is a vital network implicated in imitation and is thus applicable in social cognition in the understanding of others' intentions and emotions<sup>248</sup>. It is an important framework of study in autism research with vast structural and functional neuroimaging evidence showing dysfunctions of the MNS in ASD patients. The literature has also shown decreased activity in the MNS during social tasks as well as correlations between symptom severity and MNS functionality in patients, thus further supporting the hypothesis that MNS dysfunction is a core morphometric deficit in autistic patients<sup>248</sup>. All of these results are in line with functions and differences reported in ASD, despite the usual variability throughout the literature. We continue to show here that by dividing autistic patients into subgroups, we isolate the origin of reported variabilities and offer an explanation on why such results are encountered randomly (or by chance) across studies conducting case-control comparisons.

Our results suggest that variations on the cortex surface in areas associated with autism, but reported sporadically, may be related to whether or not a patient carries certain HLA alleles. Specifically, our results show that the cortical morphology of the frontal and temporal regions in ASD could vary according to whether or not they are carriers of the DQB1\*06 allele. The same principle applies for the DRB1\*11 allele, which could cause the inferiorparietal structure to vary whether or not a patient is a carrier. We suggest that the DQB1\*06 allele impedes the development of neuroanatomical alterations in ASD patients since the presence of this allele eliminates cortical changes typically reported in ASD patients, while its absence restores these changes. Indeed, it may seem unintuitive that alleles having a higher frequency in autistic patients could be associated with the impediment of cortical abnormalities, but this simply

revisits the dilemma of ASD heterogeneity and that perhaps not all disease risks run in parallel. Moreover, this enforces the necessity to subgroup in order to explain widespread differences in the cortical morphology of autistic patients. Perhaps we are witnessing a subgroup expressing cortical ASD signatures that are dependent on the absence of this allele. We further suggest that the DRB1\*11 allele presents a risk factor in cortical alterations in autistic patients since those with this allele exhibit structural differences in another well-known implicated region, the inferiorparietal area. Moreover, a link to dimensional quantities that are known to describe ASD was sought. We ran a post-hoc comparison between ASD patients with and without the DQB1\*06 allele, and ASD patients with and without the DRB1\*11 allele on an array of behavioral and clinical assessments, but observed no differences. This further supports the idea that DQB1\*06 the DRB1\*11 HLA alleles affect the expression of cortical morphology features but not behavioral and clinical presentations, thus explaining part of the vast variability observed in the cortical structure of autistic patients.

In conclusion, this is the first study to the best of our knowledge that aims to delineate biological traits in autistic patients based on HLA allele status. According to our results, we are convinced that heterogeneity in ASD extends into the immune function of patients and that this immune function is linked to cortical alterations. Subgrouping based on DQB1\*06 and DRB1\*11 HLA alleles enforces the idea that these alleles are implicated in neuroanatomical mechanisms in autistic patients, but offer no explanations on their behavioral features. This proposes that disease risk varies according to dimensional constructs. Studies often assume that structural changes are directly associated to behavioral differences, but perhaps the current investigation proves that this may not always be the case, again partially explaining the vast variability in the literature. Further studies with imaging-genetics are necessary in order to shed light on the causal direction of the link between cortical morphology differences and HLA variants.



# Conclusion

The work conducted in this thesis confirms the widespread heterogeneity reported in autism and further illustrates the necessity to apply dimensional and subgrouping approaches to delineate important biological subtypes, as suggested by the Research Domain Criteria (RDoC). RDoC has recently gained interest as an approach to solve the issue of characterization by offering a framework to investigate psychiatric disorders that proposes the integration of several levels of information (including genetic, clinical assessments, biological, etc.), in order to explore dimensions spanning from normal to abnormal human behavior. This is particularly relevant in the study of ASD, which is both a highly variable and a highly continuous disorder (i.e. ASD symptoms extend into the general population, with no clear gap between diagnosed and undiagnosed individuals). We showed that running case-control comparisons generated no, or limited, results and that stratifying ASD populations offers a plausible route in the understanding of etiology and underlying biological mechanisms. Therefore, though this thesis work is far from being comprehensive, the overarching principle was to extract behavioral, genetic and cortical biomarkers of autism to offer first stepping-stones into the isolation of proper subtypes.

The first study conducted in this thesis '*Cortical signatures in behaviorally clustered autistic traits subgroups: a population-based study*', succeeded in isolating meaningful autistic traits subgroups presenting unique behaviors. Furthermore, we showed that each of these subgroups have unique brain signatures. However most importantly, upon combining all autistic traits subgroups into one and running a case-control comparison, no neuroanatomical differences were revealed. This is a classic example of how variability associated with ASD traits can lead to negative results. We subsequently replicated the methodology of this first project in a separate cohort in a study entitled '*In search of consistent behavioral and neuroimaging biomarkers in three samples with different autistic criteria*'. Here we were able to generate one

very important behavioral subgroup that was also reported in the initial study: one with high autistic and ADHD-like traits. Moreover, in both studies, this ADHD-like subgroup exhibited alterations in the precentral region. The only difference was that the hemisphere and direction of the surface area alterations were opposite, specifically, an increase versus a decrease in surface area, and the left versus the right hemisphere. Indeed these differences in direction and hemisphere offer limitations to our claims, though we can still conclude that the precentral region is linked to a reproducible subgroup of autistic traits that exhibit ADHD-like traits. It could be considered a drawback that only one subgroup was reproduced (in the strict sense, as indeed there were other *similar* subgroups), however we do not expect to solve the issue of heterogeneity using only two cohorts with relatively limited subjects. The fact that we observed one subtype with the same altered cortical region in both cohorts can already indicate a potentially stable finding. Moreover, the reproducibility of this subgroup may initiate a starting point in the development of therapies, specifically for the autistic subgroup exhibiting ADHD-like behavior. This is of interest since a large proportion, up to 55%, of autistic patients have exhibited ADHD-like traits. Lastly, the third study conducted in this thesis work entitled '*HLA alleles: a stratifying factor for cortical signatures in Autism Spectrum Disorders*', showed that certain HLA alleles are associated with cortical patterns in the brains of autistic patients. In this third study, we first conducted a case-control comparison on several ROIs throughout the brain and found differences in frontal and temporal regions, which are often affected in autistic patients. Upon stratifying ASD patients according to HLA allele status, we found a subgroup exhibiting the same results seen in the case-control comparison, in addition to a few unique cortical alterations. Alternatively, we reported another ASD HLA allele subgroup exhibiting unique cortical changes without any of the neuroanatomical results observed in the case-control comparison. In addition to offering insight into the immune function of autistic patients, this study also illustrates the classic case of observing a cortical difference in case-control

comparisons that can be attributed to a subtype of autism, versus being endemic to all autistic subjects (since we found case-control neuroimaging results in one HLA allele subtype, but not the other). In summary, these studies collectively enforce the necessity to subtype using dimensional features in order to isolate the origin of several autistic biological features that are *sporadically* reported in the literature.

### *Limitations*

The studies in this thesis were conducted on older participants (with an average age between 11 and 12 years old in both cohorts), which poses a limitation since autism typically manifests at a much younger age (identifiable as early as 12 to 18 months of age). Furthermore, since autism is a neurodevelopmental disorder, early intervention is crucial in its outcome. Therefore, though identifying later-age stable subgroups is still beneficial in the comprehension of the disorder, it is vital to understand the biological mechanisms as early as possible in order to track development and outcome. The isolation of early biomarkers is important since many biological traits have variable developmental trajectories. In addition to this, and again due to the fact that autism is a neurodevelopmental disorder, a longitudinal study design would also be crucial to study. Perhaps with further advancements in big data, several time points will eventually be acquired offering the ability to conduct a sound longitudinal analysis in autism. Also, considering that we had access to so many different datatypes, these studies could have benefitted from integrative clustering analyses that included all features into one model. A final major limitation within the discussed studies was the lack of cortical investigations outside of structural MRI. Other attributes including functional and connectivity neuroimaging features would also provide alternative angles in the understanding of the autistic brain and its underlying mechanisms. This factor will however also improve with time as big data initiatives expand their cohort and data collection.

## *Perspectives*

An Autism Spectrum Disorders diagnosis is currently based solely on behavioral presentations according to the DSM-5, which does not take into account the vast variability in biological manifestations, nor the blurred boundary between diagnosed and non-diagnosed individuals. This practice may contribute to the widespread negative results, necessitating the need for new approaches. It is for this reason that we promote the RDoC ideology as a framework to follow in the investigation of ASD as it supports the dimensional integration of several types of data including neuroimaging, genetic, molecular, cellular and clinical. Furthermore, the current rise in available big data cohorts creates the perfect infrastructure to apply dimensional and subgrouping strategies on multimodal datasets with a large amount of subjects. Thus, we advocate for the continued multivariate exploration of ASD in order to overcome the overwhelming lack of consistent biological results.

## *Concluding Remarks*

This thesis work provides evidence of autistic and autistic-like subtypes, and confirms the necessity of applying different methodological approaches in order to extract meaningful features. We illustrate the importance of studying autism outside the realm of its diagnostic status by incorporating several levels of information, and hope these approaches can eventually reduce heterogeneity within the disorder and pave the way to better understanding ASD etiology and developmental pathways.

## Résumé du Projet en Français

Les troubles du spectre autistique (TSA) se caractérisent par des déficits sociaux et de communication ainsi que par des comportements restreints et répétitifs. Le taux de prévalence actuellement admis est de 1 enfant sur 100, avec un rapport garçons/filles très déséquilibré de 4:1<sup>2-4</sup>. En 1943, le pédopsychiatre et médecin Leo Kanner a fourni la toute première description de l'autisme, initialement appelé « *troubles autistiques du contact affectif* »<sup>5</sup>. Pendant une longue période suivant ces premiers rapports, l'autisme a été considéré comme une forme précoce de schizophrénie. Cependant, malgré plusieurs similitudes entre les deux troubles, l'autisme a finalement été classé comme un trouble psychiatrique à part entière, avec une composante génétique potentiellement forte. Par rapport à sa classification initiale, l'augmentation constante de l'incidence des TSA pourrait être liée à des effets environnementaux<sup>6</sup>, toutefois, des explications plus probables incluent des changements dans les informations et les critères de diagnostic, une sensibilisation et une reconnaissance accrues, des méthodes de diagnostic améliorées et un âge moyen général plus jeune pour le diagnostic<sup>7-9</sup>.

### **Le dilemme de l'hétérogénéité**

L'hétérogénéité des TSA constitue un défi important dans la compréhension de ces troubles, puisque jusqu'à 70 % des patients déclarent souffrir également d'autres troubles psychiatriques, médicaux ou génétiques, ainsi qu'une variabilité importante en matière de neuro-imagerie, de facteurs génétiques et immunitaires, ce qui complique considérablement la compréhension de l'étiologie des TSA<sup>10-22</sup>. Cette large hétérogénéité a entraîné un manque de reproductibilité dans l'identification des biomarqueurs, probablement en raison des

méthodologies expérimentales cas-témoins largement utilisées qui combinent dans un même groupe, des patients atteints de TSA présentant des profils comportementaux, génétiques et/ou cliniques différents. Bien que le diagnostic dépende des manifestations comportementales (à savoir les déficits sociaux et de communication et les comportements répétitifs), le niveau élevé de variation observé dans plusieurs autres facteurs biologiques et cliniques a grandement contribué à la complexité de la symptomatologie et de l'étiologie des TSA. Ces facteurs ont collectivement rendu difficile le développement de thérapies appropriées.

### **Solutions pour résoudre le dilemme de l'hétérogénéité**

Le fait de cibler uniquement les patients présentant un diagnostic confirmé pour qu'ils participent à des études exclut ceux qui représentent les différentes extrémités du spectre fonctionnel, ce qui, en fin de compte, entrave la recherche sur l'autisme. De telles catégories de diagnostic peuvent séparer des individus dont le comportement est en fait proche, comme c'est le cas pour un patient juste en-deçà du seuil et un autre juste au-delà. Ces deux patients peuvent en fait partager des symptômes et des mécanismes physiologiques sous-jacents qui sont plus similaires que deux patients dans les limites du diagnostic. Ces pratiques peuvent donc brouiller les informations sur le développement progressif de la psychopathologie et sur l'étude des facteurs de risque prodromiques<sup>23</sup>. Une façon de démêler la variabilité dans les études sur les TSA consiste à utiliser des approches dimensionnelles, qui se concentrent sur le type et le degré de plusieurs symptômes. Cette méthode serait supérieure aux approches qui comparent des individus séparés en deux groupes suivants qu'ils sont TSA ou contrôle. En effet les évaluations diagnostiques sont trop imparfaites actuellement pour rendre compte de la caractérisation biologique des patients et rendent caduques les approches cas/contrôle. En promouvant les approches dimensionnelles, les critères de domaine de recherche (RDoC) ont récemment gagné en intérêt comme approche pour résoudre ce problème de caractérisation.

Le RDoC est un cadre d'étude des troubles psychiatriques qui propose l'intégration de plusieurs niveaux d'information (y compris la génétique, les évaluations cliniques, la biologie, etc.) pour explorer des dimensions allant du comportement humain normal à l'anormal.

Cependant, l'application d'approches dimensionnelles ne résout pas entièrement le problème des résultats incohérents dans l'autisme. Étant donné que l'hétérogénéité étiologique, développementale et biologique est néanmoins présente, la compréhension et la distinction de ce trouble constituent toujours un énorme défi. Il est donc nécessaire de caractériser davantage des sous-groupes distincts de TSA dans le cadre d'un paradigme dimensionnel. À cette fin, la combinaison des approches dimensionnelles avec des stratégies de sous-groupage s'avère la plus pertinente pour résoudre le problème des résultats incohérents dans la littérature sur l'autisme.

### **Objectifs de la thèse**

L'objectif global de ce travail est de mieux caractériser les patients autistes, ce qui est crucial pour l'avancement des traitements et thérapies appropriés. Notre proposition est double : pour réussir à traiter et à démêler la variabilité présente chez les patients autistes, il est crucial de 1) appliquer un paradigme dimensionnel pour reconnaître le trouble à un niveau continu et d'inclure plusieurs types différents de données ; et 2) de sous-grouper les patients en fonction des similitudes intragroupes au sein de ce paradigme dimensionnel. En appliquant ces méthodes à la recherche sur les TSA, nous pouvons améliorer considérablement la compréhension de ces troubles au niveau biologique afin de mieux traiter les patients autistes, que ce soit à titre préventif ou après l'apparition des troubles.



**Résumé de l'étude 1 : *signatures corticales dans les sous-groupes de traits autistiques regroupés en fonction du comportement : une étude basée sur la population.***

Nous avons appliqué une approche dimensionnelle, telle que suggérée par l'initiative RDoC, pour mener une étude effectuant une analyse de regroupement sur des sujets présentant des traits autistiques différents. Comme indiqué précédemment, nous avons également émis l'hypothèse que l'isolement de sous-types homogénéisés pourrait permettre de comprendre les mécanismes biologiques sous-jacents et de mieux comprendre les TSA. Plus précisément, cette étude visait à délimiter les différences comportementales entre les sujets atteints de TSA et à observer ensuite les signatures corticales uniques des sous-groupes identifiés.

Nous avons utilisé une cohorte composée de personnes à risque de développer des troubles psychiatriques, ainsi que de participants ayant un développement typique. Les traits autistiques fondamentaux, tels que mesurés par le SRS (Social Responsiveness Scale) sous la forme de déficits sociaux, sont souvent concomitants avec des traits comportementaux dont le type et le degré varient (ce qui indique fréquemment la présence d'un trouble psychiatrique comorbide), ce qui complique encore la caractérisation des TSA. À cette fin, tous les participants ont été soumis à un partitionnement non supervisé sur les dimensions comportementales, qui a finalement mis en évidence trois sous-groupes distincts. D'autres analyses corticales ont ensuite été menées pour déterminer si ces sous-groupes distincts étaient liés à des signatures cérébrales structurelles uniques. Tout d'abord, nous avons combiné les trois sous-groupes en un seul groupe hétérogène « de type TSA » et l'avons comparé aux témoins, ce qui n'a révélé aucune différence significative. Cependant, en comparant chaque sous-groupe aux témoins, des signatures corticales uniques ont été observées dans des régions du cerveau souvent signalées comme étant altérées chez les patients autistes. Cette étude a renforcé la nécessité de sous-grouper afin d'isoler les

différences biologiques en montrant que le sous-groupage dans un paradigme dimensionnel est nécessaire pour découvrir les signatures corticales associées aux traits autistiques.

**Résumé de l'étude 2 : à la recherche de biomarqueurs comportementaux et de neuro-imagerie cohérents dans trois populations se chevauchant et présentant des critères autistiques différents**

Dans le prolongement du plan expérimental de l'étude précédente, nous avons cherché à reproduire une partie des méthodes en utilisant une autre cohorte. L'objectif de cette seconde étude était double. Tout d'abord, nous avons effectué un regroupement indépendant dans chacune des trois populations se chevauchant afin d'observer si des sous-groupes comportementaux étaient similaires entre les populations. Ensuite, nous avons étudié la neuro-imagerie et les scores polygéniques (PGS) des sous-groupes pour voir si nous observions des traits similaires parmi les populations.

En effet, nous avons trouvé des sous-groupes comportementaux homogènes dans les trois populations, et de plus, ces sous-groupes ont montré des résultats similaires dans les analyses corticales ultérieures et les scores de risque polygénique. De plus, comme dans l'étude précédente, en combinant tous les sous-groupes de type autistique en un seul par population pour effectuer une comparaison cas-témoins, les résultats étaient extrêmement limités. Cette investigation a également illustré la nécessité de faire des sous-groupes afin d'isoler les différences biologiques associées aux traits autistiques. Nous avons également été en mesure de générer des sous-groupes homogènes avec des traits biologiques et génétiques cohérents, ce qui plaide substantiellement en faveur de l'importance d'isoler des sous-groupes dans la recherche de phénotypes et de génotypes fiables.

### **Résumé de l'étude 3 : *le statut des allèles HLA est associé à des altérations structurelles dans le cerveau des patients autistes***

La coexistence de l'inflammation, de l'infection et de l'auto-immunité dans les TSA indique que des marqueurs de susceptibilité potentiellement essentiels peuvent être voisins ou situés dans le superlocus hautement polymorphe de l'antigène leucocytaire humain (HLA)<sup>25</sup>. Cette troisième étude consiste en une analyse d'association d'allèles sur les allèles HLA, testant la différence de leurs fréquences entre les patients autistes et les témoins. Tous les allèles dont la fréquence diffère entre les deux groupes ont été utilisés comme marqueurs de stratification au sein de la population autiste, dans l'idée de fractionner davantage les patients autistes pour isoler des sous-groupes affinis. Plus précisément, nous avons examiné si la présence ou l'absence de ces allèles exerçait un effet sur la morphologie du cerveau des patients autistes en comparant les groupes d'autistes avec et sans l'allèle d'intérêt à des témoins. En effet, nous avons trouvé plusieurs caractéristiques intéressantes associées au statut de l'allèle sur la surface corticale, ce qui prouve que les marqueurs immunitaires peuvent offrir un aperçu de la physiopathologie des TSA. Sur la base des résultats que nous avons observés, nous sommes convaincus que l'hétérogénéité des TSA s'étend à la fonction immunitaire des patients et que cette fonction immunitaire est liée aux altérations corticales.

### **Conclusion**

Ce travail de thèse montre qu'une approche dimensionnelle combinée à l'extraction de sous-groupes permet une meilleure reproductibilité des analyses de neuro-imagerie anatomique. Ce manuscrit illustre l'importance d'étudier l'autisme en dehors de son statut diagnostique en

incorporant plusieurs niveaux d'information, notamment des données comportementales, de neuro-imagerie, cliniques, génétiques et immunologiques. Nous espérons que ces approches permettront finalement de réduire l'hétérogénéité de la maladie et d'ouvrir la voie à une meilleure compréhension des étiologies des TSA et des voies de développement, ce qui conduira à la mise au point de thérapies et d'interventions améliorées.

## Bibliography

1. KANNER, L. Autistic disturbances of affective contact. *Nerv. Child* **2**, 217–250 (1943).
2. Asperger, H. *‘Autistic psychopathy’ in childhood*. 92 (Cambridge University Press, 1991). doi:10.1017/CBO9780511526770.002.
3. Wing, L. Asperger’s syndrome: a clinical account. *Psychol. Med.* **11**, 115–129 (1981).
4. Spitzer, R. L., Md, K. K. & Williams, J. B. W. Diagnostic and Statistical Manual of Mental Disorders, Third Edition. *Am. Psychiatr. Assoc.* (1980).
5. King, B. H., Navot, N., Bernier, R. & Webb, S. J. Update on diagnostic classification in autism. *Curr. Opin. Psychiatry* **27**, 105–109 (2014).
6. DSM-IV: Diagnostic and Statistical Manual of Mental Disorders | JAMA | JAMA Network. <https://jamanetwork.com/journals/jama/article-abstract/379036>.
7. Association, A. P. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. (American Psychiatric Pub, 2013).
8. Zwaigenbaum, L. *et al.* Early Intervention for Children With Autism Spectrum Disorder Under 3 Years of Age: Recommendations for Practice and Research. *Pediatrics* **136**, S60–S81 (2015).
9. Lord, C., Rutter, M. & Le Couteur, A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* **24**, 659–685 (1994).
10. Lord, C. *et al.* The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *J. Autism Dev. Disord.* **30**, 205–223 (2000).
11. Wing, L., Leekam, S. R., Libby, S. J., Gould, J. & Larcombe, M. The Diagnostic Interview for Social and Communication Disorders: background, inter-rater reliability and clinical use. *J. Child Psychol. Psychiatry* **43**, 307–325 (2002).
12. Dawkins, T., Meyer, A. T. & Van Bourgondien, M. E. The Relationship Between the Childhood Autism Rating Scale: Second Edition and Clinical Diagnosis Utilizing the DSM-IV-TR and the DSM-5. *J. Autism Dev. Disord.* **46**, 3361–3368 (2016).
13. Skuse, D. *et al.* The Developmental, Dimensional and Diagnostic Interview (3di): A Novel Computerized Assessment for Autism Spectrum Disorders. *J. Am. Acad. Child Adolesc. Psychiatry* **43**, 548–558 (2004).
14. Rutter, M. Incidence of autism spectrum disorders: Changes over time and their meaning\*. *Acta Paediatr.* **94**, 2–15 (2005).
15. Lotter, V. Epidemiology of autistic conditions in young children. *Soc. Psychiatry* **1**, 124–135 (1966).
16. Fombonne, E. Editorial: The rising prevalence of autism. *J. Child Psychol. Psychiatry* **59**, 717–720 (2018).
17. Mattila, M.-L. *et al.* Autism spectrum disorders according to DSM-IV-TR and comparison with DSM-5 draft criteria: an epidemiological study. *J. Am. Acad. Child Adolesc. Psychiatry* **50**, 583–592.e11 (2011).

- 18.Saemundsen, E., Magnússon, P., Georgsdóttir, I., Egilsson, E. & Rafnsson, V. Prevalence of autism spectrum disorders in an Icelandic birth cohort. *BMJ Open* **3**, e002748 (2013).
- 19.Karimi, P., Kamali, E., Mousavi, S. M. & Karahmadi, M. Environmental factors influencing the risk of autism. *J. Res. Med. Sci. Off. J. Isfahan Univ. Med. Sci.* **22**, (2017).
- 20.Elsabbagh, M. *et al.* Global Prevalence of Autism and Other Pervasive Developmental Disorders. *Autism Res.* **5**, 160–179 (2012).
- 21.Fombonne, E. Epidemiology of pervasive developmental disorders. *Pediatr. Res.* **65**, 591–598 (2009).
- 22.Baio, J. *et al.* Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *Morb. Mortal. Wkly. Rep. Surveill. Summ. Wash. DC 2002* **67**, 1–23 (2018).
- 23.Fombonne, E. Epidemiology of autistic disorder and other pervasive developmental disorders. *J. Clin. Psychiatry* **66 Suppl 10**, 3–8 (2005).
- 24.Christensen, D. L. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *MMWR Surveill. Summ.* **65**, (2016).
- 25.Idring, S. *et al.* Changes in prevalence of autism spectrum disorders in 2001–2011: findings from the Stockholm youth cohort. *J. Autism Dev. Disord.* **45**, 1766–1773 (2015).
- 26.Frazier, T. W., Georgiades, S., Bishop, S. L. & Hardan, A. Y. Behavioral and Cognitive Characteristics of Females and Males With Autism in the Simons Simplex Collection. *J. Am. Acad. Child Adolesc. Psychiatry* **53**, 329–340.e3 (2014).
- 27.Nag, H. E., Nordgren, A., Anderlid, B.-M. & Nærlund, T. Reversed gender ratio of autism spectrum disorder in Smith-Magenis syndrome. *Mol. Autism* **9**, 1 (2018).
- 28.Ozonoff, S. *et al.* A prospective study of the emergence of early behavioral signs of autism. *J. Am. Acad. Child Adolesc. Psychiatry* **49**, 256–266.e1–2 (2010).
- 29.Baron-Cohen, S. *et al.* Why Are Autism Spectrum Conditions More Prevalent in Males? *PLOS Biol.* **9**, e1001081 (2011).
- 30.McClure, I. The Essential Difference: Men, Women and the Extreme Male Brain. *BMJ* **327**, 57 (2003).
- 31.Baron-Cohen, S. The extreme male brain theory of autism. *Trends Cogn. Sci.* **6**, 248–254 (2002).
- 32.Luders, E., Toga, A. W. & Thompson, P. M. Why size matters: Differences in brain volume account for apparent sex differences in callosal anatomy: The sexual dimorphism of the corpus callosum. *NeuroImage* **84**, 820–824 (2014).
- 33.Barnea-Goraly, N. *et al.* A preliminary longitudinal volumetric MRI study of amygdala and hippocampal volumes in autism. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **48**, 124–128 (2014).
- 34.Lau, Y. C. *et al.* Autism traits in individuals with agenesis of the corpus callosum. *J. Autism Dev. Disord.* **43**, 1106–1118 (2013).
- 35.Auyeung, B. *et al.* Prenatal versus postnatal sex steroid hormone effects on autistic traits in children at 18 to 24 months of age. *Mol. Autism* **3**, 17 (2012).

36. Gidaya, N. B. *et al.* In Utero Exposure to Selective Serotonin Reuptake Inhibitors and Risk for Autism Spectrum Disorder. *J. Autism Dev. Disord.* **44**, 2558–2567 (2014).
37. Wood, A. G. *et al.* Prospective assessment of autism traits in children exposed to antiepileptic drugs during pregnancy. *Epilepsia* **56**, 1047–1055 (2015).
38. Dodds, L. *et al.* The Role of Prenatal, Obstetric and Neonatal Factors in the Development of Autism. *J. Autism Dev. Disord.* **41**, 891–902 (2011).
39. Sørensen, M. J. *et al.* Antidepressant exposure in pregnancy and risk of autism spectrum disorders. *Clinical Epidemiology* <https://www.dovepress.com/antidepressant-exposure-in-pregnancy-and-risk-of-autism-spectrum-disorder-peer-reviewed-article-CLEP> (2013).
40. Gardner, R. M. *et al.* Maternal body mass index during early pregnancy, gestational weight gain, and risk of autism spectrum disorders: Results from a Swedish total population and discordant sibling study. *Int. J. Epidemiol.* **44**, 870–883 (2015).
41. Guinchat, V. *et al.* Pre-, peri- and neonatal risk factors for autism. *Acta Obstet. Gynecol. Scand.* **91**, 287–300 (2012).
42. Croen, L. A., Grether, J. K., Yoshida, C. K., Odouli, R. & Van de Water, J. Maternal Autoimmune Diseases, Asthma and Allergies, and Childhood Autism Spectrum Disorders: A Case-control Study. *Arch. Pediatr. Adolesc. Med.* **159**, (2005).
43. Atladóttir, H. Ó. *et al.* Association of Family History of Autoimmune Diseases and Autism Spectrum Disorders. *Pediatrics* **124**, 687–694 (2009).
44. Allen, J. L. *et al.* Developmental neurotoxicity of inhaled ambient ultrafine particle air pollution: Parallels with neuropathological and behavioral features of autism and other neurodevelopmental disorders. *NeuroToxicology* **59**, 140–154 (2017).
45. Oudin, A. *et al.* Prenatal exposure to air pollution as a potential risk factor for autism and ADHD. *Environ. Int.* **133**, 105149 (2019).
46. Meyer, U., Feldon, J., Schedlowski, M. & Yee, B. K. Towards an immuno-precipitated neurodevelopmental animal model of schizophrenia. *Neurosci. Biobehav. Rev.* **29**, 913–947 (2005).
47. Gilmore, J. H., Jarskog, L. F. & Vadlamudi, S. Maternal poly I:C exposure during pregnancy regulates TNF alpha, BDNF, and NGF expression in neonatal brain and the maternal-fetal unit of the rat. *J. Neuroimmunol.* **159**, 106–112 (2005).
48. Meltzer, A. The Role of the Immune System in Autism Spectrum Disorder. 15.
49. Shi, L., Fatemi, S. H., Sidwell, R. W. & Patterson, P. H. Maternal Influenza Infection Causes Marked Behavioral and Pharmacological Changes in the Offspring. *J. Neurosci.* **23**, 297–302 (2003).
50. Gardener, H., Spiegelman, D. & Buka, S. L. Perinatal and Neonatal Risk Factors for Autism: A Comprehensive Meta-analysis. *Pediatrics* **128**, 344–355 (2011).
51. Kogan, M. D. *et al.* Prevalence of Parent-Reported Diagnosis of Autism Spectrum Disorder Among Children in the US, 2007. *Pediatrics* **124**, 1395–1403 (2009).
52. Mandell, D. S. *et al.* Racial/Ethnic Disparities in the Identification of Children With Autism Spectrum Disorders. *Am. J. Public Health* **99**, 493–498 (2009).

53. Pinborough-Zimmerman, J. *et al.* Sociodemographic risk factors associated with autism spectrum disorders and intellectual disability. *Autism Res.* **4**, 438–448 (2011).
54. Bilder, D., Pinborough-Zimmerman, J., Miller, J. & McMahon, W. Prenatal, Perinatal, and Neonatal Factors Associated With Autism Spectrum Disorders. *Pediatrics* **123**, 1293–1300 (2009).
55. Bhasin, T. K. & Schendel, D. Sociodemographic Risk Factors for Autism in a US Metropolitan Area. *J. Autism Dev. Disord.* **37**, 667–677 (2007).
56. Ronald, A. & Hoekstra, R. A. Autism spectrum disorders and autistic traits: a decade of new twin studies. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* **156B**, 255–274 (2011).
57. Hallmayer, J. *et al.* Genetic Heritability and Shared Environmental Factors Among Twin Pairs With Autism. *Arch. Gen. Psychiatry* **68**, 1095–1102 (2011).
58. Sharma, S. R., Gonda, X. & Tarazi, F. I. Autism Spectrum Disorder: Classification, diagnosis and therapy. *Pharmacol. Ther.* **190**, 91–104 (2018).
59. Bolton, P. *et al.* A Case-Control Family History Study of Autism. *J. Child Psychol. Psychiatry* **35**, 877–900 (1994).
60. DiGuseppi, C. *et al.* Screening for Autism Spectrum Disorders in Children With Down Syndrome. *J. Dev. Behav. Pediatr. JDBP* **31**, 181–191 (2010).
61. Hall, S. S., Lightbody, A. A. & Reiss, A. L. Compulsive, Self-Injurious, and Autistic Behavior in Children and Adolescents With Fragile X Syndrome. *Am. J. Ment. Retard.* **113**, 44–53 (2008).
62. Simonoff, E. *et al.* Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *J. Am. Acad. Child Adolesc. Psychiatry* **47**, 921–929 (2008).
63. Hofvander, B. *et al.* Psychiatric and psychosocial problems in adults with normal-intelligence autism spectrum disorders. *BMC Psychiatry* **9**, 35 (2009).
64. Lugnegård, T., Hallerbäck, M. U. & Gillberg, C. Psychiatric comorbidity in young adults with a clinical diagnosis of Asperger syndrome. *Res. Dev. Disabil.* **32**, 1910–1917 (2011).
65. Hartman, C. A., Geurts, H. M., Franke, B., Buitelaar, J. K. & Rommelse, N. N. J. Changing ASD-ADHD symptom co-occurrence across the lifespan with adolescence as crucial time window: Illustrating the need to go beyond childhood. *Neurosci. Biobehav. Rev.* **71**, 529–541 (2016).
66. Ung, D. *et al.* Clinical characteristics of high-functioning youth with autism spectrum disorder and anxiety. *Neuropsychiatry* **3**, (2013).
67. Rapoport, J., Chavez, A., Greenstein, D., Addington, A. & Gogtay, N. Autism Spectrum Disorders and Childhood-Onset Schizophrenia: Clinical and Biological Contributions to a Relation Revisited. *J. Am. Acad. Child Adolesc. Psychiatry* **48**, 10–18 (2009).
68. Vannucchi, G. *et al.* Bipolar disorder in adults with Asperger’s Syndrome: A systematic review. *J. Affect. Disord.* **168**, 151–160 (2014).
69. Chandrasekhar, T. & Sikich, L. Challenges in the diagnosis and treatment of depression in autism spectrum disorders across the lifespan. *Dialogues Clin. Neurosci.* **17**, 219–227 (2015).



70. Hill, A. P. *et al.* Aggressive Behavior Problems in Children with Autism Spectrum Disorders: Prevalence and Correlates in a Large Clinical Sample. *Res. Autism Spectr. Disord.* **8**, 1121–1133 (2014).
71. Gotham, K., Brunwasser, S. M. & Lord, C. Depressive and anxiety symptom trajectories from school age through young adulthood in samples with autism spectrum disorder and developmental delay. *J. Am. Acad. Child Adolesc. Psychiatry* **54**, 369–376.e3 (2015).
72. Schreck, K. A. & Richdale, A. L. Sleep problems, behavior, and psychopathology in autism: inter-relationships across the lifespan. *Curr. Opin. Psychol.* **34**, 105–111 (2020).
73. Rao, P. A. & Landa, R. J. Association between severity of behavioral phenotype and comorbid attention deficit hyperactivity disorder symptoms in children with autism spectrum disorders. *Autism Int. J. Res. Pract.* **18**, 272–280 (2014).
74. Bruchhage, M. M. K., Bucci, M.-P. & Becker, E. B. E. Cerebellar involvement in autism and ADHD. *Handb. Clin. Neurol.* **155**, 61–72 (2018).
75. Deshpande, G., Wang, P., Rangaprakash, D. & Wilamowski, B. Fully Connected Cascade Artificial Neural Network Architecture for Attention Deficit Hyperactivity Disorder Classification From Functional Magnetic Resonance Imaging Data. *IEEE Trans. Cybern.* **45**, 2668–2679 (2015).
76. Becker, E. B. E. & Stoodley, C. J. Autism spectrum disorder and the cerebellum. *Int. Rev. Neurobiol.* **113**, 1–34 (2013).
77. Shephard, E. *et al.* Resting-State Neurophysiological Activity Patterns in Young People with ASD, ADHD, and ASD + ADHD. *J. Autism Dev. Disord.* **48**, 110–122 (2018).
78. Gargaro, B. A., Rinehart, N. J., Bradshaw, J. L., Tonge, B. J. & Sheppard, D. M. Autism and ADHD: How far have we come in the comorbidity debate? *Neurosci. Biobehav. Rev.* **35**, 1081–1088 (2011).
79. Wang, K. *et al.* Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **459**, 528–533 (2009).
80. Bora, E. & Pantelis, C. Meta-analysis of social cognition in attention-deficit/hyperactivity disorder (ADHD): comparison with healthy controls and autistic spectrum disorder. *Psychol. Med.* **46**, 699–716 (2016).
81. Craig, F. *et al.* Overlap Between Autism Spectrum Disorders and Attention Deficit Hyperactivity Disorder: Searching for Distinctive/Common Clinical Features. *Autism Res.* **8**, 328–337 (2015).
82. Purkayastha, P., Malapati, A., Yogeewari, P. & Sriram, D. A Review on GABA/Glutamate Pathway for Therapeutic Intervention of ASD and ADHD. *Curr. Med. Chem.* **22**, 1850–59 (2015).
83. Naaijen, J., Lythgoe, D. J., Amiri, H., Buitelaar, J. K. & Glennon, J. C. Fronto-striatal glutamatergic compounds in compulsive and impulsive syndromes: a review of magnetic resonance spectroscopy studies. *Neurosci. Biobehav. Rev.* **52**, 74–88 (2015).
84. Uddin, L. Q., Dajani, D. R., Voorhies, W., Bednarz, H. & Kana, R. K. Progress and roadblocks in the search for brain-based biomarkers of autism and attention-deficit/hyperactivity disorder. *Transl. Psychiatry* **7**, e1218 (2017).
85. Leyfer, O. T. *et al.* Comorbid psychiatric disorders in children with autism: interview development and rates of disorders. *J. Autism Dev. Disord.* **36**, 849–861 (2006).

86. Davis, T. E. *et al.* Anxiety symptoms across the lifespan in people diagnosed with Autistic Disorder. *Res. Autism Spectr. Disord.* **5**, 112–118 (2011).
87. van Steensel, F. J. A., Bögels, S. M. & Perrin, S. Anxiety disorders in children and adolescents with autistic spectrum disorders: a meta-analysis. *Clin. Child Fam. Psychol. Rev.* **14**, 302–317 (2011).
88. McDougle, C. J. *et al.* A case-controlled study of repetitive thoughts and behavior in adults with autistic disorder and obsessive-compulsive disorder. *Am. J. Psychiatry* **152**, 772–777 (1995).
89. Joshi, G. *et al.* The heavy burden of psychiatric comorbidity in youth with autism spectrum disorders: a large comparative study of a psychiatrically referred population. *J. Autism Dev. Disord.* **40**, 1361–1370 (2010).
90. Kaat, A. J. & Lecavalier, L. Disruptive behavior disorders in children and adolescents with autism spectrum disorders: A review of the prevalence, presentation, and treatment. *Res. Autism Spectr. Disord.* **7**, 1579–1594 (2013).
91. Fernández, M., Mollinedo-Gajate, I. & Peñagarikano, O. Neural Circuits for Social Cognition: Implications for Autism. *Neuroscience* **370**, 148–162 (2018).
92. Hutchins, T. L. & Prelock, P. A. Using Story-Based Interventions to Improve Episodic Memory in Autism Spectrum Disorder. *Semin. Speech Lang.* **39**, 125–143 (2018).
93. Bozzi, Y., Provenzano, G. & Casarosa, S. Neurobiological bases of autism-epilepsy comorbidity: a focus on excitation/inhibition imbalance. *Eur. J. Neurosci.* **47**, 534–548 (2018).
94. O'Hare, A. & Bremner, L. Management of developmental speech and language disorders: Part 1. *Arch. Dis. Child.* **101**, 272–277 (2016).
95. Kasari, C., Paparella, T., Freeman, S. & Jahromi, L. B. Language outcome in autism: randomized comparison of joint attention and play interventions. *J. Consult. Clin. Psychol.* **76**, 125–137 (2008).
96. Rigby, S. N., Stoesz, B. M. & Jakobson, L. S. Empathy and face processing in adults with and without autism spectrum disorder. *Autism Res. Off. J. Int. Soc. Autism Res.* **11**, 942–955 (2018).
97. Trevisan, D. A., Roberts, N., Lin, C. & Birmingham, E. How do adults and teens with self-declared Autism Spectrum Disorder experience eye contact? A qualitative analysis of first-hand accounts. *PLoS ONE* **12**, (2017).
98. Harmsen, I. E. Empathy in Autism Spectrum Disorder. *J. Autism Dev. Disord.* **49**, 3939–3955 (2019).
99. Nebel, M. B. *et al.* Disruption of functional organization within the primary motor cortex in children with autism. *Hum. Brain Mapp.* **35**, 567–580 (2012).
100. Premack, D. & Woodruff, G. Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* **1**, 515–526 (1978).
101. Zelazo, P. D., Craik, F. I. M. & Booth, L. Executive function across the life span. *Acta Psychol. (Amst.)* **115**, 167–183 (2004).
102. Lai, M.-C., Lombardo, M. V. & Baron-Cohen, S. Autism. *The Lancet* **383**, 896–910 (2014).
103. Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T. & Chakrabarti, B. Talent in autism: hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 1377–1383 (2009).

104. Bauman, M. L. Medical comorbidities in autism: Challenges to diagnosis and treatment. *Neurotherapeutics* **7**, 320–327 (2010).
105. Erickson, C. A. *et al.* Gastrointestinal Factors in Autistic Disorder: A Critical Review. *J. Autism Dev. Disord.* **35**, 713 (2005).
106. Pardo, C. A., Vargas, D. L. & Zimmerman, A. W. Immunity, neuroglia and neuroinflammation in autism. *Int. Rev. Psychiatry* **17**, 485–495 (2005).
107. Zafeiriou, D. I., Ververi, A. & Vargiami, E. Childhood autism and associated comorbidities. *Brain Dev.* **29**, 257–272 (2007).
108. Bolton, P. F. *et al.* Epilepsy in autism: features and correlates. *Br. J. Psychiatry* **198**, 289–294 (2011).
109. Accardo, J. A. & Malow, B. A. Sleep, epilepsy, and autism. *Epilepsy Behav. EB* **47**, 202–206 (2015).
110. Souders, M. C. *et al.* Sleep in Children with Autism Spectrum Disorder. *Curr. Psychiatry Rep.* **19**, 34 (2017).
111. de Theije, C. G. M. *et al.* Pathways underlying the gut-to-brain connection in autism spectrum disorders as future targets for disease management. *Eur. J. Pharmacol.* **668 Suppl 1**, S70-80 (2011).
112. McElhanon, B. O., McCracken, C., Karpen, S. & Sharp, W. G. Gastrointestinal symptoms in autism spectrum disorder: a meta-analysis. *Pediatrics* **133**, 872–883 (2014).
113. Hallett, M. Tourette Syndrome: Update. *Brain Dev.* **37**, 651–655 (2015).
114. Voineagu, I. *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474**, 380–384 (2011).
115. Abrahams, B. S. & Geschwind, D. H. Advances in autism genetics: on the threshold of a new neurobiology. *Nat. Rev. Genet.* **9**, 341–355 (2008).
116. Krumm, N. *et al.* Excess of rare, inherited truncating mutations in autism. *Nat. Genet.* **47**, 582–588 (2015).
117. De Rubeis, S. *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**, 209–215 (2014).
118. Gonatopoulos-Pournatzis, T. & Blencowe, B. J. Microexons: at the nexus of nervous system development, behaviour and autism spectrum disorder. *Curr. Opin. Genet. Dev.* **65**, 22–33 (2020).
119. Folstein, S. & Rutter, M. Infantile Autism: A Genetic Study of 21 Twin Pairs. *J. Child Psychol. Psychiatry* **18**, 297–321 (1977).
120. DeMyer, M. K., Hingtgen, J. N. & Jackson, R. K. Infantile Autism Reviewed: A Decade of Research. *Schizophr. Bull.* **7**, 388–451 (1981).
121. Langan, M. Parental voices and controversies in autism. *Disabil. Soc.* **26**, 193–205 (2011).
122. Zhao, X. *et al.* A unified genetic theory for sporadic and inherited autism. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 12831–12836 (2007).

123. Ronemus, M., Iossifov, I., Levy, D. & Wigler, M. The role of de novo mutations in the genetics of autism spectrum disorders. *Nat. Rev. Genet.* **15**, 133–141 (2014).
124. Ozonoff, S. *et al.* Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. *Pediatrics* **128**, e488–495 (2011).
125. Zafeiriou, D. I., Ververi, A., Dafoulis, V., Kalyva, E. & Vargiami, E. Autism spectrum disorders: the quest for genetic syndromes. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* **162B**, 327–366 (2013).
126. Ornoy, A., Weinstein-Fudim, L. & Ergaz, Z. Genetic Syndromes, Maternal Diseases and Antenatal Factors Associated with Autism Spectrum Disorders (ASD). *Front. Neurosci.* **10**, 316 (2016).
127. Hagerman, R. J. *et al.* Fragile X syndrome. *Nat. Rev. Dis. Primer* **3**, 17065 (2017).
128. Moss, J. & Howlin, P. Autism spectrum disorders in genetic syndromes: implications for diagnosis, intervention and understanding the wider autism spectrum disorder population. *J. Intellect. Disabil. Res. JIDR* **53**, 852–873 (2009).
129. Darnell, J. C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).
130. Geschwind, D. H. Genetics of Autism Spectrum Disorders. *Trends Cogn. Sci.* **15**, 409–416 (2011).
131. Murdoch, J. D. & State, M. W. Recent developments in the genetics of autism spectrum disorders. *Curr. Opin. Genet. Dev.* **23**, 310–315 (2013).
132. Pinto, D. *et al.* Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am. J. Hum. Genet.* **94**, 677–694 (2014).
133. Sanders, S. J. *et al.* Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* **87**, 1215–1233 (2015).
134. Leppa, V. M. *et al.* Rare Inherited and De Novo CNVs Reveal Complex Contributions to ASD Risk in Multiplex Families. *Am. J. Hum. Genet.* **99**, 540–554 (2016).
135. Klei, L. *et al.* Common genetic variants, acting additively, are a major source of risk for autism. *Mol. Autism* **3**, 9 (2012).
136. Devlin, B. & Scherer, S. W. Genetic architecture in autism spectrum disorder. *Curr. Opin. Genet. Dev.* **22**, 229–237 (2012).
137. Stein, J. L., Parikshak, N. N. & Geschwind, D. H. Rare inherited variation in autism: beginning to see the forest and a few trees. *Neuron* **77**, 209–211 (2013).
138. State, M. W. & Levitt, P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nat. Neurosci.* **14**, 1499–1506 (2011).
139. Heil, K. M. & Schaaf, C. P. The genetics of Autism Spectrum Disorders—a guide for clinicians. *Curr. Psychiatry Rep.* **15**, 334 (2013).
140. Onore, C., Careaga, M. & Ashwood, P. The role of immune dysfunction in the pathophysiology of autism. *Brain. Behav. Immun.* **26**, 383–392 (2012).

141. Lee, B. K. *et al.* Maternal hospitalization with infection during pregnancy and risk of autism spectrum disorders. *Brain. Behav. Immun.* **44**, 100–105 (2015).
142. Chess, S. Autism in children with congenital rubella. *J. Autism Child. Schizophr.* **1**, 33–47 (1971).
143. Atladóttir, H. Ó., Henriksen, T. B., Schendel, D. E. & Parner, E. T. Autism after infection, febrile episodes, and antibiotic use during pregnancy: an exploratory study. *Pediatrics* **130**, e1447–1454 (2012).
144. Kaidanovich-Beilin, O., Lipina, T., Vukobradovic, I., Roder, J. & Woodgett, J. R. Assessment of social interaction behaviors. *J. Vis. Exp. JoVE* (2011) doi:10.3791/2473.
145. Bauman, M. D. *et al.* Activation of the maternal immune system during pregnancy alters behavioral development of rhesus monkey offspring. *Biol. Psychiatry* **75**, 332–341 (2014).
146. Jones, K. A. & Thomsen, C. The role of the innate immune system in psychiatric disorders. *Mol. Cell. Neurosci.* **53**, 52–62 (2013).
147. Smith, S. E. P., Li, J., Garbett, K., Mirnics, K. & Patterson, P. H. Maternal immune activation alters fetal brain development through interleukin-6. *J. Neurosci. Off. J. Soc. Neurosci.* **27**, 10695–10702 (2007).
148. Brimberg, L., Sadiq, A., Gregersen, P. K. & Diamond, B. Brain-reactive IgG correlates with autoimmunity in mothers of a child with an autism spectrum disorder. *Mol. Psychiatry* **18**, 1171–1177 (2013).
149. Braunschweig, D. *et al.* Behavioral correlates of maternal antibody status among children with autism. *J. Autism Dev. Disord.* **42**, 1435–1445 (2012).
150. Camacho, J. *et al.* Embryonic intraventricular exposure to autism-specific maternal autoantibodies produces alterations in autistic-like stereotypical behaviors in offspring mice. *Behav. Brain Res.* **266**, 46–51 (2014).
151. Singer, H. S. *et al.* Prenatal exposure to antibodies from mothers of children with autism produces neurobehavioral alterations: A pregnant dam mouse model. *J. Neuroimmunol.* **211**, 39–48 (2009).
152. Singer, H. S. *et al.* Antibrain antibodies in children with autism and their unaffected siblings. *J. Neuroimmunol.* **178**, 149–155 (2006).
153. Wills, S. *et al.* DETECTION OF AUTOANTIBODIES TO NEURAL CELLS OF THE CEREBELLUM IN THE PLASMA OF SUBJECTS WITH AUTISM SPECTRUM DISORDERS. *Brain. Behav. Immun.* **23**, 64–74 (2009).
154. Goines, P. *et al.* Autoantibodies to cerebellum in children with autism associate with behavior. *Brain. Behav. Immun.* **25**, 514–523 (2011).
155. Mazur-Kolecka, B. *et al.* Autoantibodies against neuronal progenitors in sera from children with autism. *Brain Dev.* **36**, 322–329 (2014).
156. Piras, I. S. *et al.* Anti-brain antibodies are associated with more severe cognitive and behavioral profiles in Italian children with Autism Spectrum Disorder. *Brain. Behav. Immun.* **38**, 91–99 (2014).
157. Campbell, D. B. *et al.* A genetic variant that disrupts MET transcription is associated with autism. *Proc. Natl. Acad. Sci.* **103**, 16834–16839 (2006).

158. Gupta, S. *et al.* Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. *Nat. Commun.* **5**, 5748 (2014).
159. Heuer, L. *et al.* Reduced levels of immunoglobulin in children with autism correlates with behavioral symptoms. *Autism Res. Off. J. Int. Soc. Autism Res.* **1**, 275–283 (2008).
160. Enstrom, A. M. *et al.* Altered gene expression and function of peripheral blood natural killer cells in children with autism. *Brain. Behav. Immun.* **23**, 124–133 (2009).
161. Garbett, K. *et al.* Immune transcriptome alterations in the temporal cortex of subjects with autism. *Neurobiol. Dis.* **30**, 303–311 (2008).
162. Vargas, D. L., Nascimbene, C., Krishnan, C., Zimmerman, A. W. & Pardo, C. A. Neuroglial activation and neuroinflammation in the brain of patients with autism. *Ann. Neurol.* **57**, 67–81 (2005).
163. Li, X. *et al.* Elevated immune response in the brain of autistic patients. *J. Neuroimmunol.* **207**, 111–116 (2009).
164. Haroon, E., Raison, C. L. & Miller, A. H. Psychoneuroimmunology Meets Neuropsychopharmacology: Translational Implications of the Impact of Inflammation on Behavior. *Neuropsychopharmacology* **37**, 137–162 (2012).
165. Ashwood, P. *et al.* Elevated plasma cytokines in autism spectrum disorders provide evidence of immune dysfunction and are associated with impaired behavioral outcome. *Brain. Behav. Immun.* **25**, 40–45 (2011).
166. Shiina, T., Hosomichi, K., Inoko, H. & Kulski, J. K. The HLA genomic loci map: expression, interaction, diversity and disease. *J. Hum. Genet.* **54**, 15–39 (2009).
167. Bilbo, S. D. & Schwarz, J. M. The immune system and developmental programming of brain and behavior. *Front. Neuroendocrinol.* **33**, 267–286 (2012).
168. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet Lond. Engl.* **381**, 1371–1379 (2013).
169. Trowsdale, J. & Knight, J. C. Major histocompatibility complex genomics and human disease. *Annu. Rev. Genomics Hum. Genet.* **14**, 301–323 (2013).
170. Westover, J., Sweeten, T., Benson, M., Bray-Ward, P. & Torres, A. Immune Dysfunction in Autism Spectrum Disorder. in *Autism: A Neurodevelopmental Journey from Genes to Behavior* (2011). doi:10.5772/22318.
171. Torres, A. R., Westover, J. B. & Rosenspire, A. J. HLA Immune Function Genes in Autism. *Autism Res. Treat.* **2012**, 959073 (2012).
172. Stubbs, E. G. & Magenis, R. E. HLA and autism. *J. Autism Dev. Disord.* **10**, 15–19 (1980).
173. Al-Hakbany, M., Awadallah, S. & Al-Ayadhi, L. The Relationship of HLA Class I and II Alleles and Haplotypes with Autism: A Case Control Study. *Autism Res. Treat.* **2014**, 242048 (2014).
174. Mostafa, G. A., Shehab, A. A. & Al-Ayadhi, L. Y. The link between some alleles on human leukocyte antigen system and autism in children. *J. Neuroimmunol.* **255**, 70–74 (2013).

175. Chien, Y.-L. *et al.* Association of HLA-DRB1 alleles and neuropsychological function in autism. *Psychiatr. Genet.* **22**, 46–49 (2012).
176. Guerini, F. R. *et al.* An HLA-G(\*)14bp insertion/deletion polymorphism associates with the development of autistic spectrum disorders. *Brain. Behav. Immun.* **44**, 207–212 (2015).
177. Gamliel, M., Anderson, K. L., Ebstein, R. P., Yirmiya, N. & Mankuta, D. Paternal HLA-C and Maternal Killer-Cell Immunoglobulin-Like Receptor Genotypes in the Development of Autism. *Front. Pediatr.* **4**, 76 (2016).
178. Courchesne, E. Abnormal early brain development in autism. *Mol. Psychiatry* **7 Suppl 2**, S21-23 (2002).
179. Courchesne, E., Carper, R. & Akshoomoff, N. Evidence of Brain Overgrowth in the First Year of Life in Autism. *JAMA J. Am. Med. Assoc.* **290**, 337–44 (2003).
180. Courchesne, E., Campbell, K. & Solso, S. Brain growth across the life span in autism: Age-specific changes in anatomical pathology. *Brain Res.* **1380**, 138–145 (2011).
181. Courchesne, E. *et al.* Mapping early brain development in autism. *Neuron* **56**, 399–413 (2007).
182. Gogtay, N. *et al.* Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci.* **101**, 8174–8179 (2004).
183. Hazlett, H. C. *et al.* Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. *Arch. Gen. Psychiatry* **68**, 467–476 (2011).
184. Rakic, P. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci.* **18**, 383–388 (1995).
185. Wolff, J. J. *et al.* Differences in White Matter Fiber Tract Development Present From 6 to 24 Months in Infants With Autism. *Am. J. Psychiatry* **169**, 589–600 (2012).
186. Lange, N. *et al.* Longitudinal Volumetric Brain Changes in Autism Spectrum Disorder Ages 6–35 Years. *Autism Res.* **8**, 82–93 (2015).
187. Amaral, D. G., Schumann, C. M. & Nordahl, C. W. Neuroanatomy of autism. *Trends Neurosci.* **31**, 137–145 (2008).
188. Redcay, E. & Courchesne, E. Deviant Functional Magnetic Resonance Imaging Patterns of Brain Activity to Speech in 2–3-Year-Old Children with Autism Spectrum Disorder. *Biol. Psychiatry* **64**, 589–598 (2008).
189. McAlonan, G. M. *et al.* Brain anatomy and sensorimotor gating in Asperger’s syndrome. *Brain* **125**, 1594–1606 (2002).
190. Zilbovicius, M. *et al.* Autism, the superior temporal sulcus and social perception. *Trends Neurosci.* **29**, 359–366 (2006).
191. Pelphrey, K., Adolphs, R. & Morris, J. P. Neuroanatomical substrates of social cognition dysfunction in autism. *Ment. Retard. Dev. Disabil. Res. Rev.* **10**, 259–271 (2004).
192. Zalla, T. & Sperduti, M. The amygdala and the relevance detection theory of autism: an evolutionary perspective. *Front. Hum. Neurosci.* **7**, 894 (2013).

193. Stoodley, C. J. The cerebellum and cognition: evidence from functional imaging studies. *Cerebellum Lond. Engl.* **11**, 352–365 (2012).
194. Schumann, C. M., Carter Barnes, C., Lord, C. & Courchesne, E. Amygdala Enlargement in Toddlers with Autism Related to Severity of Social and Communication Impairments. *Biol. Psychiatry* **66**, 942–949 (2009).
195. Janak, P. H. & Tye, K. M. From circuits to behaviour in the amygdala. *Nature* **517**, 284–292 (2015).
196. Gibbard, C. R., Ren, J., Skuse, D. H., Clayden, J. D. & Clark, C. A. Structural connectivity of the amygdala in young adults with autism spectrum disorder. *Hum. Brain Mapp.* **39**, 1270–1282 (2018).
197. Schumann, C. M. *et al.* Longitudinal Magnetic Resonance Imaging Study of Cortical Development through Early Childhood in Autism. *J. Neurosci.* **30**, 4419–4427 (2010).
198. Weiner, K. S. & Grill-Spector, K. The improbable simplicity of the fusiform face area. *Trends Cogn. Sci.* **16**, 251–254 (2012).
199. Waiter, G. D. *et al.* A voxel-based investigation of brain structure in male adolescents with autistic spectrum disorder. *NeuroImage* **22**, 619–625 (2004).
200. Herbert, M. R. *et al.* Abnormal asymmetry in language association cortex in autism. *Ann. Neurol.* **52**, 588–596 (2002).
201. Dziobek, I., Bahnemann, M., Convit, A. & Heekeren, H. R. The Role of the Fusiform-Amygdala System in the Pathophysiology of Autism. *Arch. Gen. Psychiatry* **67**, 397–405 (2010).
202. Hardan, A. Y., Jou, R. J., Keshavan, M. S., Varma, R. & Minshew, N. J. Increased frontal cortical folding in autism: a preliminary MRI study. *Psychiatry Res. Neuroimaging* **131**, 263–268 (2004).
203. Libero, L. E., Schaer, M., Li, D. D., Amaral, D. G. & Nordahl, C. W. A Longitudinal Study of Local Gyrification Index in Young Boys With Autism Spectrum Disorder. *Cereb. Cortex N. Y. N 1991* (2018) doi:10.1093/cercor/bhy126.
204. Wallace, G. L. *et al.* Increased gyrification, but comparable surface area in adolescents with autism spectrum disorders. *Brain J. Neurol.* **136**, 1956–1967 (2013).
205. Libero, L. E., DeRamus, T. P., Deshpande, H. D. & Kana, R. K. Surface-based morphometry of the cortical architecture of autism spectrum disorders: volume, thickness, area, and gyrification. *Neuropsychologia* **62**, 1–10 (2014).
206. Zilles, K., Palomero-Gallagher, N. & Amunts, K. Development of cortical folding during evolution and ontogeny. *Trends Neurosci.* **36**, 275–284 (2013).
207. Di, X. & Biswal, B. B. Similarly Expanded Bilateral Temporal Lobe Volumes in Female and Male Children With Autism Spectrum Disorder. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **1**, 178–185 (2016).
208. Cai, J. *et al.* Increased Left Inferior Temporal Gyrus Was Found in Both Low Function Autism and High Function Autism. *Front. Psychiatry* **9**, (2018).
209. Hadjikhani, N., Joseph, R. M., Snyder, J. & Tager-Flusberg, H. Anatomical differences in the mirror neuron system and social cognition network in autism. *Cereb. Cortex N. Y. N 1991* **16**, 1276–1282 (2006).



210. Boddaert, N. *et al.* Superior temporal sulcus anatomical abnormalities in childhood autism: a voxel-based morphometry MRI study. *NeuroImage* **23**, 364–369 (2004).
211. Barnea-Goraly, N. *et al.* White matter structure in autism: preliminary evidence from diffusion tensor imaging. *Biol. Psychiatry* **55**, 323–326 (2004).
212. Lo, Y.-C., Chen, Y.-J., Hsu, Y.-C., Tseng, W.-Y. I. & Gau, S. S.-F. Reduced tract integrity of the model for social communication is a neural substrate of social communication deficits in autism spectrum disorder. *J. Child Psychol. Psychiatry* **58**, 576–585 (2017).
213. Amaral, D. G., Schumann, C. M. & Nordahl, C. W. Neuroanatomy of autism. *Trends Neurosci.* **31**, 137–145 (2008).
214. Noordermeer, S. D. S., Luman, M. & Oosterlaan, J. A Systematic Review and Meta-analysis of Neuroimaging in Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD) Taking Attention-Deficit Hyperactivity Disorder (ADHD) Into Account. *Neuropsychol. Rev.* **26**, 44–72 (2016).
215. De Bellis, M. D. *et al.* A pilot study of amygdala volumes in pediatric generalized anxiety disorder. *Biol. Psychiatry* **48**, 51–57 (2000).
216. Just, M. A., Cherkassky, V. L., Keller, T. A. & Minshew, N. J. Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. *Brain* **127**, 1811–1821 (2004).
217. Buckner, R. L. & DiNicola, L. M. The brain’s default network: updated anatomy, physiology and evolving insights. *Nat. Rev. Neurosci.* **20**, 593–608 (2019).
218. Pereira, A. M. *et al.* Differences in Cortical Structure and Functional MRI Connectivity in High Functioning Autism. *Front. Neurol.* **9**, 539 (2018).
219. Padmanabhan, A., Lynch, C. J., Schaer, M. & Menon, V. The Default Mode Network in Autism. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2**, 476–486 (2017).
220. Jung, M. *et al.* Decreased structural connectivity and resting-state brain activity in the lateral occipital cortex is associated with social communication deficits in boys with autism spectrum disorder. *NeuroImage* **190**, 205–212 (2019).
221. Baron-Cohen, S. *et al.* Social intelligence in the normal and autistic brain: an fMRI study. *Eur. J. Neurosci.* **11**, 1891–1898 (1999).
222. Rausch, A. *et al.* Altered functional connectivity of the amygdaloid input nuclei in adolescents and young adults with autism spectrum disorder: a resting state fMRI study. *Mol. Autism* **7**, 13 (2016).
223. Tyszka, J. M., Kennedy, D. P., Paul, L. K. & Adolphs, R. Largely Typical Patterns of Resting-State Functional Connectivity in High-Functioning Adults with Autism. *Cereb. Cortex* **24**, 1894–1905 (2014).
224. Supekar, K. *et al.* Brain Hyperconnectivity in Children with Autism and its Links to Social Deficits. *Cell Rep.* **5**, 738–747 (2013).
225. Ciaramidaro, A. *et al.* Transdiagnostic deviant facial recognition for implicit negative emotion in autism and schizophrenia. *Eur. Neuropsychopharmacol. J. Eur. Coll. Neuropsychopharmacol.* **28**, 264–275 (2018).

226. Yang, Y. J. D. *et al.* Neural mechanisms of behavioral change in young adults with high-functioning autism receiving virtual reality social cognition training: A pilot study. *Autism Res. Off. J. Int. Soc. Autism Res.* **11**, 713–725 (2018).
227. Sato, W., Toichi, M., Uono, S. & Kochiyama, T. Impaired social brain network for processing dynamic facial expressions in autism spectrum disorders. *BMC Neurosci.* **13**, 99 (2012).
228. Lynn, A. C. *et al.* Functional connectivity differences in autism during face and car recognition: underconnectivity and atypical age-related changes. *Dev. Sci.* **21**, e12508 (2018).
229. Dichter, G. S., Richey, J. A., Rittenberg, A. M., Sabatino, A. & Bodfish, J. W. Reward Circuitry Function in Autism During Face Anticipation and Outcomes. *J. Autism Dev. Disord.* **42**, 147–160 (2012).
230. Berger, H. On the electroencephalogram of man. *Electroencephalogr. Clin. Neurophysiol. Suppl* 28:37+ (1969).
231. Adeli, H. & Lichtenstein, A. G. Automated EEG-based diagnosis of the neurological and psychiatric disorders. in *2011 18th International Conference on Systems, Signals and Image Processing* 1–4 (2011).
232. Djemal, R., Alsharabi, K., Ibrahim, S. & Alsuwailem, A. EEG-Based computer aided diagnosis of autism spectrum disorder using wavelet, entropy, and ANN. *BioMed Res. Int.* **2017**, (2017).
233. Alsaggaf, E. & Kamel, M. Using EEGs to diagnose autism disorder by classification algorithm. *Life Sci. J.* **11**, 305–308 (2014).
234. Keller, R. *I disturbi dello spettro autistico in adolescenza e in età adulta: Aspetti diagnostici e proposte di intervento.* (Edizioni Centro Studi Erickson, 2016).
235. Stafstrom, C. E. & Benke, T. A. Autism and Epilepsy: Exploring the Relationship Using Experimental Models: Autism and Epilepsy: Exploring the Relationship. *Epilepsy Curr.* **15**, 206–210 (2015).
236. Maquet, P. The role of sleep in learning and memory. *Science* **294**, 1048–1052 (2001).
237. Goodlin-Jones, B. L., Tang, K., Liu, J. & Anders, T. F. Sleep patterns in preschool-age children with autism, developmental delay, and typical development. *J. Am. Acad. Child Adolesc. Psychiatry* **47**, 930–938 (2008).
238. Elrod, M. G. & Hood, B. S. Sleep Differences Among Children With Autism Spectrum Disorders and Typically Developing Peers: A Meta-analysis. *J. Dev. Behav. Pediatr.* **36**, 166–177 (2015).
239. Singh, K. & Zimmerman, A. W. Sleep in Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder. *Semin. Pediatr. Neurol.* **22**, 113–125 (2015).
240. Green, H. L., Edgar, J. C., Matsuzaki, J. & Roberts, T. P. L. Magnetoencephalography Research in Pediatric Autism Spectrum Disorder. *Neuroimaging Clin. N. Am.* **30**, 193–203 (2020).
241. Wilson, T. W., Rojas, D. C., Reite, M. L., Teale, P. D. & Rogers, S. J. Children and Adolescents with Autism Exhibit Reduced MEG Steady-State Gamma Responses. *Biol. Psychiatry* **62**, 192–197 (2007).

242. Yoshimura, Y. *et al.* Atypical brain lateralisation in the auditory cortex and language performance in 3- to 7-year-old children with high-functioning autism spectrum disorder: a child-customised magnetoencephalography (MEG) study. *Mol. Autism* **4**, 38 (2013).
243. Orekhova, E. V. *et al.* Auditory cortex responses to clicks and sensory modulation difficulties in children with autism spectrum disorders (ASD). *PloS One* **7**, e39906 (2012).
244. Lameka, K., Farwell, M. D. & Ichise, M. Positron Emission Tomography. *Handb. Clin. Neurol.* **135**, 209–227 (2016).
245. Lefevre, A. *et al.* Oxytocin Fails to Recruit Serotonergic Neurotransmission in the Autistic Brain. *Cereb. Cortex N. Y. N 1991* **28**, 4169–4178 (2018).
246. Cai, Z., Li, S., Matuskey, D., Nabulsi, N. & Huang, Y. PET Imaging of Synaptic Density: A New Tool for Investigation of Neuropsychiatric Diseases. *Neurosci. Lett.* **691**, 44–50 (2019).
247. Rizzolatti, G. & Craighero, L. The Mirror-Neuron System. *Annu. Rev. Neurosci.* **27**, 169–192 (2004).
248. Williams, J. H. G., Whiten, A., Suddendorf, T. & Perrett, D. I. Imitation, mirror neurons and autism. *Neurosci. Biobehav. Rev.* **25**, 287–295 (2001).
249. Iacoboni, M. *et al.* Cortical Mechanisms of Human Imitation. *Science* **286**, 2526–2528 (1999).
250. Allison, null, Puce, null & McCarthy, null. Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* **4**, 267–278 (2000).
251. Villalobos, M. E., Mizuno, A., Dahl, B. C., Kemmotsu, N. & Müller, R.-A. Reduced functional connectivity between V1 and inferior frontal cortex associated with visuomotor performance in autism. *NeuroImage* **25**, 916–925 (2005).
252. Dapretto, M. *et al.* Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat. Neurosci.* **9**, 28–30 (2006).
253. Williams, J. H. G. *et al.* Neural mechanisms of imitation and ‘mirror neuron’ functioning in autistic spectrum disorder. *Neuropsychologia* **44**, 610–621 (2006).
254. Iacoboni, M. & Dapretto, M. The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* **7**, 942–951 (2006).
255. Saxe, R. & Kanwisher, N. People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *NeuroImage* **19**, 1835–1842 (2003).
256. Singer, T. & Lamm, C. The Social Neuroscience of Empathy. *Ann. N. Y. Acad. Sci.* **1156**, 81–96 (2009).
257. Bzdok, D. *et al.* Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Struct. Funct.* **217**, 783–796 (2012).
258. Bowman, L. & Wellman, H. Neuroscience Contributions to Childhood Theory of Mind Development. in *Contemporary Perspectives on Research in Theories of Mind in Early Childhood Education* 195–223 (2014).
259. Shamsi, F., Hosseini, S., Tahamtan, M. & Bayat, M. Methodology Report: The Impaired Theory of Mind in Autism Spectrum Disorders and the Possible Remedative Role of Transcranial Direct Current Stimulation. *J. Adv. Med. Sci. Appl. Technol.* **3**, 175–178 (2017).

260. Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., MRC AIMS Consortium & Baron-Cohen, S. Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *NeuroImage* **56**, 1832–1838 (2011).
261. Libero, L. E. *et al.* The role of mirroring and mentalizing networks in mediating action intentions in autism. *Mol. Autism* **5**, 50 (2014).
262. Ilzarbe, D. *et al.* Neural Correlates of Theory of Mind in Autism Spectrum Disorder, Attention-Deficit/Hyperactivity Disorder, and the Comorbid Condition. *Front. Psychiatry* **11**, (2020).
263. Kana, R. K. *et al.* Aberrant functioning of the theory-of-mind network in children and adolescents with autism. *Mol. Autism* **6**, 59 (2015).
264. Corbetta, M., Patel, G. & Shulman, G. L. The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron* **58**, 306–324 (2008).
265. Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277 (2006).
266. Lynch, C. J. *et al.* Default mode network in childhood autism: posteromedial cortex heterogeneity and relationship with social deficits. *Biol. Psychiatry* **74**, 212–219 (2013).
267. Kennedy, D. P. & Courchesne, E. The intrinsic functional organization of the brain is altered in autism. *NeuroImage* **39**, 1877–1885 (2008).
268. Lombardo, M. V. *et al.* Atypical neural self-representation in autism. *Brain J. Neurol.* **133**, 611–624 (2010).
269. Murdaugh, D. L., Nadendla, K. D. & Kana, R. K. Differential role of temporoparietal junction and medial prefrontal cortex in causal inference in autism: an independent component analysis. *Neurosci. Lett.* **568**, 50–55 (2014).
270. Haar, S., Berman, S., Behrmann, M. & Dinstein, I. Anatomical Abnormalities in Autism? *Cereb. Cortex N. Y. N 1991* **26**, 1440–1452 (2016).
271. Valk, S. L., Di Martino, A., Milham, M. P. & Bernhardt, B. C. Multicenter mapping of structural network alterations in autism. *Hum. Brain Mapp.* **36**, 2364–2373 (2015).
272. David, N. *et al.* Right temporoparietal gray matter predicts accuracy of social perception in the autism spectrum. *J. Autism Dev. Disord.* **44**, 1433–1446 (2014).
273. Schaer, M., Kochalka, J., Padmanabhan, A., Supekar, K. & Menon, V. Sex differences in cortical volume and gyrification in autism. *Mol. Autism* **6**, 42 (2015).
274. Doyle-Thomas, K. A. R. *et al.* Effects of age and symptomatology on cortical thickness in autism spectrum disorders. *Res. Autism Spectr. Disord.* **7**, 141–150 (2013).
275. Lin, H.-Y., Ni, H.-C., Lai, M.-C., Tseng, W.-Y. I. & Gau, S. S.-F. Regional brain volume differences between males with and without autism spectrum disorder are highly age-dependent. *Mol. Autism* **6**, 29 (2015).
276. Kaplan, J. T. *et al.* Processing Narratives Concerning Protected Values: A Cross-Cultural Investigation of Neural Correlates. *Cereb. Cortex N. Y. N 1991* **27**, 1428–1438 (2017).
277. Insel, T. *et al.* Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* **167**, 748–751 (2010).

278. Constantino, J. N. *et al.* Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *J. Autism Dev. Disord.* **33**, 427–433 (2003).
279. Kamio, Y. *et al.* Quantitative autistic traits ascertained in a national survey of 22 529 Japanese schoolchildren. *Acta Psychiatr. Scand.* **128**, 45–53 (2013).
280. Bölte, S., Poustka, F. & Constantino, J. N. Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Res.* **1**, 354–363 (2008).
281. Aldridge, F. J., Gibbs, V. M., Schmidhofer, K. & Williams, M. Investigating the Clinical Usefulness of the Social Responsiveness Scale (SRS) in a Tertiary Level, Autism Spectrum Disorder Specific Assessment Clinic. *J. Autism Dev. Disord.* **42**, 294–300 (2012).
282. Blanken, L. M. E. *et al.* Cortical morphology in 6- to 10-year old children with autistic traits: a population-based neuroimaging study. *Am. J. Psychiatry* **172**, 479–486 (2015).
283. Prigge, M. B. D. *et al.* Social Responsiveness Scale (SRS) in Relation to Longitudinal Cortical Thickness Changes in Autism Spectrum Disorder. *J. Autism Dev. Disord.* **48**, 3319–3329 (2018).
284. Page, J. *et al.* Quantitative autistic trait measurements index background genetic risk for ASD in Hispanic families. *Mol. Autism* **7**, 39 (2016).
285. Frazier, T. W. *et al.* Quantitative autism symptom patterns recapitulate differential mechanisms of genetic transmission in single and multiple incidence families. *Mol. Autism* **6**, 58 (2015).
286. Blanken, L. M. E. *et al.* White matter microstructure in children with autistic traits. *Psychiatry Res. Neuroimaging* **263**, 127–134 (2017).
287. Tu, P.-C. *et al.* Structural and functional correlates of a quantitative autistic trait measured using the social responsive scale in neurotypical male adolescents. *Autism Res.* **9**, 570–578 (2016).
288. Wallace, G. L. *et al.* Distinct Cortical Correlates of Autistic versus Antisocial Traits in a Longitudinal Sample of Typically Developing Youth. *J. Neurosci.* **32**, 4856–4860 (2012).
289. Ota, M. *et al.* Relationship between Autistic Spectrum Trait and Regional Cerebral Blood Flow in Healthy Male Subjects. *Psychiatry Investig.* **15**, 956–961 (2018).
290. Jakab, A. *et al.* Autistic traits in neurotypical adults: correlates of graph theoretical functional network topology and white matter anisotropy patterns. *PLoS One* **8**, e60982 (2013).
291. Di Martino, A. *et al.* Relationship between cingulo-insular functional connectivity and autistic traits in neurotypical adults. *Am. J. Psychiatry* **166**, 891–899 (2009).
292. Vinkhuyzen, A. A. E. *et al.* Gestational vitamin D deficiency and autism-related traits: the Generation R Study. *Mol. Psychiatry* **23**, 240–246 (2018).
293. Oulhote, Y. *et al.* Gestational Exposures to Phthalates and Folic Acid, and Autistic Traits in Canadian Children. *Environ. Health Perspect.* **128**, 27004 (2020).
294. Verhoeff, M. E. *et al.* The bidirectional association between sleep problems and autism spectrum disorder: a population-based cohort study. *Mol. Autism* **9**, 8 (2018).
295. Wakeford, S., Hinvest, N., Ring, H. & Brosnan, M. Autistic characteristics in adults with epilepsy and perceived seizure activity. *Epilepsy Behav.* **52**, 244–250 (2015).

296. Morris, S. M. *et al.* Disease Burden and Symptom Structure of Autism in Neurofibromatosis Type 1: A Study of the International NF1-ASD Consortium Team (INFACT). *JAMA Psychiatry* **73**, 1276–1284 (2016).
297. Carrasco, M., Volkmar, F. R. & Bloch, M. H. Pharmacologic Treatment of Repetitive Behaviors in Autism Spectrum Disorders: Evidence of Publication Bias. *Pediatrics* **129**, e1301–e1310 (2012).
298. Turner, M. A. & Hammond, N. Cognitive behavioural therapy in the treatment of social skills deficits and social phobia in a man with an autism spectrum disorder: a single-case study. *Cogn. Behav. Ther.* **9**, (2016).
299. Lombardo, M. V., Lai, M.-C. & Baron-Cohen, S. Big data approaches to decomposing heterogeneity across the autism spectrum. *Mol. Psychiatry* **24**, 1435–1450 (2019).
300. Ellegood, J. *et al.* Clustering autism: using neuroanatomical differences in 26 mouse models to gain insight into the heterogeneity. *Mol. Psychiatry* **20**, 118–125 (2015).
301. Stevens, E. *et al.* Identification and analysis of behavioral phenotypes in autism spectrum disorder via unsupervised machine learning. *Int. J. Med. Inf.* **129**, 29–36 (2019).
302. Mihailov, A. *et al.* Cortical signatures in behaviorally clustered autistic traits subgroups: a population-based study. *Transl. Psychiatry* **10**, (2020).
303. Alexander, L. M. *et al.* An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci. Data* **4**, 170181 (2017).
304. Kessler, R. C. *et al.* Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 593–602 (2005).
305. Loth, E. *et al.* The EU-AIMS Longitudinal European Autism Project (LEAP): design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. *Mol. Autism* **8**, (2017).
306. Cho, G., Yim, J., Choi, Y., Ko, J. & Lee, S.-H. Review of Machine Learning Algorithms for Diagnosing Mental Illness. *Psychiatry Investig.* **16**, 262–269 (2019).
307. Strauss, J. S., Bartko, J. J. & Carpenter, W. T. The Use of Clustering Techniques for the Classification of Psychiatric Patients. *Br. J. Psychiatry* **122**, 531–540 (1973).
308. Jolliffe, I. T. & Cadima, J. Principal component analysis: a review and recent developments. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **374**, 20150202 (2016).
309. Ringnér, M. What is principal component analysis? *Nat. Biotechnol.* **26**, 303–304 (2008).
310. Abdi, H. & Williams, L. J. Principal Component Analysis. *Princ. Compon. Anal.* **47**.
311. Likas, A., Vlassis, N. & J. Verbeek, J. The global k-means clustering algorithm. *Pattern Recognit.* **36**, 451–461 (2003).
312. Duda, R. & Hart, P. Pattern classification and scene analysis. in *A Wiley-Interscience publication* (1973). doi:10.2307/2286028.
313. Park, H.-S. & Jun, C.-H. A simple and fast algorithm for K-medoids clustering. *Expert Syst. Appl.* **36**, 3336–3341 (2009).

314. Page, J. T., Liechty, Z. S., Huynh, M. D. & Udall, J. A. BamBam: genome sequence analysis tools for biologists. *BMC Res. Notes* **7**, 829 (2014).
315. Day, W. & Edelsbrunner, H. Efficient algorithms for agglomerative hierarchical clustering methods. *J. Classif.* **1**, 7–24 (1984).
316. Tullis, T. & Albert, B. Chapter 9 - Special Topics. in *Measuring the User Experience (Second Edition)* (eds. Tullis, T. & Albert, B.) 209–236 (Morgan Kaufmann, 2013). doi:10.1016/B978-0-12-415781-1.00009-1.
317. Fukunaga, K. & Hostetler, L. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Inf. Theory* **21**, 32–40 (1975).
318. Yizong Cheng. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **17**, 790–799 (1995).
319. Zivkovic, Z. Improved adaptive Gaussian mixture model for background subtraction. in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004*. vol. 2 28–31 Vol.2 (2004).
320. Carr, M. W. & Grey, M. L. Magnetic Resonance Imaging: Overview, risks, and safety measures. *AJN Am. J. Nurs.* **102**, 26–33 (2002).
321. Fisher, A. R. & Siegelman, E. S. Magnetic resonance imaging: Techniques. *Clin. Liver Dis.* **6**, 53–72 (2002).
322. Chen, J. E. & Glover, G. H. Functional Magnetic Resonance Imaging Methods. *Neuropsychol. Rev.* **25**, 289–313 (2015).
323. Poldrack, R. A., Nichols, T. & Mumford, J. *Handbook of Functional MRI Data Analysis*. (Cambridge University Press, 2011). doi:10.1017/CBO9780511895029.
324. Friston, K. J. *et al.* Spatial registration and normalization of images. *Hum. Brain Mapp.* **3**, 165–189 (1995).
325. Worsley, K. J. & Friston, K. J. Analysis of fMRI time-series revisited--again. *NeuroImage* **2**, 173–181 (1995).
326. Fischl, B. *et al.* Automatically Parcellating the Human Cerebral Cortex. *Cereb. Cortex* **14**, 11–22 (2004).
327. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage* **9**, 179–194 (1999).
328. Ashburner, J. & Friston, K. J. Voxel-based morphometry--the methods. *NeuroImage* **11**, 805–821 (2000).
329. Fischl, B., Sereno, M. I., Tootell, R. B. & Dale, A. M. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* **8**, 272–284 (1999).
330. Wang, J. *et al.* Age estimation using cortical surface pattern combining thickness with curvatures. *Med. Biol. Eng. Comput.* **52**, 331–341 (2014).
331. Schaer, M. *et al.* A surface-based approach to quantify local cortical gyrification. *IEEE Trans. Med. Imaging* **27**, 161–170 (2008).

332. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).
333. Destrieux, C., Fischl, B., Dale, A. & Halgren, E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage* **53**, 1–15 (2010).
334. Ségonne, F., Grimson, E. & Fischl, B. Topological Correction of Subcortical Segmentation. in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003* (eds. Ellis, R. E. & Peters, T. M.) 695–702 (Springer, 2003). doi:10.1007/978-3-540-39903-2\_85.
335. Fortin, J.-P. *et al.* Harmonization of multi-site diffusion tensor imaging data. *NeuroImage* **161**, 149–170 (2017).
336. Fortin, J.-P. *et al.* Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* **167**, 104–120 (2018).
337. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11**, 499–511 (2010).
338. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype Imputation. *Annu. Rev. Genomics Hum. Genet.* **10**, 387–406 (2009).
339. Couzin, J. Human genome. HapMap launched with pledges of \$100 million. *Science* **298**, 941–942 (2002).
340. Chang, B.-L. *et al.* Fine mapping association study and functional analysis implicate a SNP in MSMB at 10q11 as a causal variant for prostate cancer risk. *Hum. Mol. Genet.* **18**, 1368–1375 (2009).
341. Marsh, S. G. E. *et al.* Nomenclature for factors of the HLA system, 2010. *Tissue Antigens* **75**, 291–455 (2010).
342. Fan, Y. & Song, Y.-Q. PyHLA: tests for the association between HLA alleles and diseases. *BMC Bioinformatics* **18**, 90 (2017).
343. Robinson, J. *et al.* The IMGT/HLA database. *Nucleic Acids Res.* **41**, D1222–D1227 (2013).
344. Lai, M.-C., Lombardo, M. V. & Baron-Cohen, S. Autism. *The Lancet* **383**, 896–910 (2014).
345. Salazar, F. *et al.* Co-occurring Psychiatric Disorders in Preschool and Elementary School-Aged Children with Autism Spectrum Disorder. *J. Autism Dev. Disord.* **45**, 2283–2294 (2015).
346. Goodman, R., Ford, T., Simmons, H., Gatward, R. & Meltzer, H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Br. J. Psychiatry* **177**, 534–539 (2000).
347. Birmaher, B. *et al.* The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale Construction and Psychometric Characteristics. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 545–553 (1997).
348. Stringaris, A. *et al.* The Affective Reactivity Index: a concise irritability scale for clinical and research settings. *J. Child Psychol. Psychiatry* **53**, 1109–1117 (2012).
349. Achenbach, T. M., Edelbrock, C. & Howell, C. T. Empirically based assessment of the behavioral/emotional problems of 2- and 3- year-old children. *J. Abnorm. Child Psychol.* **15**, 629–650 (1987).



350. Jain, A. K. & Dubes, R. C. *Algorithms for Clustering Data*. (Prentice-Hall, Inc., 1988).
351. Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **6**, 461–464 (1978).
352. Fischl, B. *et al.* Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
353. Rosen, A. F. G. *et al.* Quantitative assessment of structural image quality. *NeuroImage* **169**, 407–418 (2018).
354. Fortin, J.-P. *et al.* Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* **167**, 104–120 (2018).
355. Coon, H. *et al.* Genome-wide linkage using the Social Responsiveness Scale in Utah autism pedigrees. *Mol. Autism* **1**, 8 (2010).
356. Pan, P.-Y., Tammimies, K. & Bölte, S. The Association Between Somatic Health, Autism Spectrum Disorder, and Autistic Traits. *Behav. Genet.* (2019) doi:10.1007/s10519-019-09986-3.
357. Tehrani-Doost, M., Shahrivar, Z., Pakbaz, B., Rezaie, A. & Ahmadi, F. Normative Data and Psychometric Properties of the Child Behavior Checklist and Teacher Rating Form in an Iranian Community Sample. *Iran. J. Pediatr.* **21**, 331–342 (2011).
358. Achenbach, T. M. *Manual for the child behavior checklist/4-18 and 1991 profile*. (Burlington, VT : Dept. of Psychiatry, University of Vermont, 1991).
359. Carballo, J. J., Rodríguez-Blanco, L., García-Nieto, R. & Baca-García, E. Screening for the ADHD Phenotype Using the Strengths and Difficulties Questionnaire in a Clinical Sample of Newly Referred Children and Adolescents. *J. Atten. Disord.* **22**, 1032–1039 (2018).
360. Schaer, M. *et al.* Decreased frontal gyrification correlates with altered connectivity in children with autism. *Front. Hum. Neurosci.* **7**, (2013).
361. Utevsky, A. V., Smith, D. V. & Huettel, S. A. Precuneus Is a Functional Core of the Default-Mode Network. *J. Neurosci.* **34**, 932–940 (2014).
362. Bruner, E., Preuss, T. M., Chen, X. & Rilling, J. K. Evidence for expansion of the precuneus in human evolution. *Brain Struct. Funct.* **222**, 1053–1060 (2017).
363. Margulies, D. S. *et al.* Precuneus shares intrinsic functional architecture in humans and monkeys. *Proc. Natl. Acad. Sci.* **106**, 20069–20074 (2009).
364. Jackson, R. L., Bajada, C. J., Rice, G. E., Cloutman, L. L. & Lambon Ralph, M. A. An emergent functional parcellation of the temporal cortex. *NeuroImage* **170**, 385–399 (2018).
365. Wallace, G. L. *et al.* Cortical and Subcortical Abnormalities in Youths with Conduct Disorder and Elevated Callous-Unemotional Traits. *J. Am. Acad. Child Adolesc. Psychiatry* **53**, 456–65.e1 (2014).
366. Hyatt, C. J., Haney-Caron, E. & Stevens, M. C. Cortical thickness and folding deficits in conduct-disordered adolescents. *Biol. Psychiatry* **72**, 207–214 (2012).
367. Zhu, W., Zhou, X. & Xia, L.-X. Brain structures and functional connectivity associated with individual differences in trait proactive aggression. *Sci. Rep.* **9**, 7731 (2019).
368. Di Martino, A. *et al.* Relationship Between Cingulo-Insular Functional Connectivity and Autistic Traits in Neurotypical Adults. *Am. J. Psychiatry* **166**, 891–899 (2009).

369. Nagahama, Y. *et al.* Transient neural activity in the medial superior frontal gyrus and precuneus time locked with attention shift between object features. *NeuroImage* **10**, 193–199 (1999).
370. Thach, W. T. Correlation of neural discharge with pattern and force of muscular activity, joint position, and direction of intended next movement in motor cortex and cerebellum. *J. Neurophysiol.* **41**, 654–676 (1978).
371. Georgopoulos, A. P., Kalaska, J. F., Caminiti, R. & Massey, J. T. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. Neurosci. Off. J. Soc. Neurosci.* **2**, 1527–1537 (1982).
372. Chouinard, P. A. & Paus, T. What have We Learned from ‘Perturbing’ the Human Cortical Motor System with Transcranial Magnetic Stimulation? *Front. Hum. Neurosci.* **4**, 173 (2010).
373. Buckner, R. L., Andrews-Hanna, J. R. & Schacter, D. L. The brain’s default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* **1124**, 1–38 (2008).
374. Bhat, A. N., Galloway, J. C. & Landa, R. J. Relation between early motor delay and later communication delay in infants at risk for autism. *Infant Behav. Dev.* **35**, 838–846 (2012).
375. Brian, J. *et al.* Clinical assessment of autism in high-risk 18-month-olds. *Autism Int. J. Res. Pract.* **12**, 433–456 (2008).
376. Landa, R. & Garrett-Mayer, E. Development in infants with autism spectrum disorders: a prospective study. *J. Child Psychol. Psychiatry* **47**, 629–638 (2006).
377. Kohli, J. S. *et al.* Local Cortical Gyrification is Increased in Children With Autism Spectrum Disorders, but Decreases Rapidly in Adolescents. *Cereb. Cortex N. Y. N 1991* (2018) doi:10.1093/cercor/bhy111.
378. Pierce, K., Müller, R. A., Ambrose, J., Allen, G. & Courchesne, E. Face processing occurs outside the fusiform ‘face area’ in autism: evidence from functional MRI. *Brain J. Neurol.* **124**, 2059–2073 (2001).
379. Zhao, Y. *et al.* Altered spontaneous neural activity in the occipital face area reflects behavioral deficits in developmental prosopagnosia. *Neuropsychologia* **89**, 344–355 (2016).
380. Rausch, M., Spengler, F. & Eysel, U. T. Proprioception acts as the main source of input in human S-I activation experiments: a functional MRI study. *Neuroreport* **9**, 2865–2868 (1998).
381. Chudler, E. H., Anton, F., Dubner, R. & Kenshalo, D. R. Responses of nociceptive SI neurons in monkeys and pain sensation in humans elicited by noxious thermal stimulation: effect of interstimulus interval. *J. Neurophysiol.* **63**, 559–569 (1990).
382. Green, S. A. *et al.* Overreactive brain responses to sensory stimuli in youth with autism spectrum disorders. *J. Am. Acad. Child Adolesc. Psychiatry* **52**, 1158–1172 (2013).
383. Liss, M., Saulnier, C., Fein, D. & Kinsbourne, M. Sensory and attention abnormalities in autistic spectrum disorders. *Autism Int. J. Res. Pract.* **10**, 155–172 (2006).
384. Cascio, C. J. *et al.* Perceptual and neural response to affective tactile texture stimulation in adults with autism spectrum disorders. *Autism Res. Off. J. Int. Soc. Autism Res.* **5**, 231–244 (2012).

385. Redcay, E. The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. *Neurosci. Biobehav. Rev.* **32**, 123–142 (2008).
386. Iidaka, T., Miyakoshi, M., Harada, T. & Nakai, T. White matter connectivity between superior temporal sulcus and amygdala is associated with autistic trait in healthy humans. *Neurosci. Lett.* **510**, 154–158 (2012).
387. Mensen, V. T. *et al.* Development of cortical thickness and surface area in autism spectrum disorder. *NeuroImage Clin.* **13**, 215–222 (2016).
388. Loomes, R., Hull, L. & Mandy, W. P. L. What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 466–474 (2017).
389. Li, G., Lin, W., Gilmore, J. H. & Shen, D. Spatial Patterns, Longitudinal Development, and Hemispheric Asymmetries of Cortical Thickness in Infants from Birth to 2 Years of Age. *J. Neurosci.* **35**, 9150–9162 (2015).
390. Li, G. *et al.* Mapping Longitudinal Development of Local Cortical Gyrification in Infants from Birth to 2 Years of Age. *J. Neurosci.* **34**, 4228–4238 (2014).
391. Klopper, F., Testa, R., Pantelis, C. & Skafidas, E. A cluster analysis exploration of autism spectrum disorder subgroups in children without intellectual disability. *Res. Autism Spectr. Disord.* **36**, 66–78 (2017).
392. Ousley, O. & Cermak, T. Autism Spectrum Disorder: Defining Dimensions and Subgroups. *Curr. Dev. Disord. Rep.* **1**, 20–28 (2014).
393. DuPaul, G. J., Power, T. J., Anastopoulos, A. D. & Reid, R. *ADHD Rating Scale—IV: Checklists, norms, and clinical interpretation*. viii, 79 (Guilford Press, 1998).
394. Bodfish, J. W., Symons, F. J., Parker, D. E. & Lewis, M. H. Varieties of Repetitive Behavior in Autism: Comparisons to Mental Retardation. *J. Autism Dev. Disord.* **30**, 237–243 (2000).
395. Goodman, R., Ford, T., Richards, H., Gatward, R. & Meltzer, H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J. Child Psychol. Psychiatry* **41**, 645–655 (2000).
396. Choi, S. W. & O'Reilly, P. F. PRSice-2: Polygenic Risk Score software for biobank-scale data. *GigaScience* **8**, (2019).
397. Antezana, L. *et al.* Gender differences in restricted and repetitive behaviors and interests in youth with autism. *Autism Res. Off. J. Int. Soc. Autism Res.* **12**, 274–283 (2019).
398. Esbensen, A. J., Seltzer, M. M., Lam, K. S. L. & Bodfish, J. W. Age-related Differences in Restricted Repetitive Behaviors in Autism Spectrum Disorders. *J. Autism Dev. Disord.* **39**, 57–66 (2009).
399. McGoey, K. E., DuPaul, G. J., Haley, E. & Shelton, T. L. Parent and Teacher Ratings of Attention-Deficit/Hyperactivity Disorder in Preschool: The ADHD Rating Scale-IV Preschool Version. *J. Psychopathol. Behav. Assess.* **29**, 269–276 (2007).
400. Goodman, A., Heiervang, E., Collishaw, S. & Goodman, R. The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. *Soc. Psychiatry Psychiatr. Epidemiol.* **46**, 521–532 (2011).

401. Poquérousse, J., Pastore, L., Dellantonio, S. & Esposito, G. Alexithymia and Autism Spectrum Disorder: A Complex Relationship. *Front. Psychol.* **9**, (2018).
402. Li, M. *et al.* Abnormal cortical thickness in heroin-dependent individuals. *NeuroImage* **88**, 295–307 (2014).
403. Li, Y. *et al.* Reduced frontal cortical thickness and increased caudate volume within fronto-striatal circuits in young adult smokers. *Drug Alcohol Depend.* **151**, 211–219 (2015).
404. Sliwinska, M. W. W., Khadilkar, M., Campbell-Ratcliffe, J., Quevenco, F. & Devlin, J. T. Early and Sustained Supramarginal Gyrus Contributions to Phonological Processing. *Front. Psychol.* **3**, (2012).
405. Molent, C. *et al.* Reduced cortical thickness and increased gyrification in generalized anxiety disorder: a 3 T MRI study. *Psychol. Med.* **48**, 2001–2010 (2018).
406. Pedapati, E. V. *et al.* Motor cortex facilitation: a marker of attention deficit hyperactivity disorder co-occurrence in autism spectrum disorder. *Transl. Psychiatry* **9**, 298 (2019).
407. Kern, J. K., Geier, D. A., Sykes, L. K., Geier, M. R. & Deth, R. C. Are ASD and ADHD a Continuum? A Comparison of Pathophysiological Similarities Between the Disorders. *J. Atten. Disord.* **19**, 805–827 (2015).
408. Banker, L. & Tadi, P. Neuroanatomy, Precentral Gyrus. in *StatPearls* (StatPearls Publishing, 2020).
409. Estes, A. *et al.* Long-Term Outcomes of Early Intervention in 6-Year-Old Children With Autism Spectrum Disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **54**, 580–587 (2015).
410. Mitchell, R. H. B. & Goldstein, B. I. Inflammation in Children and Adolescents With Neuropsychiatric Disorders: A Systematic Review. *J. Am. Acad. Child Adolesc. Psychiatry* **53**, 274–296 (2014).
411. Finegold, S. M. Desulfovibrio species are potentially important in regressive autism. *Med. Hypotheses* **77**, 270–274 (2011).
412. Bennabi, M. *et al.* HLA-class II haplotypes and Autism Spectrum Disorders. *Sci. Rep.* **8**, 7639 (2018).
413. Zheng, X. *et al.* HIBAG—HLA genotype imputation with attribute bagging. *Pharmacogenomics J.* **14**, 192–200 (2014).
414. Hanaie, R. *et al.* White matter volume in the brainstem and inferior parietal lobule is related to motor performance in children with autism spectrum disorder: A voxel-based morphometry study. *Autism Res. Off. J. Int. Soc. Autism Res.* **9**, 981–992 (2016).
415. Wieber, K., Zimmer, C. L. & Hertl, M. Detection of autoreactive CD4+ T cells by MHC class II multimers in HLA-linked human autoimmune diseases. *J. Clin. Invest.* **131**, (2021).
416. Sette, A. & Sidney, J. HLA supertypes and supermotifs: a functional perspective on HLA polymorphism. *Curr. Opin. Immunol.* **10**, 478–482 (1998).
417. Olerup, O. & Hillert, J. HLA class II-associated genetic susceptibility in multiple sclerosis: A critical evaluation. *Tissue Antigens* **38**, 1–15 (1991).
418. Noble, J. A. *et al.* HLA Class I and Genetic Susceptibility to Type 1 Diabetes: Results From the Type 1 Diabetes Genetics Consortium. *Diabetes* **59**, 2972–2979 (2010).

419. Johnson, W. G. *et al.* HLA-DR4 as a risk allele for autism acting in mothers of probands possibly during pregnancy. *Arch. Pediatr. Adolesc. Med.* **163**, 542–546 (2009).
420. Torres, A. R. *et al.* Common Genetic Variants Found in HLA and KIR Immune Genes in Autism Spectrum Disorder. *Front. Neurosci.* **10**, 463 (2016).
421. Bassetti, C. L. A. *et al.* Narcolepsy - clinical spectrum, aetiopathophysiology, diagnosis and treatment. *Nat. Rev. Neurol.* **15**, 519–539 (2019).
422. Miyagawa, T. & Tokunaga, K. Genetics of narcolepsy. *Hum. Genome Var.* **6**, 4 (2019).
423. Zivadinov, R. *et al.* HLA-DRB1\*1501, -DQB1\*0301, -DQB1\*0302, -DQB1\*0602, and -DQB1\*0603 alleles are associated with more severe disease outcome on MRI in patients with multiple sclerosis. *Int. Rev. Neurobiol.* **79**, 521–535 (2007).
424. Posar, A., Visconti, P., Blunda, V., Pizza, F. & Plazzi, G. Autism Spectrum Disorder and Narcolepsy: A Possible Connection That Deserves to Be Investigated. *Front. Psychiatry* **11**, (2020).
425. Cryan, J. F., O’Riordan, K. J., Sandhu, K., Peterson, V. & Dinan, T. G. The gut microbiome in neurological disorders. *Lancet Neurol.* **19**, 179–194 (2020).
426. Nagano, T. *et al.* Novel cancer therapy targeting microbiome. *OncoTargets Ther.* **12**, 3619–3624 (2019).
427. Kočovská, E., Gaughran, F., Krivoy, A. & Meier, U.-C. Vitamin-D Deficiency As a Potential Environmental Risk Factor in Multiple Sclerosis, Schizophrenia, and Autism. *Front. Psychiatry* **8**, 47 (2017).
428. Burkill, S. *et al.* The DQB1\*03:02 Genotype and Treatment for Pain in People With and Without Multiple Sclerosis. *Front. Neurol.* **11**, (2020).
429. Moutsianas, L. *et al.* Class II HLA interactions modulate genetic risk for multiple sclerosis. *Nat. Genet.* **47**, 1107–1113 (2015).
430. Ramgopal, S. *et al.* Interaction of HLA-DRB1\* alleles and CTLA4 (+49 AG) gene polymorphism in Autoimmune Thyroid Disease. *Gene* **642**, 430–438 (2018).
431. Uddin, L. Q. & Menon, V. The anterior insula in autism: under-connected and under-examined. *Neurosci. Biobehav. Rev.* **33**, 1198–1203 (2009).

**Titre :** Approches multidimensionnelles pour la caractérisation de l'hétérogénéité des troubles du spectre autistique

**Mots clés :** Neuroimagerie, IRM Structurale, Analyses Multivariées, Autisme, Psychiatrie, Comportement

**Résumé :** Le trouble du spectre autistique (TSA) est un trouble neuro-développemental caractérisé par une altération des interactions sociales et de la communication, ainsi que par des intérêts restreints et répétitifs. Un défi important dans la compréhension du TSA et l'identification de biomarqueurs tient à son hétérogénéité. En effet, près de 70% des patients présentent une condition psychiatrique ou médicale supplémentaire. Ces conditions se caractérisent aussi par des altérations génétiques, immunitaires ou des différences dans la neuro-imagerie cérébrale. Cette hétérogénéité est souvent observable dans les études cas-témoins regroupant des patients dont les profils comportementaux, génétiques et / ou cliniques sont variables au sein d'un même groupe. L'objectif global de ce travail de thèse est de mieux caractériser les

patients autistes, ce qui est vital pour l'avancement du diagnostic et de thérapies appropriés. Pour ce faire, nous avons utilisé des méthodes statistiques et des données multidimensionnelles. Nous montrons que les comparaisons cas-témoins dans les populations autistes ne parviennent pas à fournir des résultats cohérents et significatifs, et qu'une combinaison d'approches dimensionnelles associée à un sous-groupage de patients s'avère plus précieuse dans la compréhension du TSA. Ce travail de thèse fournit des preuves de sous-types d'autisme et confirme la nécessité d'appliquer des approches dimensionnelles et des stratégies de sous-groupes afin d'extraire des traits significatifs chez les patients atteints de TSA.

**Title :** Applying Multidimensional Approaches to Disentangle Autism Spectrum Disorder Heterogeneity

**Keywords :** Neuroimaging, Structural MRI, Multivariate Analysis, Autism, Psychiatry, Behavior

**Abstract :** Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social, communication and restricted and repetitive behaviors. A significant challenge in understanding ASD lies in its heterogeneity, with up to 70% of patients reporting comorbid psychiatric, medical or genetic conditions, as well as vast differences in neuroimaging, genetic and immune factors. This variability has hindered biomarker isolation, possibly due to widely used case-control experimental designs combining patients varying in behavioral, genetic and/or clinical profiles into one

group. The global aim of this thesis work is to *better characterize autistic patients*, which is vital in the advancement of appropriate treatments and therapies. To achieve this, we used statistical methods and multidimensional data. We show that case-control comparisons in autistic populations fail to elicit consistent and meaningful results, and that a combination of dimensional and subgrouping approaches proves most valuable in the understanding of ASD. In the process, we isolated important and consistent behavioral and neuroimaging autistic traits subgroups.

