Investigating Individual Variability of Morphometric Covariance Patterns in Autism Spectrum Disorder

Justine Ziolkowski

Integrated Program in Neuroscience
McGill University, Montreal

November 5, 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Table of Contents

Abstract

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by substantial phenotypic and neuroanatomical variation. Atypicalities in multiple macrostructural properties have been implicated in ASD. However, canonical approaches to modelling cortical architecture consist of assessing case-control differences in discrete neuroanatomical metrics. We therefore evaluated individual variability in morphometric covariance patterns across four cortical measures in ASD and neurotypical subjects. These patterns were then related to behavioural characteristics that affect phenotypic heterogeneity. A total of 1097 participants were pooled across three datasets. Following quality control and matching of diagnostic groups for age and sex, 486 subjects aged 4-52 remained (243 ASD; mean age = 19.88 ± 9.6; 166 females). Vertex-wise cortical thickness (CT), surface area (SA), local gyrification index (LGI) and mean curvature (MC) estimates were extracted. The measures were corrected for batch effects in a multi-parameter framework with the CovBat algorithm. Non-negative matrix factorization decomposed cortical covariance into 16 components that describe the spatial location of covarying vertices and subject-wise loading coefficients that indicate each metric's contribution to the structural components. Partial least squares correlation was used to associate macrostructural covariance with behavioural measures. The main latent variable (LV; 87.4% covariance explained; $p<0.0001$) describes a significant correlation between younger age, female sex, and lower fullscale IQ, with a cortical profile characterized by increased whole-brain contribution of CT and LGI, as well as decreased contribution of SA to covariance patterns. Diagnostic group does not reach significance, although ASD subjects with reduced social abilities and elevated sensory sensitivities display a stronger expression of the LV brain-behaviour phenotype. These results suggest a neuroanatomical covariance signature that relates robustly to developmental and cognitive domains but has limited clinical relevance.

Résumé

Les troubles du spectre de l'autisme (TSA) sont des maladies neurodéveloppementales caractérisées par des variations phénotypiques et neuroanatomiques considérables. Des anomalies au niveau de multiples propriétés macrostructurelles ont été impliquées dans le TSA. Toutefois, les approches standards modélisant l'architecture corticale consistent à l'évaluation des variables neuroanatomiques individuelles par rapport à des cas-contrôles. Par conséquent, nous avons évalué la variabilité individuelle des motifs de covariance morphométrique au niveau de quatre mesures corticales au sein de sujets neurotypiques et TSA. Ces motifs ont ensuite été reliés à des caractéristiques comportementales qui affectent l'hétérogénéité phénotypique. Un total de 1097 participants ont été sélectionnés à travers trois bases de données. Suite à un contrôle de qualité et à la correspondance des groupes diagnostiques selon l'âge et le sexe, 486 sujets âgés de 4 à 52 ans ont été retenus (243 TSA; âge moyen = 19.88 ± 9.6; 166 femmes). À chaque vertex, les estimations de l'épaisseur corticale (EC), de l'aire de surface (AS), de l'indice de gyrification local (IGL) et la courbure moyenne (CM) ont été extraites. Les mesures ont été corrigées pour l'effet de site selon une approche multiparamétrique avec l'algorithme CovBat. La factorisation matricielle non-négative a décomposé la covariance corticale en 16 composants qui décrivent la localisation spatiale des vertex qui covarient ensemble et les coefficients de pondération qui indiquent l'ampleur de la contribution de chaque métrique sur les composants structurels. L'analyse de corrélation de la régression des moindres carrés partiels a été utilisée afin d'associer les profils de covariances macrostructurelles et les mesures comportementales. Les principales variables latentes (VL; 87.4% covariance expliquée; $p<0.0001$) décrivent une corrélation significative entre le jeune âge, le sexe féminin et un quotient intellectuel global bas associé à un profil cortical caractérisé par une hausse de la contribution totale du cerveau pour l'EC et l'IGL, ainsi qu'une baisse de la contribution de l'AS aux motifs de covariance. Le groupe diagnostique n'a pas atteint de résultats significatifs, bien que les sujets TSA arborant des habiletés sociales réduites et une sensibilité sensorielle élevée présentent une expression forte de la VL du phénotype cerveau-comportement. Ces résultats suggèrent une signature neuroanatomique covariant de façon robuste avec les variables développementales et cognitives, mais ayant une pertinence plus limitée d'un point de vue clinique.

Acknowledgements

This academic journey would not have been possible without the intellectual and moral support of many individuals. First, I am thankful for the mentorship of **Dr. Mallar Chakravarty**. Mallar, thank you for being incredibly generous with your time, energy, and brain power. I am grateful for the care you put into creating an inclusive lab environment, for prioritizing our wellbeing, and looking out for our future selves before we have the foresight to do so. This experience would have been infinitely more challenging without your kindness and enthusiasm.

I am deeply appreciative of the friendship and collegiality of the CoBrA lab members with whom I've had the pleasure of working. In particular, a big thank you to my NMF Buddies, **Dr. Raihaan Patel**, **Alyssa Dai**, and **Dr. Hadis Kalantar-Hormozi**. Thank you for bearing my endless Slack messages with good humour, for being my 24/7 helpline, and most importantly, for frequent doses of positivity during the dreary days of COVID lockdowns. I am profoundly grateful to **Manuela Costantino** for your help with my sanity and QC; to **Mila Urosevic** for being my thesis partner, physics guru, and tea companion; to **Corinne Robert** and **David MacDonald** for your statistical insights; to **Gabriel Desrosiers-Grégoire** and **Dr. Gabriel Devenyi** for your patience in explaining technicalities; to **Olivier Parent** for feedback on drafts of my writing; and to **Lani Cupo**, **Cindy Lucero Garcia**, and **Annie Phan** for being rays of sunshine.

I am thankful for the academic guidance of my Committee Members, **Dr. Emma Duerden**, **Dr. Sébastien Jacquemont**, and **Dr. Yashar Zeighami**, as well as my IPN mentor, **Dr. Joseph Rochford,** who have infused my thinking with new perspectives. I am indebted to all participants, research staff, and collaborators at SickKids Hospital, POND, and UK AIMS who have made this work possible. Thank you to my funding bodies, the **Natural Sciences and Engineering Research Council of Canada, Transforming Autism Care Consortium**, and **Québec Bio-Imaging Network**.

Last, but certainly not least, I am fortunate to have the support of family, friends, and housemates, especially **Barbara, Jozef, Margaret**, **John**, **Vince**, **Paulina**, **Konrad**, **Adam**, **Sara**, **Christina**, **Alyssa, Catherine**, **Camille**, **Starr**, **Éléa**, **Annika**, and **Jericha**. Thank you for reminding me to turn off my brain from time to time, for supplementing my minimalist diet, and for sending care packages and encouragement from afar.

Contribution of Authors

The original work presented in this thesis was produced independently by the author, **Justine Ziolkowski,** who performed data processing, quality control, analysis, visualization, interpretation, and writing. Several co-authors contributed significantly in the following ways:

- **Dr. Mallar Chakravarty** conceptualized and supervised the project, provided guidance with analysis and interpretation of results, and reviewed drafts of the thesis.
- **Dr. Raihaan Patel** provided code for NMF analyses and visualization, as well as assistance with NMF implementation and interpretation.
- **Alyssa Dai** provided code and methodological support for optimization of NMF analyses and visualization.
- **Dr. Hadis Kalantar-Hormozi** provided methodological support for NMF optimization and interpretation.
- **Manuela Costantino** assisted with quality control for a subset of Province of Ontario Neurodevelopmental Disorders Network (POND) subjects.
- **Dr. Evdokia Anagnostou**, **Dr. Jason P. Lerch,** and **Dr. Margot J. Taylor** are collaborators who shared data from the Hospital for Sick Children (Toronto, ON, Canada) dataset.
- **Dr. Edward T. Bullmore**, **Dr. Lindsay R. Chura**, **Dr. Simon Baron-Cohen**, **Dr. Dorothea L. Floris**, **Dr. Rosemary J. Holt**, **Dr. Meng-Chuan Lai**, **Dr. Michael V. Lombardo**, **Dr. Amber N.V. Ruigrok** at the University of Cambridge (Cambridge, UK) are collaborators who shared data from the Cambridge Site of the UK Medical Research Council Autism Imaging Multisite Study (UK AIMS) dataset.
- **Dr. Michael C. Craig**, **Dr. Christine Ecker** and **Dr. Declan G. M. Murphy** at King's College London (London, UK) are collaborators who shared data from the King's College London site of the UK AIMS dataset.
- **Dr. Evdokia Anagnostou**, **Dr. Muhammad Ayub, Dr. Jennifer Crosbie, Dr. Stelios Georgiades, Dr. Jessica Jones, Dr. Elizabeth Kelley, Dr. Jason P. Lerch**, **Dr. Xudong Liu**, **Dr. Russell Schachar,** and **Dr. Margot J. Taylor** are collaborators who shared data from the POND dataset. **Dr. Alana Iaboni** managed the POND dataset.
- **Dr. Russell Shinohara** provided code and technical consultation for CovBat implementation.

List of Figures

List of Tables

List of Abbreviations

ADHD: Attention Deficit Hyperactivity Disorder

ADI: Autism Diagnostic Interview

ADOS: Autism Diagnostic Observation Schedule

ASD: Autism Spectrum Disorder:

BEaST: Brain Extraction based on nonlocal Segmentation Technique

BSR: Bootstrap Ratio

CLASP: Constrained Laplacian Anatomical Segmentation using Proximities

ComBat: Combating Batch Effects when Combining Batches

CovBat: Correcting Covariance Batch Effects

CSF: Cerebrospinal Fluid

CT: Cortical Thickness

DSM: Diagnostic and Statistical Manual of Mental Disorders

FA: Flip Angle

FDR: False Discovery Rate

FIQ: Fullscale Intelligence Quotient

fMRI: Functional Magnetic Resonance Imaging

FOV: Field of View

FWHM: Full-Width at Half-Maximum

GI: Gyrification Index

GM: Grey Matter

GMV: Grey Matter Volume

ICA: Independent Component Analysis

IQ: Intelligence Quotient

LGI: Local Gyrification Index

LV: Latent Variable

MC: Mean Curvature

MNI: Montreal Neurological Institute

MPRAGE: Magnetization Prepared Rapid Gradient Echo

MRI: Magnetic Resonance Imaging

N3: Nonparametric Nonuniform Intensity Normalization Correction

NMF: Non-negative Matrix Factorization

NNDSVD: Non-negative Double Singular Value Decomposition

NT: Neurotypical

PCA: Principal Component Analysis

PLSC: Partial Least Squares Correlation

PLSR: Partial Least Squares Regression

POND: Province of Ontario Neurodevelopmental Disorders Network

PVE: Partial Volume Effects

QC: Quality Control

RF: Radiofrequency

RSE: Residual Standard Error

SA: Surface Area

SVD: Singular Value Decomposition

T: Tesla

T1w: T1-weighted

T2w: T2-weighted

TBV: Total Brain Volume

TE: Echo Time

TI: Inversion Time

TR: Repetition Time

WASI: Wechsler Abbreviated Scale of Intelligence

WISC: Wechsler Intelligence Scale for Children

WM: White Matter

WPPSI: Wechsler Preschool & Primary Scale of Intelligence

UK AIMS: UK Medical Research Council Autism Imaging Multicentre Study

1. Introduction

Autism spectrum disorder (ASD) is a highly heterogeneous neurodevelopmental disorder that manifests variably according to age, sex, and intelligence quotient (IQ; Daniels & Mandell, 2014; Ruigrok & Lai, 2020; Ryland et al., 2014). Perturbations in neurodevelopmental processes disrupt brain maturation in ASD (Courchesne et al., 2011; Marchetto et al., 2017; Stoner et al., 2014; Toma et al., 2012), which is reflected across features of cortical macrostructure (Bedford et al., 2020; Hammill et al., 2021; Hazlett et al., 2011; Khundrakpam et al., 2017; Piven et al., 1990; van Rooij et al., 2018; Yang et al., 2016). However, decades of neuroimaging studies have failed to identify a reproducible cortical profile of ASD, despite growing sample sizes, accessibility of deeply phenotyped datasets, and methodological improvements in cortical modelling software. The inability to reconcile neuroanatomical substrates with clinical phenotypes poses a fundamental challenge to conceptualizing ASD and impedes progress in identifying neurobiological markers. These discrepancies may stem, in part, from the predominance of case-control paradigms, which yield limited subject-wise specificity. Additionally, traditional mass-univariate analyses assessing diagnostic effects on cortical measures are necessarily constrained to model a single outcome variable, which limits the understanding of interdependencies between multiple structural properties that jointly comprise the cortex. Together, these challenges suggest that standard analytical approaches may be ill-suited to capture the complexity of cortical organization in ASD. Statistical frameworks that model individual variation at the intersection of neurobiological and clinical signatures of ASD may yield more representative and clinically robust models.

Therefore, the overarching aim of this thesis is to employ a complement of multivariate analyses to evaluate how individual variation in the structural covariance patterns of four cortical measures relates to each subject's unique set of demographic and behavioural features known to influence ASD heterogeneity. More specifically, the aims are as follows: 1) evaluate individual variability in structural covariance patterns of cortical thickness (CT), surface area (SA), local gyrification index (LGI), and mean curvature (MC); and 2) relate these covariance patterns to diagnosis, age, sex, and fullscale IQ (FIQ). To address Aim 1, orthogonal projective Non-negative Matrix Factorization (NMF) is used to identify spatial patterns where the four cortical indices co-occur in consistent ways across individuals. This method segregates whole-brain covariance patterns into spatially discrete parcels of lower dimensionality (Lee & Seung,

1999; Sotiras et al., 2015) such that each individual's unique macrostructural profile is represented for a given component. Importantly, performing this analysis on a pooled group of ASD and neurotypical (NT) subjects situates individuals along a continuum of variation to model covariance patterns dimensionally, being unconstrained by *a priori* diagnostic labels. To address Aim 2, Partial Least Squares Correlation (PLSC) analysis is used to identify latent variables (LVs) that link each subject's macrostructural covariance profile (identified in Aim 1) to individual variation in phenotypic measures. This associative technique finds underlying latent dimensions that represent a composite of maximally associated brain and behavioural elements (Krishnan et al., 2011). Together, this work aims to offer a more comprehensive and individualized characterization of cortical architecture in ASD.

2. Comprehensive Literature Review

2.1 Normative Brain Development

Anatomical origins of the human brain are anchored in prenatal neurodevelopment. The precise sequence and timing of early maturational events shape neuroanatomy across spatiotemporal scales to influence lifelong cognitive and health outcomes. Characterizing elements of normative brain growth is therefore essential to understand how developmental dysfunctions across different levels of neural organization may culminate in a disorder such as ASD.

2.1.1 Mechanisms of Prenatal Neurodevelopment

Neurodevelopment involves a cascade of highly coordinated cellular processes that occur at specific phases of embryonic and fetal maturation. These mechanisms are tightly regulated by genetic and environmental influences, underpinning both the structural framework of the central nervous system and its functional capacities.

Cortical morphogenesis begins during the third week of gestation with neurulation, when the neural plate is inducted after the formation of the notochord. The neural groove emerges around 23 days post conception, with folds arising and then fusing on each side of the plate as it transforms into the neural tube (Silbereis et al., 2016). During the fourth week of gestation, neuroepithelial cell populations in the ventricular zone divide symmetrically to

produce daughter progenitors, establishing a pool of multipotent founder progenitors (Budday et al., 2015). Neurogenesis begins at 40 days post conception, as radial glial cells divide to generate postmitotic neurons and intermediate progenitor cells, which themselves yield neuronal progeny. A subset of intermediate progenitors delaminate from the ventricular zone to form the subventricular proliferative region, where their continued division contributes to the expanding neuronal population. At approximately seven to eight weeks post conception, neurons born in both proliferative compartments migrate sequentially along radial glial fibers to form the layered cortical plate, where they undergo differentiation and form connections with neighbouring neurons through the emergence of synaptogenesis, dendritic/axonal arborization, and myelination (Budday et al., 2015; Jiang & Nardelli, 2016; Kowalczyk et al., 2009; Rakic et al., 1988; Silbereis et al., 2016). Cell proliferation, migration, and differentiation persist throughout prenatal development, under temporal control of key molecular and signaling factors (Jiang & Nardelli, 2016; Sun & Hevner, 2014).

These organizational principles unfold along both radial and tangential axes of the cortex, giving rise to its distinctive laminar and columnar organization. Six neocortical layers are arranged parallel to the pial surface, with each lamina characterized by distinct cellular density and configuration of synaptic and axonal projections (Wagstyl & Lerch, 2018). Perpendicularly, cortical columns function as processing units, spanning the layers along the vertical aspect (Molnar & Rockland, 2020; Mountcastle, 1997; Rakic, 2009). The cerebral cortex ultimately forms a neuronal sheet encasing the cerebrum, composed of projection and interneuron cell bodies, in addition to neuropil, vasculature, extracellular space, free water, glial cells, and mineral deposits (Rakic et al., 2009; Wagstyl & Lerch, 2018). Although the production of new neurons is almost entirely finished at parturition, ongoing changes in microstructure, pruning, apoptosis, and gliogenesis support cortical change during postnatal development (Silbereis et al., 2016).

2.1.2 Cellular Properties of Cortical Organization

Cytoarchitecture scaffolds cortical macrostructure, which is commonly quantified with measures of thickness, area, and folding. According to the radial unit hypothesis, macrostructural indices stem from different neurodevelopmental origins, each arising from distinct progenitors and neurogenic phases (Jalbrzykowski et al., 2013). Specifically, the radial

dimension of the cortex (i.e., CT) is determined by the number of asymmetric divisions of neural progenitors within each cortical column, which govern the number of constituent cells (Mountcastle et al., 1997; Pontious et al., 2008; Rakic et al., 1988). The tangential dimension (i.e., SA) is determined by the number of radial units formed by symmetric divisions of neuroepithelial progenitors (Mountcastle, 1997; Pontious et al., 2008; Rakic et al., 1988, 2009).

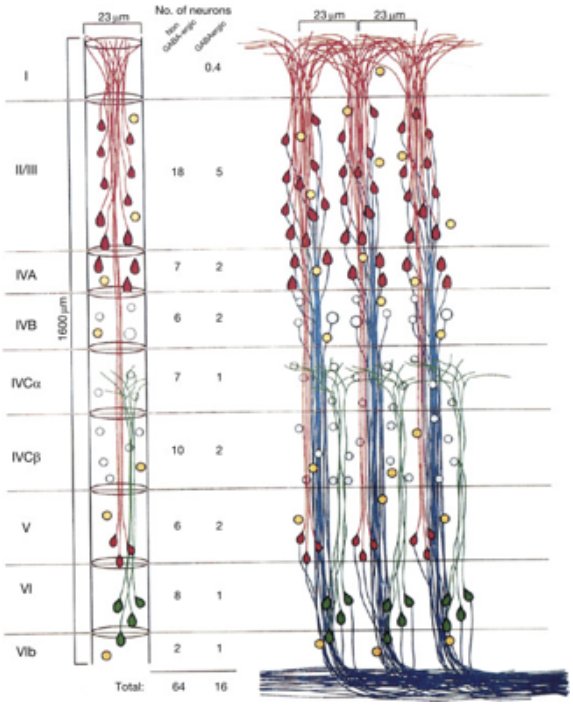


Figure 2.1. Cortical columns spanning laminae. Reproduced with permission from Peters and Sethares (1996).

As the cortex expands laterally, folds begin to appear. Primary fissures arise as early as the eighth to 10th gestational week, and sulci first emerge by the 14th gestational week. During the third trimester, gyrification expands significantly, transforming the brain into its distinctive gyrencephalic form (Chi et al., 1977). The predominantly tangential direction of cortical expansion enables the cortical sheet to expand to roughly three times the inner cranial surface, while maintaining efficient wiring organization (White et al., 2010). As a result, cortical volume growth is attributable in large part to gyrification and SA expansion rather than increasing thickness (Hogstrom et al., 2013; White et al., 2010). Theories about the processes that shape cortical convolution have varied considerably over the decades. Earlier conjectures focused on the influence of forces such as friction, cranial pressure, and volumetric constraint (Le Gros

Clark, 1945; Papez, 1929). Subsequent theories focused more so on biological properties of the cortex, including differential radial expansion of laminae (Richman et al., 1975), and the tensile pulling of axon fibers (Van Essen, 1997). In more recent years, computational simulations have assessed its elasto-plastic (Toro & Burnod, 2005), adhesive (Tallinen et al., 2014), and stress-dependent (Bayly et al., 2013) physical features. There have been growing efforts to integrate these theories into a cohesive framework, while also accounting for the manner in which cellular processes at the prenatal (e.g., neurogenesis, migration, axon guidance) and postnatal (e.g., synaptogenesis, dendritic arborization, cortical myelination) stages of development underlie the precise timing and placement of folds (Fernandez et al., 2016; Hill et al., 2010; Ronan & Fletcher, 2015).

2.2 Magnetic Resonance Imaging (MRI)

Structural magnetic resonance imaging (MRI) has been used extensively to map brain morphology. The safety of this technique, along with its excellent spatial resolution and test-retest reliability (Madan & Kensinger, 2017) allow for non-invasive *in vivo* characterization of cortical structure.

2.2.1. MRI Physics

MRI signal acquisition operates on the principle of energy transfer, as the spin and charge properties of hydrogen particles interact with a magnetic field (Dale et al., 2015). The core elements of an MRI scanner, including the magnet, radiofrequency (RF) coils, and gradient coils manipulate magnetization of protons in the body, such that the detected signal depends on the manner in which a given tissue type reacts to this manipulation. Hydrogen is prevalent in the body due to its hydrous composition. Hydrogen contains one positively charged proton that rotates about its own axis, inducing a magnetic moment. The magnetic moments of a group of protons are null at equilibrium due to mutual cancellation of their random spin orientations (Elmaoğlu & Çelik, 2011). When a strong static magnetic field (B_0) is applied, the hydrogen protons within the body rotate around the B_0 axis and align their energy states either parallel or anti-parallel to it. Since less energy is required to align with B_0, the majority of protons orient themselves along the main magnetic field and net magnetization is positioned in the longitudinal plane (z-axis). As the protons' magnetic field interacts with B_0, magnetic resonance is created, whereby the precessional frequency is proportional to the magnetic field

strength, B_0, known as the Larmor frequency. The net magnetic vector of all the protons generates a combined magnetic vector known as the total magnetization, M_0 (Currie et al., 2013; Elmaoğlu & Çelik, 2011).

The RF transmit coil broadcasts an excitation pulse (B_1 field) at the Larmor frequency to rotate the protons out of alignment with B_0, such that the strength and duration of the pulse determines the angle of the flip (Pai et al., 2021). A 90° flip angle (FA), for instance, tilts the magnetization vector M_0 from the longitudinal plane (z-axis) to the transverse plane (xy-axis), M_{xy}, as energy is induced to the protons, and then re-emitted at an equivalent frequency (Currie et al., 2013; Dale et al., 2015).

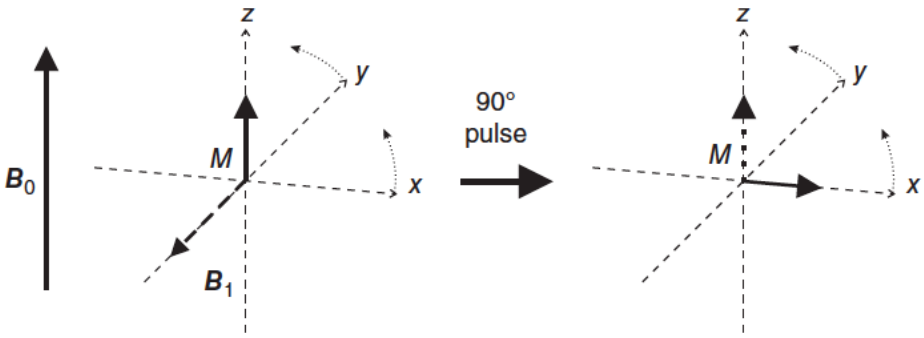


Figure 2.2. Following a radiofrequency (RF) pulse, magnetization is tilted to the transverse plane perpendicular to B_0, such that the degree of rotation is proportional to the RF pulse flip angle. Reproduced with permission from Dale et al. (2015).

Protons reorient back to the z-axis, and transverse magnetization decreases (Currie et al., 2013). A brief voltage is induced in the RF receiver coil placed perpendicular to M_{xy} from the changing transverse magnetization vector while it precesses back to alignment with B_0. This voltage constitutes the magnetic resonance signal. Over time, this signal decays as protons release their absorbed energy and the coherence of proton spins diminishes (Dale et al., 2015).

The spin properties of the protons during this state are quantified with T1 and T2 relaxation. T1 represents the rate at which protons exchange energy with their surroundings as magnetization realigns along the longitudinal plane, whereas T2 represents the rate at which transverse magnetization decreases after RF excitation, as protons dephase due to spin-spin interactions and magnetic field inhomogeneity (Currie et al., 2013; Lerch et al., 2017). T1 and T2 values are governed by the interaction of protons' magnetic moments, most being hydrogen

protons in water and macromolecules in the surrounding tissue. Their biological composition determines the strength of the hydrogen's bond with a given molecule, as well as the density of hydrogen protons within a voxel. These factors affect the interactions of their respective spin properties with the electromagnetic pulses. In turn, relaxation times differ for different tissues, yielding variation in signal intensities and contrasts between tissue types (Currie et al., 2013; Lerch et al., 2017; Pai et al., 2021). Manipulation of RF and gradient pulse parameters affects T1 and T2 relaxation times, and by extension, signal intensities of tissue for T1-weighted (T1-w) and T2-weighted (T2-w) sequences. In T1-w scans, the period of time between successive pulses (repetition time; TR), and the time between an initial RF pulse and its apex (echo time; TE) are typically shorter relative to T2-w scans.

Magnetic gradient coils positioned in each of the x, y, and z axes are used to locate the signal emitted by protons after RF radiation. Short electromagnetic pulses are superimposed onto $B_{0,}$ to generate small oscillations in the precession speed of protons away from their resonance at Larmor frequency, as a function of their position along the plane. These fluctuations cause dephasing along the gradient's direction, allowing the MR signal to be spatially encoded (Currie et al., 2013; McRobbie et al., 2017).

2.3. Image Processing

The three-dimensional structure of the brain is represented as approximately one million voxels, typically of $0.5mm^3$ - $1mm^3$ resolution. Each voxel contains a representation of MRI signal specific to the intensity of the tissue located within its spatial dimensions. Several pipelines, analytical tools, and software packages exist to process T1-w structural MRI scans. Although the precise sequence of the steps may differ, their main features involve transforming individual scans into a common format, removing artifacts that confound MRI signal, extracting the brain from surrounding non-brain tissue, and classifying the brain into different tissue and surface types.

2.3.1 Image Registration

Inter-scan variation in head positioning, cranial dimensions, anatomical morphology, and voxel properties are resolved with image registration. This process aligns scans from multiple participants to an atlas template in stereotaxic coordinate space to ensure spatial

correspondence of anatomical structures on a voxel-wise basis (Collins et al., 1994). Typically, transformations are estimated by optimizing intensity-based similarity measures that seek to minimize differences between two images. Algorithms can iteratively revise resolution, smoothing, or transformation parameters to improve convergence of similarity estimates (Dadar et al., 2018; Zhang et al., 2019). Transformations are linear or non-linear, across varying degrees of freedom. Linear registration (i.e. six-parameter rigid and 12-parameter affine) manipulates the source scan through translation, rotation, scaling, and shearing. Non-linear transformation involves more elastic deformations that are suitable for addressing distortions on a regional level. The transformations are typically defined on a vector field that can be used to derive scalar values indicating the extent to which the volume of a source image expands or contracts to fit the target (Chung et al., 2001, 2003; Collins et al., 1994; Maintz & Viergever, 1998; Song et al., 2017). The BestLinReg method is a widely used linear registration technique whereby images are initially blurred, then linear transformation estimates are hierarchically optimized using intensity-based gradient magnitude across increasing resolutions and degrees of freedom (Dadar et al., 2018; Lepage et al., 2017; Robbins et al., 2004).

2.3.2. Bias Field Correction

Spatial variation in the uniformity of the main magnetic field causes smoothly fluctuating tissue intensity across the MR image, known as bias field inhomogeneity or intensity non-uniformity. This non-biological variance is a confound that can affect other processing steps that are reliant upon image intensity (Sled et al., 1998; Vovk et al., 2007). Correction of bias field inhomogeneity is therefore a standard step of image processing. State-of-the-art algorithms include nonparametric nonuniform intensity normalization correction (N3; Sled et al., 1998) and its updated variant, N4ITK (N4; Tutison et al., 2010). N3 uses high frequency maximization to estimate the non-corrupted tissue intensity. The bias field, modelled as a multiplicative factor, is then smoothed with a cubic b-spline and removed from the scan, such that the uncorrupted image is used in successive iterations of the process (Boyes et al., 2008; Kahali et al., 2016; Sled et al., 1998). N4 incorporates improvements in the b-spline approximation (e.g. multivariate and non-cubic extensions), and hierarchically fits the bias field across increasing surface resolutions to optimize convergence of non-uniformity estimates (Tutison et al., 2010).

2.3.3 Brain Extraction

Accurate estimation of cortical measurements relies on the precise removal of the brain from surrounding non-brain tissue. Certain attributes of non-brain tissue can challenge brain extraction due to their physical proximity and similar signal intensities. The dura mater, for example, closely circumscribes the cortex and often confounds measurements of cortical volume and thickness due to overestimation or underestimation of grey matter (GM; Eskildsen et al., 2012; Lerch et al., 2017). Brain Extraction based on the nonlocal Segmentation Technique (BEaST; Eskildsen et al., 2012) is a commonly used tool that crops out the skull, dura mater, fat, muscles, skin, eyes, bone, as well as exterior blood vessels and nerves. Surrounding patches of a target voxel are compared with known patches in a set of priors and assigned a label based on their likeness using sum of squared distance to estimate similarity. A brain mask is ultimately generated, whereby voxels are tagged with a label indicating whether they contain brain or non-brain tissue (Eskildsen et al., 2012).

2.3.4. Tissue Classification

Voxels comprising the brain mask are then classified into GM, white matter (WM), and cerebrospinal fluid (CSF). Classification techniques typically involve intensity-based algorithms to estimate tissue labels. When used alone, however, they are susceptible to inaccuracies due to signal confounds from bias field inhomogeneity and partial volume effects (PVE), a phenomenon whereby a given voxel contains more than one tissue type with overlapping intensities (Nakamura & Fisher, 2009; Zhang et al., 2019). Therefore, these models are often supplemented with other approaches to improve performance. Some strategies include the incorporation of *a priori* anatomical information as spatial probabilities (e.g. Gaussian Mixture Model, Bayesian frameworks; Ashburner & Friston, 2003, 2005), contextual information from surrounding voxels (e.g. hidden Markov random fields; Zhang et al., 2001), allowing partial membership to more than one tissue cluster (e.g. fuzzy C-means; Chuang et al., 2006; Zhou et al., 2009), or modelling PVE estimates for different tissue classes (e.g., mixed PVE models; Choi et al., 1991; Tohka et al., 2004).

2.3.5 Surface Extraction

Cortical surface reconstruction consists of extracting and measuring surfaces belonging to different tissues in a manner that preserves their geometric and topological complexities (Lee

et al., 2006). A popular class of surface extraction methods involve deformation-based iterative morphing techniques (Liu et al., 2008). In particular, marching cubes algorithms have been used extensively to derive surface information from three-dimensional brain volumes. In brief, an isosurface is fitted to the cubic dimensions of a voxel, such that the intersection of the surface with a voxel edge creates a triangular patch dividing its spatial features. Adjacent patches are connected across voxels to form a tessellated mesh for a surface class. The vertices and adjoining edges of the triangular facets can be used to further characterize structural properties of the surface on a vertex-wise basis with measurements of thickness, area, and folding (Lorenson & Cline, 1987; Shattuck & Leahy, 2001). Constrained Laplacian Anatomical Segmentation using Proximities (CLASP) algorithm for surface extraction is commonly integrated into image processing pipelines (Kim et al., 2005). A three-dimensional spherical mesh is deformed to the GM/WM boundary, and expanded from the WM surface to the outer cortical boundary at the GM/CSF interface to form the pial surface. A vertex-to-vertex correspondence is maintained between inner and outer surfaces, with GM enclosed in between (Evans, 2005; Kim et al., 2005; Lee et al., 2006; MacDonald et al., 2000). Following extraction, surface-based registration can be performed to align surfaces of an individual to an average template so that vertex-wise cortical measurements computed at the subject's surface can be resampled to the reference template for inter-subject comparison (Lepage et al., 2021).

2.4. Normative Developmental Trajectories of Macrostructural Properties

Structural properties of the cortical mantle can be quantified with measures of thickness, area, gyrification, and curvature, which vary in their spatial distribution across the brain and undergo different maturational trajectories.

CT is quantified as the distance between corresponding points on WM and pial boundaries. CT follows a non-linear trajectory throughout development, increasing rapidly from birth, followed by decline (Fjell et al., 2015; Raznahan et al., 2011; Shaw et al., 2008) that may start as early as 1-2 years of age (Brown et al., 2012; Gilmore et al., 2021; Li et al., 2015; Remer et al., 2017; Wang et al., 2019). The spatial distribution of age-related changes is variable along both axes of cortical organization: sulcal depths experience more extensive thinning relative to gyral crowns (Vandekar et al., 2015), whereas across the cortical sheet, lower-order primary and sensorimotor areas mature earlier than association cortices (Brown et

al., 2012; Gilmore et al., 2021; Sowell et al., 2004; Syndor et al., 2021), with regionally-specific developmental trajectories (Ducharme et al., 2016; Mutlu et al., 2013; Shaw et al., 2008; Wang et al., 2019). Together, these maturational changes are thought to reflect axonal and dendritic remodelling, myelination, apoptosis, trophic glial and vascular changes, as well as synaptic pruning (Hogstrom et al., 2013; Jeon et al., 2015; Pontious et al., 2008).

Vertex-wise SA is typically computed by taking one third of the surface of triangular facets adjoining a given vertex (Lyttelton et al., 2009). Being strongly influenced by genetics, (Jha et al., 2018; Panizzon et al., 2009; Strike et al., 2019) changes to SA throughout the lifespan are less dynamic, with most of the developmental changes occurring during infancy (Jha et al., 2018; Raznahan et al., 2012; Yu et al., 2016). Fetal and neonatal studies have shown that cortical development accelerates in the second half of gestation, with the enlargement of SA and emergence of folds (Batchelor et al., 2002; Clouchoux et al., 2012; Dubois et al., 2008). This rapid expansion continues in the first two years of life (Gilmore et al., 2012; Hill et al., 2010; Li et al., 2013; Lyall et al., 2015; Remer et al., 2017) and into early childhood, increasing between 20-108% from 1-6 years of age, with regional dependencies (Gilmore et al., 2020; Remer et al., 2017). Notably, 73-83% of individual variation in SA at 6 years of age has been found to be present by the first year of life, relative to merely 45% of variation in CT at 2 years of age (Gilmore et al., 2020). SA continues to expand until later in childhood, then remains relatively stable or marginally declines thereafter (Ducharme et al., 2015; Lyall et al., 2015; Raznahan et al., 2011). These global changes are likely influenced by density of dendritic and synaptic architecture, gliogenesis, cortical myelination, and underlying WM development (Cafiero et al., 2019; Hill et al., 2010; Hogstrom et al., 2013; Rakic, 1995).

Approaches that quantify folding commonly involve the use of cortical SA as a reference. The gyrification index (GI; Elias & Schwartz, 1969; Zilles et al., 1988) represents a ratio of the total surface to the area of the convex hull around the brain (Zilles et al., 1997), such that a value of 1 describes a lissencephalic contour, and increasing values represent more cortex enfolded in sulcal crevices. Local gyrification index (LGI) is an extension of this measurement, and involves computing this ratio within a small circular region of interest on a cortical mesh, offering a finer grained measurement suited to three-dimensional topography (Schaer et al., 2008). Early studies characterizing GI found it to increase substantially during late gestation, and more subtly during the neonatal period, reaching an adult-like phenotype shortly after birth (Zilles et al., 1997). GI has been found to increase by 16% during the first

year of life and by 6.6% in the second year, with some evidence suggesting that the association cortex exhibits the most elevated growth rate (Li et al., 2014). This age-related increase in gyrification is believed to be predominantly driven by the dramatic increase in synaptic density, dendritic length and dendritic spine density that occurs in the perinatal period (Hill et al., 2010). GI and LGI decline in childhood and into adulthood (Forde et al., 2017; Hogstrom et al., 2013; Klein et al., 2014; Mutlu et al., 2013; Raznahan et al., 2011), generally exhibiting a linear decrease (Forde et al., 2017; Hogstrom et al., 2013; Mutlu et al., 2013), although region-specific non-linear trajectories have also been identified (Hogstrom et al., 2013; Klein et al., 2014; Mutlu et al., 2013). While gyri and sulci jointly comprise the folded surface of the mantle, they differ anatomically in their cellular organization and density, vascular properties, and neuropil concentration (Hilgetag & Barbas, 2005; Sun & Hevner, 2014; Wagstyl & Lerch, 2018).

Curvature is equivalent to the reciprocal of the radius of an inscribed circle, measured on points along the curve. Extrinsic curvature, often measured using MC, describes how the shape of the cortex is geometrically embedded in space (King et al., 2016; Schaer et al., 2008; White et al., 2010). On a surface, osculating planes define curvatures in principal directions, from which MC is computed. This measure of cortical complexity represents change in the normal direction of the curve relative to a tangent reference surface (do Carmo et al., 1976; Kim et al., 2006; King et al., 2016; Luders et al., 2006; MacDonald, 1998; White et al., 2010), yielding positive values in concave regions such as gyral crowns and negative values in convex regions such as sulcal folds, with a larger absolute value denoting sharper curvature (King et al., 2016). Studies characterizing developmental trajectories of cortical curvature in healthy samples have been sparse. MC of primary sulci and gyri follow variable trajectories between 20-28 weeks of gestation. For instance, linear decreases in MC has been reported in the circular sulcus and parieto-occipital sulcus, curvilinear decreases in superior temporal and central sulcus, and non-linear increases in the fronto-parietal operculum, as well as post-central and lingual gyri (Habas et al., 2012). Generally, age-related changes in cortical curvature are fairly modest in early life, increasing 1-20% from 1-6 years of age (Remer et al., 2017), and decline starting in early childhood (Pienaar et al., 2008). Hemispheric asymmetries have also been reported in curvature patterns during the first two years of life (Li et al., 2014). The mechanisms underlying age-related changes in curvature are believed to be largely influenced by cortical myelin and WM development, as proportions of GM and WM shift in infancy. Indeed, regionally specific increases in WM volume in sulcal depths have been associated with

increasing MC (King et al., 2016). This suggests that synaptic pruning and apoptosis contributing to cortical volume reduction in early life are accompanied by a concomitant increase in myelination that may result in the expansion of both WM and curvature (King et al., 2016).

Collectively, cortical indices vary spatially in their relative proportions (Hogstrom et al., 2013), have distinct neurodevelopmental origins (Mountcastle 1997; Rakic, 1995), genetic etiologies (Pazzinon et al., 2009), and maturational time courses (Raznahan, 2011). Given the divergent genesis of individual cortical metrics, the natural history of the brain can be inferred by studying multiple measures together, with each offering unique yet complementary biological information. By virtue of this phenomenon, perturbations in a specific measure are informative of key mechanisms of developmental programming that may be disrupted in neurodevelopmental disorders, such as ASD (Gharehgazlou et al., 2021; Jalbrzikowski et al., 2013).

2.5 Autism Spectrum Disorder

2.5.1. Diagnostic Criteria

ASD is characterized by impairments in social interactions and communication, as well as repetitive behaviours and interests. According to the Diagnostic and Statistical Manual of Mental Disorders - 5th edition (DSM-V; American Psychiatric Association, 2013), diagnostic criteria are as follows:

1) Social communication deficiencies in each of the following social domains:
    a) Social emotional reciprocity
    b) Non-verbal communication
    c) Formation and maintenance of relationships
2) Restricted and repetitive behaviours and interests in a minimum of two of the following domains:
    a) Stereotyped or persistent movements, object use, or speech patterns
    b) Attachment to sameness, routines, or ritualized behaviour/speech patterns
    c) Extremely restricted interests or fixations
    d) Sensory atypicalities (i.e. hyper- or hyposensitivity or preoccupation with sensory elements of surroundings)

These symptoms must endure from early development, incur significant functional impairment in various facets of daily life, and are not attributable to a primary diagnosis of intellectual disability or global developmental delay. Additional specifications are provided by indicating the presence of language or intellectual impairment, medical/genetic conditions, other neurodevelopmental or psychiatric disorders, and catatonia (American Psychiatric Association, 2013).

ASD is commonly diagnosed by licensed mental health or medical professionals with specialized training. The diagnostic procedure involves clinical observation and the administration of standardized diagnostic assessment tools (e.g., Autism Diagnostic Observation Schedule [ADOS; Lord et al., 2000; 2012] or Autism Diagnostic Interview-Revised [ADI-R; Lord et al., 1994]), as well as comprehensive medical and family history, caregiver interviews, and sensory testing (Zwaigenbaum et al., 2019).

Although not included in the core symptom domains, co-occurring psychiatric and medical conditions are highly prevalent, with approximately 70% of autistic individuals presenting with at least one comorbid mental health diagnosis and almost 50% with multiple (de Bruin et al., 2007; Lai et al., 2019). Among the most common are attention deficit hyperactivity disorder (ADHD), anxiety, obsessive compulsive disorder, bipolar disorder, oppositional defiant disorders, depression, and schizophrenia (Buck et al., 2014; Croen et al., 2015; de Bruin et al., 2007; Elder et al., 2017; Lai et al., 2019). Other medical disorders are also over-represented, including gastrointestinal disorders, sleep disorders, epilepsy, hypertension, obesity, diabetes, autoimmune conditions, asthma, allergies, and hormonal dysregulation (Al-Beltagi, 2021; Bauman, 2010; Croen et al., 2015). Co-occurring conditions contribute to impairment and negatively affect developmental progression, posing additional challenges to accurate and timely diagnosis (Bauman 2010; Lord et al., 2022).

2.5.2 Prevalence

ASD affects 1 in 66 Canadians aged 5-17 years old (Ofner et al., 2018). It can be reliably diagnosed by 2 years of age (Zwaigenbaum et al., 2019), with 56% of Canadian children diagnosed by 6 years of age (Ofner et al., 2018). A child's symptom severity affects the age of diagnosis. Greater symptom severity and presence of language delay or cognitive deficits are likely to result in an earlier diagnosis (Avlund et al., 2021; Daniels & Mandell, 2013), while milder symptoms, and a history of other developmental or psychiatric conditions are likely to

result in a later diagnosis (Avlund et al., 2021). Certain sociodemographic factors are associated with diagnostic delays, including lower parental education and socioeconomic status, belonging to an ethnic minority group, female sex, and living in a rural geographic location. However, many of these factors are difficult to parse and conflicting findings do exist (Avlund et al., 2021; Brett et al., 2016; Daniels & Mandell, 2013; Giarelli et al., 2010; Hrdlicka et al., 2016).

Importantly, sex differences exist in the prevalence of ASD. Indeed, males are diagnosed approximately 3.5 times as often as females (Loomes et al., 2017; Posserud et al., 2021), and about twice as often in individuals with lower IQ (Fombonne, 2009; Loomes et al., 2017; Volkmar et al., 1993). These discrepancies may stem from sex differences in core symptom domains. For instance, autistic females tend to have milder restricted and repetitive behaviours relative to males (Antezana et al., 2018; Hartley & Sikora, 2009; Hiller et al., 2014; Van Wijngaarden-Cremers et al., 2014), and more expressive social skills, including social emotional reciprocity, non-verbal communication (Hiller et al., 2014), displays of context-appropriate behaviour (de Giambattista et al., 2021), and friendship quality (Head et al., 2014), although IQ (Ryland et al., 2014) and age (Mahendiran et al., 2019) are important mediators of this variation. Females also have greater prevalence of co-occurring internalizing disorders (Supekar et al., 2017), more often display lower IQ (Werling & Geschwind, 2013), and have distinct behavioural and cognitive profiles (Carter et al., 2007; Lemon et al., 2011), suggestive of a divergent sex-specific phenotype.

The source of these sex differences is contested. Some have argued that discrepancies in prevalence are attributable to the fact that females are more proficient at camouflaging symptoms and using compensatory coping strategies that may obscure overt ASD characteristics (Hull et al., 2017; Lai et al., 2017). Moreover, the traditional characterization of ASD has been formed from decades of predominantly male-centred research studies (Lai et al., 2015) which may have biased diagnostic criteria, making it more likely for typically female symptoms to be overlooked (Hull et al., 2020). Others have argued that clinically observed sex differences originate from sex differences in pathophysiology. The Female Protective Effect theory suggests that females require more significant genetic burden and biological disruptions to meet diagnostic thresholds, with some studies finding a greater quantity of highly penetrant rare copy number variants in females that may act as ASD risk factors (Jacquemont et al., 2014; Levy et al., 2011). The theory posits that this protection results in more subtle behavioural symptoms in females when genetic load is equal to their male counterparts. In parallel, greater

genetic variability due to male sex chromosomes (XY) confers enhanced genetic vulnerability in males (Hull et al., 2020; May et al., 2019).

2.5.3 Prognosis

Due to the lifelong course of ASD and complexity of core and co-morbid symptoms, autistic individuals experience functional impairments in many spheres of life (Lenroot & Yueng, 2013; Masi et al., 2017). Although long-term trajectories are variable (Fountain et al., 2012; Levy & Perry, 2011), autistic individuals are generally at greater risk of higher mortality rates (Farley & McMahon, 2014; Hirvikoski et al., 2018), poor social (Farley & McMahon, 2014; Magiati et al., 2014) and health outcomes (Benevides et al., 2020; Bishop-Fitzpatrick & Kind, 2016), in addition to limited academic and vocational achievement (Dudley & Emery, 2014; Farley & McMahon, 2014), autonomy (Billstedt et al., 2005), and lower reports of life satisfaction (Burgess & Gutstein, 2007).

Research has shown that symptomatology remains stable (Lord et al., 2006) or modestly improves with age (Fecteau et al., 2003; McGovern & Sigman, 2005; Shattuck et al., 2007; Woodman et al., 2014). Age-related changes in comorbidities have also been observed, with co-occurring conditions such as epilepsy and ADHD decreasing appreciably throughout development, whereas the prevalence of schizophrenia rises (Supekar et al., 2017). Overall, outcomes depend on a host of factors including symptom severity, cognitive ability, comorbidities, socioeconomic status, family environment, and timing of intervention onset (Avlund et al., 2021; Fountain et al., 2012; Matson & Smith, 2008; Taylor & Seltzer, 2010; Woodman et al., 2014). It is understood that achieving normative functioning should not be considered the ultimate benchmark for positive long-term outcomes. Rather, both self-advocates and researchers endorse the importance of health, adaptive living skills, independence, community integration, education, employment, and meaningful relationships as key elements of optimal outcomes (Eigsti et al., 2022; McCauley et al., 2020).

2.5.4 Treatment Approaches

Most treatment efforts involve early and intensive behavioural and psychoeducational interventions. These approaches aim to enhance adaptive behaviours and independence by improving language, social communication, self-regulation, and cognitive skills, while reducing challenging behaviours (Posar & Visconti, 2019; Vismara & Rogers, 2010). Other

common treatments including speech-language, occupational (Hebert et al., 2014), and physical therapy (Atun-Einy et al., 2013; Mieres et al., 2012), in addition to sensory processing interventions (Baranek et al., 2002; Case-Smith et al., 2014). Comorbid conditions are often addressed through counselling and pharmacological treatments (Genovese & Butler, 2020). While these approaches produce favourable responses, the pervasive and complex nature of symptomatology requires prolonged treatments which incur significant financial cost. Indeed, the lifelong cost of caring for an individual with ASD in Canada has been estimated to be up to 5.5 million dollars (Dudley & Emery, 2014). The source of these expenses stem from the intensity and frequency of professional services and therapies required, in addition to other medication, equipment, accommodations, and specialized support. Indirect costs accrued from lost income on the part of both caregivers and ASD individuals are a source of indirect financial loss (Dudley & Emery, 2014; Rogge & Janssen, 2019).

2.5.5 Etiology

Despite the discovery of numerous biological, environmental, and epigenetic risk factors, a comprehensive etiological framework of ASD causation has not been established, which has hampered the identification of robust biomarkers.

ASD has strong genetic liability. Heritability estimates are approximately 80%, with hundreds of rare and common genetic variants implicated in ASD susceptibility (Bai et al., 2019). Genetic abnormalities include single nucleotide polymorphisms (e.g., CDH10 to CDH9 genes at 5p14.1 locus), copy number variations (e.g., at 15q11-13, 16p11, and 22q11-13 regions), de novo mutations (e.g., R12C and L68P point mutations of SHANK3), single gene conditions (e.g. Fragile X and Rett syndromes), and numerous gene variants (Bill & Geschwind, 2009; Bucher et al., 2021; Genovese & Butler, 2020; Grove et al., 2019; Waye & Cheng, 2018; Weiss et al., 2009). Studies have profiled the effect of ASD risk genes on various aspects of prenatal brain development, finding perturbations in synaptogenesis, calcium binding, cell adhesion, regulation of tissue and cell development, ion transport, axon guidance, neurotransmission, among many others (Grove et al., 2019; Rylaarsdam & Guemez-Gamboa, 2019; Toma et al., 2012; Weiss et al., 2009). While abundant candidate genes have been recognized, most contribute modestly to overall genetic risk and many confer susceptibility to multiple disorders, some of which are comorbid with ASD (Genovese & Butler, 2020). Given the complexity of biological systems affected by genetic abnormalities, linking genotype to

clinical phenotype across multiple levels of biological dysfunction has been challenging. It is likely that candidate genes interact via both epistatic and pleiotropic effects through mechanistic pathways with shared functions (Tordjman et al., 2018).

Other lines of investigation have explored alternative biological signatures. Studies of sex chromosome abnormalities have found that ASD is significantly overrepresented in males with Y chromosome aneuploidies (Tartaglia et al., 2017) as well as X-linked genetic conditions such as Fragile X, Turner, and Klinefelter syndromes (Baron-Cohen, 2011). Furthermore, mitochondrial dysfunction and co-occurring mitochondrial diseases are elevated in ASD (Giulivi et al., 2010; Rossignol & Frye, 2011), which affect energy metabolism and synaptic neurotransmission (Rossignol & Frye, 2011). Neuroendocrine dysregulation has also been proposed in ASD etiology, including in utero exposure to fetal and maternal androgens (Baron-Cohen et al., 2015) and thyroid (Khan et al., 2014) hormones, as well as deficiencies in oxytocin receptors (Gregory et al., 2009; Pobbe et al., 2012), though their causal roles are somewhat contested.

Environmental factors affecting brain development likewise play an important role in conferring risk. Advanced parental age, toxic metabolites (e.g., from pollutants, heavy metals, certain drugs/mediations), nutrient deficiencies, obstetric complications, as well as maternal infection, stress, hypertension, diabetes, and obesity have all been implicated in ASD risk (Hodges et al., 2021; Gore et al., 2014; Worsham et al., 2021). These factors can directly affect sensitive biological systems in utero (e.g., immune, oxidative, and endocrine function; Graham et al., 2021) and indirectly influence epigenetic mechanisms via DNA methylation, chromatin structure, and histone acetylation (Waye & Cheng, 2018) to modify gene expression involved in key neurodevelopmental pathways. Together, these environmental stressors exert both cumulative and interactive effects on neurodevelopmental programming during critical maturational windows and contribute to differential susceptibility to psychopathology (Buss et al., 2012; Entringer et al., 2016; Graham et al., 2021; Rice & Barone, 2000).

2.5.6 Cortical Histology Findings

The post-mortem literature suggests abnormal columnar organization in ASD, such that minicolumns are more numerous, smaller (Casanova et al., 2002) and atypical in width (Casanova et al., 2006; McKavanaugh et al., 2015), particularly in frontal and temporal cortices (Casanova et al., 2002). More rapid proliferation of neural progenitors and reduced

synaptogenesis have also been found in ASD children aged 2-5 years in the dorsolateral prefrontal cortex (Marchetto et al., 2017), in addition to increased neuronal density in the prefrontal cortex (Courchesne et al., 2011), abnormal laminar architecture in prefrontal and temporal cortices (Stoner et al., 2014), increased CSF fluid in frontal lobes (Shen et al., 2013), and abnormal cellular architecture at the GM-WM interface (Avino & Hutsler, 2010). There has also been evidence of atypical microglial activation in the prefrontal cortex as well as greater microglial somal volume in WM and microglial density in GM (Morgan et al., 2010). Collectively, irregular cellular patterning appears to converge in prefrontal regions. Together, these alterations point to multifactorial cellular mechanisms that may influence both the composition and developmental course of macrostructural properties.

2.5.7 Case-Control Differences in MRI-Derived Cortical Measures

One of the earliest and most replicated findings is of early brain overgrowth, which has been shown with measures of enlarged head circumference (Courchesne et al., 2003), increased brain weight (Bailey et al., 1993) and brain volume (Courchesne et al., 2001; Hazlett et al., 2006, 2011; Ismail et al., 2016; Li et al., 2017; Schumann et al., 2010). However, beyond this early developmental stage, a consensus regarding region-specific trends remains elusive. For instance, one review article reported that 41% of studies found significant increases in frontal lobe GM volume (GMV) in ASD subjects relative to NT controls, 24% reported a decrease, and 35% reported no statistical difference (Pagnozzi et al., 2018).

2.5.7.1 Cortical Thickness and Surface Area

These discrepancies extend to volume's subcomponents as well. Indeed, conflicting findings point to both decreases (Ecker et al., 2013, 2014; Hadjikhani et al., 2006; Hyde et al., 2010; Jiao et al., 2010; Scheel et al., 2011), and increases (Haar et al., 2016; Hardan et al., 2006; Hyde et al., 2010; Khundrakpam et al., 2017; Yang et al., 2016) in CT in ASD subjects compared to controls, as well as reduced (Ecker et al., 2013, 2014; Mensen et al., 2017), and enlarged SA (Doyle-Thomas et al., 2013; Hazlett et al., 2011). Null findings are also prevalent (Mak-Fan et al., 2011; Raznahan et al., 2011; Wallace et al., 2013, 2015; Yang et al., 2016). Longitudinal studies have been conducted to address the often non-overlapping age ranges in cross-sectional reports. Evidence points to a relative absence of normative age-related thickness decline in ASD subjects from early to late childhood (Smith et al., 2016), as well as steeper

thickness decrease in temporal and occipital regions in older childhood (Hardan et al., 2009). Between adolescence and early adulthood, greater reductions have been reported in parietal (Wallace et al., 2010, 2015; Zielinski et al., 2014), frontal, and occipital lobes (Zielinski et al., 2014), with slower thinning in adulthood (Zielinski et al., 2014). With respect to SA, some longitudinal studies do not report developmental differences (Smith et al., 2016; Wallace et al., 2015), while others have observed greater SA enlargement between 2-5 years of age in ASD subjects (Hazlett et al., 2011), as well as a lack of age-related reduction between late childhood and adulthood in superior temporal, postcentral, and supramarginal partitions (Mensen et al., 2017).

2.5.7.2 Case-Findings from Large-Scale Studies

Another approach to resolving discrepant cross-sectional findings has involved amassing larger subject pools. In recent years, there has been mounting concern about low statistical power in small- and moderately- sized neuroimaging studies that may reduce the detection of true effects, while also increasing the likelihood of false positives and artificially inflated effect sizes (Postema et al., 2020). Several recent studies have therefore probed morphometric case-control differences with well-powered sample sizes. Some examples include a study by Bedford et al. (2020) employing a vertex-wise meta-analytic approach (n=3145) and an ENIGMA Consortium study by van Rooij et al. (2018), using a mega-analysis (n=3222). Overall, the location, direction, and magnitude of case-control differences fails to replicate across studies, apart from increased CT in ASD subjects within frontal and cingulate regions (Bedford et al., 2020; van Rooij et al., 2018). Unreplicated findings include ASD-specific CT decreases in the temporal pole, entorhinal cortex, and parahippocampus (van Rooij et al., 2018), as well as increases in occipital and temporal poles, postcentral gyrus, and precuneus (Bedford et al., 2020). Neither study found significant diagnostic effects on SA.

Notably, the importance of sex, IQ, and clinical severity have been empirically shown. Bedford et al. (2020) and van Rooij et al. (2018) describe enlarged thickness in temporal and frontal cortices in ASD males, with the former finding additional regions of significance, alongside ASD female-specific patterns with larger effect sizes. Both studies also report differential spatial patterns for significant diagnosis-by-IQ trends in CT. Further, both Bedford et al. (2020) and van Rooij et al., (2018) noted that ASD symptom severity was associated with increased CT in frontal nodes, in addition to study-specific anatomical trends. Medication use

was associated with CT only in the inferior temporal region by van Rooij et al. (2018), and neither study detected a significant effect of comorbidities on cortical measures.

Unreconciled differences between large scale studies suggest that merely increasing sample size is not enough to resolve inconsistencies. Indeed, inter-study differences in participant characteristics (e.g., ascertainment practices, inclusion/exclusion criteria), data collection (e.g., scanners, acquisition sequences), processing (e.g., software), quality control, site correction, and analysis (e.g., region-of-interest vs vertex-wise approaches, statistical models, covariate management) likely contribute to variability in findings. Moreover, it is apparent that accounting for sources of heterogeneity such as sex, age, IQ, and symptom severity is essential to generate robust statistical inferences.

2.5.7.3 Folding

Measures of cortical folding have received less attention in the ASD literature. Early work using manual assessments identified developmental cortical malformations in a subset of ASD subjects, including localized regions of polymicrogyria (i.e. the presence of many abnormally small gyri; Piven et al., 1990). Other qualitative assessments of sulcal morphology have identified altered shape and depth of sulcal folds in ASD subjects within frontal, parietal, and insular clusters (Nordahl et al., 2007), the left Sylvian fissure (Brun et al., 2016), as well as the temporo-parietal junction and anterior insula (Dierker et al., 2015).

Increased GI has been found in children and adolescents with ASD within the left prefrontal cortex (Hardan et al., 2004) and right parietal lobe (Kates et al., 2009). Conversely, age-related lobar-level GI decreases in the left prefrontal and parietal cortex have also been identified in older ASD children and adolescents (Bos et al., 2015). Studies using LGI have reported mixed findings. Increases in ASD subjects relative to controls have been identified in the precuneus and occipital lobes (Wallace et al., 2013), frontal lobes (Ecker et al., 2016; Kohli et al., 2019; Libero et al., 2018; Yang et al., 2016), temporal and parietal lobes (Ecker et al., 2016; Kohli et al., 2019; Yang et al., 2016), as well as precentral gyri (Ecker et al., 2016; Kohli et al., 2019), central sulcus, postcentral gyri (Ecker et al., 2016), insula (Kohli et al., 2019), lingual gyrus, and isthmus cingulate (Yang et al., 2016). Conversely, reductions in LGI have been noted in the precentral gyrus, frontal, parietal, and parieto-occipital regions (Schaer et al., 2013), the lingual gyrus (Kohli et al., 2019), and fusiform gyri (Libero et al., 2019). Notably, a number of studies have not found any significant diagnostic effects on GI or LGI including

Schaer et al. (2015), Hirjak et al. (2016), Koolschijn and Geurts (2016), Casanova et al. (2009) and Gharehgazlou et al. (2021) as well as a meta-analysis of ~1000 participants by Gharehgazlou et al. (2020).

One longitudinal LGI study has been conducted in children 3-5 years of age, which found enhanced LGI in ASD within this age range relative to stable or declining values in NT controls (Libero et al., 2018). Cross-sectional studies have noted increasing LGI with age in ASD children (Kates et al., 2009; Yang et al., 2016), followed by a decrease in adolescence (Bos et al., 2015; Gharehgazlou et al., 2021; Kohli et al., 2019; Wallace et al., 2013), and adulthood (Koolschijn & Geurts, 2016). Some evidence suggests greater LGI in ASD females relative to males in ventromedial prefrontal and orbitofrontal cortices (Schaer et al., 2015), though many studies consist of male-only subjects (e.g. Casanova et al., 2009; Hardan et al., 2004; Libero et al., 2013; Nordhal et al., 2007; Wallace et al., 2013; Yang et al., 2016) and thus more evidence is needed to substantiate these findings. Significant diagnosis-by-IQ effects have generally not been supported (Bos et al., 2015; Gharehgazlou et al., 2021; Kates et al., 2009; Wallace et al., 2013). Interestingly, a number of studies that identified significant diagnostic differences in GI or LGI did not find any co-occurring group differences in SA (Bos et al., 2015; Kohli et al., 2019; Wallace et al., 2013; Yang et al., 2016) or volume (Schaer et al., 2013), despite the fact that folding is intrinsically tied to cortical expansion, which could suggest questionable neurobiological specificity.

These discrepant findings could be due to a number of factors. Studies vary in the global brain measure used as a covariate in statistical models of GI/LGI, which have included intracranial volume (ICV; Yang et al., 2016), total brain volume (TBV; Ecker et al., 2016; Hardan et al., 2004; Kohli et al., 2019), cortical volume (Schaer et al., 2015), SA (Ecker et al., 2016; Gharehgazlou et al., 2021), or none (Bos et al., 2015; Casanova et al., 2009; Hirjak et al., 2016; Nordhal et al., 2007; Schaer et al., 2013; Wallace et al., 2013). Some have also applied additional smoothing kernels to LGI values (Ecker et al., 2016; Hirjak et al., 2016; Koolschijn & Geurts, 2016) which may serve as a source of methodological variance. It is also possible that many studies are simply underpowered to detect reliable effects, with the majority having sample sizes of ~100 subjects or less (e.g. Bos et al., 2015; Casanova et al., 2009; Ecker et al., 2016; Hardan et al., 2004; Hirjak et al., 2016; Kates et al., 2009; Koolschijn & Geurts, 2016; Nordhal et al., 2007; Schaer et al., 2013; Wallace et al., 2013; Yang et al., 2016). Lastly, atypicalities in sulcal and gyral morphology may exist in ways that are not captured with GI or

LGI measures such as qualitative differences in the shape and placement of folds (Gharehgazlou et al., 2020). Indeed, individual variation in folding patterns may challenge accurate registration during image processing and further confound the abovementioned issues.

Few studies investigating curvature exist in the ASD literature. Of those found, there have been reports of increased Gaussian curvature in ASD subjects relative to controls, especially in the superior temporal region (Levman et al., 2018) as well as sex-specific qualitative differences in MC of ASD subjects (Hammill et al., 2021).

Taken together, ASD heterogeneity extends beyond key symptom domains to encompass variation in etiological pathways (Masi et al., 2017), onset (Ozonoff et al., 2010), comorbidities (Lai et al., 2019), developmental trajectories (Fountain et al., 2012), cognitive abilities (Mandy et al., 2015), and neuroanatomy (Bedford et al., 2020; van Rooij et al., 2018). In order to establish a cohesive understanding of interdimensional variability, it is essential to use statistical models that account for individual variation across multiple clinical and neurobiological domains.

2.6 Statistical Considerations

Clinical heterogeneity has long been considered the primary source of inconsistent neuroimaging findings in ASD. However, in recent years, there has been growing acknowledgement that limitations of common statistical approaches also contribute to ongoing replicability challenges.

2.6.1 Limitations of Univariate Models

For decades, mass-univariate models have been the main statistical method in case-control studies probing the relationship between diagnostic status and MRI-derived neuroanatomical indices. This approach fits the same Ordinary Least Squares regression model to each vertex/voxel across the brain to evaluate β weights of predictor variables (e.g. ASD diagnosis; $\beta_1 X_1$) relative to the variance of a vertex-wise neuroanatomical outcome measurement (e.g., CT; Y_i; Guillaume et al., 2018; McIntosh & Mišić, 2013). After multiple comparison correction, the analyses yield a set of vertices that represent the location of significant diagnostic differences in the average value of the neuroanatomical metric, which can be represented with statistics that convey the effect size and significance (e.g., t-values,

standardized beta coefficients, q-values; Patel, 2022). While this method has yielded many valuable findings and is favoured for its ease of interpretation (Friston et al., 1995), there are several key challenges due to the high-dimensional and multi-scaled nature of neuroimaging data and the disproportionate ratio of features to samples (i.e. thousands of vertices per subject; O'Toole et al., 2007; Wang et al., 2014).

The substantial number of vertex-wise models introduces a number of constraints. For one, statistical inferences are generally made in the absence of regression diagnostics to confirm that linear model assumptions are met, which are important for accurate variance and parameter estimation (Guillaume et al., 2018; Luo & Nichols, 2003; Zhu et al., 2007). Moreover, each vertex is assumed to be independent, which ignores spatial autocorrelation, the systematic regional variation between neighbouring vertices (Habeck, 2010). This modelling scheme obviates the possibility that an observed diagnostic effect is represented by a joint collection of interdependent vertices (McIntosh & Mišić, 2013). Further, the stringent multiple comparison correction required to limit inflation of Type I error across thousands of models produces conservative significance thresholds that reduce sensitivity to smaller effect sizes (O'Toole et al., 2007; Poline & Brett, 2012). Given the subtlety and variability of brain-behaviour associations, identifying reproducible phenotypes requires thousands of subjects (Marek et al., 2022), which is prohibitive for many researchers.

From an interpretation standpoint, case-control univariate models produce an average vertex-wise measurement for a given group, which may be unrepresentative of neurobiological variation in heterogeneous disorders such as ASD (e.g., Pua et al., 2019; Tunç et al., 2019; Zabihi et al., 2019). Compounding this issue further is the fact that case-control designs suffer from the choice of variables used as covariates, which may drive directionality or effect sizes. Indeed, it is common practice to residualize important sources of variability (e.g. sex and age) as nuisance variables or remove them as covariates to limit collinearity (Bedford et al., 2020; Bethlehem et al., 2020; Lombardo et al., 2019). Previous studies have underscored the differential effects of raw versus residualized predictors, the risk of biased coefficient estimates (Lynam et al., 2006; Wurm & Fisicaro, 2014; York, 2012), and inadequate modelling of age- and sex-specific effects (Bethlehem et al., 2020). Interpretive challenges are also created, as partialling predictors effectively isolates cortical measures from variables that are intrinsically tied to their anatomical organization. Ultimately, probing diagnostic effects without accounting

for variance attributable to sex and age may fail to create clinically and biologically robust models.

2.6.2 Structural Covariance

Modern computational techniques recognize that brain regions are organized in community structures, underpinned by common biological and functional processes. One strategy to evaluate this phenomenon is structural covariance, which measures coordinated patterns of macrostructural change between brain regions that is consistent across individuals (Alexander-Bloch et al., 2013; Lerch et al., 2006). This index represents a fundamental organizational principle of the cortex because brain regions that change together also share similarities in underlying gene expression profiles (Romero-Garcia et al., 2018), structural connectivity (Yee et al., 2018), and functional coupling (Liao et al., 2013). Moreover, a growing body of research suggests that differences in structural covariance networks exist across ages, psychiatric disorders, and cognitive abilities (Liu et al., 2017; Mitteroecker et al., 2012; Seidlitz et al., 2018; Sha et al., 2022; Sharda et al., 2016; Zielinski et al., 2012), implying functional and clinical relevance. Previous studies of structural covariance in ASD have identified differences from NT subjects, with functional implications across various socio-cognitive domains including social-emotional regulation (Zielinski et al., 2012), theory of mind (Bernhardt et al., 2013), and language functioning (Sharda et al., 2016). These studies demonstrate that covariance patterns of distinct cortical metrics may differentially influence cognitive function and clinical symptomatology.

2.6.3 Multivariate Analyses

Multivariate approaches are increasingly prevalent as counterpoints to canonical univariate techniques. These methods identify interdependencies between input variables in a sparse and low dimensional latent framework (Habeck, 2010; Habeck et al., 2010; McIntosh & Mišić, 2013), representing the fusion of individual elements that covary or correlate together (McIntosh & Lobaugh, 2004). In the context of neuroimaging data, multivariate analyses combine thousands of individual cortical vertices into structural networks of interrelated regions (McIntosh & Mišić, 2013), which are suitable for modelling the spatially distributed and interdependent nature of cortical organization (Segovia et al., 2014). These techniques are therefore more sensitive and specific to subtle interactions between a collection of cortical

regions that share commonalities in an underlying dimension (Habeck et al., 2010) and may be jointly involved in a clinical, cognitive, or biological phenomenon (Patel, 2022; Sotiras et al., 2015). Many exploratory multivariate techniques aim to identify data-driven patterns of covariation using singular value decomposition (SVD). These methods decompose input data (e.g vertex-by-subject matrix of CT values) into: 1) a set of latent components that describe the spatial location of covarying vertices and 2) subject-level loadings that describe each individual's expression of the components (McIntosh & Mišić, 2013).

Importantly, the subject-wise component loadings yield an individualized profile of each subject's covariance networks. The patterns can thus be expressed in a continuous manner, which aligns with the increasingly dimensional view of many psychopathologies (Patel, 2022). This scheme places an individual with a set of characteristics or symptoms along a continuum of variation (either within a disorder, with the normative population, or transdiagnostically) to identify patterns across individuals rather than groups (Lombardo et al., 2019). Aspects of health and illness can thus be understood across a range of functioning rather than seeking group-level differences relative to a diagnostic threshold (e.g. Research Domain Criteria framework; Cuthbert, 2015). Similarly, biological dysfunctions underpinning psychopathology are increasingly understood to transcend diagnostic classifications to fall along broader spectral dimensions (Stoyanov et al., 2020). For instance, in some cases, variations in the degree of a given dysfunction can inform treatment plans and prognostic predictions over and above a diagnostic label (McPartland, 2016).

In line with the dimensional view of ASD (American Psychiatric Association, 2013), studies seeking to understand its pathophysiology ought to similarly adopt dimensional approaches across multiple domains implicated in ASD symptomatology. Indeed, there is a growing need to understand how variability is coordinated across biological systems in an individualized manner (Lombardo et al., 2019; Vargason et al., 2020). A number of multivariate techniques can be employed to achieve this aim, two of which are covered in detail below.

2.6.3.1 Non-negative Matrix Factorization (NMF)

Orthogonal projective NMF (Lee & Seung, 1999, 2000; Yang & Oja, 2010) is an unsupervised matrix decomposition technique that models dominant patterns of covariance in the data, whereby an input matrix X (m x n) is factorized into a product of two matrices W (m x k) and H (k x n) that jointly recover the raw data patterns through a linear combination of

latent components and subject-wise loadings (Sotiras et al., 2015). Here, m represents cortical vertices, n represents subjects, and k represents user-selected components. NMF is well-suited for neuroimaging data because it identifies data-driven covariance patterns of lower dimensionality that can be mapped to subjects on an individual basis. Since the elements of the input and output matrices are constrained to be non-negative, it allows for an additive reconstruction of components to recapitulate the covariance patterns of the original data in a sparse and parts-based way (Sotiras et al., 2015; 2017).

NMF seeks to approximate the factorization of X≈WH by solving the following minimization problem (Lee & Seung, 1999; Sotrias et al., 2015):

(1) $$\min ||X - WH||^2_F$$

whereby W ≥ 0, H ≥ 0 and $||\ .||^2_F$ represents the squared Frobenius norm, the cost function that is minimized.

The projective property (Yuan & Oja, 2005), allows for the subject-by-metric loadings of the H matrix to be estimated as the projection of the data matrix X to the components W:

(2) $$H = W^T X$$

Such that H is now replaced in the approximation: $X \approx WW^T X$ (Yang & Oja, 2010).

Further, the orthogonality constraint on W modifies the minimization problem as follows (Sotiras et al., 2015; Robert et al., 2022):

(3) $$\min_W ||X - WW^T X||^2_F$$

whereby $WW^T = I$, W ≥ 0, such that I represents the (orthogonal) identity matrix, and $||\ .||^2_F$ represents the squared Frobenius norm.

The spatial location of covariance patterns is represented in the W matrix, also known as the basis matrix, which is organized with m vertices by k components. This matrix describes the components' spatial properties by indicating the weightings of each vertex onto every component, represented with a W score. Through a multiplicative update rule, W is iteratively revised to minimize the reconstruction error while satisfying orthogonality constraints (Varikuti et al., 2018; Yang & Oja, 2010):

(4) $$W'_{ij} = W_{ij} \frac{(XX^T W)_{ij}}{(WW^T XX^T W)_{ij}}$$

where i = 1.. number of vertices and j = 1.. number of components.

The orthogonal parameter of orthogonal projective NMF ensures that components are non-overlapping, which improves specificity and interpretability, while the projective property ensures that all components participate in the reconstruction of the entire data sample, which improves sparsity (Sotiras et al., 2015). Additionally, the orthogonality constraint renders the NMF minimization mathematically nearly identical to that of k-means clustering, which allows a soft clustering of vertices with a winner-take-all approach (Ding et al., 2005). Thus, although the W matrix contains component-wise weights for each vertex, the clustering property allows the allocation of a given vertex to a single component with the highest W weight. This way, each vertex is assigned to the component to which it has the strongest membership, thereby improving sparsity and interpretability.

The H matrix stores each subject's loading coefficient or weighting onto each component. If a multi-metric input is used, the H matrix shows each subject's loading onto each cortical metric comprising every component. It is organized with k components by n subject-metric combinations. Due to the projective property, the subject-by-metric loadings are estimated as the projection of the data matrix X to the components: $H = W^T X$ (Sotiras et al., 2015; Yang & Oja, 2010). The magnitude of a loading coefficient therefore indicates the extent to which a given metric is expressed within each component's covariance pattern for a given participant, indicating subject-level macrostructural variation.

NMF is gaining traction in the ASD neuroimaging literature. Tang et al. (2019) used joint-symmetrical NMF to identify modules of functional MRI (fMRI) data with network clustering, then trained ASD classifiers with the extracted features. Similarly, Pua et al. (2021) used automatic relevance determination and projective NMF to link individual variation in fMRI subnetworks to symptom severity. With electroencephalography data, Zhou et al (2020b) identified frequency-specific networks that differ between ASD and NT individuals, as well as age-related changes in connectivity (Zhou et al., 2020a). NMF has also been implemented with diffusion tensor imaging data to identify ASD-specific components and their developmental features using a projective variant, together with graph embedding (Ghanbari et al., 2014). With structural measures, a recent use case (Shan et al., 2022) involved ℓ^0–sparse NMF to identify GMV components in ASD and NT participants. Subject-level loadings were then used to evaluate individual deviations in a normative modelling framework, followed by clustering to identify ASD subtypes. ASD-specific components of gene expression have also been detected

across common and cell-type specific patterns with integrative regularized NMF (Guan et al., 2022).

Outside of ASD research, there has also been a growing interest in using NMF for multimodal analyses. One approach has relied on topic modelling, which has been used to incorporate fMRI, structural, and phenotypic data in ADHD participants (Anderson et al., 2014), as well as in imaging genetics using multi-view NMF with sparsity constraints and prior knowledge use (Won et al., 2022). Multimodal fusion has also been performed within a joint NMF framework, which factorizes more than one input matrix simultaneously (Zhang et al., 2012). Wang et al. (2020) used a variant of this approach, group sparse joint NMF, to fuse fMRI, single nucleotide polymorphism, and DNA methylation data to identify latent multimodal patterns in schizophrenia (Wang et al., 2020). Lastly, within the orthogonal projective NMF framework, multimodal input matrices have been used to decompose covariance patterns across structural and diffusion MRI measures to create parcellation maps of the hippocampus (Patel et al., 2020) and striatum (Robert et al., 2022).

2.6.3.2 Partial Least Squares Correlation (PLSC)

PLSC is an associative multivariate technique that finds a linear combination of variables that yields maximally correlated orthogonal LVs relating two input matrices, each containing different data types from the same set of participants (McIntosh & Mišić, 2013). This method can be used to identify interdependencies between variables of interest by inferring underlying multivariate dimensions linking brain and behavioural data. The input consists of a matrix X (n x p) organized n subjects by p brain variables and a matrix Y (n x q) organized n subjects by q behavioural variables. The X and Y matrices are standardized across columns. A cross covariance matrix R (q x p) is generated where

$$R = Y^{T}X \tag{5}$$

Since the relationship between X and Y is measured by the dot product, standardizing the input yields a correlation value between corresponding columns of the two matrices. This cross-correlation matrix is then subjected to SVD (Krishnan et al., 2011), such that

$$R = U\Delta V^{T} \tag{6}$$

The SVD decomposes X and Y into left and right singular vectors (U [p x l] and V [q x l]; i.e. saliences) and a diagonal matrix of singular values (Δ) whereby the rows of U and V

correspond to input matrices X and Y, respectively, whereas their columns represent novel components. The ith columns of U and V weigh the contribution of each individual input feature to the LV such that they maximally correlate (Krishnan et al., 2011; McIntosh & Mišić, 2013; Zeighami et al., 2019). The ith singular value is proportional to the covariance explained by a LV and is computed as the ratio of the squared singular value to the sum of all squared singular values (McIntosh & Lobaugh, 2004; Zeighami et al., 2019). Therefore, the ith left singular vector (U), ith right singular vector (V), and ith singular value jointly comprise a LV (Zeighami et al., 2019). Additionally, brain and behaviour scores are computed by projecting the original subject-wise matrices onto the weighted patterns U and V as

(7) $$L_x = XU \text{ and } L_y = YV$$

to quantify the extent to which a given subject maps onto the LV (Zeighami et al., 2019).

PLS has been useful for relating neuroimaging and behavioural variables in ASD. Previous studies have associated GMV patterns with social skill measures (Sato et al., 2017) and diagnostic membership (Ecker et al., 2012). It has also been used in combination with classification algorithms to identify diagnosis- and sex-specific phenotypes (Gorriz et al., 2019). PLS has been employed with functional data to associate age-related changes in resting state fMRI with diagnostic status and symptom severity (Morgan et al., 2019) and to link magnetoencephalography-derived connectivity with executive functioning skills (Mišić et al., 2015).

PLS Regression (PLSR) variants have also been implemented in ASD to link diagnostic differences in CT measurements (Romero-Garcia et al., 2018) and functional connectivity networks (Long et al., 2022) to gene expression profiles, to identify resting state functional brain networks related to behavioural metrics (Wong et al., 2016), and to account for site effects in CT (Moradi et al., 2017) and resting state fMRI data (Bhaumik et al., 2018).

Taken together, as researchers look beyond case-control paradigms and mass-univariate testing, there has been growing interest in using large and representative samples, multi-metric approaches, and data-driven techniques to more comprehensively characterize individual variation of cortical anatomy in ASD. The main objective of this thesis is therefore to employ multivariate analyses to link individualized multi-metric covariance patterns with common sources of phenotypic heterogeneity within a dimensional framework of ASD.

3. Methods

Broadly, the methodological workflow is as follows: First, T1-w scans were processed, and cortical derivatives were extracted. Case-control differences in global and vertex-wise CT, SA, LGI, and MC measures were assessed with linear mixed effects models and general linear models, respectively. The CovBat algorithm was implemented to remove batch effects in the mean, variance, and covariance of the four measures in a cross-metric framework. NMF was subsequently used to model joint covariance patterns between the CovBat-corrected cortical indices. To confirm that methodological decisions did not bias our results, we performed NMF on two decomposition granularities, on matched and unmatched subjects, and on GMV-residualized and non-residualized input measures. Univariate analyses were then used to examine diagnostic differences in NMF-derived covariance loadings. Lastly, PLSC was used to identify LVs that describe brain-behaviour interrelations. The most significant and stable LV was qualitatively related to additional cognitive and clinical variables.

3.1 Participant and Acquisition Details

The sample consists of 1097 individuals aged 2-65 years (452 ASD; 42.9% female; mean age=17.11 y/o) from three sources: the Hospital for Sick Children (SickKids Hospital; n=520) in Toronto, the Province of Ontario Neurodevelopmental Disorders (POND) Network (n=330) <https://www.braincode.ca/content/open-data-releases>, and the UK Medical Research Council Autism Imaging Multicentre Study (UK AIMS) Consortium (n=247). Demographic information can be found in Table 4.1.

3.1.1 SickKids Hospital

A total of 520 cross-sectional T1-w scans acquired between November 2010 and May 2016 at the Taylor laboratory at SickKids Hospital in Toronto, Ontario are included in the study. Scanning was performed with a Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence on a 3T Siemens Tim Trio (MAGNETOM, Siemens AG, Erlangen, Germany) with a 12 channel head coil and $1mm^3$ resolution; FOV= 192 mm x 240 mm x 256 mm; TR/TE/TI/FA = 2300ms/2.96ms/900ms/9°. Foam padding around the head was used to limit movement of participants, who watched a movie during image acquisition using an MRI-safe audio/visual system.

Research ethics approval was received from the Hospital for Sick Children Research Ethics Board. Informed written consent was obtained from adult participants and parents of participating children. Informed verbal assent was obtained from children. ASD diagnosis was ascertained by expert clinical judgment through medical diagnostic reports as well as ADOS General and ADOS-II assessments (Lord et al., 2000; 2012) in accordance with DSM-IV or DSM-V criteria (American Psychiatric Association, 2000, 2013; Leung et al., 2019; Safar et al., 2020, 2021; Vogan et al., 2019; Yuk et al., 2020). A subset of 437 subjects had FIQ scores available, assessed with age-appropriate versions of the Wechsler Intelligence Scale (Wechsler, 1999, 2011).

3.1.2 POND

A total of 330 baseline cross-sectional T1-w scans acquired between July 6, 2016 and March 8, 2021 are included in the study. Scanning was performed at the SickKids Hospital (n=260) in Toronto, Ontario on a 3T Siemens Prisma Fit and at Queen's University in Kingston, Ontario (n=70) on either a 3T Siemens Tim Trio (n=67) or 3T Siemens Prisma Fit (n=3). Acquisition sequences are as follows. At SickKids Hospital, Toronto, Ontario: MPRAGE sequence on a Siemens Prisma Fit with 20 channel head coil and 0.8mm^3 resolution; FOV=240mm x 300mm x 320mm; TR/TE/IT/FA=1870ms/3.14ms/945ms/9°. At Queen's University, Kingston, Ontario: i) MPRAGE sequence on a 3T Siemens Tim Trio with 32 channel head coil and 1mm^3 resolution; FOV= 224mm x 300mm x 320mm; TR/TE/TI/FA= 2300ms/2.98ms/900ms/9°. ii) MPRAGE sequence on a 3T Siemens Prisma Fit with 20 channel head coil and 0.8 mm^3 resolution; FOV= 240mm x 300mm x 320mm; TR/TE/TI/FA= 1870ms/3.1ms/945ms/9°.

Research ethics approval was received from each site's institutional research ethics board. Informed written consent was obtained from adult participants and parents of participating children. Informed verbal assent was obtained from children. ASD subjects have a primary diagnosis of ASD in accordance with DSM-IV or V criteria (American Psychiatric Association, 2000; 2013). Clinical ascertainment was performed with the ADOS-II (Lord et al., 2012) and ADI-R (Lord et al. 1994) by experienced clinicians. NT control subjects have no history of psychiatric, neurodevelopmental, or neurological diagnoses, premature birth (<35 weeks), nor family history of neurodevelopmental disorders in first-degree relatives (Gharehgazlou et al., 2021; Hammill et al., 2021; Jacobs et al., 2021; Kushki et al., 2021; Nakua

et al., 2022; Safar et al., 2022). FIQ was available for 291 subjects, estimated using age-appropriate standardized assessments including the Stanford-Binet Intelligence Scales (Roid & Pomplun, 2012) or Wechsler Intelligence Scales (WASI, WASI-II; WISC-IV; WPPSI-III, Wechsler 1999, 2003, 2005, 2011, 2012).

3.1.3 UK AIMS

A total of 247 cross-sectional T1-w scans from the UK AIMS Consortium are included in the study. Images were acquired at the Institute of Psychiatry, Psychology and Neuroscience at King's College London (IoP; n=121) and the Autism Research Centre, University of Cambridge (n=126). Scanning was conducted using 3T GE HDx Signa Systems scanners (General Electric, Milwaukee, USA) with an 8 channel receive-only RT head coil. Standardized acquisition protocols were implemented across sites, with a Driven Equilibrium Single Pulse Observation of T1 sequence that simulated T1-w inversion recovery images; $1mm^3$ resolution; FOV= 176mm x 256mm x 256mm; TR/TI/FA = 1800ms/50ms/20° (Bedford et al., 2020; Ecker et al., 2012, 2017; Lai et al., 2012).

Ethical approval was obtained from the National Research Ethics Committee, Suffolk, England. Subjects with ASD met diagnostic criteria of either International Classification of Diseases-10 (WHO, 1992) or DSM-IV guidelines (American Psychiatric Association, 2000). Diagnosis was ascertained with the ADI-R (Lord et al. 1994) by experienced clinicians. Exclusion criteria for all subjects include a history of psychiatric disorders, substance use disorders, severe head injury, ASD-related genetic disorders, or neurological conditions that affect brain structure and function. NT control subjects did not have a family history of ASD (Deoni et al., 2008; Ecker et al., 2012, 2016, 2017; Lai et al., 2011, 2012, 2013). Intelligence was assessed with the WASI (Wechsler, 1999).

3.2 Image processing

3.2.1 Preprocessing

Raw T1-w scans were preprocessed in dataset-specific batches with the iterativeN4 Multispectral pipeline <https://github.com/CoBrALab/iterativeN4_multispectral >. Briefly, the pipeline performs N4 correction of bias field inhomogeneities (Tutison et al., 2010) with an iterative correction and tissue estimation algorithm. Target voxels were successively refined using intensity thresholding by optimizing cross-correlation measures across increasing

resolutions. Scans underwent affine registration to the MNI ICBM NLIN SYM 09c model and patch-based brain segmentation using the BeAST algorithm (Eskildsen et al., 2012), followed by an additional stage of registration performed non-linearly, and tissue classification. The final outputs include an inhomogeneity corrected image, a corrected non-local-means denoised image, a brain mask, and converged tissue classification. The pipeline was run with the default settings for tissue priors, convergence threshold (0.01), maximum iterations (10), prior weights for classification probabilities (0.25), and cropping parameters.

3.2.2 Processing

The bias field-corrected version of T1-w scans along with each scan's generated brain mask were then submitted to the CIVET 2.1.1 pipeline in dataset-specific batches (Montreal Neurological Institute; <https://www.bic.mni.mcgill.ca/ServicesSoftware/CIVET>; Ad-Dab'bagh et al., 2006). T1-w images were linearly registered to MNI ICBM152 average template (Kim et al., 2005; Lepage et al., 2017; MacDonald et al., 2000; Mazziotta et al., 2001; Tohka et al., 2004; Zijdenbos et al., 1998). Voxels were classified into WM, cortical and subcortical GM, CSF, and background. Cortical surface reconstruction was performed for each hemisphere with the CLASP algorithm and PVE estimation (Kim et al., 2005; Lepage et al., 2017; MacDonald et al., 2000; Tohka et al., 2004; Zijdenbos et al., 1998). Extracted surfaces were nonlinearly co-registered to surface templates based on sulcal morphometry using spherical registration, while maintaining vertex-wise mesh alignment between surfaces classes (Lepage et al., 2017; 2021; Robbins et al., 2004).

The pipeline was implemented with default CIVET parameters in addition to the following specifications: 12-parameter affine alignment to template model, inclusion of MC maps, hippocampus masking, disabling of N3 correction (Sled et al., 1998), iterative PV estimates, tlaplace method for CT estimates (Lerch & Evans, 2005) and Automated Anatomical Labeling parcellation (Tzourio-Mazoyer et al., 2002).

3.2.3 Cortical Derivatives

Cortical measurements were derived at 81,924 vertices, resampled to the MNI ICBM152 surface model. Vertex-wise CT was computed as the distance between WM and pial boundaries with automated three-dimensional extraction via Laplacian distance and PVE classification (Lerch & Evans, 2005; Tohka et al., 2004) and blurred with a 30mm full-width

at half-maximum (FWHM) surface-based diffusion smoothing kernel. Vertex-wise SA was computed by imposing a Voronoi parcellation and calculating one third of the area of each adjacent triangular facet adjoining a vertex, midway between GM and WM surfaces (Lyttelton et al., 2009), and smoothing with a 40mm FWHM surface-based diffusion kernel. Vertex-wise LGI represents the quantity of cortex buried in sulci relative to the convex hull, computed with a post-processing script using CIVET-derived surfaces <https://github.com/r03ert0/surfaceratio> and a sphere radius of 20mm (Toro et al., 2008). Vertex-wise mean surface curvature was computed as the average of principal curvatures midway between GM and WM surfaces, then smoothed with a 30mm FWHM surface-based diffusion kernel (do Carmo et al., 1976; Kim et al., 2006; King et al., 2016; Luders et al., 2006).

The 'averge_surface' CIVET command was used to compute population surface average brain maps for visualizing statistical outputs. Vertices belonging to the non-cortical midline of both hemispheres were masked, leaving a total of 77,122 vertex-wise measures per cortical metric to be used in analyses.

3.3 Quality Control (QC)

Step-wise quality control (QC) was implemented at three stages of the workflow. Rating was performed by the author across all QC stages for SickKids and UK AIMS datasets, and by the author and a co-rater for the POND dataset. Raters established reliability by reviewing criteria outlined in QC guidelines, rating the same set of out-of-sample scans, cross-referencing scores, resolving discrepancies, and repeating the process if necessary. For POND QC, raters split subjects in half, rated their respective sets, and discussed borderline scans to arrive at a consensus. All scores were reviewed by the author to ensure consistency with rating criteria used for other datasets. Intra-rater reliability was enforced by re-rating a minimum of half of the scans at each stage, with particular attention given to ambiguous cases. Raters were blind to phenotypic data, including diagnosis.

3.3.1 Motion QC

First, raw T1-w scans were inspected for the presence of visible motion artifacts such as ringing and blurring. The scans were rated on a four-point scale according to the following criteria <https://github.com/CoBrALab/documentation/wiki/Motion-Quality-Control-(QC)-Manual> with 0.5 ratings assigned for additional specificity. Scores were allocated based on

the proportion of the brain affected and depth of artifacts. A stringent cut-off rating of ≥2 was used to exclude scans as failures (n=265). Failed ratings were generally reserved for ringing and blurring that were non-superficial, extended throughout several slices, and/or were accompanied by notable acquisition artifacts. For POND participants with multiple baseline scans, the highest quality scan was kept. If equivalent ratings were obtained for multiple scans from a given session, those with the earliest acquisition date were kept.

3.3.2. iN4 QC

The QC images generated by the iterativeN4 Multispectral pipeline were rated on a five-point scale according to the following criteria <https://docs.google.com/presentation/d/1vTqy9KDop_2QD7p6-NIYV-rL5IdkZyvQRPc69-kWDlA/edit#slide=id.ga9ce5ca1b4_0_10>. QC scores were penalized if residual bias field, brain masking errors (e.g., brain mask contamination with non-brain tissue, mask underestimation, image registration inaccuracies), and/or misclassification of tissue classes were identified, with ratings dependent on the extent and severity of these artifacts. Scans with a score of >3 were considered failures and excluded (n=19). Failed ratings were reserved for significant misclassification or underestimation of brain mask, exclusion of brainstem or cerebellum, and/or significant regions of uncorrected bias field. Since the tissue classification produced by the iterativeN4 Multispectral pipeline was not used for further data processing or analyses, the presence of any related errors in the iN4 QC images were noted but not counted towards a failed rating.

3.3.3. CIVET QC

QC images (i.e., verify.png and clasp.png) generated by the CIVET pipeline were rated on a three-point scale according to the following criteria: <https://github.com/CoBrALab/documentation/wiki/CIVET-Quality-Control-Guidelines>. Scores were dependent on the accuracy of surface classification and extraction, with penalties assigned for segmentation errors of tissue classes (i.e., overestimations, underestimations, misclassifications). Scans with white matter bridges, blood vessels, and misclassified ventricles were flagged and examined in greater detail. Specifically, raw images were visualized with GM and WM surface objects overlaid to assess extent of the errors. Other CIVET QC images were

examined to corroborate scan quality (e.g., laplace.png, converge.png, surfsurf.png). Scans with a score of 0 were considered failures (n=51). CIVET failures were generally reserved for sizable under-, over- or misclassifications of tissue/surface classes, perfuse WM bridges, and/or ventricle misclassifications, or a combination of such errors that significantly diminished the quality of tissue surfaces.

Scans that passed all three stages of QC and had complete data for the main variables of interest were included in statistical analyses.

3.4 Batch Correction

Differences in scanning equipment and acquisition protocols between sites are known to induce unwanted variability in the data that can systematically bias cortical measures and reduce reliability (Fortin et al., 2017, 2018). A common strategy to resolve this issue involves residualizing the site or scanner from cortical measures of interest and using the residuals in statistical analyses. However, this approach may remove variance attributable to behavioural variables that are confounded with site or scanner, in cases where systematic demographic differences exist. Additionally, this approach fails to account for site effects in the covariance structure of the measures, which is of particular interest in multivariate analyses. Therefore, to address these concerns, raw brain measures were harmonized across sites with the Correcting Covariance Batch Effects (CovBat) batch correction tool <https://github.com/andy1764/CovBat_Harmonization> (Chen et al., 2022), a multivariate extension of the original Combating Batch effects when Combining Batches (ComBat) package (Fortin et al., 2017, 2018). We opted to use CovBat instead of ComBat because the former removes site effects from mean, variance, and covariance of brain data, whereas the latter removes site effects solely from mean and variance. Since the aim of this thesis is to characterize covariance patterns *between* multiple cortical indices, we sought to remove site effects from the combined covariance structure of all four cortical measures rather than batch-correcting each metric individually, as is typically done.

In brief, CovBat uses Empirical Bayes to adjust site-specific location and scaling parameters by shifting within-site mean and variance in the marginal distributions of cortical measures to that of the pooled data (Chen et al., 2022; Fortin et al., 2017, 2018). Within-site covariance is adjusted by performing principal component analysis (PCA) on the ComBat-corrected residuals and shifting the covariance matrix of each site to that of the combined

covariance structure. The adjusted PC scores are then projected back from eigenspace to yield vertex-wise corrected values in their original units (Chen et al., 2022).

3.5 Statistical Analyses

3.5.1 Descriptive Statistics

Participants who passed QC (n=762; Table 4.1) were characterized with descriptive statistics using the base 'stats' package in R v.3.5.1 <https://www.r-project.org/>. Non-parametric tests were used for assessing group differences in continuous variables to account for non-normal data distribution.

A chi-square test was used to assess diagnostic group differences (i.e. ASD vs NT control) in the proportion of males to females. A non-parametric Wilcoxon-Mann-Whitney test was used to assess diagnostic group differences in age and FIQ.

Next, chi-square tests were used to assess QC group differences (i.e. included vs excluded) in the proportion of ASD to NT control subjects and males to females. Wilcoxon-Mann-Whitney tests were used to assess QC differences in age, FIQ, and clinical severity

Finally, chi-square tests were used to assess dataset differences in the proportion of included to excluded subjects, ASD to NT control subjects, and males to females. A Kruskal-Wallis test was performed to assess dataset differences in age and FIQ, followed by the pairwise Wilcoxon Rank Sum test with Benjamini-Hochberg correction for multiple comparisons.

3.5.2 Univariate Analyses

Associations between diagnostic status and cortical measures were evaluated with linear mixed effects models and general linear models.

3.5.2.1 Linear Models: Global Measures

To inspect diagnostic differences in raw global brain measures, linear mixed effects models were performed with the 'lme4' package in R v.3.5.1 on subjects who passed QC and had available diagnosis, age, and sex data (n=749). The models were fit with diagnosis, age, sex, and TBV as fixed effects, global cortical measure as outcome variable, and site as a random effect. A version of each model was also run without including TBV as a covariate. The variance inflation factor was used to confirm the absence of collinear predictors. Regression diagnostic tests were performed on model residuals and the random effect structure to confirm

that linear model assumptions were met. Results are reported with a Bonferroni correction for six models (α of 0.05/6 = padj < 0.008):

$$Mean\_CT \sim Diagnosis + Age + Sex + TBV + (1|Site)$$

$$Total\_SA \sim Diagnosis + Age + Sex + TBV + (1|Site)$$

$$Mean\_LGI \sim Diagnosis + Age + Sex + TBV + (1|Site)$$

$$Mean\_MC \sim Diagnosis + Age + Sex + TBV + (1|Site)$$

$$GMV \sim Diagnosis + Age + Sex + TBV + (1|Site)$$

$$TBV \sim Diagnosis + Age + Sex + (1|Site)$$

3.5.2.2 Linear Models: Vertex-wise Measures

Vertex-wise general linear models were performed with the base 'stats' package in R v.3.5.1 on CovBat-corrected measures (see section 3.5.2 for CovBat implementation details) for subjects who passed QC and had available diagnosis, age, sex, and FIQ data (n=668; Supplementary Table 2). Statistical models were fit for each vertex-wise cortical metric with diagnosis, age, sex, and the global equivalent of each vertex-wise measure set as predictor terms. Correction for multiple comparisons was done with false discovery rate (FDR; Benjamini & Hochberg, 1995) using the 'p.adjust' function. The statistical models are as follows:

$$CT_i \sim Diagnosis + Age + Sex + Mean\_CT$$

$$SA_i \sim Diagnosis + Age + Sex + Total\_SA$$

$$LGI_i \sim Diagnosis + Age + Sex + Mean\_LGI$$

$$MC_i \sim Diagnosis + Age + Sex + Mean\_MC$$

Masked midline vertices were added back to map the data to its original dimensions (81,924 vertices). FDR-corrected t-values were visualized using functions in MINC-toolkit v.1.9.18 and MINC-toolkit-extras v.1.0 with a modified create_civet_image.sh script <https://github.com/CoBrALab/minc-toolkit-extras/blob/master/create_civet_image.sh> on population surface average .obj files.

3.5.3 CovBat

3.5.3.1 CovBat Implementation

Vertex-wise cortical matrices were extracted for subjects that passed QC and had complete data for diagnosis, age, sex, and FIQ (n=668; Supplementary Table 2). This was done

for every cortical measure (CT, SA, LGI, and MC) to yield four matrices, each of dimensions (77,122 vertices by 668 subjects). Each column contained all cortical vertices for a given subject, and each row contained cortical values for a given vertex across subjects. The matrices were concatenated together vertically into a tall matrix of dimensions (77,122 vertices * 4 measures = 308,488 vertices by 668 subjects). Each subject's unique combination of dataset, site, and scanner was coded as a 'batch', representing the smallest grouping level. The list of batches are as follows: SickKids_Toronto_Trio, POND_Toronto_Prisma, POND_Queens_Trio, POND_Queens_Prisma, UKAIMS_IoP_GE, and UKAIMS_Cambridge_GE. The three subjects belonging to the POND_Queens_Prisma designation did not pass motion QC, so this batch was excluded from CovBat correction. The default Empirical Bayes parameters were implemented in CovBat package v.1.0 <https://github.com/andy1764/CovBat_Harmonization>, which uses functions from <https://github.com/brentp/combat.py> in Python v.3.6.8 <https://www.python.org/>. The number of components accounting for 95% covariance explained was selected, as recommended by the CovBat developers (Chen et al., 2022). Variability attributable to batches was removed from cortical measures, while inter-site biological variability in diagnostic status, sex, age, and FIQ was preserved. The multi-metric CovBat-corrected output matrix was then split into each of the four measures, yielding four matrices of CovBat-adjusted residuals, each of dimension (77,122 vertices by 668 subjects).

Due to this novel multi-metric CovBat implementation, we sought to confirm that: 1) batch-specific mean, variance, and covariance were removed from CovBat-corrected measures and 2) batch-specific demographic associations with cortical measures were retained. For validation analyses, both the raw pre-CovBat (77,122 vertices by 668 subjects) and corrected post-CovBat (77,122 vertices by 668 subjects) matrices were used for each of the four measures. To ensure that diagnosis, age, and sex were not confounded with batches, these three variables were residualized from raw and CovBat-corrected measures prior to the analyses outlined in Aim 1.

3.5.3.2 CovBat Validation

All CovBat validation analyses were performed with the base 'stats' package in R v.3.5.1. Batch effects in mean vertex-wise cortical measures were examined before and after

CovBat correction. Vertex-wise one-way ANOVAs were performed separately on pre- and post-CovBat values:

$RawCorticalMeasure_i\_Vertex_j \sim Batch$

$CovBatCorrectedCorticalMeasure_i\_Vertex_j \sim Batch$

The resulting p-values were FDR-corrected across 77,122 vertices with the 'p.adjust' function. The proportion of vertices with significant batch effects were computed.

Batch effects in the variance of median cortical measures were examined before and after CovBat correction with Bartlett's Sphericity test:

$RawMedianCorticalMeasure_i \sim Batch$

$CovBatCorrectedMedianCorticalMeasure_i \sim Batch$

Batch effects in covariance patterns between the cortical measures was assessed by performing a PCA on scaled median CT, SA, LGI, and MC values with the 'prcomp' function. Batch effects on PC weights were assessed with a one-way ANOVA:

$RawMedianPC_i\_Weights \sim Batch$

$CovBatCorrectedMedianPC_i\_Weights \sim Batch$

Tukey's Honestly Significant Difference test was used to assess pairwise differences.

To confirm that diagnostic, age, and sex relationships with cortical measures were retained after batch correction, general linear models were implemented on pre- and post-CovBat median values:

$RawMedianCorticalMeasure_i \sim Diagnosis + Age + Sex$

$CovBatCorrectedMedianCorticalMeasure_i \sim Diagnosis + Age + Sex$

Improvement in overall model fit was assessed by comparing the adjusted R^2 and residual standard error (RSE) values. The effect size and significance of each model predictor was assessed by comparing t-values and p-values.

3.5.4 Covariate Matching

To ensure that multivariate analyses are not driven by the disproportionate ratio of ASD to NT control subjects, the sample of 668 subjects that passed QC without missing data for the four main variables of interest were matched across diagnostic groups for age and sex. The

'Matchit' package v.4.4 (Ho et al., 2011) was used in R v.4.1.2, using functionalities from the 'Matching' package (Diamond & Sekhon, 2013; Sekhon, 2011). Improvement in covariate balance was assessed with standardized mean difference, variance ratio, and empirical cumulative density functions. Genetic matching with Mahalanobis distance yielded the most robust balance statistics and was therefore chosen as the optimal matching method. A total of 486 participants remained after the matching procedure (Supplementary Table 3).

3.5.5 Multivariate Analyses

3.5.5.1 Non-negative Matrix Factorization (NMF)

For a given cortical measure, vertex-wise metrics were stacked across both hemispheres to create a single column vector of data per subject. A vertex-by-subject (m x n) matrix for the 486 matched subjects was created for each of the four cortical measures, with participants organized by ascending age within each metric block (77,122 vertices by 486 subjects). Next, raw vertex-wise cortical matrices were CovBat corrected as described above (section 3.5.2) in CovBat Run 1. In brief, the four metric blocks were concatenated vertically, creating a multi-metric input matrix (308,488 vertices by 486 subjects) that was CovBat corrected to remove variability in cortical measures attributable to batch, and retain variability attributable to diagnosis, age, sex, and FIQ. The default Empirical Bayes parameters were implemented in CovBat package v.1.0 <https://github.com/andy1764/CovBat_Harmonization> in Python v.3.6.8.

The large batch-corrected matrix was then split into four cortical matrices each of dimension (77,122 vertices by 486 subjects) and GMV was residualized on a vertex-wise basis from each metric. This was done because preliminary analyses of non-residualized data found that each subject's NMF component loadings were strongly correlated with GMV, which suggested that covariance patterns captured by the NMF may have been driven by inter-subject differences in GMV, a measure that varied substantially across the wide age range. For completeness, associations with other global brain measures were also probed (TBV, mean CT, total SA, mean LGI, mean MC), but GMV displayed the strongest relationship. Therefore, a matrix of vertex-wise GMV (77,122 vertices x 486 subjects) was extracted and CovBat corrected in CovBat Run 2 as a single metric input matrix, with batch variability in demographic variables preserved in a similar manner. GMV was computed for each subject by

summing across all CovBat-corrected GMV vertices. This measure was then residualized from each subject's vertex-wise CovBat-corrected cortical metrics that were computed in Run 1.

The batch-corrected and GMV-residualized matrices were then z-scored across columns of subjects to account for differences in magnitude between the cortical measures. This axis of standardization was chosen to ensure that individual differences in a given vertex-wise measure (columns) were not overshadowed by the greater within-subject variation in cortical measures across the cortex (rows). This allowed the NMF analysis to prioritize individual variation in covariance patterns rather than covariance within subjects, which would yield a parcellation-type decomposition. The CovBat-corrected, residualized, and z-scored cortical matrices were then concatenated together width-wise into a multi-metric matrix of dimensions (77,122 vertices by 486 subjects * 4 metrics = 1944 columns), and shifted by the minimum value to induce non-negativity. Each subject was therefore represented a total of four times, once within each cortical metric block.

3.5.5.1.1 NMF Implementation

The final matrix created in section 3.5.5.1 was used as input for NMF <https://github.com/CoBrALab/cobra-nmf>, using publicly available code <https://github.com/asotiras/brainparts> with Octave v.5.2.0 <https://docs.octave.org/v5.2.0/> functions, and the 'sklearn' v.0.23.1 package in Python v.3.6.8. The analysis was initialized with non-negative double singular value decomposition (NNDSVD; Boutsidis & Gallopoulos, 2008) with maximum iteration of 100,000 and tolerance of = 0.00001. This initialization strategy is advantageous relative to the randomized alternative because it has been shown to diminish residual error more efficiently, prioritize sparsity, and converge on a deterministic solution (Boustidis & Gallopoulos, 2008; Sotiras et al., 2015). The algorithm converges on a solution when the tolerance threshold for the error value is met. For completeness, NMF decompositions were also performed on non-GMV-residualized input matrices and on unmatched subjects (n=668; Supplementary Table 2).

3.5.5.1.2 Component Selection

The reliability and accuracy of decompositions were assessed for a range of 2-30 NMF component solutions with split-half stability analysis and change in reconstruction error. The number of components ultimately chosen for constructing the low-rank matrix approximation

balanced: 1) maximal stability, such that similar component outputs were produced when NMF was performed across random resamplings of subjects; and 2) minimal reconstruction error, such that the residual error between the reconstructed covariance pattern (W x H) and the original covariance structure of the input data (X) was attenuated.

3.5.5.1.3 NMF Stability Analysis

A split-half stability analysis was performed to quantify the degree of spatial overlap between components when NMF was applied to random subsamples of subjects. The 486 subjects were split into two subgroups, na=243 and nb=243 with stratified random sampling by diagnosis, sex, and median values for age, and FIQ. This was done across 10 unique splits of participants for a total of 20 subgroups. Multi-metric input matrices Xa and Xb (77,122 vertices by 243 * 4 measures = 972 subjects) were generated for each subgroup of each split after residualizing GMV on a vertex-wise basis from each cortical measure. NMF was performed for each subgroup's input matrix for every other granularity between 2-30 components. Together, this yielded a total of 20 subgroups x 15 granularities = 300 NMF runs. Each combination of split and granularity therefore had two component matrices Wa and Wb (m x k), and two weight matrices Ha and Hb (k x n), with one for each subgroup.

For each component matrix, a similarity matrix CWa and CWb (m x m) was generated by computing the cosine similarity between component scores of each vertex with every other vertex (i.e. between rows of W matrices) within each subgroup, such that higher similarity scores are indicative of two vertices sharing membership in the same component (Kalantar-Hormozi et al., 2022; Patel et al., 2020; Robert et al., 2022). Corresponding rows of the similarity matrices CWa and CWb were then correlated to quantify the degree to which a given vertex is grouped with the same set of vertices when NMF is applied to different subgroups. Correlation coefficients were then averaged across all vertices to yield a stability measure for k granularity. This was repeated for a total of 10 splits of data across 15 component granularities, yielding a stability coefficient for each granularity (Kalantar-Hormozi et al., 2022; Patel et al., 2020; Robert et al., 2022). A higher correlation coefficient is indicative of more stable component solutions across subsets of the sample.

3.5.5.1.4 NMF Accuracy Testing

The reconstruction error for each split of data was computed as such:

Reconstruction error A = $|| Xa - WaHa||^2_F$

Reconstruction error B = $|| Xb - WbHb||^2_F$

such that Xa and Xb are input matrices of two subgroups of a split (Patel et al., 2020; Robert et al., 2022). The change in reconstruction error at granularity k to granularity k+2 is computed from 2-30 component solutions, representing change in accuracy as additional component granularities are added. For each granularity, reconstruction error matrices are generated by performing element-wise subtraction of the original from the reconstructed input, such that each element of the input matrix has an error value. The Frobenius norm is computed for each component's reconstruction error matrix to derive a single error value for a given decomposition k, which is then averaged across components and splits of data.

3.5.5.1.5 Linear Models: Case-Control Differences

The association between demographic variables and NMF subject-by-metric loadings for each component were examined using general linear models with the base 'stats' package in R v.3.5.1. This statistical approach assessed the effect of diagnosis on subject loadings for each of the measures comprising the covariance patterns. The following model was implemented, with Bonferroni correction applied across the total number of models:

$Component_i\_Metric_j\_Weights \sim Diagnosis + Age + Sex + FIQ$

3.5.6.1 PLSC

3.5.6.1.1 PLSC Implementation

PLSC was used to relate each subjects' loadings onto the NMF-derived components with a set of behavioural variables using the 'pyls' package v.0.1.6 <https://github.com/rmarkello/pyls> in Python v.3.8.3. The X matrix contained n subjects by p component-by-metric combinations such that a given element represents a subject's loading onto a metric within an NMF component (486 subjects by k components * 4 metric loadings). The Y matrix contained n subjects by q demographic variables (diagnosis, age, sex, and FIQ) of dimensions (486 subjects by 4 behavioural variables). Symmetrical SVD was used to decompose the cross-covariance matrix to yield a set of LVs that represent the association between subjects' loadings onto the multi-metric covariance components and their demographic information.

3.5.6.1.2 PLSC Permutation Testing

The statistical significance of the LVs was assessed through non-parametric permutation testing of singular values. Rows of the X matrix were shuffled to eliminate the relationship between brain and behaviour variables. SVD was applied to the newly generated correlations, such that a new singular value of the LVs was calculated for each permutation (n=10,000) to construct a null distribution (McIntosh & Lobaugh, 2004; McIntosh & Mišić, 2013). A threshold of $p < 0.05$ was considered significant.

3.5.6.1.3 Bootstrap Resampling

The reliability of each brain and behavioural variable's contribution to the LVs were assessed with bootstrap resampling. Sampling with replacement was done to generate a sampling distribution for each singular vector weight (Zeighami et al., 2019), then the contribution of each variable to the LV was reassessed (n=10,000). A bootstrap ratio (BSR) was calculated by dividing the vector weight of a given LV by the standard error to produce a reliability estimate (Krishnan et al., 2011; McIntosh & Lobaugh, 2004; McIntosh & Mišić, 2013). A brain salience weight of 1.96 was used to threshold significance as it is comparable to a z-score in normally distributed BSRs, thus corresponding to $p<0.05$; 95% confidence interval (Efron & Tibshirani, 1986; McIntosh & Lobaugh, 2004).

3.5.6.1.4 PLSC Latent Variable (LV) Stability Analysis

The reliability of the brain-behaviour associations captured by each LV was assessed with split-half stability using the procedure outlined in Kovacevic et al. (2013). This approach assessed whether stability of the associations between left and right singular vectors is high for different subsets of the sample. Additionally, it ensures that the LVs are not disproportionately driven by very strong associations between variables in one input matrix in cases when the other input matrix has weaker associations. In brief, the 486 subjects were resampled 200 times into split halves of (na=243 and nb=243) to create a total of 400 split half folds. Cross-correlation matrices Ra and Rb were generated for each split, which were decomposed with SVD to yield left and right singular vectors (Ua, Ub, Va, Vb). Correlations were then computed between subgroups of each split, (i.e. Ua, Ub and Va,Vb), yielding a mean correlation across split halves (Ucorr, Vcorr), which was then compared against a null distribution of randomly permuted data (pUcorr, pVcorr; Kovacevic et al., 2013).

All subject resampling for permutation and bootstrapping analyses were performed with Procrustes rotation to limit the possibility of altering the order of LVs and salience signs during SVD (McIntosh & Lobaugh, 2004).

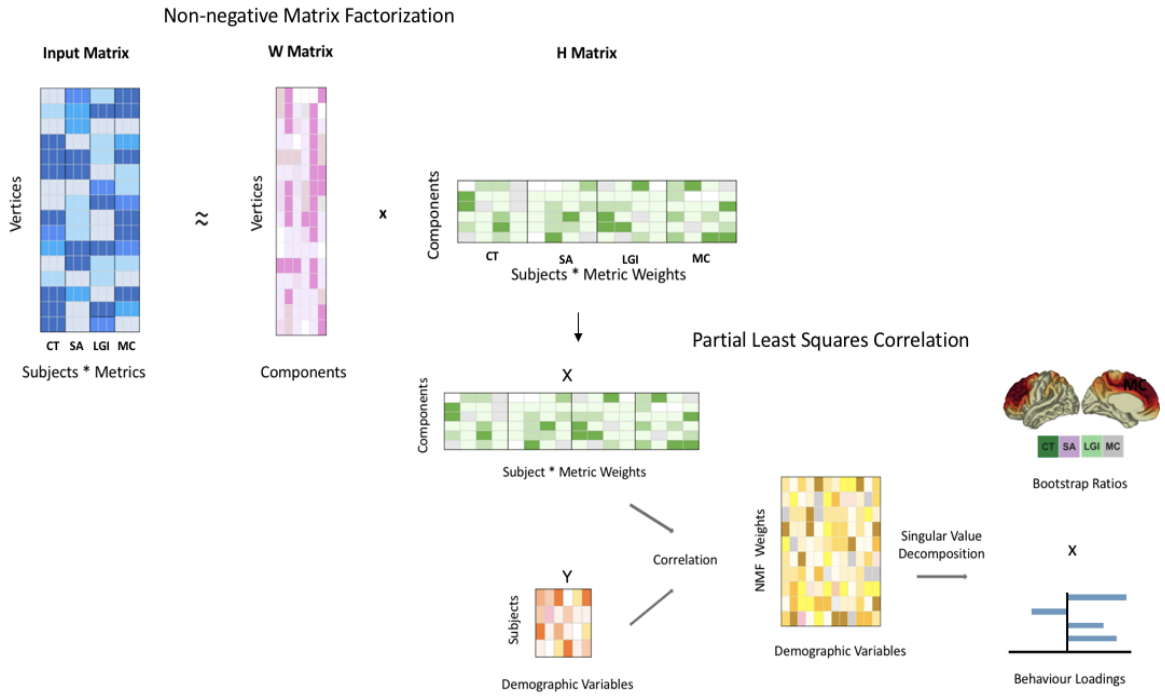


Figure 3.1. Multivariate workflow. An input matrix containing subjects' vertex-wise cortical values is decomposed with Non-negative Matrix Factorization to yield a component matrix representing the spatial location of covarying vertices and a subject matrix containing subjects' loadings onto each of the metrics comprising the components. The subject loadings are used as input into the Partial Least Squares Correlation analysis, along with a set of demographic variables. These features are cross-correlated and decomposed into latent variables, which represent a composite of maximally correlated brain and behavioural variables. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature.

4. Results

4.1. Descriptive and Univariate Statistics

4.1.1 Demographic Information

From an initial sample of 1097 subjects, a total of 762 remained after three stages of QC (Table 4.1). The subjects that passed QC were aged 4-65 years with 262 ASD (88 female/174 male) and 500 NT control subjects (246 female/254 male); SickKids (n=402), UK AIMS (n=211) and POND (n=149). Of these, a total of 668 passed subjects had complete data for the main demographic variables of interest (diagnosis, age, sex, and FIQ). Specifically, 13 subjects (SickKids) had missing ages and 85 had missing FIQ scores (n=69 SickKids, n=16 POND). For a breakdown of subjects excluded at each stage of QC, see Figure 4.2.

Table 4.1. Demographic information of subjects before and after quality control. ASD = Autism Spectrum Disorder. NT = Neurotypical. FIQ = Fullscale IQ.

Before QC (After QC)	**POND** N=330 (N=149)	**SickKids** N=520 (N=402)	**UK AIMS** N=247 (N=211)	**Total** N=1097 (N=762)
Female_ASD	47(22)	25(20)	54(46)	126(88)
Male_ASD	149(54)	106(66)	71(54)	326(174)
Female_NT	58(36)	194(165)	49(45)	301(246)
Male_NT	76(37)	195(151)	73(66)	344(254)
ASD: NT	196:134(76:73)	131:389(86:316)	120:122(100:111)	452:645(262:500)
Age (years) Range [median]	2-21.96[11.92] (4-21.96[13.78])	4-65[13] NA=16 (4-65[14]) NA=13	18-52[25] (18-52[25.76])	2-65[15] NA=16 (4-65[16.42]) NA=13
FIQ Range [median]	40-144[105] NA=39 (44-142[106]) NA=12	69-149[111] NA=83 (69-149[111]) NA=69	73-137[116] (73-137[117])	40-149[109] NA=122 (44-149[110]) NA=81

4.1.1.1 Diagnostic Differences in Demographic Variables

Diagnostic groups differ significantly in the proportion of males to females ($\chi 2$ (df =1) =17.02, p<0.001). There is no significant diagnostic difference in age, $U(N_{ASD}=262, N_{control}=500)=68033$, z=1.50, p=0.136), however ASD subjects have significantly lower FIQ (*Mdn*= 107.0) than NT control subjects (*Mdn*=114.0), $U(N_{ASD}=262, N_{control}=500)= 39370$, z = -5.63, p<0.001).

4.1.1.2 QC Differences in Demographic Variables

Included and excluded participants differ significantly in the proportion of ASD to NT control subjects ($\chi 2$(df=1)=47.90, p<0.001) and males to females ($\chi 2$(df=1)=25.28, p<0.001). Age of excluded subjects (*Mdn*= 11.0) is significantly lower than included subjects (*Mdn*=16.42), $U(N_{included}=762, N_{excluded}=335)=725566$, z= -10.93, p<0.001), as is FIQ of excluded subjects (*Mdn*=105.0), relative to included subjects (*Mdn*=112.0), $U(N_{included}=762, N_{excluded}=335)=70932$, z= -7.232, p<0.001). Conversely, clinical severity measured with the ADOS-G Total score, is significantly higher in excluded (*Mdn* = 11.0) relative to included subjects (*Mdn*=9.0), $U(N_{included}=100, N_{excluded}=25)=1437$, z= -0.64, p=0.03).

4.1.1.3 Dataset Differences in Demographic Variables

Datasets differ significantly in the proportion of included to excluded subjects ($\chi 2$(df =2)=136.7, p<0.001), ASD to NT control subjects ($\chi 2$(df =2)=109.11, p<0.001) and males to females ($\chi 2$(df=2)=10.038, p=0.006). Age differs significantly by dataset, ($\chi 2$ (df=2)=423.12, p<0.001; Figure 4.1B), including SickKids and UK AIMS (p<0.001), SickKids and POND (p<0.001), and UK AIMS and POND (p<0.001). FIQ also differs significantly by dataset ($\chi 2$(df=2)=71.292, p<0.001; Figure 4.1D), including SickKids and POND (p<0.001), SickKids and UK AIMS (p<0.001), and UK AIMS and POND (p<0.001).

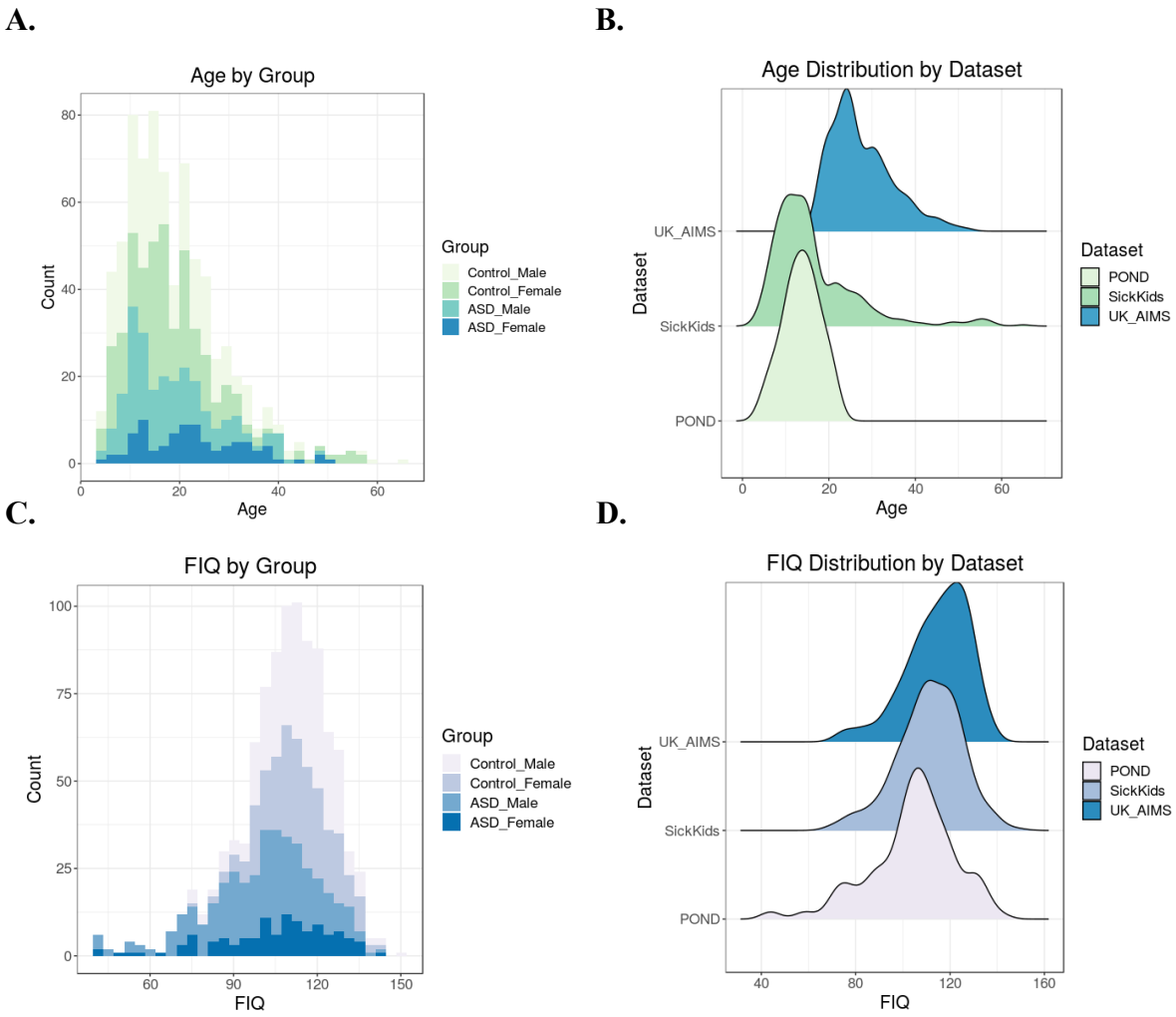


Figure 4.1. Histograms showing demographic characteristics for subjects who passed QC. **A.** Age distribution by group. **B**. Age distribution by dataset. **C**. Fullscale IQ distribution by group. **D.** Fullscale IQ distribution by dataset. FIQ = Fullscale IQ. POND = Province of Ontario Neurodevelopmental Disorders Network. UK_AIMS = UK Medical Research Council Autism Imaging Multicentre Study.

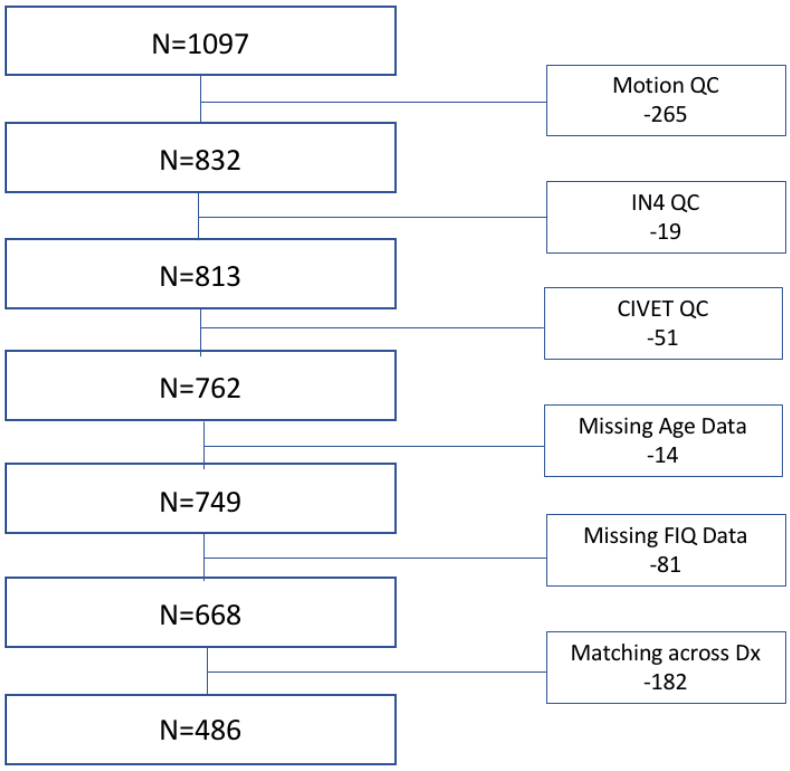


Figure 4.2. Stepwise exclusion numbers for each stage of quality control. Passed subjects with no missing age data (n=749; Table 4.1) were used in univariate analyses of global measures. Passed subjects with no missing age or FIQ data (n=668; Supplementary Table 2) were used in vertex-wise univariate analyses and CovBat validation analyses. A subset of these subjects matched across diagnosis for age and sex were used for multivariate analyses (n=486; Supplementary Table 3). QC = Quality Control. FIQ = Fullscale IQ. Dx = Diagnosis.

4.1.2 Brain Measures

4.1.2.1 Global Brain Measures: Linear Mixed Effects Models

Results of the global linear mixed effects models show that relative to NT controls, ASD subjects have insignificantly greater mean CT (β=0.0150 (SE=0.0063), t=2.36, p=0.018), total SA (β=157.9 (SE=448.9), t=0.35, p=0.725), mean LGI (β=0.0077 (SE=0.0062), t=1.22, p=0.219), GMV (β=3467 (SE=1606), t=2.15, p=0.031), and TBV (β=7909 (SE=9142), t=0.86, p=0.387), and insignificantly lower average MC (β= -0.000065 (SE=0.00010), t= -0.65 p=0.51; Figure 4.3) at the Bonferroni-adjusted significance threshold. Removing TBV as a covariate does not change the significance of any model: mean CT (β=0.0164 (SE=0.0069), t=2.36, p=0.0182), total SA (β=1139.350 (SE=1098.701), t=1.03, p=0.3), mean LGI (β=0.0089 (SE=0.006482), t=1.38, p=0.168), GMV (β=6891.921 (SE=3846.416), t=1.79, p=0.073), and mean MC (β= -0.0001.34 (SE=0.00011), t= -1.13, p=0.258).

4.1.2.2. Vertex-wise Measures: General Linear Models

Results of the vertex-wise general linear models show that there is no significant effect of diagnosis on any of the vertex-wise cortical measures at 1%-5% FDR. At a more relaxed threshold of 10% FDR, a significant decrease in CT within the ASD group is observed in the right precentral gyrus (Supplementary Figure 1). No other models yielded significant diagnostic differences. Brain maps of vertex-wise standard deviation of the four cortical measures reveal metric-specific patterns of regional variation (Supplementary Figure 2).

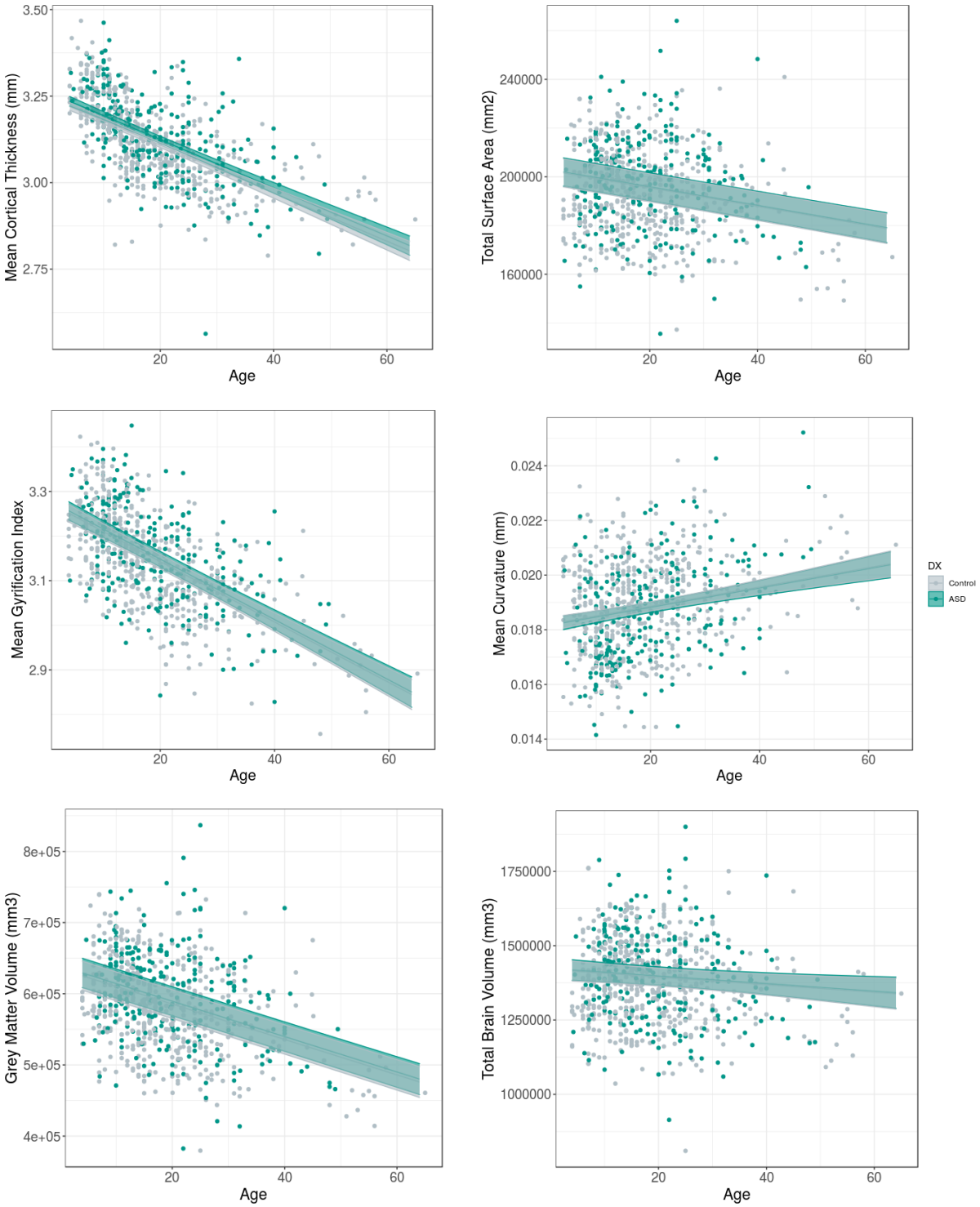


Figure 4.3. Results of linear mixed effects models showing the main effect of diagnostic group on global brain measures (y-axis), with age (x-axis), sex, and total brain volume as fixed effects and site as a random effect. Age is modelled as a linear term. Subjects are coloured by diagnosis. DX = Diagnosis. ASD = Autism Spectrum Disorder.

4.2 CovBat Validation Results

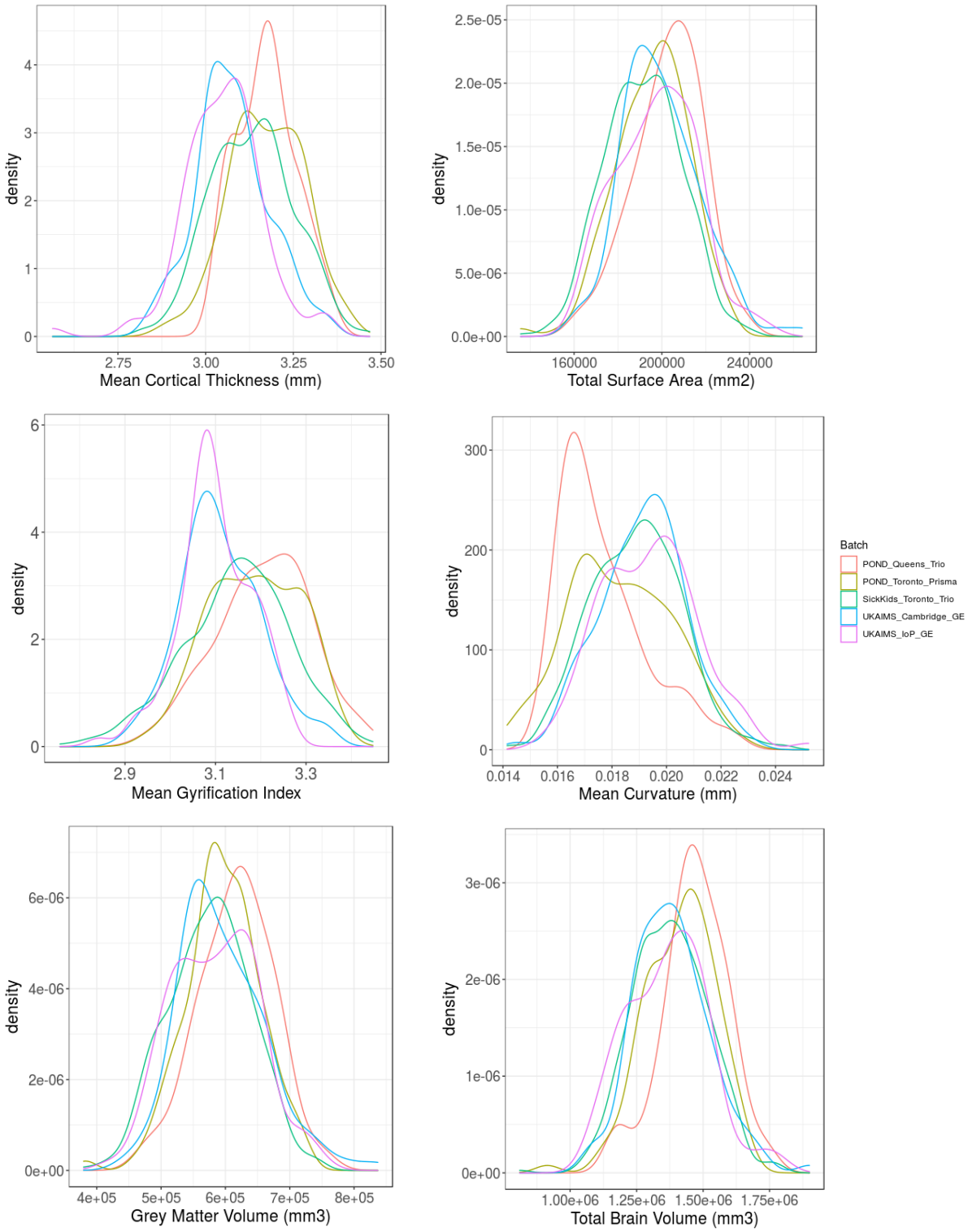


Figure 4.4. Density plots depict raw global measures coloured by batch. POND = Province of Ontario Neurodevelopmental Disorders Network. UKAIMS = UK Medical Research Council Autism Imaging Multicentre Study. IoP = Institute of Psychiatry, Psychology and Neuroscience. GE = General Electric.

4.2.1 Mean

Raw global measures differ in batch-specific variability (Figure 4.4). Results of the vertex-wise ANOVA indicate that 86.2% of raw CT vertices, 84.17% of raw SA vertices, 31.9% of raw LGI vertices, and 25.9% of raw MC vertices have significant batch effects. Following CovBat correction, no vertices remain with significant batch effects in any of the four measures.

4.2.2 Variance

Results of Bartlett's Sphericity test reveal that raw median CT values demonstrate a significant difference in variance attributable to batch (K^2=16.42, p=0.002), whereas after CovBat-correction, this difference is removed (K^2=0.73, p=0.94). Other raw cortical values do not display significant inhomogeneities of variance attributable to batch, including median SA (K^2=4.66, p=0.32; [post-CovBat (K^2=0.57, p=0.96)]); median LGI (K^2=3.61, p=0.46; [post-CovBat (K^2=1.86, p=0.76)]), as well as median MC (K^2=5.08, p=0.27; [post-CovBat (K^2=3.42, p=0.48)]).

4.2.3 Covariance

Results of the PCA performed on raw median CT, SA, LGI, and MC values identified four principal components (PCs), explaining: PC1 (45.5%), PC2 (30.12%), PC3 (16.2%) and PC4 (8.1%) of the covariance. PC1 weights differ significantly between batches (F(4)=13.54, p<0.001). Specifically, the mean value of PC1 weights is significantly different between POND_Toronto_Prisma and POND_Queens_Trio (p=0.03), SickKids_Toronto_Trio and POND_Queens_Trio (p<0.001), UKAIMS_Cambridge_GE and POND_Toronto_Prisma (p=0.004), UKAIMS_Cambridge_GE and SickKids_Toronto_Trio (p<0.001), and UKAIMS_IoP_GE and UKAIMS_Cambridge_GE (p=0.01). Additionally, PC3 weights differ significantly between batches (F(4)=2.618, p=0.03). The mean value of PC3 weights is significantly different between UKAIMS_Cambridge_GE and POND_Toronto_Prisma (p=0.04). There are no significant batch differences in PC2 weights (F(4)=0.30, p=0.877), nor PC4 weights (F(4)=0.936, p=0.442).

Although not homologous to the previous analysis, results of the PCA performed on CovBat-corrected median CT, SA, LGI, and MC identified four PCs explaining: PC1 (43.5%), PC2 (31.5%), PC3 (16.3%), and PC4 (8.5%) of the covariance. No significant batch effects

were found on PC1 weights (F(4)=0.162, p=0.958), PC2 weights (F(4)=0.046, p=0.996), PC3 weights (F(4)=0.188, p=0.945), nor PC4 weights (F(4)=0.107, p=0.98).

4.2.4 Demographic Associations

Results of general linear models used to assess demographic relationships with median cortical values are reported below in Tables 4.2-4.5. Across metrics, goodness-of-fit measures including the adjusted R^2 and RSE improve after CovBat correction. The CovBat-corrected models predict a greater proportion of variance in average cortical measures relative to the raw models, with marginally decreased RSE. The largest improvements are seen in CT, the outcome measure with the greatest batch effect. The variance explained of the model increases from 39% in raw measures to 50% in batch-corrected measures: Median CT (raw: $F(3,664)=147.7$, $p<0.001$, $R^2=0.39$, RSE=0.089; [CovBat-corrected: $F(3,664)=224.5$, $p<0.001$, $R^2=0.50$, RSE=0.086]). The parameters of other measures are as follows: Median SA (raw: $F(3,664)=89.21$, $p<0.001$, $R^2=0.28$, RSE=0.187; [CovBat-corrected: $F(3,664)=120.2$, $p<0.001$, $R^2=0.34$, RSE=0.177]); Median LGI (raw: $F(3,664)=120.5$, $p<0.001$, $R^2=0.34$, RSE=0.0086; [CovBat-corrected: $F(3,664)=155.9$, $p<0.001$, $R^2=0.41$, RSE=0.084]). Median MC (raw: $F(3,664)=37.01$, $p<0.001$, $R^2=0.13$, RSE=0.001; [CovBat-corrected: $F(3,664)=41.32$, $p<0.001$, $R^2=0.15$, RSE=0.001]).

Individual predictors within the models maintain their associations with the outcome measures (Tables 4.2-4.5), with two exceptions. One is the ‘Age’ covariate in the median SA models, which reaches significance ($p<0.001$) following CovBat correction, in contrast to its insignificant status ($p=0.08$) prior to adjustment (Table 4.3). The second is the ‘Diagnosis’ covariate in the median MC models, which is significant ($p=0.03$) prior to CovBat correction, but insignificant ($p=0.07$) after adjustment (Table 4.5).

Table 4.2. Results of linear models assessing the effect of diagnosis, age, and sex on median cortical thickness before and after CovBat correction. Degrees of freedom, F-statistic, and adjusted R^2 are shown for each model. Unstandardized beta coefficient, t-value, and p-value are shown for each predictor. CT = Cortical Thickness. DX= Diagnosis.

Statistical Model: CT	**Degrees of Freedom**	**F-statistic**	**Adjusted R^2**	**Predictor**	**β Coefficient [95% CI]**	**T-value**	**P-value**
Raw_Median_CT ~ DX + Age + Sex	3,664	147.7	0.39	DX	0.022 [0.008, 0.037]	3.11	0.001
				Age	-0.007 [-0.008, -0.006]	-19.73	p<0.001
				Sex	-0.038 [-0.052, -0.024]	-5.47	p<0.001
CovBat_Corrected_Median_CT ~ DX + Age + Sex	3,664	224.5	0.50	DX	0.018 [0.005, 0.032]	2.69	0.007
				Age	-0.008 [-0.009, -0.008]	-24.77	p<0.001
				Sex	-0.038 [-0.051, -0.025]	-5.63	p<0.001

Table 4.3. Results of linear models assessing the effect of diagnosis, age, and sex on median surface area before and after CovBat correction. Degrees of freedom, F-statistic, and adjusted R^2 are shown for each model. Unstandardized beta coefficient, t-value, and p-value are shown for each predictor. SA =Surface Area. DX= Diagnosis.

Statistical Model: SA	**Degrees of Freedom**	**F-statistic**	**Adjusted R^2**	**Predictor**	**β Coefficient [95% CI]**	**T-value**	**P-value**
Raw_Median_SA ~ DX + Age + Sex	3,664	89.21	0.28	DX	-0.026 [-0.003, 0.056]	1.74	0.081
				Age	-0.001 [-0.002,0.0002]	-1.70	0.088
				Sex	-0.233 [-0.262, -0.203]	-15.692	p<0.001
CovBat_Corrected_Median_SA ~ DX + Age + Sex	3,664	120.2	0.34	DX	0.017 [-0.010, 0.046]	1.21	0.2227
				Age	-0.005 [-0.007, -0.004]	-7.50	p<0.001
				Sex	-0.233 [-0.261, -0.206]	-16.58	p<0.001

Table 4.4. Results of linear models assessing the effect of diagnosis, age, and sex on median local gyrification index before and after CovBat correction. Degrees of freedom, F-statistic, and adjusted R^2 are shown for each model. Unstandardized beta coefficient, t-value, and p-value are shown for each predictor. LGI = Local Gyrification Index. DX= Diagnosis.

Statistical Model: LGI	**Degrees of Freedom**	**F-statistic**	**Adjusted R^2**	**Predictor**	**β Coefficient [95% CI]**	**T-value**	**P-value**
Raw_Median_LGI ~ DX + Age + Sex	3,664	120.5	0.34	DX	0.010 [-0.003, 0.024]	1.44	0.148
				Age	-0.006 [-0.006, -0.005]	-16.55	p<0.001
				Sex	-0.054 [-0.068, -0.041]	-7.92	p<0.001
CovBat_Corrected_Median_LGI ~ DX + Age + Sex	3,664	155.91	0.41	DX	0.006 [-0.007, 0.020]	0.929	0.353
				Age	-0.006 [-0.007, -0.006]	-19.23	p<0.001
				Sex	-0.055 [-0.068, -0.042]	-8.25	p<0.001

Table 4.5. Results of linear models assessing the effect of diagnosis, age, and sex on median mean curvature before and after CovBat correction. Degrees of freedom, F-statistic, and adjusted R^2 are shown for each model. Unstandardized beta coefficient, t-value, and p-value are shown for each predictor. MC = Mean Curvature. DX= Diagnosis.

Statistical Model: MC	**Degrees of Freedom**	**F-statistic**	**Adjusted R^2**	**Predictor**	**β Coefficient [95% CI]**	**T-value**	**P-value**
Raw_Median_MC ~ DX + Age + Sex	3,664	37.01	0.13	DX	-0.0002 [-0.0005, -0.00002]	-2.13	0.0331
				Age	0.00003 [0.00001, 0.00004]	4.59	p<0.001
				Sex	0.001 [0.0008, 0.001]	8.67	p<0.001
CovBat_Corrected_Median_MC ~ DX + Age + Sex	3,664	41.32	0.15	DX	-0.0002 [-0.0004, 0.00002]	-1.78	0.0749
				Age	0.00003 [0.00002, 0.00004]	5.19	p<0.001
				Sex	0.001 [0.0009, 0.001]	9.09	p<0.001

CovBat validation results confirm that the novel multi-metric implementation was successful in removing batch effects in the mean, variance, and covariance of cortical measures, while retaining biological variability attributable to demographic variables.

4.3. NMF Results

4.3.1. Component Selection

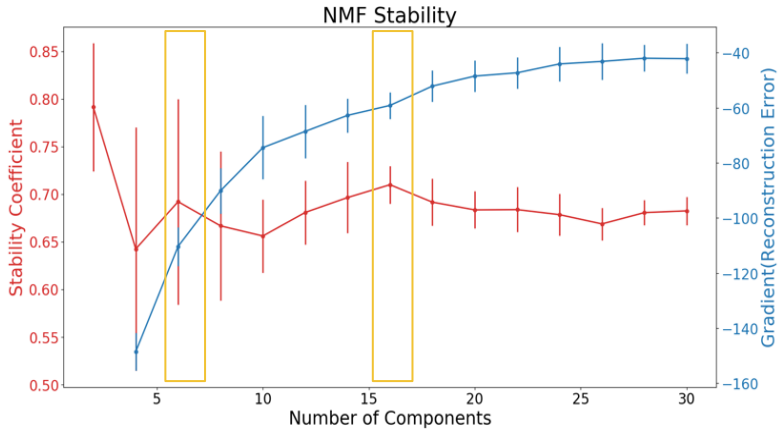


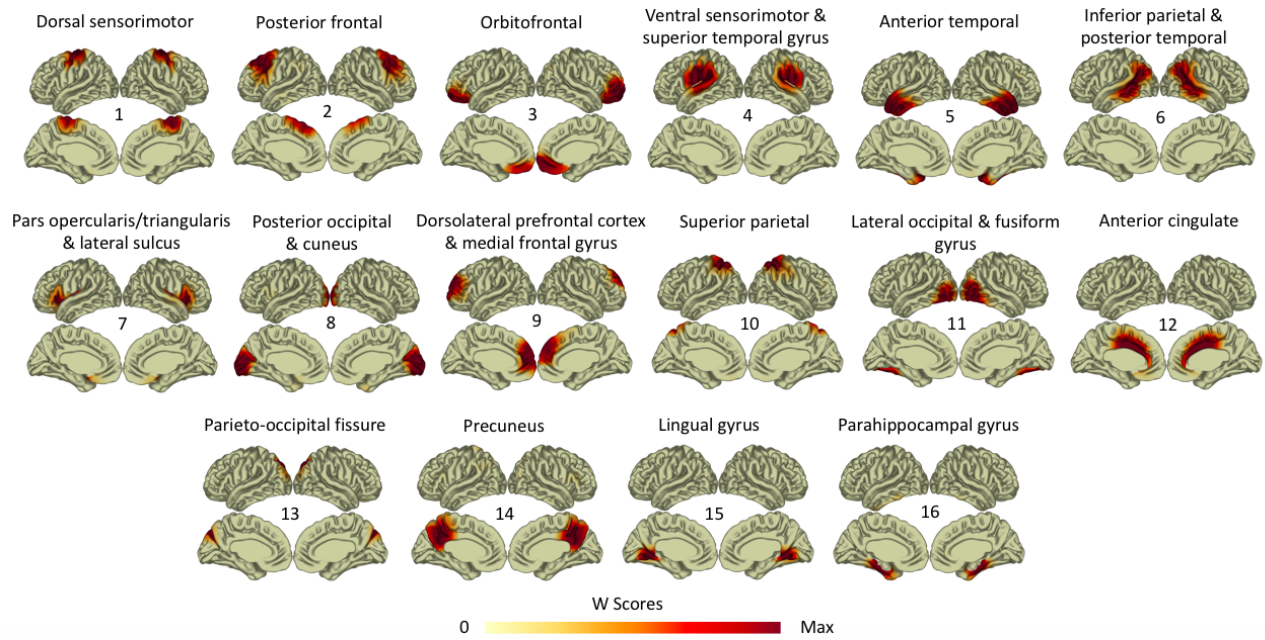
Figure 4.5. Stability coefficient (red) and change in reconstruction error (blue) across 2-30 component granularities. Granularities with optimal balance of stability and accuracy are indicated in orange boxes.

Results of the split-half stability analysis are shown in Figure 4.5. The stability coefficient and reconstruction error demonstrate inverse trends across increasing component granularities. Stability declines, with certain exceptions, suggesting that the most common sources of covariance among subjects are captured with fewer components, while smaller and more localized parcels yield less generalizable decompositions. On the other hand, accuracy of the reconstructions improves at higher resolutions, suggesting that a larger number of components yield more localized and accurate approximations of the covariance patterns.

Specifically, stability peaks at k=6 and k=16 solutions, with smaller error bars in the latter granularity. The greatest improvement in accuracy occurs when moving from k=4 to k=6 components, and a more stable plateau is reached after k=10 components. Ultimately, k=6 and k=16 were chosen as the optimal resolutions. The k=6 solution demonstrates strong stability and follows the largest gain in accuracy, while k=16 has marginally higher stability and comparatively lower change in reconstruction error. Therefore, the k=16 decomposition is shown as the main result and the k=6 decomposition can be found in Supplementary Figure 6.

4.3.2 16 Component NMF Decomposition

A.



B.

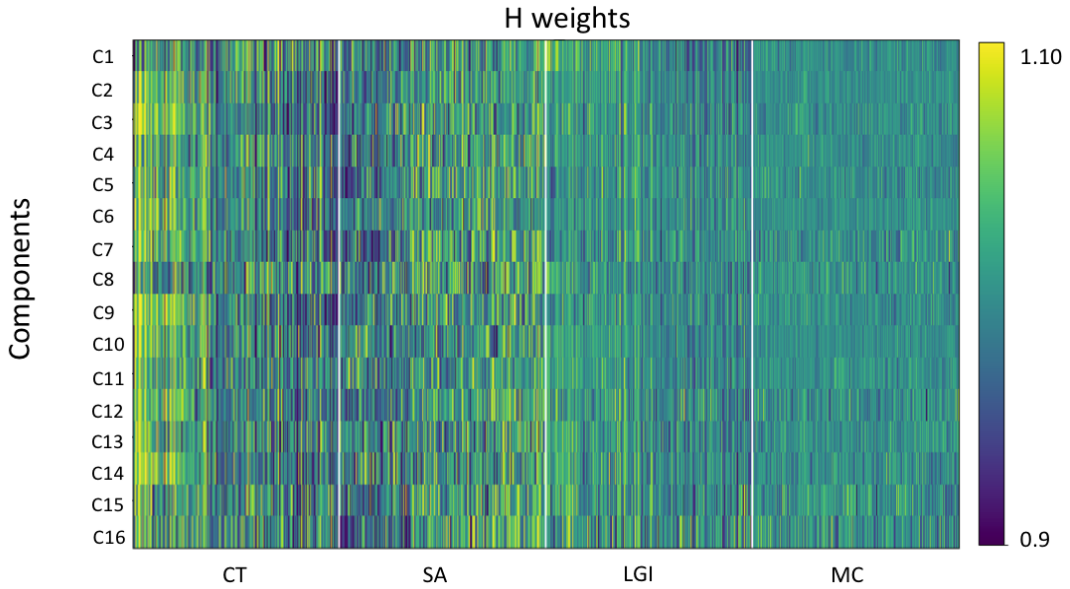


Figure 4.6. NMF Decomposition **A.** Spatial maps for a 16 component decomposition showing covariance patterns of the four cortical measures. Darker orange hues represent vertices with high W scores and therefore strong membership in a given pattern. Anatomical descriptors corresponding to each component are displayed above the brain maps. **B.** The H matrix shows participants arranged along the x-axis and components along the y-axis. Brighter colours indicate a higher participant-by-metric loading onto a given component, suggesting greater contribution of that metric to the component's covariance pattern. C = Component. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature.

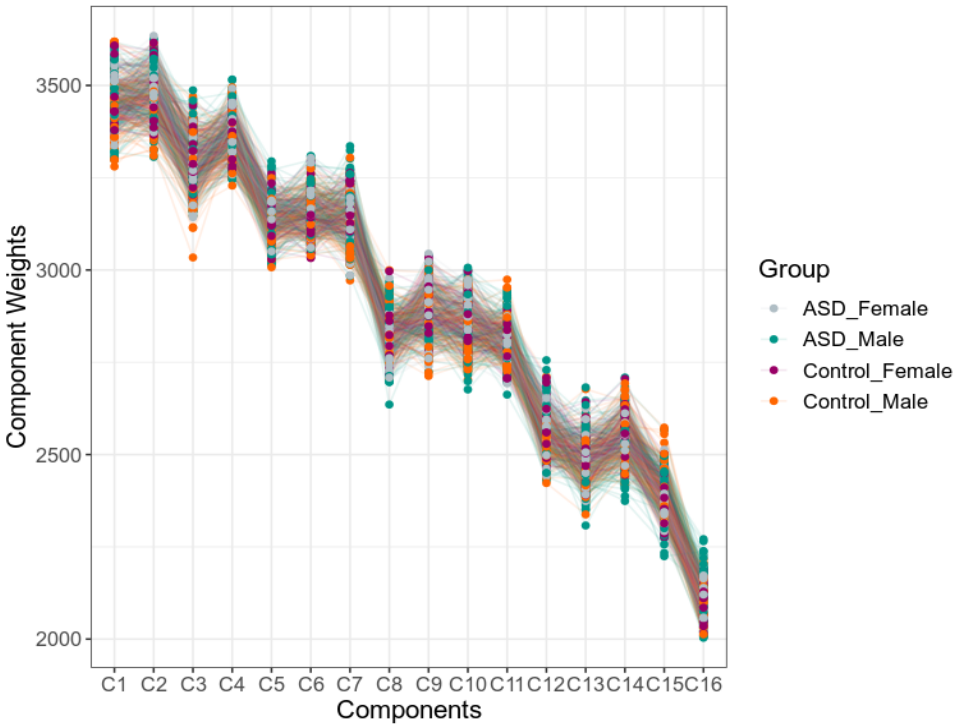


Figure 4.7. Total component weights are shown for each subject and component. Subjects are represented as dots connected by lines across components, coloured by diagnostic and sex group. C = Component. ASD = Autism Spectrum Disorder.

Results of the 16 component NMF decomposition are shown in Figure 4.6. Spatial maps depict vertex-wise W scores for each component (Figure 4.6A), indicating patterns of multi-metric covariance across individuals. Covarying vertices are found in anatomically proximal and bilaterally symmetric clusters within known cortical partitions. The H matrix depicts subject-wise weightings onto each component-by-metric combination (Figure 4.6B), indicating cortical covariance profiles. Cortical measures display variation in the magnitude and spatial topography of subject weights. Together, the W scores and H weights describe individual variability in multi-metric covariance at a vertex-wise scale.

Total component weights summed across metrics are shown in Figure 4.7. Irrespective of group, subjects load most strongly onto covariance patterns in C2 (posterior frontal cortex). Component weights are not clearly separated by diagnostic or sex characteristics.

4.3.3. Univariate Case-Control Differences in NMF Weights

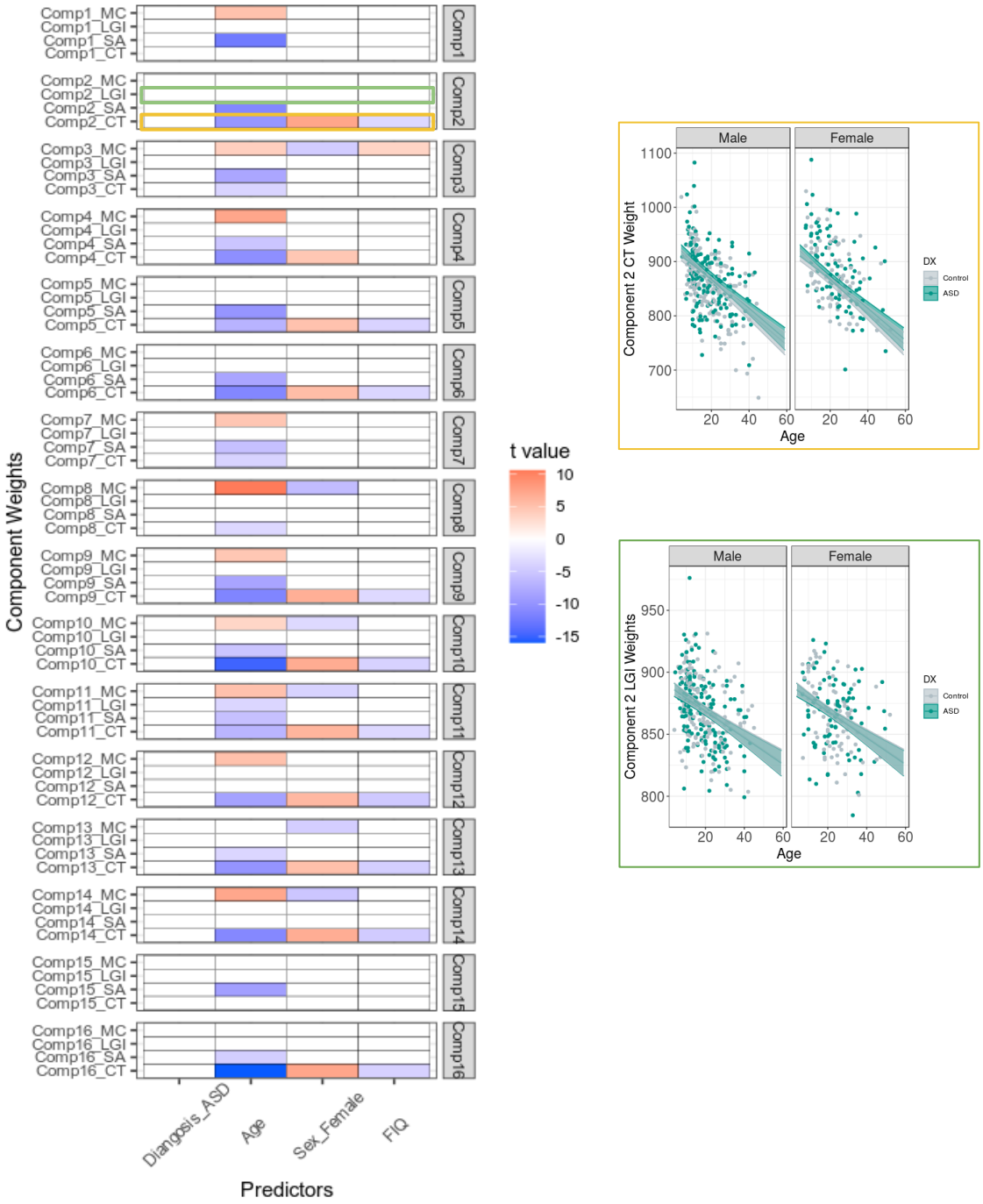


Figure 4.8. Heatmap on the left shows results of linear models investigating the main effect of diagnosis on subjects' component-by-metric weights, covarying for age, sex, and fullscale IQ. Each row represents a statistical model with model terms along x-axis and corresponding response variables along y-axis. T-values are shown for predictors that remained significant after multiple comparison correction (α of 0.05/64 models = padj <0.0007). Selected plots of the highest weighted component, C2, are depicted on the right (CT=yellow box; LGI=green box). C= Component. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature. ASD = Autism Spectrum Disorder. FIQ = Fullscale IQ.

Univariate results reveal no significant effect of diagnosis on individual variability in covariance patterns for any of the component-by-metric weight combinations (Figure 4.8). Older subjects display a global decrease in CT (C2-14, C16) and SA (C1-7, C9-11, C13, C15-16), alongside widespread increases in MC (C1, C3-4, C7-12, C14). Females demonstrate a perfuse increase in CT (C2, C4, C5-6, C9-14, C16) and localized decreases in MC relative to males in occipital (C8, C11), superior parietal (C10), and orbitofrontal (C3) regions, as well as the cuneus (C8), precuneus (C14), fusiform gyrus (C11), and parieto-occipital fissure (C13). Individuals with higher FIQ show a widespread decrease in CT (C2, C5-6, C9-14, C16) and a localized increase in MC in the orbitofrontal cortex (C3).

4.4 PLSC Results

4.4.1 Permutation Testing Results

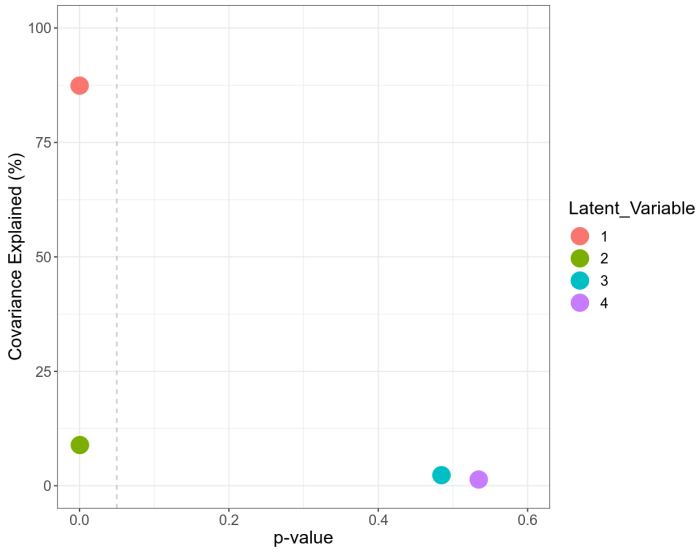
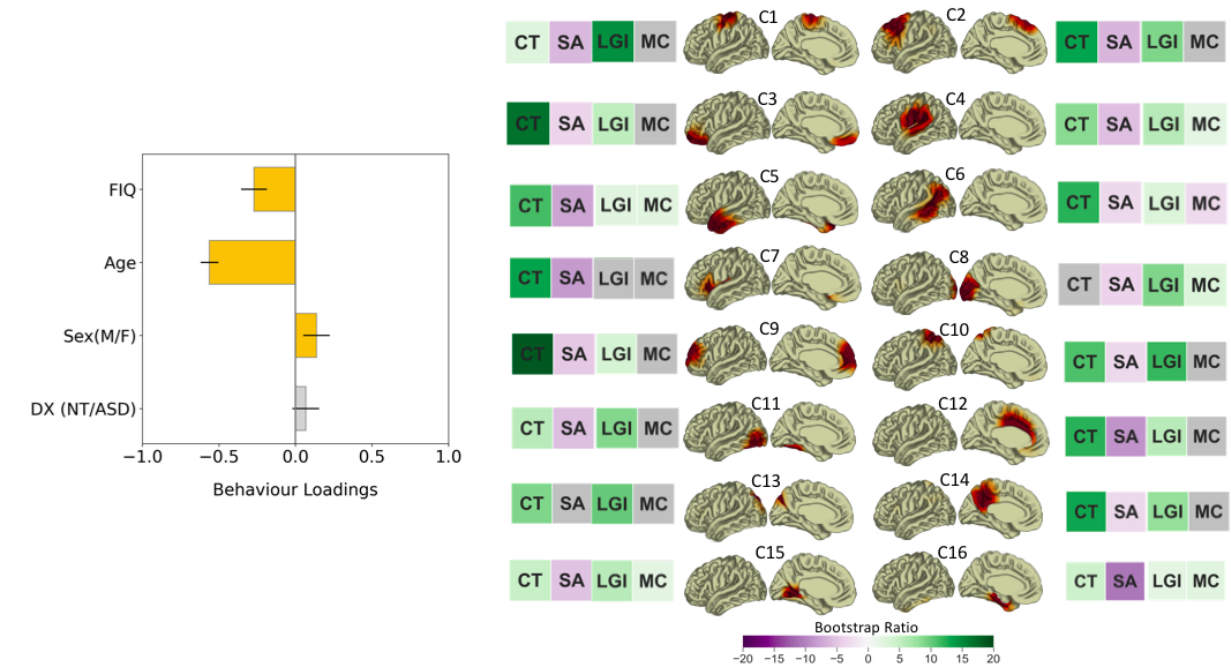


Figure 4.9. Scree plot showing significance and covariance explained of each latent variable (LV). LV1 (red) and LV2 (green) remain significant after permutation testing. Grey dotted line represents p=0.05.

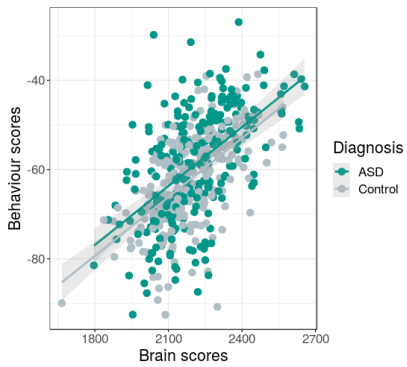
Results of PLSC permutation testing are displayed in Figure 4.9. Two LVs remain significant. LV1 ($p<0.0001$) explains 87.4% covariance and LV2 ($p<0.001$) explains 8.9% covariance.

4.4.2 Significant LVs

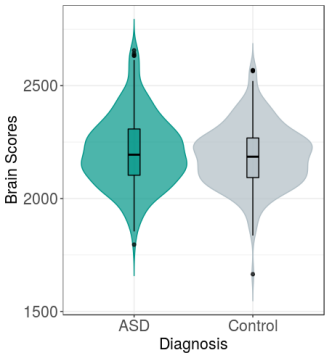
A. Latent Variable 1



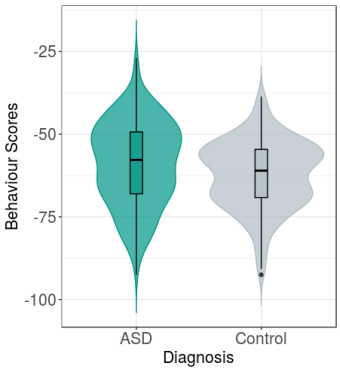
B.



C.



D.



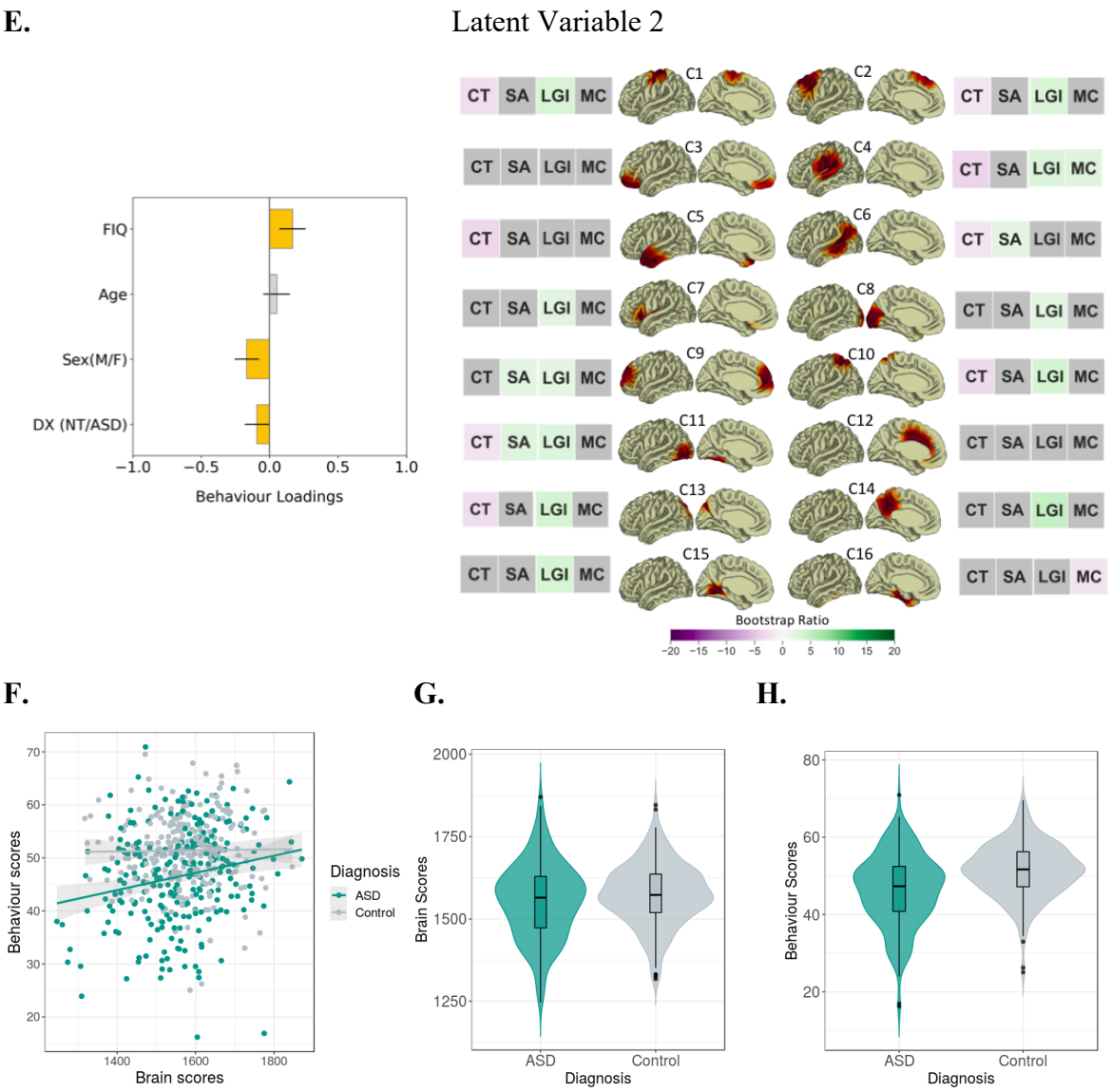


Figure 4.10. Partial least squares correlation (PLSC) results showing significant latent variables (LVs) that capture the relationship between subjects' NMF covariance weights and behavioural characteristics. **A**. Results of LV1. Barplots show each behavioural measure's loading onto the LV, expressed as a correlation coefficient. Orange bars are indicative of measures that contribute significantly to the LV. Brain maps on the right depict the 16 component NMF solution, with adjacent heat maps showing bootstrap ratios (BSR) of the NMF component-by-metric loadings. A BSR of 1.96 (95% CI) was used to threshold significance. Coloured squares (non-grey) indicate cortical measures within each covariance pattern that contribute significantly to the LV. Positive BSRs are displayed in green (denoting an increase of a given metric in the adjacent component) and negative in purple (denoting a decrease). **B.** Correlation of brain scores (brain patterns of LV projected onto individual subjects) and behaviour scores (behaviour patterns of LV projected onto individual subjects) coloured by diagnostic group ($r=0.56$; $p<0.0001$). **C.** Brain scores grouped by diagnosis. **D.** Behaviour scores grouped by diagnosis. **E.** Results of LV2. **F**. Correlation of brain and behaviour scores for LV2 ($r=0.15$; $p<0.001$). **G.** Brain scores grouped by diagnosis for LV2. **H.** Behaviour scores grouped by diagnosis for LV2. LV = Latent Variable. FIQ = Fullscale IQ. M = Male. F = Female. DX = Diagnosis. NT = Neurotypical. ASD = Autism Spectrum Disorder. C = Component. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature.

4.4.2.1 LV Brain-Behaviour Patterns

Results of the PLSC analysis show that behavioural variables that contribute significantly to LV1 include younger age (r= -0.56), lower FIQ (r= -0.27), and female sex (r= 0.13), which are significantly correlated with increased whole-brain contribution of CT (except C8; posterior occipital cortex & cuneus) and LGI (except C7; pars opercularis /triangularis & lateral sulcus) as well as decreased whole-brain contribution of SA to covariance patterns (except C13; parieto-occipital fissure). MC displays an increase in C4 (ventral sensorimotor & superior temporal gyrus), C5 (anterior temporal), C8 (posterior occipital cortex & cuneus), C15 (lingual gyrus), and C16 (parahippocampal gyrus), as well as a decrease in C6 (inferior parietal & posterior temporal pole; Figure 4.10 A-D).

Behavioural variables that contribute significantly to LV2 include higher FIQ (r = 0.16), male sex (r = -0.16), and NT control status (r = -0.09), which is significantly correlated with lower CT in C1 (dorsal sensorimotor), C2 (posterior frontal), C4 (ventral sensorimotor & superior temporal gyrus), C5 (anterior temporal), C6 (inferior parietal & posterior temporal pole), C10 (superior parietal cortex), C11 (lateral occipital cortex & fusiform gyrus), C13 (parieto-occipital fissure); higher SA in C6 (inferior parietal & posterior temporal pole), C9 (dorsolateral prefrontal cortex & medial frontal gyrus) and C11 (lateral occipital cortex & fusiform gyrus); higher LGI in C1 (dorsal sensorimotor), C2 (posterior frontal), C4 (ventral sensorimotor & superior temporal gyrus), C7 (pars opercularis/triangularis & lateral sulcus), C8 (posterior occipital cortex & cuneus), C9 (dorsolateral prefrontal cortex & medial frontal gyrus), C10 (superior parietal cortex), C11 (lateral occipital cortex & fusiform gyrus), C13 (parieto-occipital fissure), C14 (precuneus), C15 (lingual gyrus); and lower MC in C16 (parahippocampal gyrus; Figure 4.10 E-H).

4.4.2.2 LV Split-half Stability Results

Stability results of LV1 reveal that left and right singular vector correlations are significant across split halves (rUcorr = 0.95, pUcorr <0.001; rVcorr = 0.98, pVcorr <0.001), indicating stability of brain-behaviour associations captured by LV1 across subgroups of subjects. However, stability results of LV2 show that left and right singular vector correlations are not significant across split halves (rUcorr = 0.62, pUcorr <0.001; rVcorr = 0.84, pVcorr = 0.16), indicating that the brain-behaviour patterns captured by LV2 are not generalizable across subgroups of subjects.

4.4.2.3 Metric-Specific Regional Variation

Across metrics, CT is the cortical measure with the greatest overall contribution to covariance patterns in LV1 (Supplementary Table 5). The largest absolute CT BSRs are observed in frontal regions, specifically C9 (dorsolateral prefrontal cortex & medial frontal gyrus), C3 (orbitofrontal cortex), and C7 (pars opercularis/triangularis & lateral sulcus). Across components, CT is the largest contributor to covariance patterns in frontal (C2, C3, C7, C9) and temporal (C4, C5, C6) parcels, as well as inferior parietal lobule (C6), anterior cingulate (C12), and precuneus (C14). LGI is the dominant contributor to covariance patterns in the dorsal sensorimotor cortex (C1), cuneus (C8), superior parietal cortex (C10), ventral occipital cortex (C11), parieto-occipital fissure (C13), and lingual gyrus (C15). SA is the dominant contributor to the parahippocampal gyrus (C16), showing an inverse relationship with the other measures.

4.4.2.4 Effects of Methodological Variability

Methodological variation in the statistical workflow was found to affect both NMF and PLSC results. Specifically, performing NMF on unmatched subjects (n=668; Supplementary Figure 4), and non-GMV-residualized input measures (n=668; Supplementary Figure 5) affected: 1) optimal component granularity of NMF stability analyses; 2) the anatomical location of spatial components; and 3) direction of brain-behaviour associations captured by PLSC LVs, relative to the results shown in Figure 4.10. However, when the analyses were performed on two different NMF decomposition granularities (i.e. six and 16 components), while keeping constant subject matching and GMV-residualization, PLSC results were remarkably similar (n=486, Supplementary Figure 6).

4.4.3 Relating LV1 to Other Clinical and Cognitive Variables

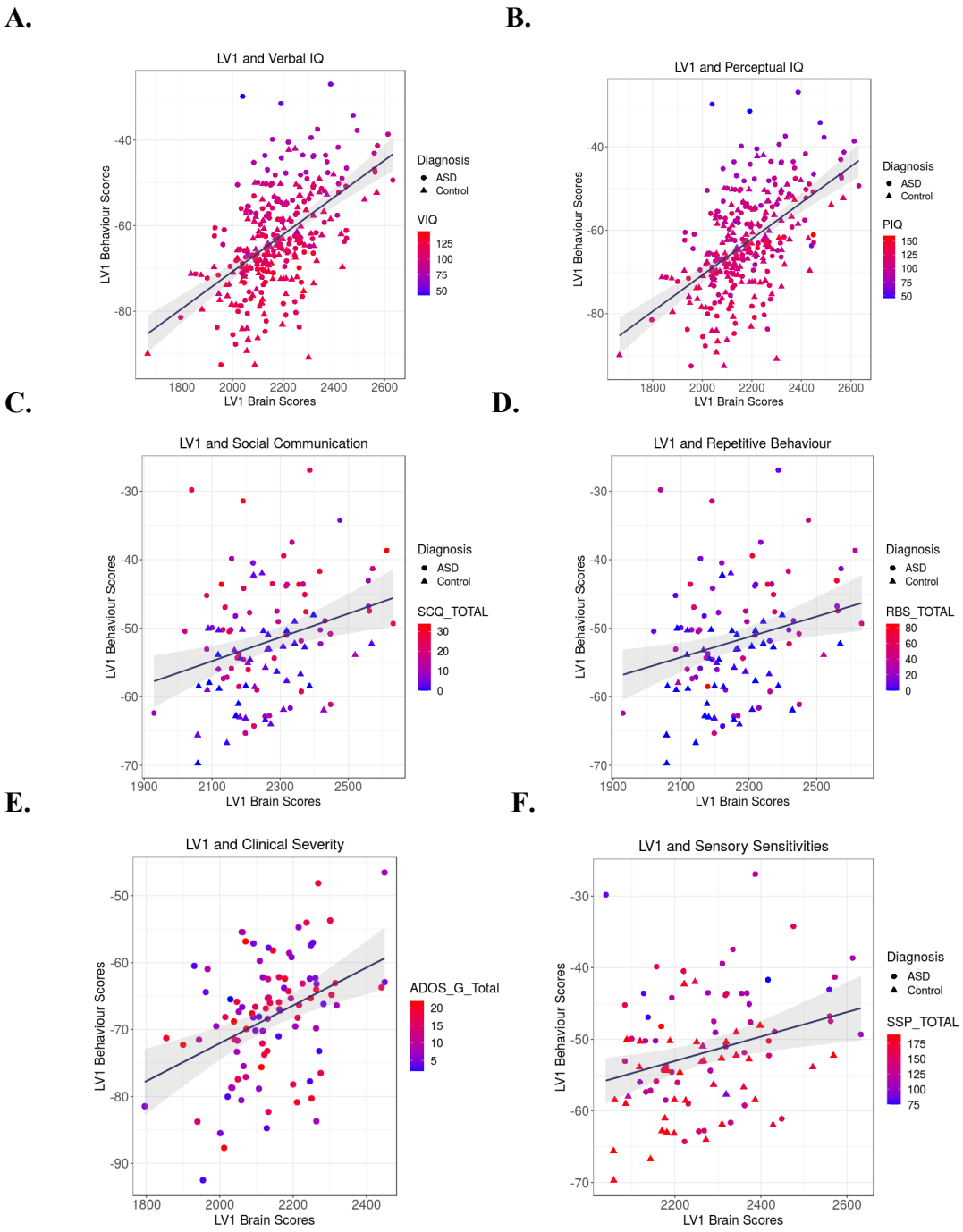


Figure 4.11. Correlations of LV1 brain and behaviour scores are shown for ASD (circles) and NT (triangle) subjects. Individuals are coloured by scores on cognitive and clinical assessments, shown separately in each plot. Hotter colours represent higher IQ for plots A-B, more items endorsed for plots C-E (signifying more symptoms or clinical severity), and fewer items endorsed for plot F (signifying less symptoms; reverse grading). **A.** Subjects are coloured by Verbal IQ scores (r=0.53). **B.** Perceptual IQ scores (r=0.53). **C.** Social Communication Questionnaire Total scores (r=0.29). **D.** Repetitive Behaviour Scale Total scores (r=0.25). **E.** Autism Diagnostic Observation Schedule General Total scores (r=0.38). **F.** Short Sensory Profile Total Scores (r= 0.27). ASD = Autism Spectrum Disorder. NT = Neurotypical.

LV1 brain and behaviour scores were correlated (Figure 4.11) for a subset of individuals included in the multivariate analyses that had additional cognitive and clinical data available (Supplementary Table 4). Each individual was coloured by their respective assessment values to produce a qualitative characterization of subjects who express the brain-behaviour profile captured by LV1.

The most striking relationship was seen for cognitive measures. Specifically, subjects in both diagnostic groups with lower Verbal IQ (VIQ) and Perceptual IQ (PIQ) express the LV1 phenotype more strongly relative to those with higher IQ values. To a lesser extent, ASD subjects with higher Social Communication Questionnaire scores (i.e. greater social impairment; Figure 4.11C) and lower Short Sensory Profile total scores (i.e. greater sensory sensitivities; Figure 4.11F) map onto LV1 to a greater degree, relative to subjects with lower social impairments and sensory sensitivities. Weaker differentiation is evident for Repetitive Behaviour Scale Total scores (i.e. stereotyped behaviours; Figure 4.11D) and ADOS-G Total scores (i.e. clinical severity; Figure 4.11E). These results reaffirm the main PLSC findings that cognitive ability has stronger relevance to the LV1 brain-behaviour phenotype than clinical variables.

5. Discussion

The lack of reproducible ASD neurophenotypes may stem from two fundamental limitations of neuroimaging studies. The first is the predominance of case-control analytical approaches, which seek to find average group differences in neuroanatomical measures based on *a priori* diagnostic labels. This conceptual scheme is unrepresentative of biological and clinical heterogeneity captured by the broad ASD umbrella, nor does it account for the possibility that neuroanatomical variability may exist along a continuum of variation with the normative population. The second is the paucity of statistical approaches that robustly model shared variance among age, sex, and IQ, as they pertain to variability in neuroanatomical properties and phenotypic presentation of ASD.

To this end, we used NMF to decompose individual variation in cross-metric covariance patterns in a data-driven dimensional framework. Specifically, covariance was modelled in a pooled group of ASD and NT subjects to yield transdiagnostic spatial patterns, in addition to subject-specific loadings that describe each individual's unique macrostructural profile. We

then used PLSC to relate these covariance patterns to diagnosis, age, sex, and FIQ. We find that structural networks of covariation are predominantly associated with age and general cognitive ability, with no significant diagnostic differences. These findings suggest a limited neuroanatomical signature corresponding to the clinical ASD construct.

5.1. Univariate Results

Conventional univariate analyses were first employed to assess diagnostic differences in global and vertex-wise cortical measures in order to characterize the data. ASD subjects had greater mean CT relative to NT controls, but this relationship did not survive multiple comparison correction (Figure 4.3). Similarly, at the vertex-level, no group differences were found in any of the cortical indices at 5% FDR; although at a more lenient significance threshold, decreased CT was observed in the right precentral gyrus of ASD subjects (Supplementary Figure 1). Interestingly, metric-specific regional variation was evident in vertex-wise standard deviation of the four measures (Supplementary Figure 2). This may suggest that although there were no statistically significant group differences in mean vertex-wise measures, more subtle and localized regional variation exists in morphometric organization.

Previous studies in the literature have also reported null findings for case-control differences in CT and SA (Mak-Fan et al., 2011; Mensen et al., 2017; Wallace et al., 2013, 2015; Yang et al., 2016), as well as GI/LGI (Casanova et al., 2009; Gharehgazlou et al., 2020, 2021; Hirjak et al., 2016; Koolschijn & Geurts, 2016; Schaer et al., 2015). Likewise, decreased CT has been reported in ASD relative to NT participants (Ecker et al., 2013, 2014; Hadjikhani et al., 2006; Hyde et al., 2010; Jiao et al., 2010; Scheel et al., 2011; van Rooij et al., 2018). Notably, Hammill et al (2021) observed lower CT in the right precentral gyrus of ASD subjects at the same FDR threshold (see Supplementary Figure 9, Hammill et al., 2021), alongside regional CT increases. Spatial patterns of case-control CT differences reported here do not replicate those found in previous work by our group (see Bedford et al., 2020). These discrepancies may stem from several factors. First, the pool of participants and their demographic characteristics differ between studies. Preprocessing (i.e. bpipe versus iN4 Multispectral) and processing (CIVET 1.1 versus CIVET 2.1.1) pipelines are also sources of disparity. Quality control was conducted by different raters, although effort was made to standardize approaches by using guidelines developed in the lab (see section 3.3). Importantly,

Bedford et al (2020) used a random-effects meta-analytic approach, which pooled effect sizes across sites, while in the current work, CovBat was used to correct batch effects across multiple sites and measures, followed by mass-univariate general linear models. The latter approach modelled age as a linear term and covaried for the global brain measure corresponding to the vertex-wise outcome metric. It is possible that using Akaike Information Criterion to select the best-fitting model terms and removing the global covariate may yield results that align more closely with previous work. In a similar vein, applying CovBat (or ComBat) to each metric independently may have been a more appropriate strategy for univariate analyses, rather than correcting for batch effects across all four measures concurrently. Indeed, the present use of CovBat may have altered the mean and variance of individual metrics, relative to what would be expected if batch correction was performed for each index separately. Empirically testing how different CovBat applications affect downstream statistics could be explored in future studies.

5.2 CovBat Validation Results

A series of validation analyses were performed to confirm the success of a novel multi-metric implementation of the CovBat batch correction algorithm. This technique removes batch effects from mean, variance, and covariance patterns of brain data, while retaining biological relationships in specified behavioural variables. Instead of applying CovBat to a single cortical measure as was done in the proof-of-concept paper (Chen et al., 2022), we applied the algorithm to four cortical measures simultaneously to align with our overarching analytical aim. To this end, CovBat-corrected vertex-wise measures were analyzed to confirm the absence of batch effects in the mean, variance, and covariance of brain metrics (section 4.2). We also demonstrated that relationships between demographic and cortical variables were retained after CovBat correction (Tables 4.2-4.5). Together, these results validate the effectiveness of this application.

Certain CovBat-related considerations are worth acknowledging. CovBat requires users to specify the behavioural variables whose associations with brain features are to be preserved (i.e., diagnosis, age, sex, and FIQ in the current work). These relationships are removed with linear approaches, which may not be suitable for modelling certain brain-behaviour relationships such as age (Chen et al., 2022; Pomponio et al., 2020). Furthermore, CovBat assumes that the optimal covariance structure is found in the covariance matrix of the combined

dataset, so covariance patterns of individual batches are shifted to that of the pooled data. However, this approach may be insufficient to remove batch effects in more complex cases when the covariance of individual batches diverge from that of the pooled data (e.g., do not lie along PC directions). Alternatively, overcorrection of PCs may occur in cases when individual batches have marginal or nonexistent batch effects in covariance patterns (Chen et al., 2022). Other applications of CovBat should be attempted to systematically assess the relative benefits and drawbacks of different strategies for multi-metric or multimodal data. For instance, we considered batch correcting NMF component weights instead of raw measures but were concerned that NMF would capture noise attributable to batch effects. For validation purposes, NMF components from one batch could be projected onto another batch to confirm reliability of spatial patterns in the pooled analyses. This could be accompanied by the use of more advanced classification techniques (e.g., support vector machines) for detection of batch and demographic variables in CovBat-adjusted residuals. Lastly, it may be worthwhile to apply CovBat to individual cortical measures rather than a multi-metric input to confirm that the latter approach is a more effective use case for multi-metric multivariate analyses.

Increasingly, novel approaches are emerging for retrospectively harmonizing structural brain data across batches. Some have extended existing ComBat capabilities by controlling for non-linear age effects with generalized additive models (Pomponio et al., 2020) or enabling meta-analytic implementations (Radua et al., 2020). Other approaches eliminate technical variability during data processing rather than from cortical derivatives (e.g., the Removal of Artificial Voxel Effect by Linear regression method; Fortin et al., 2016). Alternative techniques have made use of source-based morphometry for voxel-wise data (Chen et al., 2014), hierarchical Bayes for normative modelling (Bayer et al., 2021), and generative adversarial models to eliminate *a priori* scan or site labels (Liu et al., 2021). More rigorous comparisons of these strategies are needed to evaluate their efficacy for a variety of participant characteristics and downstream statistical analyses.

5.3 NMF Results

NMF was used to investigate the shared covariance structure between multiple macrostructural measures, each offering distinct yet complementary biological information.

5.3.1 Selecting Component Granularity

Using orthogonal projective NMF (Lee & Seung, 1999, 2000; Yang & Oja, 2010) we found both six and 16 component solutions optimally balanced generalizability across subgroups and minimal residual error in the reconstructed low-rank matrices (Figure 4.5). Spatially, the six component decomposition (Supplementary Figure 6) yielded broader and more non-specific patterns of variation that were bilaterally symmetric and adhered to lobular anatomy. In comparison, the finer grained 16 component decomposition (Figure 4.6 A) produced more focal and regionally differentiated patterns.

Selecting the optimal decomposition granularity impacts the spatial variation captured by components as well as downstream statistics performed on component weights. Previous studies investigating cortical covariance patterns with NMF have likewise evaluated stability and accuracy of decompositions to arrive at a granularity decision. For instance, Sotiras et al. (2015) evaluated GM covariance in a human aging dataset with NMF, PCA, and Independent Component Analysis (ICA) across 2-100 component resolutions. The authors found that as the number of NMF components increased, so did sparsity and coherence, while reconstruction error declined. Split-half generalizability measured with the inner product varied across resolutions, with a decline observed at ~2-10 components, followed by a slight increase at ~25 components, and a subsequent plateau. Sotiras et al. (2017) used NMF to map CT covariance patterns in adolescent participants. The authors found peaks in split-half stability at two, seven, and 18 component granularities, which closely approximate the stability peaks seen in our work at two, six, and 16 components (Figure 4.5). NMF has also been used to assess covariance patterns in a clinical framework. Sanfelici et al. (2021) evaluated LGI covariance patterns in a group of patients with recent onset depression, recent onset psychosis, and clinical high risk, relative to healthy controls. An 18 component solution was selected after assessing change in reconstruction error and split-half stability in the group of control participants. The authors validated this granularity in two held-out sites, before applying it to the group of clinical subjects. Similarly, Shan et al. (2022) used NMF to identify GMV covariance patterns in ASD and NT participants. Changes in reconstruction error and sparsity were assessed from 2-20 components in two samples of NT subjects, then the degree of spatial overlap was computed for a smaller range of components between the NT groups to select the granularity with the highest reproducibility. A six component solution was chosen and used for separate NMF decompositions in ASD and NT groups. Taken together, although these model selection

approaches produce quantitative values, it is important to bear in mind that the decomposition granularity is ultimately chosen by the user, even in cases where the optimal combination of favorable parameters may be ambiguous.

Lastly, previous work has confirmed that incorporating multiple modalities or metrics yields more robust decompositions relative to unimodal analyses performed independently on each modality, likely owing to the fact that more information is available at each vertex to guide consistent and accurate decompositions (Patel et al., 2020; Robert et al., 2022). Indeed, our multi-metric solutions achieved high stability across subjects (r= ~0.7) and a plateau in reconstruction error slope (Figure 4.5).

5.3.2 Other Methodological Considerations

Outside of granularity selection, methodological decisions at other stages of the workflow affect spatial features of components. Batch correction approaches vary between studies, with some researchers opting to residualize site or scanner variables from raw data (e.g., Patel et al., 2022; Shan et al., 2022), while others apply algorithms such as ComBat (e.g., Sanfelici et al., 2021) prior to NMF implementation. These techniques differ with respect to the amount of batch-specific noise remaining in cortical measures, as well as the manner in which associations between brain and behavioural measures are affected if they are confounded with site (Fortin et al., 2018). Importantly, neither approach removes batch effects from covariance patterns of brain measures, which is of particular concern for multivariate analyses (Chen et al., 2022).

Secondly, if substantial age differences exist in a pool of participants, it is possible that covariance patterns capture age-related changes in global cortical measures (e.g., mean CT, total SA, GMV, TBV) that are not explicitly modeled in the NMF. To this end, we tested whether global brain measures may be conflated with localized metric-specific spatial patterns. Indeed, we found that GMV exhibited the strongest correlation with NMF H weights when raw/non-residualized indices were used as NMF input (Supplementary Figure 3). GMV was therefore residualized on a vertex-wise basis from every cortical measure prior to NMF implementation to ensure that vertex-wise covariance patterns were not driven by individual variation in GMV. This step ultimately affected the split-half stability results and spatial patterns of components, relative to unresidualized analyses (Supplementary Figure 5).

Furthermore, when constructing the multi-metric NMF input, it is necessary to

standardize each metric block to account for varying magnitudes of cortical measures. Standardization approaches should be considered in relation to the variance that a user wishes to prioritize, given that the magnitude of a vertex-wise measure differs according to the matrix axis. For instance, in a vertex-by-subject matrix of CT values, greater variation exists in the magnitude of CT measures within a given subject (i.e. between rows/vertices) than between two subjects' CT measures for a given vertex (i.e. between columns). Since the aim of our analysis was to capture individual variability in morphological covariance patterns, we standardized the measures across columns to ensure that the more subtle variation in CT magnitude between individuals would not be overshadowed. For an alternative standardization approach that yields parcellation-type decompositions, see Patel et al. (2020), Robert et al. (2022), and Kalantar-Hormozi et al. (2022).

Lastly, inter-study variation in NMF parameters contribute to variability in results. This includes differences in initialization strategies (e.g., randomized, NNDSVD), numerical solvers (e.g., multiplicative update, coordinate descent), constraints (e.g., orthogonality, sparsity), and cost functions.

5.4 PLSC Results

PLSC analysis was used to associate individual variation in morphometric covariance patterns with demographic and behavioural characteristics that affect phenotypic heterogeneity in ASD.

5.4.1. Age and Cognitive Correlates of Structural Covariance

Across both decomposition granularities, younger age, female sex, and lower FIQ contributed significantly to LV1, with age having the strongest correlation (r= -0.56). These variables were significantly associated with increased whole-brain contribution of CT, decreased whole-brain contribution of SA, widespread increase in LGI, and more subtle and regionally nuanced MC patterns. The relationships were stable across subsets of subjects, as seen in the LV1 split-half stability analysis (section 4.4.2.2), suggesting that the brain-behaviour patterns represent primarily age-driven, whole-brain changes that are robust across subjects and NMF resolutions.

Across measures, the largest absolute BSRs were observed for CT in frontal lobe components, specifically dorsolateral prefrontal cortex, medial frontal gyrus, orbitofrontal

cortex, pars opercularis/triangularis and the lateral sulcus. Across components, CT was the largest contributor to covariance patterns in association areas primarily involved in higher order cognition, including frontal and temporal parcels, as well as the inferior parietal lobule, anterior cingulate, and precuneus. LGI was the dominant contributor in regions involved in sensory processing and integration, including dorsal sensorimotor cortex, cuneus, superior parietal cortex, ventral occipital cortex, parieto-occipital fissure, and lingual gyrus. This pattern of brain-behaviour results suggests metric-specific functional delineations, with CT exhibiting the most robust relationship with age and cognitive ability.

Previous work assessing the spatial patterning of age-related maturation have found that cortical development follows regionally heterochronous trajectories whereby lower-order primary sensory and motor areas mature before higher-order association areas involved in language and cognition (Amso and Casey, 2006). This maturational profile has been recapitulated at both local (Gogtay et al., 2004) and network (Khundrakpam et al., 2013; Raznahan et al., 2011; Zielinski et al., 2010) scales, and aligns with the developmental staging of functional processes (Alexander-Bloch et al., 2013). Inter-regional communication transitions from local to distributed topologies throughout development (Alexander-Bloch et al., 2013; Zielinski et al., 2010), with network-specific growth trajectories (Liu et al., 2017), and evidence of sexual differentiation in temporal (Zielinski et al., 2010) and spatial (Raznahan et al., 2011) patterning.

This hierarchical organization is likewise seen in CT, with developmental changes occurring spatially from sensorimotor to association axes (Brown et al., 2012; Gilmore et al., 2021; Sowell et al., 2004; Syndor et al., 2021). Importantly, relative to other cortical measures, CT is more amenable to experience-induced plasticity (Gilmore et al., 2020), and is thought to facilitate cortical refinement through adaptations to functional demands in the environment (Lerch et al., 2011).

Studies investigating the structural substrates of general intelligence have also identified relationships between CT and FIQ, predominantly in higher-order association cortices. These regions reside in the temporal lobe (Choi et al., 2008; Karama et al., 2009; Narr et al., 2007), cingulate (Karama et al., 2009; Menay et al., 2013), as well as prefrontal (Menary et al., 2013; Narr et al., 2007), extra-striate (Narr et al., 2007), and parietal (Karama et al., 2009; Menary et

al., 2013) cortices. Sex differences in spatial patterns have been observed, with female-specific CT-FIQ relationships reported in prefrontal and temporal association cortices, and male-specific patterns in temporal-occipital association cortices (Narr et al., 2006). Additionally, sex differences have been reported in the valence of brain-cognition links in prefrontal compartments (Goh et al., 2011). Another line of evidence suggests an age-related inversion of the relationship between CT and FIQ, with negative associations in early childhood (< eight years old), and positive associations in frontal and temporal regions in later childhood (> eight years old; Shaw et al. 2006). When examining this relationship in a subscale-specific manner, Menary et al. (2013) found more widespread, positive coupling between CT and FIQ in a sample of 9-16 year old subjects, whereas in the older sample of 16-24 year-olds, this was found only for non-verbal IQ scores.

CT structural covariance patterns have also been investigated in relation to intelligence. In one study, Khundrakpam et al. (2017) related interregional correlations in CT with IQ subscales, finding divergent regions of cognitive associations in higher versus lower VIQ and PIQ. Graph-theoretic analyses revealed that higher VIQ subjects demonstrated more optimal topological organization relative to those with lower VIQ. Specifically, these subjects displayed higher global efficiency and modularity, alongside lower local efficiency. Differences in regional efficiency in higher VIQ versus lower VIQ individuals were also found. In another study, Seidlitz et al. (2018) developed Morphometric Similarity Networks to evaluate pairwise interregional correlation of multiple macro- and micro-structural indices. PLSR was used to assess the effect of inter-individual variability in node degree for each subject's MSN on VIQ and PIQ scores. The dominant LV described a significant association between high node degree of language-related cortical regions in left frontal and temporal cortices, with higher VIQ and PIQ. Together, these findings suggest greater integration of cortical regions involved in higher order intellectual function, with some spatial specificity pertaining to cognitive domains.

Our work adds to this literature by demonstrating that in younger subjects (median age=18.42) with lower cognitive abilities (median FIQ=111.5), CT has the greatest contribution to cortical covariance patterns, with the highest values observed in association cortices. These results underscore the biological importance of CT in supporting developmental and cognitive processes.

5.4.2 Clinical Correlates of Structural Covariance

We did not find a significant contribution of diagnostic status to LV1 in the PLSC analysis. However, since ASD status approached significance (Figure 4.10), we sought to further characterize the subjects that expressed the LV1 phenotype in relation to additional cognitive and clinical variables. In line with our PLSC results, we observed that LV1 was more robustly related to cognitive rather than clinical measures (Figure 4.11). Indeed, both ASD and NT subjects with lower VIQ and PIQ exhibited the LV1 brain-behaviour profile more strongly, relative to those with higher VIQ and PIQ. Additionally, ASD subjects with greater social impairment and sensory sensitivities mapped more strongly onto the LV1 patterns relative to those with lower social impairment and sensory sensitivities. In comparison, repetitive behaviour symptomatology and ASD clinical severity showed less discernible relationships with LV1. These findings suggest that individual differences in neuroanatomical organization are more representative of developmental, cognitive, and behavioral variation than a diagnostic label.

It is possible that part of the variance attributable to diagnosis was subsumed by FIQ due to significant case-control differences in FIQ (see Section 4.1.1.1). Indeed, during the matching procedure, ASD and NT subjects were paired for age and sex, but not FIQ. This decision was undertaken for a number of reasons. FIQ represents a general measure of cognition that is closely related to ASD diagnosis, level of functioning, and phenotypic presentation (Rao et al., 2015), thus it is difficult to dissociate ASD as a diagnostic label from IQ as a cognitive construct. In a similar vein, some have argued that valid assessment of IQ in ASD is often confounded by symptoms of ASD and its common comorbidities (e.g., impairments in social communication, ritualized behaviors/speech, fixed interests, limited attention span) since prolonged engagement of verbal, comprehension, and social skills are required during the administration of standardized cognitive assessments (Rao et al., 2015; Wolff et al., 2022). Moreover, the use of FIQ as a composite score of cognition may be unrepresentative of the variable and nuanced IQ profiles observed in autistic individuals (Wolff et al., 2022). Thus, matching subjects on FIQ may ultimately further bias clinical samples (Jarrold & Brock, 2004).

Contrary to our results, previous studies investigating case-control differences in structural covariance patterns have found diagnostic differences related to symptom dimensions, developmental trajectories, and structural properties of covariance networks.

Sharda et al. (2016) found that ASD subjects with lower language functioning had lower structural covariance in CT but not SA in left fronto-temporal areas when using language-based seed regions. Similarly, Bernhardt et al. (2013) observed that subjects with ASD had lower covariance in CT structural networks occupying regions involved in Theory of Mind, including the dorsal medial prefrontal cortex and temporal parietal junction, which corresponded to decreased self-reported perspective-taking abilities. Conversely, increased GMV structural covariance patterns between subcortical and subcortical-cortical regions have been linked to elevated insistence of sameness scores on the Repetitive Behaviour Scale assessment (Eisenberg et al., 2015).

From a developmental perspective, Wang et al. (2022) found age-related alterations in structural covariance profiles of ASD infants, with significant case-control differences by 24 months of age in SA covariance networks, accompanied by significant differences in network segregation, integration, and small-worldness, particularly in fronto-parietal and occipital compartments. Leveraging a sliding window approach, Cai et al. (2021) assessed changes in CT covariance patterns from 7-45 years of age, finding an inverted U-shaped trajectory in both ASD subjects and NT controls, with the ASD group showing a sustained lag relative to the NT group.

Other organizational properties of structural covariance networks also show case-control distinctions. For instance, Zielinski et al. (2012) identified a more restricted GM intensity salience network across diffuse regions in the ASD group, alongside more nuanced region-specific diagnostic differences in the default mode network. Sha et al. (2022) reported predominantly right-lateralized asymmetry of CT structural covariance patterns of ASD subjects in the fusiform gyrus and subdivisions of the frontal cortex, with no significant effect of clinical severity, medication use, IQ, or gender on asymmetry measures. Bethlehem et al. (2017) found that hubs in CT and LGI covariance networks displayed lower node degree in ASD subjects relative to NT controls, suggesting weaker integration with other nodes in the network.

Using SVD-based decomposition techniques that are more comparable with our approach, other studies have identified clusters of cortical regions that covary together. For instance, Shan et al. (2022) used a sparse-NMF variant to decompose voxel-wise GMV in ASD and NT controls separately, then used subject loadings to evaluate individual deviations from normative trajectories, followed by a clustering to identify ASD subtypes. The authors found

group differences in subject weights in five of six components occupying most cortical and subcortical parcels. Covariance patterns in the medial frontal cortex and parietal lobe were significantly associated with ADOS communication scores, whereas those in the cerebellum, parietal lobe, frontal lobe, and limbic system were significantly associated with ADOS social scores. Lastly, Mei et al. (2020) used ICA to map GM covariance patterns. General linear models were first used to relate ICA-derived subject weights with diagnostic status, finding that ASD-specific decreases in volume of insula and aspects of the frontal lobe, in addition to an increase in the caudate nucleus. Significantly decreased volume of amygdala, hippocampus, and parahippocampal gyrus in the ASD group were also found. Since case-control differences were observed, ICA was then performed solely on the ASD subjects, and Canonical Correlation Analysis was used to relate ASD subjects' ICA component weights with cognitive and clinical assessments. A significant canonical variate related volume of thalamus, putamen, motor cortex, and cerebellum, in addition to distributed regions of frontal, parietal, and occipital divisions with greater repetitive/stereotyped behaviors and sensory sensitivities.

Differences between our results and those reviewed above may be due to a number of factors. While our PLSC analysis included age, sex, and IQ as main variables, other studies have either set these variables as covariates (Cai et al., 2021; Eisenberg et al., 2015; Sha et al., 2022; Sharda et al., 2016; Zielinski et al., 2012), residualized them from raw data (Shan et al., 2022; Wang et al., 2022), or matched for age, IQ (and/or gender) but did not explicitly model these variables statistically (Bernhardt et al., 2013; Bethlehem et al., 2017). Matching strategies also vary among studies, with some pairing subjects based on IQ (Bernhardt et al., 2013; Bethlehem et al., 2017; Zielinski et al., 2012), or additional cognitive variables (Bernhardt et al., 2013). We performed our analyses on both matched and unmatched subjects to assess putative differences in diagnostic effects that may be attributable to the disproportionate ratio of ASD subjects to NT controls. Since we found that the direction of diagnostic results differed in either case, we performed our main analyses in the matched subset of subjects to limit the possibility of group size discrepancy driving our results.

Accounting for inter-subject differences in global cortical measures was also performed inconsistently, with some studies covarying or residualizing TBV (Sharda et al., 2016; Zielinski et al., 2012), ICV (Eisenberg et al., 2015), mean GMV (Cai et al., 2021), or the global equivalent of the local measure being used (e.g. mean CT or total SA; Bernhardt et al. 2013;

Wang et al., 2022). In our work, we found that covariance patterns expressed the strongest correlation with GMV among the other measures we tested (TBV, mean CT, total SA, mean LGI, mean MC) and thus we decided to residualize GMV from our vertex-wise measures before modelling structural covariance. In light of these findings, it may be useful for researchers to quantitatively test a range of global measures on covariance networks to arrive at an empirically-based decision.

Lastly, across studies, covariance patterns are modelled using a variety of statistical techniques, including seed-based interregional associations (Bernhardt et al., 2013; Eisenberg et al., 2015; Sharda et al., 2016; Zielinski et al., 2012), graph theoretical network analyses (Bethlehem et al., 2017; Cai et al., 2021; Wang et al., 2022), and decomposition-based multivariate techniques (Shan et al., 2022; Mei et al., 2021). Each approach quantifies covariance differently and varies with respect to subject-wise specificity. Interestingly, the other multivariate techniques reviewed here modelled structural covariance separately in ASD and NT control subjects, presumably to relate subject weights from the clinical decomposition to additional ASD-specific assessments. However, this approach assumes that ASD and NT subjects have inherently divergent covariance profiles and creates challenges with selecting the optimal component granularity for both group that would allow for homologous comparisons. Instead of adopting a case-control paradigm, we opted to identify patterns of covariance in a pooled group of ASD and NT controls to model covariance patterns within a dimensional framework. Since significant diagnostic differences were not observed, we did not perform follow-up decompositions on cases and controls separately.

5.5 Interdependencies Between Cortical Measures

When considering the cortical BSRs in LV1, CT demonstrated an inverse relationship with SA, and positive relationships with LGI and MC. CT and SA represent orthogonal axes of cortical organization, and thus it is not surprising that they are independent in their spatial organization, considering their divergent genetic (Pazzinon et al., 2009) and neurodevelopmental (Raznahan et al., 2011) profiles.

Conversely, CT and MC describe cortical properties along the same spatial dimension, with CT representing the distance between WM and pial boundaries, and MC representing the convexity or concavity of the cortex along its radial aspect. Previous research has shown that sulcal depths are thinner than gyral crowns (Vandekar et al., 2015), with sulci and gyri also

showing different cellular and vascular properties (Hiletag & Barbas, 2005; Sun & Hevner, 2014; Wagstyl & Lerch, 2018) thereby confirming shared regional variation between these two measures.

Interestingly, LGI shows a negative association with SA. This inverse relationship is somewhat unexpected given that LGI is partly derived from SA, representing the ratio of SA to the convex hull within a region of interest. These findings could suggest limited anatomical specificity of LGI as a measure of folding, possibly stemming from methodological challenges related to quantifying cortical convolutions, and individual variability in folding patterns that may affect quality of image registration. Furthermore, the SA-LGI relationship reported here may be specific to the particular set of brain and behavioural variables included in the PLSC analysis. Indeed, LV1 associates globally reduced SA and elevated LGI with younger age, lower IQ, and female sex. An equally valid way to interpret LV1 results is that globally elevated SA and reduced LGI are significantly associated with older age, higher IQ, and male sex.

By incorporating multiple measures to characterize cortical covariance patterns, we are able to represent more information at each vertex, relative to modelling covariance of each measure separately. However, this approach comes with certain interpretational drawbacks. Primarily, the relative covariance of four cortical properties must be interpreted within the context of each measure's individual variance across the mantle, in terms of both spatial extent and absolute magnitude (Supplementary Figure 2).

Relatedly, the precision of spatial specificity also varies for each metric. For instance, while vertex-wise CT represents the distance between corresponding points on WM and pial boundaries along the vertical dimension, vertex-wise SA represents one third of the SA of triangular facets adjoining a vertex along the horizontal dimension, thereby comprising a larger spatial field. The large size of blurring kernels used in CIVET software (i.e. 30mm for CT and MC; 40mm for SA) may also play a role in reducing spatial specificity of each measure, particularly for measures of cortical complexity. Conceivably, the importance of reducing random noise by imposing a Gaussian distribution may be less relevant for multivariate than univariate analyses, and thus future studies may benefit from assessing the effect of smaller kernel sizes.

Further still, cortical measures vary in their measurement reliability. Indeed, one study reported lower reproducibility of CT-derived covariance patterns across different subsets of

age- and sex-matched subjects than SA and GMV metrics, particularly with smaller sample sizes and multi-site data (Carmon et al., 2020). This is likely due to the fact that CT is derived from signal intensity, therefore being more susceptible to measurement error stemming from scans of different resolutions, subjects of different ages, as well as software used to process scans and derive cortical measurements (Carmon et al., 2020).

5.6 Univariate versus Multivariate Results of Covariance Weights

Univariate models were used to complement the multivariate results with the aim of evaluating the associations between behavioural variables and variation in morphometric covariance. Linear models confirmed the absence of diagnostic effects on the covariance weights (Figure 4.8). Across analyses, age and IQ effects were pronounced, with CT showing inverse relationships with both variables. Main trends of global CT increases in females relative to males were consistent across modelling schemes. Interestingly, LGI failed to reach significance in any univariate model, relative to its profuse effects seen in the PLSC analysis. Indeed, LGI results differed substantially between the analytic approaches, with a lack of significant effects in all univariate results. This could suggest that individual behavioural predictors have greater effects on MC as a measure of cortical complexity, whereas the interdependencies between age, sex, and FIQ differentially impact LGI's contribution to cross-metric covariance. Future work could use interaction terms to model dependencies between predictor variables in univariate models.

5.7 Limitations

The findings presented here should be interpreted in the context of several limitations. First, many subjects were excluded due to QC and the matching protocol (only 44% of baseline subjects were included in matched multivariate analyses). Excluded individuals consisted of a greater proportion of ASD subjects, were younger, and had lower IQ and higher clinical severity relative to included subjects. Furthermore, the lack of standardized clinical and cognitive variables is a major drawback of unharmonized datasets. Indeed, only four variables (diagnosis, age, sex, FIQ) were complete across datasets, which limited our ability to quantitatively relate the multivariate results to additional dimensions of ASD symptomatology. Similarly, datasets differ in their ascertainment practices, exclusion criteria, and acquisition

protocols. Although we statistically accounted for these differences, they may nonetheless affect our findings. Lastly, age-related changes are modelled from cross-sectional data, and are represented as correlative linear relationships. Given the substantial age range of participants, it is likely that age-related trends may in fact be non-linear. The use of univariate approaches to model non-linear age terms in relation to covariance weights would be beneficial to supplement the current set of multivariate results.

6. Conclusion

We evaluated interdependencies between morphometric covariance patterns and subjects' behavioural characteristics in ASD and NT individuals, while accounting for several sources of methodological variation. We found that sexually differentiated macrostructural networks are largely driven by younger age and lower cognitive ability, with no significant diagnostic differences in the main LV. Our brain-behaviour results suggest that CT is the most robust contributor to covariance patterns, being primarily anchored in association areas involved in higher order cognitive processes. We show that individual variation in structural covariance has metric-specific functional delineations but limited clinical relevance.

References

1. Ad-Dab'bagh Y, Lyttelton O, Muehlboeck J, Lepage C, Einarson D, Mok K, et al. (2006). The CIVET image-processing environment: a fully automated comprehensive pipeline for anatomical neuroimaging research. *Proceedings of the 12th Annual Meeting of the Organization for Human Brain Mapping: Florence, Italy.*
2. Al-Beltagi, M. (2021). Autism medical comorbidities. *World Journal of Clinical Pediatrics*, *10*(3), 15–28.
3. Alexander-Bloch, A., Giedd, J. N., & Bullmore, E. (2013). Imaging structural co-variance between human brain regions. *Nature Reviews. Neuroscience*, *14*(5), 322–336.
4. American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition Text Revision* (DSM-IV-TR). Washington DC: American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890423349
5. American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. https://doi.org/10.1176/appi.books.9780890425596
6. Amso, D., & Casey, B. J. (2006). Beyond What Develops When: Neuroimaging May Inform How Cognition Changes With Development. *Current Directions in Psychological Science*, *15*(1), 24–29.
7. Anderson, A., Douglas, P. K., Kerr, W. T., Haynes, V. S., Yuille, A. L., Xie, J., Wu, Y. N., Brown, J. A., & Cohen, M. S. (2014). Non-negative matrix factorization of multimodal MRI, fMRI and phenotypic data reveals differential changes in default mode subnetworks in ADHD. *NeuroImage, 102*, 207–219. https://doi.org/10.1016/j.neuroimage.2013.12.015
8. Antezana, L., Factor, R. S., Condy, E. E., Strege, M. V., Scarpa, A., & Richey, J. A. (2019). Gender differences in restricted and repetitive behaviors and interests in youth with autism. *Autism Research: Official Journal of the International Society for Autism Research*, *12*(2), 274–283.
9. Ashburner, J. & Friston. K.J. (2003). Image segmentation. In R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, K.J. Friston, C.J. Price, S. Zeki, J. Ashburner, and W.D. Penny (Eds.), *Human Brain Function*. Academic Press, 2nd edition.
10. Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851.
11. Atun-Einy, O., Lotan, M., Harel, Y., Shavit, E., Burstein, S., & Kempner, G. (2013). Physical Therapy for Young Children Diagnosed with Autism Spectrum Disorders–

Clinical Frameworks Model in an Israeli Setting. *Frontiers in Pediatrics*, *1*. https://doi.org/10.3389/fped.2013.00019

12. Avino, T. A., & Hutsler, J. J. (2010). Abnormal cell patterning at the cortical gray–white matter boundary in autism spectrum disorders. *Brain Research*, *1360*, 138–146.
13. Avlund, S. H., Thomsen, P. H., Schendel, D., Jørgensen, M., Carlsen, A. H., & Clausen, L. (2021). Factors Associated with a Delayed Autism Spectrum Disorder Diagnosis in Children Previously Assessed on Suspicion of Autism. *Journal of Autism and Developmental Disorders*, *51*(11), 3843–3856.
14. Bai, D., Yip, B. H. K., Windham, G. C., Sourander, A., Francis, R., Yoffe, R., Glasson, E., Mahjani, B., Suominen, A., Leonard, H., Gissler, M., Buxbaum, J. D., Wong, K., Schendel, D., Kodesh, A., Breshnahan, M., Levine, S. Z., Parner, E. T., Hansen, S. N., … Sandin, S. (2019). Association of Genetic and Environmental Factors With Autism in a 5-Country Cohort. *JAMA Psychiatry* , *76*(10), 1035–1043.
15. Bailey, A., Luthert, P., Bolton, P., Le Couteur, A., Rutter, M., & Harding, B. (1993). Autism and megalencephaly. *The Lancet*, *341*(8854), 1225–1226. https://doi.org/10.1016/0140-6736(93)91065-t
16. Baranek, G. T. (2002). Efficacy of sensory and motor interventions for children with autism. *Journal of Autism and Developmental Disorders*, *32*(5), 397–422.
17. Baron-Cohen, S., Auyeung, B., Nørgaard-Pedersen, B., Hougaard, D. M., Abdallah, M. W., Melgaard, L., Cohen, A. S., Chakrabarti, B., Ruta, L., & Lombardo, M. V. (2015). Elevated fetal steroidogenic activity in autism. *Molecular Psychiatry*, *20*(3), 369–376.
18. Baron-Cohen, S., Lombardo, M. V., Auyeung, B., Ashwin, E., Chakrabarti, B., & Knickmeyer, R. (2011). Why are autism spectrum conditions more prevalent in males? *PLoS Biology*, *9*(6), e1001081.
19. Batchelor, P. G., Castellano Smith, A. D., Hill, D. L. G., Hawkes, D. J., Cox, T. C. S., & Dean, A. F. (2002). Measures of folding applied to the development of the human fetal brain. *IEEE Transactions on Medical Imaging*, *21*(8), 953–965.
20. Bauman, M. L. (2010). Medical comorbidities in autism: challenges to diagnosis and treatment. *Neurotherapeutics: The Journal of the American Society for Experimental NeuroTherapeutics*, *7*(3), 320–327.
21. Bayer, J. M. M., Dinga, R., Kia, S. M., Kottaram, A. R., Wolfers, T., Lv, J., Zalesky, A., Schmaal, L., & Marquand, A. (2021). Accommodating site variation in neuroimaging

data using normative and hierarchical Bayesian models. In *bioRxiv* (p. 2021.02.09.430363). https://doi.org/10.1101/2021.02.09.430363

22. Bayly, P. V., Okamoto, R. J., Xu, G., Shi, Y., & Taber, L. A. (2013). A cortical folding model incorporating stress-dependent growth explains gyral wavelengths and stress patterns in the developing brain. *Physical Biology*, *10*(1), 016005.
23. Bedford, S. A., Park, M. T. M., Devenyi, G. A., Tullo, S., Germann, J., Patel, R., Anagnostou, E., Baron-Cohen, S., Bullmore, E. T., Chura, L. R., Craig, M. C., Ecker, C., Floris, D. L., Holt, R. J., Lenroot, R., Lerch, J. P., Lombardo, M. V., Murphy, D. G. M., Raznahan, A., … Chakravarty, M. M. (2020). Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. *Molecular Psychiatry*, *25*(3), 614–628.
24. Benevides, T. W., Shore, S. M., Andresen, M.-L., Caplan, R., Cook, B., Gassner, D. L., Erves, J. M., Hazlewood, T. M., King, M. C., Morgan, L., Murphy, L. E., Purkis, Y., Rankowski, B., Rutledge, S. M., Welch, S. P., & Wittig, K. (2020). Interventions to address health outcomes among autistic adults: A systematic review. *Autism: The International Journal of Research and Practice*, *24*(6), 1345–1359.
25. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300.
26. Bernhardt, B. C., Valk, S. L., Silani, G., Bird, G., Frith, U., & Singer, T. (2014). Selective disruption of sociocognitive structural brain networks in autism and alexithymia. *Cerebral Cortex*, *24*(12), 3258–3267.
27. Bethlehem, R. A. I., Seidlitz, J., Romero-Garcia, R., Trakoshis, S., Dumas, G., & Lombardo, M. V. (2020). A normative modelling approach reveals age-atypical cortical thickness in a subgroup of males with autism spectrum disorder. *Communications Biology*, *3*(1), 1–10.
28. Bhaumik, R., Pradhan, A., Das, S., & Bhaumik, D. K. (2018). Predicting Autism Spectrum Disorder Using Domain-Adaptive Cross-Site Evaluation. *Neuroinformatics,16*(2), 197–205. https://doi.org/10.1007/s12021-018-9366-0
29. Bill, B. R., & Geschwind, D. H. (2009). Genetic advances in autism: heterogeneity and convergence on shared pathways. *Current Opinion in Genetics & Development*, *19*(3), 271–278.

30. Billstedt, E., Gillberg, C., & Gillberg, C. (2005). Autism after Adolescence: Population-based 13- to 22-year Follow-up Study of 120 Individuals with Autism Diagnosed in Childhood. In *Journal of Autism and Developmental Disorders*, *35*(3), 351–360. https://doi.org/10.1007/s10803-005-3302-5
31. Bos, D. J., Merchán-Naranjo, J., Martínez, K., Pina-Camacho, L., Balsa, I., Boada, L., Schnack, H., Oranje, B., Desco, M., Arango, C., Parellada, M., Durston, S., & Janssen, J. (2015). Reduced Gyrification Is Related to Reduced Interhemispheric Connectivity in Autism Spectrum Disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(8), 668–676.
32. Boutsidis, C., & Gallopoulos, E. (2008). SVD based initialization: A head start for nonnegative matrix factorization. *Pattern Recognition*, *41*(4), 1350–1362. https://doi.org/10.1016/j.patcog.2007.09.010
33. Boyes, R. G., Gunter, J. L., Frost, C., Janke, A. L., Yeatman, T., Hill, D. L. G., Bernstein, M. A., Thompson, P. M., Weiner, M. W., Schuff, N., Alexander, G. E., Killiany, R. J., DeCarli, C., Jack, C. R., & Fox, N. C. (2008). Intensity non-uniformity correction using N3 on 3-T scanners with multichannel phased array coils. *NeuroImage, 39*(4), 1752–1762. https://doi.org/10.1016/j.neuroimage.2007.10.026
34. Brett, D., Warnell, F., McConachie, H., & Parr, J. R. (2016). Factors Affecting Age at ASD Diagnosis in UK: No Evidence that Diagnosis Age has Decreased Between 2004 and 2014. *Journal of Autism and Developmental Disorders*, *46*(6), 1974–1984.
35. Brown, T. T., Kuperman, J. M., Chung, Y., Erhart, M., McCabe, C., Hagler, D. J., Jr, Venkatraman, V. K., Akshoomoff, N., Amaral, D. G., Bloss, C. S., Casey, B. J., Chang, L., Ernst, T. M., Frazier, J. A., Gruen, J. R., Kaufmann, W. E., Kenet, T., Kennedy, D. N., Murray, S. S., … Dale, A. M. (2012). Neuroanatomical assessment of biological maturity. *Current Biology: CB*, *22*(18), 1693–1698.
36. Brun, L., Auzias, G., Viellard, M., Villeneuve, N., Girard, N., Poinso, F., Da Fonseca, D., & Deruelle, C. (2016). Localized Misfolding Within Broca's Area as a Distinctive Feature of Autistic Disorder. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *1*(2), 160–168.
37. Bucher, M., Niebling, S., Han, Y., Molodenskiy, D., Hassani Nia, F., Kreienkamp, H.-J., Svergun, D., Kim, E., Kostyukova, A. S., Kreutz, M. R., & Mikhaylova, M. (2021). Autism-associated SHANK3 missense point mutations impact conformational

fluctuations and protein turnover at synapses. *eLife*, *10*. https://doi.org/10.7554/eLife.66165

38. Buck, T. R., Viskochil, J., Farley, M., Coon, H., McMahon, W. M., Morgan, J., & Bilder, D. A. (2014). Psychiatric comorbidity and medication use in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(12), 3063–3071.
39. Budday, S., Steinmann, P., & Kuhl, E. (2015). Physical biology of human brain development. *Frontiers in Cellular Neuroscience*, *9*, 257.
40. Burgess, A. F., & Gutstein, S. E. (2007). Quality of Life for people with autism: Raising the standard for evaluating successful outcomes. *Child and Adolescent Mental Health*, *12*(2), 80–86.
41. Buss, C., Entringer, S., & Wadhwa, P. D. (2012). Fetal programming of brain development: intrauterine stress and susceptibility to psychopathology. *Science Signaling*, *5*(245), t7.
42. Cafiero, R., Brauer, J., Anwander, A., & Friederici, A. D. (2019). The Concurrence of Cortical Surface Area Expansion and White Matter Myelination in Human Brain Development. *Cerebral Cortex*, *29*(2), 827–837.
43. Cai, S., Wang, X., Yang, F., Chen, D., & Huang, L. (2021). Differences in Brain Structural Covariance Network Characteristics in Children and Adults With Autism Spectrum Disorder. *Autism Research: Official Journal of the International Society for Autism Research*, *14*(2), 265–275.
44. Carmon, J., Heege, J., Necus, J. H., Owen, T. W., Pipa, G., Kaiser, M., Taylor, P. N., & Wang, Y. (2020). Reliability and comparability of human brain structural covariance networks. *NeuroImage*, *220*, 117104.
45. Carter, A. S., Black, D. O., Tewani, S., Connolly, C. E., Kadlec, M. B., & Tager-Flusberg, H. (2007). Sex differences in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(1), 86–97.
46. Casanova, M. F., Buxhoeveden, D. P., Switala, A. E., & Roy, E. (2002). Minicolumnar pathology in autism. *Neurology*, *58*(3), 428–432.
47. Casanova, M. F., El-Baz, A., Mott, M., Mannheim, G., Hassan, H., Fahmi, R., Giedd, J., Rumsey, J. M., Switala, A. E., & Farag, A. (2009). Reduced gyral window and corpus callosum size in autism: possible macroscopic correlates of a minicolumnopathy. *Journal of Autism and Developmental Disorders*, *39*(5), 751–764.

48. Casanova, M. F., van Kooten, I. A. J., Switala, A. E., van Engeland, H., Heinsen, H., Steinbusch, H. W. M., Hof, P. R., Trippe, J., Stone, J., & Schmitz, C. (2006). Minicolumnar abnormalities in autism. *Acta Neuropathologica*, *112*(3), 287–303.
49. Case-Smith, J., Weaver, L. L., & Fristad, M. A. (2015). A systematic review of sensory processing interventions for children with autism spectrum disorders. *Autism: The International Journal of Research and Practice*, *19*(2), 133–148.
50. Chen, A. A., Beer, J. C., Tustison, N. J., Cook, P. A., Shinohara, R. T., Shou, H., & Alzheimer's Disease Neuroimaging Initiative. (2022). Mitigating site effects in covariance for machine learning in neuroimaging data. *Human Brain Mapping*, *43*(4), 1179–1195.
51. Chen, J., Liu, J., Calhoun, V. D., Arias-Vasquez, A., Zwiers, M. P., Gupta, C. N., Franke, B., & Turner, J. A. (2014). Exploration of scanning effects in multi-site structural MRI studies. *Journal of Neuroscience Methods*, *230*, 37–50.
52. Chi, J. G., Dooling, E. C., & Gilles, F. H. (1977). Gyral development of the human brain. In *Annals of Neurology*, *1*(1), 86–93. https://doi.org/10.1002/ana.410010109
53. Choi, H. S., Haynor, D. R., & Kim, Y. (1991). Partial volume tissue classification of multichannel magnetic resonance images-a mixel model. *IEEE Transactions on Medical Imaging*, *10*(3), 395–407.
54. Choi, Y. Y., Shamosh, N. A., Cho, S. H., DeYoung, C. G., Lee, M. J., Lee, J.-M., Kim, S. I., Cho, Z.-H., Kim, K., Gray, J. R., & Lee, K. H. (2008). Multiple bases of human intelligence revealed by cortical thickness and neural activation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(41), 10323–10329.
55. Chuang, K.-S., Tzeng, H.-L., Chen, S., Wu, J., & Chen, T.-J. (2006). Fuzzy c-means clustering with spatial information for image segmentation. *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society*, *30*(1), 9–15.
56. Chung, M. K., Worsley, K. J., Paus, T., Cherif, C., Collins, D. L., Giedd, J. N., Rapoport, J. L., & Evans, A. C. (2001). A unified statistical approach to deformation-based morphometry. *NeuroImage*, *14*(3), 595–606.
57. Chung, M. K., Worsley, K. J., Robbins, S., Paus, T., Taylor, J., Giedd, J. N., Rapoport, J. L., & Evans, A. C. (2003). Deformation-based surface morphometry applied to

gray matter deformation. *NeuroImage*, *18*(2), 198–213. https://doi.org/10.1016/s1053-8119(02)00017-4

58. Clouchoux, C., Kudelski, D., Gholipour, A., Warfield, S. K., Viseur, S., Bouyssi-Kobar, M., Mari, J.-L., Evans, A. C., du Plessis, A. J., & Limperopoulos, C. (2012). Quantitative in vivo MRI measurement of cortical development in the fetus. *Brain Structure & Function*, *217*(1), 127–139.
59. Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, *18*(2), 192–205.
60. Courchesne, E., Carper, R., & Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. *JAMA: The Journal of the American Medical Association*, *290*(3), 337–344.
61. Courchesne, E., Karns, C. M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D., Chisum, H. J., Moses, P., Pierce, K., Lord, C., Lincoln, A. J., Pizzo, S., Schreibman, L., Haas, R. H., Akshoomoff, N. A., & Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: an MRI study. *Neurology*, *57*(2), 245–254.
62. Courchesne, E., Mouton, P. R., Calhoun, M. E., Semendeferi, K., Ahrens-Barbeau, C., Hallet, M. J., Barnes, C. C., & Pierce, K. (2011). Neuron number and size in prefrontal cortex of children with autism. *JAMA: The Journal of the American Medical Association*, *306*(18), 2001–2010.
63. Croen, L. A., Zerbo, O., Qian, Y., Massolo, M. L., Rich, S., Sidney, S., & Kripke, C. (2015). The health status of adults on the autism spectrum. *Autism: The International Journal of Research and Practice*, *19*(7), 814–823.
64. Currie, S., Hoggard, N., Craven, I. J., Hadjivassiliou, M., & Wilkinson, I. D. (2013). Understanding MRI: basic MR physics for physicians. *Postgraduate Medical Journal*, *89*(1050), 209–223.
65. Cuthbert, B. N. (2015). Research Domain Criteria: toward future psychiatric nosologies. *Dialogues in Clinical Neuroscience*, *17*(1), 89–97.
66. Dadar, M., Fonov, V. S., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2018). A comparison of publicly available linear MRI stereotaxic registration techniques. *NeuroImage*, *174*, 191–200.

67. Dale, B. M., Brown, M. A., & Semelka, R. C. (2015). *MRI: Basic Principles and Applications*. John Wiley & Sons.
68. Daniels, A. M., & Mandell, D. S. (2014). Explaining differences in age at autism spectrum disorder diagnosis: a critical review. *Autism: The International Journal of Research and Practice*, *18*(5), 583–597.
69. de Bruin, E. I., Ferdinand, R. F., Meester, S., de Nijs, P. F. A., & Verheij, F. (2007). High rates of psychiatric co-morbidity in PDD-NOS. *Journal of Autism and Developmental Disorders*, *37*(5), 877–886.
70. de Giambattista, C., Ventura, P., Trerotoli, P., Margari, F., & Margari, L. (2021). Sex Differences in Autism Spectrum Disorder: Focus on High Functioning Children and Adolescents. *Frontiers in Psychiatry/Frontiers Research Foundation*, *12*. https://doi.org/10.3389/fpsyt.2021.539835
71. Deoni, S. C. L., Williams, S. C. R., Jezzard, P., Suckling, J., Murphy, D. G. M., & Jones, D. K. (2008). Standardized structural magnetic resonance imaging in multicentre studies using quantitative T1 and T2 imaging at 1.5 T. *NeuroImage*, *40*(2), 662–671.
72. Diamond, & Sekhon. (2012). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *The Review of Economics and Statistics*. https://direct.mit.edu/rest/article-abstract/95/3/932/5810
73. Dierker, D. L., Feczko, E., Pruett, J. R., Jr, Petersen, S. E., Schlaggar, B. L., Constantino, J. N., Harwell, J. W., Coalson, T. S., & Van Essen, D. C. (2015). Analysis of cortical shape in children with simplex autism. *Cerebral Cortex*, *25*(4), 1042–1051.
74. do Carmo, M.P. (1976). *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Inc.
75. Doyle-Thomas, K. A. R., Kushki, A., Duerden, E. G., Taylor, M. J., Lerch, J. P., Soorya, L. V., Wang, A. T., Fan, J., & Anagnostou, E. (2013). The effect of diagnosis, age, and symptom severity on cortical surface area in the cingulate cortex and insula in autism spectrum disorders. *Journal of Child Neurology*, *28*(6), 732–739.
76. Dubois, J., Benders, M., Borradori-Tolsa, C., Cachia, A., Lazeyras, F., Ha-Vinh Leuchter, R., Sizonenko, S. V., Warfield, S. K., Mangin, J. F., & Hüppi, P. S. (2008). Primary cortical folding in the human newborn: an early marker of later functional development. *Brain: A Journal of Neurology*, *131*(8), 2028–2041.

77. Ducharme, S., Albaugh, M. D., Nguyen, T.-V., Hudziak, J. J., Mateos-Pérez, J. M., Labbe, A., Evans, A. C., & Karama, S. (2015). Trajectories of cortical surface area and cortical volume maturation in normal brain development. *Data in Brief*, *5*, 929–938. https://doi.org/10.1016/j.dib.2015.10.044

78. Ducharme, S., Albaugh, M. D., Nguyen, T.-V., Hudziak, J. J., Mateos-Pérez, J. M., Labbe, A., Evans, A. C., & Karama, S. (2016). Trajectories of cortical thickness maturation in normal brain development — The importance of quality control procedures. In *NeuroImage*, *125*, 267–279. https://doi.org/10.1016/j.neuroimage.2015.10.010

79. Dudley, C., & Emery, J. C. H. (2014). *The Value of Caregiver Time: Costs of Support and Care for Individuals Living with Autism Spectrum Disorder*. https://doi.org/10.2139/ssrn.237963

80. Ecker, C., Andrews, D., Dell'Acqua, F., Daly, E., Murphy, C., Catani, M., Thiebaut de Schotten, M., Baron-Cohen, S., Lai, M. C., Lombardo, M. V., Bullmore, E. T., Suckling, J., Williams, S., Jones, D. K., Chiocchetti, A., the MRC AIMS Consortium, Murphy, D. G. M., Bailey, A. J., Baron-Cohen, S., … Wilson, C. E. (2016). Relationship Between Cortical Gyrification, White Matter Connectivity, and Autism Spectrum Disorder. *Cerebral Cortex*, *26*(7), 3297–3309.

81. Ecker, C., Andrews, D. S., Gudbrandsen, C. M., Marquand, A. F., Ginestet, C. E., Daly, E. M., Murphy, C. M., Lai, M.-C., Lombardo, M. V., Ruigrok, A. N. V., Bullmore, E. T., Suckling, J., Williams, S. C. R., Baron-Cohen, S., Craig, M. C., Murphy, D. G. M., & for the Medical Research Council Autism Imaging Multicentre Study (MRC AIMS) Consortium. (2017). Association Between the Probability of Autism Spectrum Disorder and Normative Sex-Related Phenotypic Diversity in Brain Structure. *JAMA Psychiatry*,*74*(4), 329. https://doi.org/10.1001/jamapsychiatry.2016.3990

82. Ecker, C., Ginestet, C., Feng, Y., Johnston, P., Lombardo, M. V., Lai, M.-C., Suckling, J., Palaniyappan, L., Daly, E., Murphy, C. M., Williams, S. C., Bullmore, E. T., Baron-Cohen, S., Brammer, M., Murphy, D. G. M., & MRC AIMS Consortium. (2013). Brain surface anatomy in adults with autism: the relationship between surface area, cortical thickness, and autistic symptoms. *JAMA Psychiatry*, *70*(1), 59–70.

83. Ecker, C., Shahidiani, A., Feng, Y., Daly, E., Murphy, C., D'Almeida, V., Deoni, S., Williams, S. C., Gillan, N., Gudbrandsen, M., Wichers, R., Andrews, D., Van Hemert, L., & Murphy, D. G. M. (2014). The effect of age, diagnosis, and their interaction on

vertex-based measures of cortical thickness and surface area in autism spectrum disorder. *Journal of Neural Transmission*, *121*(9), 1157–1170.

84. Edwards, L. J., Kirilina, E., Mohammadi, S., & Weiskopf, N. (2018). Microstructural imaging of human neocortex in vivo. *NeuroImage*, *182,* 184–206. https://doi.org/10.1016/j.neuroimage.2018.02.055
85. Eigsti, I., Fein, D., & Larson, C. (2022). Editorial Perspective: Another look at "optimal outcome" in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*. https://doi.org/10.1111/jcpp.13658
86. Eisenberg, I. W., Wallace, G. L., Kenworthy, L., Gotts, S. J., & Martin, A. (2015). Insistence on sameness relates to increased covariance of gray matter structure in autism spectrum disorder. *Molecular Autism*, *6*(1). https://doi.org/10.1186/s13229-015-0047-7
87. Elder, J., Kreider, C., Brasher, S., & Ansell, M. (2017). Clinical impact of early diagnosis of autism on the prognosis and parent-child relationships. In *Psychology Research and Behavior Management* (Vol. 10, pp. 283–292). https://doi.org/10.2147/prbm.s117499
88. Elias, H., & Schwartz, D. (1969). Surface areas of the cerebral cortex of mammals determined by stereological methods. *Science*, *166*(3901), 111–113.
89. Elmaoğlu, M., & Çelik, A. (2012). Fundamentals of Magnetic Resonance Imaging. In M. Elmaoğlu & A. Çelik (Eds.), *MRI Handbook: MR Physics, Patient Positioning, and Protocols* (pp. 7–23). Springer US.
90. Entringer, S., Buss, C., & Wadhwa, P. D. (2015). Prenatal stress, development, health and disease risk: A psychobiological perspective-2015 Curt Richter Award Paper. *Psychoneuroendocrinology*, *62*, 366–375.
91. Eskildsen, S. F., Coupé, P., Fonov, V., Manjón, J. V., Leung, K. K., Guizard, N., Wassef, S. N., Østergaard, L. R., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2012). BEaST: brain extraction based on nonlocal segmentation technique. *NeuroImage*, *59*(3), 2362–2373.
92. Evans, A. C. (2005). Large-scale morphometric analysis of neuroanatomy and neuropathology. *Anatomy and Embryology*, *210*(5-6), 439–446.
93. Farley, M., & McMahon, B. (2014). Range of Outcomes and Challenges in Middle and Later Life. In *Adolescents and Adults with Autism Spectrum Disorders* (pp. 211–238). https://doi.org/10.1007/978-1-4939-0506-5_11

94. Fava, L., & Strauss, K. (2014). Response to Early Intensive Behavioral Intervention for autism—An umbrella approach to issues critical to treatment individualization. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, *39*, 49–58.

95. Fecteau, S., Mottron, L., Berthiaume, C., & Burack, J. A. (2003). Developmental Changes of Autistic Symptoms. *Autism*, *7*(3), 255–268. https://doi.org/10.1177/13623613030073003

96. Fernández, V., Llinares-Benadero, C., & Borrell, V. (2016). Cerebral cortex expansion and folding: what have we learned? *The EMBO Journal*, *35*(10), 1021–1044.

97. Fjell, A. M., Grydeland, H., Krogsrud, S. K., Amlien, I., Rohani, D. A., Ferschmann, L., Storsve, A. B., Tamnes, C. K., Sala-Llonch, R., Due-Tønnessen, P., Bjørnerud, A., Sølsnes, A. E., Håberg, A. K., Skranes, J., Bartsch, H., Chen, C.-H., Thompson, W. K., Panizzon, M. S., Kremen, W. S., … Walhovd, K. B. (2015). Development and aging of cortical thickness correspond to genetic organization patterns. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(50), 15462–15467.

98. Forde, N. J., Ronan, L., Zwiers, M. P., Schweren, L. J. S., Alexander-Bloch, A. F., Franke, B., Faraone, S. V., Oosterlaan, J., Heslenfeld, D. J., Hartman, C. A., Buitelaar, J. K., & Hoekstra, P. J. (2017). Healthy cortical development through adolescence and early adulthood. *Brain Structure & Function*, *222*(8), 3653–3663.

99. Fortin, J.-P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., Adams, P., Cooper, C., Fava, M., McGrath, P. J., McInnis, M., Phillips, M. L., Trivedi, M. H., Weissman, M. M., & Shinohara, R. T. (2018). Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, *167*, 104–120.

100. Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., Schultz, R. T., Verma, R., & Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, *161*, 149–170.

101. Fortin, J.-P., Sweeney, E. M., Muschelli, J., Crainiceanu, C. M., Shinohara, R. T., & Alzheimer's Disease Neuroimaging Initiative. (2016). Removing inter-subject technical variability in magnetic resonance imaging studies. *NeuroImage*, *132*, 198–212.

102. Fountain, C., Winter, A. S., & Bearman, P. S. (2012). Six developmental trajectories characterize children with autism. *Pediatrics*, *129*(5), e1112–e1120.

103. Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210.

104. Genovese, A., & Butler, M. G. (2020). Clinical Assessment, Genetics, and Treatment Approaches in Autism Spectrum Disorder (ASD). *International Journal of Molecular Sciences, 21*(13), 4726. https://doi.org/10.3390/ijms21134726

105. Ghanbari, Y., Smith, A. R., Schultz, R. T., & Verma, R. (2014). Identifying group discriminative and age regressive sub-networks from DTI-based connectivity via a unified framework of non-negative matrix factorization and graph embedding. *Medical Image Analysis*, *18*(8), 1337–1348. https://doi.org/10.1016/j.media.2014.06.006

106. Gharehgazlou, A., Freitas, C., Ameis, S. H., Taylor, M. J., Lerch, J. P., Radua, J., & Anagnostou, E. (2021). Cortical Gyrification Morphology in Individuals with ASD and ADHD across the Lifespan: A Systematic Review and Meta-Analysis. *Cerebral Cortex*, *31*(5), 2653–2669.

107. Gharehgazlou, A., Vandewouw, M., Ziolkowski, J., Wong, J., Crosbie, J., Schachar, R., Nicolson, R., Georgiades, S., Kelley, E., Ayub, M., Hammill, C., Ameis, S. H., Taylor, M. J., Lerch, J. P., & Anagnostou, E. (2021). Cortical Gyrification Morphology in ASD and ADHD: Implication for Further Similarities or Disorder-Specific Features? *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhab326

108. Giarelli, E., Wiggins, L. D., Rice, C. E., Levy, S. E., Kirby, R. S., Pinto-Martin, J., & Mandell, D. (2010). Sex differences in the evaluation and diagnosis of autism spectrum disorders among children. *Disability and Health Journal*, *3*(2), 107–116.

109. Gilmore, J. H., Langworthy, B., Girault, J. B., Fine, J., Jha, S. C., Kim, S. H., Cornea, E., & Styner, M. (2020). Individual Variation of Human Cortical Structure Is Established in the First Year of Life. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *5*(10), 971–980.

110. Gilmore, J. H., Shi, F., Woolson, S. L., Knickmeyer, R. C., Short, S. J., Lin, W., Zhu, H., Hamer, R. M., Styner, M., & Shen, D. (2012). Longitudinal development of cortical and subcortical gray matter from birth to 2 years. *Cerebral Cortex*, *22*(11), 2478–2485.

111. Giulivi, C., Zhang, Y.-F., Omanska-Klusek, A., Ross-Inta, C., Wong, S., Hertz-Picciotto, I., Tassone, F., & Pessah, I. N. (2010). Mitochondrial dysfunction in autism. *JAMA: The Journal of the American Medical Association*, *304*(21), 2389–2396.

112. Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., 3rd, Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., & Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(21), 8174–8179.

113. Goh, S., Bansal, R., Xu, D., Hao, X., Liu, J., & Peterson, B. S. (2011). Neuroanatomical correlates of intellectual ability across the life span. *Developmental Cognitive Neuroscience*, *1*(3), 305–312.

114. Gore, A. C., Martien, K. M., Gagnidze, K., & Pfaff, D. (2014). Implications of Prenatal Steroid Perturbations for Neurodevelopment, Behavior, and Autism. *Endocrine Reviews*, *35*(6), 961–991. https://doi.org/10.1210/er.2013-1122

115. Górriz, J. M., Ramírez, J., Segovia, F., Martínez, F. J., Lai, M.-C., Lombardo, M. V., Baron-Cohen, S., Suckling, J., & MRC AIMS Consortium. (2019). A Machine Learning Approach to Reveal the NeuroPhenotypes of Autisms. *International Journal of Neural Systems, 29*(7). https://doi.org/10.1142/s0129065718500582

116. Graham, A. M., Marr, M., Buss, C., Sullivan, E. L., & Fair, D. A. (2021). Understanding Vulnerability and Adaptation in Early Brain Development using Network Neuroscience. *Trends in Neurosciences*, *44*(4), 276–288.

117. Gregory, S. G., Connelly, J. J., Towers, A. J., Johnson, J., Biscocho, D., Markunas, C. A., Lintas, C., Abramson, R. K., Wright, H. H., Ellis, P., Langford, C. F., Worley, G., Delong, G. R., Murphy, S. K., Cuccaro, M. L., Persico, A., & Pericak-Vance, M. A. (2009). Genomic and epigenetic evidence for oxytocin receptor deficiency in autism. *BMC Medicine*, *7*, 62.

118. Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O. A., Anney, R., Awashti, S., Belliveau, R., Bettella, F., Buxbaum, J. D., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Christensen, J. H., … Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*(3), 431–444.

119. Guan, J., Zhuang, Y., Kang, Y., & Ji, G. (2022). Shared and Cell-Type-Specific Gene Expression Patterns Associated With Autism Revealed by Integrative Regularized Non-Negative Matrix Factorization. *Frontiers in Genetics*,*13*. https://doi.org/10.3389/fgene.2022.865371

120. Guillaume, B., Wang, C., Poh, J., Shen, M. J., Ong, M. L., Tan, P. F., Karnani, N., Meaney, M., & Qiu, A. (2018). Improving mass-univariate analysis of neuroimaging data by modelling important unknown covariates: Application to Epigenome-Wide Association Studies. *NeuroImage*, *173*, 57–71.

121. Haar, S., Berman, S., Behrmann, M., & Dinstein, I. (2016). Anatomical Abnormalities in Autism? *Cerebral Cortex*, *26*(4), 1440–1452.

122. Habas, P. A., Scott, J. A., Roosta, A., Rajagopalan, V., Kim, K., Rousseau, F., Barkovich, A. J., Glenn, O. A., & Studholme, C. (2012). Early folding patterns and asymmetries of the normal human brain detected from in utero MRI. *Cerebral Cortex*, *22*(1), 13–25.

123. Habeck, C. G. (2010). Basics of Multivariate Analysis in Neuroimaging Data. *Journal of Visualized Experiments*(41). https://doi.org/10.3791/1988

124. Habeck, C., the Alzheimer's Disease Neuroimaging Initiative, & Stern, Y. (2010). Multivariate Data Analysis for Neuroimaging Data: Overview and Application to Alzheimer's Disease. *Cell Biochemistry and Biophysics,58*(2), 53–67. https://doi.org/10.1007/s12013-010-9093-0

125. Hadjikhani, N., Joseph, R. M., Snyder, J., & Tager-Flusberg, H. (2006). Anatomical differences in the mirror neuron system and social cognition network in autism. *Cerebral Cortex*, *16*(9), 1276–1282.

126. Hammill, C., Lerch, J. P., Taylor, M. J., Ameis, S. H., Chakravarty, M. M., Szatmari, P., Anagnostou, E., & Lai, M.-C. (2021). Quantitative and Qualitative Sex Modulations in the Brain Anatomy of Autism. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *6*(9), 898–909.

127. Hardan, A. Y., Jou, R. J., Keshavan, M. S., Varma, R., & Minshew, N. J. (2004). Increased frontal cortical folding in autism: a preliminary MRI study. *Psychiatry Research*, *131*(3), 263–268.

128. Hardan, A. Y., Muddasani, S., Vemulapalli, M., Keshavan, M. S., & Minshew, N. J. (2006). An MRI study of increased cortical thickness in autism. *The American Journal of Psychiatry*, *163*(7), 1290–1292.

129. Hartley, S. L., & Sikora, D. M. (2009). Sex differences in autism spectrum disorder: an examination of developmental functioning, autistic symptoms, and coexisting behavior problems in toddlers. *Journal of Autism and Developmental Disorders*, *39*(12), 1715–1722.

130. Hasegawa, M., Houdou, S., Mito, T., Takashima, S., Asanuma, K., & Ohno, T. (1992). Development of myelination in the human fetal and infant cerebrum: a myelin basic protein immunohistochemical study. *Brain & Development*, *14*(1), 1–6.

131. Hazlett, H. C., Poe, M. D., Gerig, G., Smith, R. G., & Piven, J. (2006). Cortical gray and white brain tissue volume in adolescents and adults with autism. *Biological Psychiatry*, *59*(1), 1–6.

132. Hazlett, H. C., Poe, M. D., Gerig, G., Styner, M., Chappell, C., Smith, R. G., Vachet, C., & Piven, J. (2011). Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. *Archives of General Psychiatry*, *68*(5), 467–476.

133. Head, A. M., McGillivray, J. A., & Stokes, M. A. (2014). Gender differences in emotionality and sociability in children with autism spectrum disorders. *Molecular Autism*, *5*(1), 19.

134. Hébert, M. L. J., Kehayia, E., Prelock, P., Wood-Dauphinee, S., & Snider, L. (2014). Does occupational therapy play a role for communication in children with autism spectrum disorders? *International Journal of Speech-Language Pathology*,16(6), 594–602. https://doi.org/10.3109/17549507.2013.876665

135. Hilgetag, C. C., & Barbas, H. (2006). Role of mechanical factors in the morphology of the primate cerebral cortex. *PLoS Computational Biology*, *2*(3), e22.

136. Hill, J., Inder, T., Neil, J., Dierker, D., Harwell, J., & Van Essen, D. (2010). Similar patterns of cortical expansion during human development and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(29), 13135–13140.

137. Hiller, R. M., Young, R. L., & Weber, N. (2014). Sex differences in autism spectrum disorder based on DSM-5 criteria: evidence from clinician and teacher reporting. *Journal of Abnormal Child Psychology*, *42*(8), 1381–1393.

138. Hirjak, D., Wolf, R. C., Paternoga, I., Kubera, K. M., Thomann, A. K., Stieltjes, B., Maier-Hein, K. H., & Thomann, P. A. (2016). Neuroanatomical Markers of Neurological Soft Signs in Recent-Onset Schizophrenia and Asperger-Syndrome. *Brain Topography*, *29*(3), 382–394.

139. Hirvikoski, T., Mittendorfer-Rutz, E., Boman, M., Larsson, H., Lichtenstein, P., & Bölte, S. (2016). Premature mortality in autism spectrum disorder. *The British Journal of Psychiatry: The Journal of Mental Science*, *208*(3), 232–238.

140. Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software, 42*(8). https://doi.org/10.18637/jss.v042.i08

141. Hodges, H., Fealko, C., & Soares, N. (2020). Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation. *Translational Pediatrics*, *9*(Suppl 1), S55–S65.

142. Hogstrom, L. J., Westlye, L. T., Walhovd, K. B., & Fjell, A. M. (2013). The structure of the cerebral cortex across adult life: age-related patterns of surface area, thickness, and gyrification. *Cerebral Cortex*, *23*(11), 2521–2530.

143. Hrdlicka, M., Vacova, M., Oslejskova, H., Gondzova, V., Vadlejchova, I., Kocourkova, J., Koutek, J., & Dudova, I. (2016). Age at diagnosis of autism spectrum disorders: is there an association with socioeconomic status and family self-education about autism? *Neuropsychiatric Disease and Treatment*, *12*, 1639–1644.

144. Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.-C., & Mandy, W. (2017). "Putting on My Best Normal": Social Camouflaging in Adults with Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*,*47*(8), 2519–2534. https://doi.org/10.1007/s10803-017-3166-5

145. Hull, L., Petrides, K. V., & Mandy, W. (2020). The Female Autism Phenotype and Camouflaging: a Narrative Review. *Review Journal of Autism and Developmental Disorders*, *7*(4), 306–317.

146. Hyde, K. L., Samson, F., Evans, A. C., & Mottron, L. (2010). Neuroanatomical differences in brain areas implicated in perceptual and other core features of autism revealed by cortical thickness analysis and voxel-based morphometry. *Human Brain Mapping*, *31*(4), 556–566.

147. Ismail, M. M. T., Keynton, R. S., Mostapha, M. M. M. O., ElTanboly, A. H., Casanova, M. F., Gimel'farb, G. L., & El-Baz, A. (2016). Studying Autism Spectrum Disorder with Structural and Diffusion Magnetic Resonance Imaging: A Survey. *Frontiers in Human Neuroscience*, *10*, 211.

148. Jacobs, G. R., Voineskos, A. N., Hawco, C., Stefanik, L., Forde, N. J., Dickie, E. W., Lai, M.-C., Szatmari, P., Schachar, R., Crosbie, J., Arnold, P. D., Goldenberg, A., Erdman, L., & Ameis, S. H. (2021). Integration of brain and behavior measures for identification of data-driven groups cutting across children with ASD, ADHD, or OCD. In *Neuropsychopharmacology*, *46*(3), 643–653. https://doi.org/10.1038/s41386-020-00902-6

149. Jacquemont, S., Coe, B. P., Hersch, M., Duyzend, M. H., Krumm, N., Bergmann, S., Beckmann, J. S., Rosenfeld, J. A., & Eichler, E. E. (2014). A Higher Mutational Burden in Females Supports a "Female Protective Model" in Neurodevelopmental Disorders. *American Journal of Human Genetics*, *94*(3), 415–425.

150. Jalbrzikowski, M., Jonas, R., Senturk, D., Patel, A., Chow, C., Green, M. F., & Bearden, C. E. (2013). Structural abnormalities in cortical volume, thickness, and surface area in 22q11.2 microdeletion syndrome: Relationship with psychotic symptoms. *NeuroImage: Clinical,3,* 405–415. https://doi.org/10.1016/j.nicl.2013.09.013

151. Jarrold, C., & Brock, J. (2004). To match or not to match? Methodological issues in autism-related research. *Journal of Autism and Developmental Disorders*, *34*(1), 81–86.

152. Jeon, T., Mishra, V., Ouyang, M., Chen, M., & Huang, H. (2015). Synchronous Changes of Cortical Thickness and Corresponding White Matter Microstructure During Brain Development Accessed by Diffusion MRI Tractography from Parcellated Cortex. In *Frontiers in Neuroanatomy*,*9*. https://doi.org/10.3389/fnana.2015.00158

153. Jha, S. C., Xia, K., Schmitt, J. E., Ahn, M., Girault, J. B., Murphy, V. A., Li, G., Wang, L., Shen, D., Zou, F., Zhu, H., Styner, M., Knickmeyer, R. C., & Gilmore, J. H. (2018). Genetic influences on neonatal cortical thickness and surface area. *Human Brain Mapping*, *39*(12), 4998–5013.

154. Jiang, X., & Nardelli, J. (2016). Cellular and molecular introduction to brain development. *Neurobiology of Disease*, *92*(Pt A), 3–17.

155. Jiao, Y., Chen, R., Ke, X., Chu, K., Lu, Z., & Herskovits, E. H. (2010). Predictive models of autism spectrum disorder based on brain regional cortical thickness. *NeuroImage*, *50*(2), 589–599.

156. Kahali, S., Adhikari, S. K., & Sing, J. K. (2016). On estimation of bias field in MRI images: polynomial vs Gaussian surface fitting method. *Journal of Chemometrics*, *30*(10), 602–620.

157. Kalantar-Hormozi, H., Patel, R., Dai, A., Ziolkowski, J., Dong, H.-M., Holmes, A., Raznahan, A., Devenyi, G. A., & Mallar Chakravarty, M. (2022). Human Brain Development: a cross-sectional and longitudinal study integrating multiple neuromorphological features. *bioRxiv* (p. 2022.07.21.501018). https://doi.org/10.1101/2022.07.21.501018

158. Karama, S., Ad-Dab'bagh, Y., Haier, R. J., Deary, I. J., Lyttelton, O. C., Lepage, C., & Evans, A. C. (2009). Positive association between cognitive ability and cortical thickness in a representative US sample of healthy 6 to 18 year-olds. *NeuroImage*,*47,* p. S88. https://doi.org/10.1016/s1053-8119(09)70678-0

159. Karama, S., Colom, R., Johnson, W., Deary, I. J., Haier, R., Waber, D. P., Lepage, C., Ganjavi, H., Jung, R., Evans, A. C., & Brain Development Cooperative Group. (2011). Cortical thickness correlates of specific cognitive performance accounted for by the general factor of intelligence in healthy children aged 6 to 18. *NeuroImage*, *55*(4), 1443–1453.

160. Kates, W. R., Ikuta, I., & Burnette, C. P. (2009). Gyrification patterns in monozygotic twin pairs varying in discordance for autism. *Autism Research: Official Journal of the - International Society for Autism Research*, *2*(5), 267–278.

161. Khan, A., Harney, J. W., Zavacki, A. M., & Sajdel-Sulkowska, E. M. (2014). Disrupted brain thyroid hormone homeostasis and altered thyroid hormone-dependent brain gene expression in autism spectrum disorders. *Journal of Physiology and Pharmacology: An Official Journal of the Polish Physiological Society*, *65*(2), 257–272.

162. Khundrakpam, B. S., Lewis, J. D., Kostopoulos, P., Carbonell, F., & Evans, A. C. (2017). Cortical Thickness Abnormalities in Autism Spectrum Disorders Through Late Childhood, Adolescence, and Adulthood: A Large-Scale MRI Study. *Cerebral Cortex, 27*(3), 1721–1731.

163. Khundrakpam, B. S., Lewis, J. D., Reid, A., Karama, S., Zhao, L., Chouinard-Decorte, F., Evans, A. C., & Brain Development Cooperative Group. (2017). Imaging structural covariance in the development of intelligence. *NeuroImage*, *144*, 227–240.

164. Khundrakpam, B. S., Reid, A., Brauer, J., Carbonell, F., Lewis, J., Ameis, S., Karama, S., Lee, J., Chen, Z., Das, S., Evans, A. C., & Brain Development Cooperative Group. (2013). Developmental changes in organization of structural brain networks. *Cerebral Cortex*, *23*(9), 2072–2085.

165. Kim, J. S., Singh, V., Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J. M., Kim, S. I., & Evans, A. C. (2005). Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *NeuroImage*, *27*(1), 210–221.

166. King, J. B., Lopez-Larson, M. P., & Yurgelun-Todd, D. A. (2016). Mean cortical curvature reflects cytoarchitecture restructuring in mild traumatic brain injury. *NeuroImage. Clinical*, *11*, 81–89.

167. Klein, D., Rotarska-Jagiela, A., Genc, E., Sritharan, S., Mohr, H., Roux, F., Han, C. E., Kaiser, M., Singer, W., & Uhlhaas, P. J. (2014). Adolescent Brain Maturation and Cortical Folding: Evidence for Reductions in Gyrification. In *PLoS ONE,9*(1), e84914. https://doi.org/10.1371/journal.pone.0084914

168. Kohli, J. S., Kinnear, M. K., Fong, C. H., Fishman, I., Carper, R. A., & Müller, R.-A. (2019). Local Cortical Gyrification is Increased in Children With Autism Spectrum Disorders, but Decreases Rapidly in Adolescents. *Cerebral Cortex*, *29*(6), 2412–2423.

169. Koolschijn, P. C. M. P., & Geurts, H. M. (2016). Gray Matter Characteristics in Mid and Old Aged Adults with ASD. *Journal of Autism and Developmental Disorders*, *46*(8), 2666–2678.

170. Kovacevic, N., Abdi, H., Beaton, D., & McIntosh, A. R. (2013). Revisiting PLS Resampling: Comparing Significance Versus Reliability Across Range of Simulations. *New Perspectives in Partial Least Squares and Related Methods*, 159–170.

171. Kowalczyk, T., Pontious, A., Englund, C., Daza, R. A. M., Bedogni, F., Hodge, R., Attardo, A., Bell, C., Huttner, W. B., & Hevner, R. F. (2009). Intermediate Neuronal Progenitors (Basal Progenitors) Produce Pyramidal–Projection Neurons for All Layers of Cerebral Cortex. *Cerebral Cortex*, *19*(10), 2439–2450.

172. Krishnan, A., Williams, L. J., McIntosh, A. R., & Abdi, H. (2011). Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *NeuroImage*, *56*(2), 455–475.

173. Kushki, A., Cardy, R. E., Panahandeh, S., Malihi, M., Hammill, C., Brian, J., Iaboni, A., Taylor, M. J., Schachar, R., Crosbie, J., Arnold, P., Kelley, E., Ayub, M., Nicolson, R., Georgiades, S., Lerch, J. P., & Anagnostou, E. (2021). Cross-Diagnosis Structural Correlates of Autistic-Like Social Communication Differences. *Cerebral Cortex*, 31(11), 5067–5076. https://doi.org/10.1093/cercor/bhab142

174. Lai, M.-C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., Szatmari, P., & Ameis, S. H. (n.d.). Prevalence of Co-Occurring Mental Health Diagnoses in the Autism Population: A Systematic Review and Meta-Analysis. *SRN Electronic Journal*. https://doi.org/10.2139/ssrn.3310628

175. Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/gender differences and autism: setting the scene for future research. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(1), 11–24.

176. Lai, M.-C., Lombardo, M. V., Chakrabarti, B., Ecker, C., Sadek, S. A., Wheelwright, S. J., Murphy, D. G. M., Suckling, J., Bullmore, E. T., & Baron-Cohen, S. (2012). Individual differences in brain structure underpin empathizing–systemizing cognitive styles in male adults. *NeuroImage*, *61*(4), 1347–1354. https://doi.org/10.1016/j.neuroimage.2012.03.018

177. Lai, M.-C., Lombardo, M. V., Pasco, G., Ruigrok, A. N. V., Wheelwright, S. J., Sadek, S. A., Chakrabarti, B., MRC AIMS Consortium, & Baron-Cohen, S. (2011). A behavioral comparison of male and female adults with high functioning autism spectrum conditions. *PloS One*, *6*(6), e20835.

178. Lai, M.-C., Lombardo, M. V., Ruigrok, A. N., Chakrabarti, B., Auyeung, B., Szatmari, P., Happé, F., Baron-Cohen, S., & MRC AIMS Consortium. (2017). Quantifying and exploring camouflaging in men and women with autism. *Autism: The International Journal of Research and Practice*, *21*(6), 690–702.

179. Lai, M.-C., Lombardo, M. V., Suckling, J., Ruigrok, A. N. V., Chakrabarti, B., Ecker, C., Deoni, S. C. L., Craig, M. C., Murphy, D. G. M., Bullmore, E. T., Baron-Cohen, S., & MRC AIMS Consortium. (2013). Biological sex affects the neurobiology of autism. *Brain*,*136*(9), 2799–2815. https://doi.org/10.1093/brain/awt216

180. Le Gros Clark, W. E. (1945). *Deformation Patterns in the Cerebral Cortex*. Printed at the Oxford University Press by John Johnson.

181. Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*(6755), 788–791.

182. Lee, D., & Seung, H. S. (2000). Algorithms for non-negative matrix factorization. *Advances in Neural Information Processing Systems*, *13*. https://proceedings.neurips.cc/paper/1861-algorithms-for-non-negative-matrix-factorization

183. Lee, J. K., Lee, J.-M., Kim, J. S., Kim, I. Y., Evans, A. C., & Kim, S. I. (2006). A novel quantitative cross-validation of different cortical surface reconstruction algorithms using MRI phantom. *NeuroImage*, *31*(2), 572–584.

184. Lenroot, R. K., & Yeung, P. K. (2013). Heterogeneity within Autism Spectrum Disorders: What have We Learned from Neuroimaging Studies? *Frontiers in Human Neuroscience*, *7*, 733.

185. Lemon, J. M., Gargaro, B., Enticott, P. G., & Rinehart, N. J. (2011). Brief Report: Executive Functioning in Autism Spectrum Disorders: A Gender Comparison of Response Inhibition. *Journal of Autism and Developmental Disorders*, *41*(3), 352–356.

186. Lepage, C., Lewis, L., Jeun, S., Bermudez, P., Khalili-Mahani, N., Omidyegaheh, M., Zijdenbos, A., Vincent, R. D., Adalat, R., & Evans, A. C. (2017). Human MR evaluation of cortical thickness using CIVET v2. 1. *Organization for Human Brain Mapping*. https://archive.aievolution.com/2017/hbm1701/index.cfm?do=abs.viewAbs&abs=329

187. Lepage, C., Wagstyl, K., Jung, B., Seidlitz, J., Sponheim, C., Ungerleider, L., Wang, X., Evans, A. C., & Messinger, A. (2021). CIVET-Macaque: An automated pipeline for MRI-based cortical surface generation and cortical thickness in macaques. *NeuroImage*, *227*, 117622.

188. Lerch, J. P., & Evans, A. C. (2005). Cortical thickness analysis examined through power analysis and a population simulation. *NeuroImage*, *24*(1), 163–173.

189. Lerch, J. P., van der Kouwe, A. J. W., Raznahan, A., Paus, T., Johansen-Berg, H., Miller, K. L., Smith, S. M., Fischl, B., & Sotiropoulos, S. N. (2017). Studying neuroanatomy using MRI. *Nature Neuroscience*, *20*(3), 314–326.

190. Lerch, J. P., Worsley, K., Philip Shaw, W., Greenstein, D. K., Lenroot, R. K., Giedd, J., & Evans, A. C. (2006). Mapping anatomical correlations across cerebral cortex (MACACC) using cortical thickness from MRI. *NeuroImage, 31*(3), 993–1003. https://doi.org/10.1016/j.neuroimage.2006.01.042

191. Lerch, J. P., Yiu, A. P., Martinez-Canabal, A., Pekar, T., Bohbot, V. D., Frankland, P. W., Henkelman, R. M., Josselyn, S. A., & Sled, J. G. (2011). Maze training in mice induces MRI-detectable brain shape changes specific to the type of learning. *NeuroImage*, *54*(3), 2086–2095.

192. Leung, R. C., Pang, E. W., Brian, J. A., & Taylor, M. J. (2019). Happy and Angry Faces Elicit Atypical Neural Activation in Children With Autism Spectrum Disorder. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *4*(12), 1021–1030.

193. Levman, J., Takahashi, E., Forgeron, C., MacDonald, P., Stewart, N., Lim, A., & Martel, A. (2018). A Sorting Statistic with Application in Neurological Magnetic Resonance Imaging of Autism. *Journal of Healthcare Engineering*, *2018*, 8039075.

194. Levy, A., & Perry, A. (2011). Outcomes in adolescents and adults with autism: A review of the literature. *Research in Autism Spectrum Disorders*, *5*(4), 1271–1282.

195. Levy, D., Ronemus, M., Yamrom, B., Lee, Y.-H., Leotta, A., Kendall, J., Marks, S., Lakshmi, B., Pai, D., Ye, K., Buja, A., Krieger, A., Yoon, S., Troge, J., Rodgers, L., Iossifov, I., & Wigler, M. (2011). Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*, *70*(5), 886–897.

196. Li, D., Karnath, H.-O., & Xu, X. (2017). Candidate Biomarkers in Children with Autism Spectrum Disorder: A Review of MRI Studies. *Neuroscience Bulletin*, *33*(2), 219–237.

197. Li, G., Nie, J., Wang, L., Shi, F., Lyall, A. E., Lin, W., Gilmore, J. H., & Shen, D. (2014). Mapping longitudinal hemispheric structural asymmetries of the human cerebral cortex from birth to 2 years of age. *Cerebral Cortex* , *24*(5), 1289–1300.

198. Li, G., Wang, L., Shi, F., Lyall, A. E., Lin, W., Gilmore, J. H., & Shen, D. (2014). Mapping longitudinal development of local cortical gyrification in infants from birth to 2 years of age. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(12), 4228–4238.

199. Liao, W., Zhang, Z., Mantini, D., Xu, Q., Wang, Z., Chen, G., Jiao, Q., Zang, Y.-F., & Lu, G. (2013). Relationship between large-scale functional and structural covariance networks in idiopathic generalized epilepsy. *Brain Connectivity*, *3*(3), 240–254.

200. Libero, L. E., Schaer, M., Li, D. D., Amaral, D. G., & Nordahl, C. W. (2018). A Longitudinal Study of Local Gyrification Index in Young Boys With Autism Spectrum Disorder. *Cerebral Cortex*, *29*(6), 2575–2587.

201. Liu, K., Yao, S., Chen, K., Zhang, J., Yao, L., Li, K., Jin, Z., & Guo, X. (2017). Structural Brain Network Changes across the Adult Lifespan. *Frontiers in Aging Neuroscience*, *9*, 275.

202. Liu, M., Maiti, P., Thomopoulos, S., Zhu, A., Chai, Y., Kim, H., & Jahanshad, N. (2021). Style Transfer Using Generative Adversarial Networks for Multi-Site MRI

Harmonization. *Medical Image Computing and Computer-Assisted Intervention: MICCAI... International Conference on Medical Image Computing and Computer-Assisted Intervention*, *12903*, 313–322.

203. Liu, T., Nie, J., Tarokh, A., Guo, L., & Wong, S. T. C. (2008). Reconstruction of central cortical surface from brain MRI images: method and application. *NeuroImage*, *40*(3), 991–1002.
204. Long, J., Lu, F., Yang, S., Zhang, Q., Chen, X., Pang, Y., Wang, M., He, B., Liu, H., Duan, X., Chen, H., Ye, S., & Chen, H. (2022). Different functional connectivity optimal frequency in autism compared with healthy controls and the relationship with social communication deficits: Evidence from gene expression and behavior symptom analyses. *Human Brain Mapping*. https://doi.org/10.1002/hbm.26011
205. Lombardo, M. V., Lai, M.-C., & Baron-Cohen, S. (2019.). *Big data approaches to decomposing heterogeneity across the autism spectrum*. https://doi.org/10.1101/278788
206. Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(6), 466–474.
207. Lord, C., Charman, T., Havdahl, A., Carbone, P., Anagnostou, E., Boyd, B., Carr, T., de Vries, P. J., Dissanayake, C., Divan, G., Freitag, C. M., Gotelli, M. M., Kasari, C., Knapp, M., Mundy, P., Plank, A., Scahill, L., Servili, C., Shattuck, P., … McCauley, J. B. (2022). The Lancet Commission on the future of care and clinical research in autism. *The Lancet*, *399*(10321), 271–334.
208. Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism From 2 to 9 Years of Age. In *Archives of General Psychiatry* (Vol. 63, Issue 6, p. 694). https://doi.org/10.1001/archpsyc.63.6.694
209. Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685.
210. Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. In *ACM SIGGRAPH Computer Graphics* (Vol. 21, Issue 4, pp. 163–169). https://doi.org/10.1145/37402.37422

211. Luders, E., Thompson, P. M., Narr, K. L., Toga, A. W., Jancke, L., & Gaser, C. (2006). A curvature-based approach to estimate local gyrification on the cortical surface. *NeuroImage*, *29*(4), 1224–1230.

212. Luo, W.-L., & Nichols, T. E. (2003). Diagnosis and exploration of massively univariate neuroimaging models. *NeuroImage*, *19*(3), 1014–1032.

213. Lyall, A. E., Shi, F., Geng, X., Woolson, S., Li, G., Wang, L., Hamer, R. M., Shen, D., & Gilmore, J. H. (2015). Dynamic Development of Regional Cortical Thickness and Surface Area in Early Childhood. *Cerebral Cortex*, *25*(8), 2204–2212.

214. Lynam, D. R., Hoyle, R. H., & Newman, J. P. (2006). The perils of partialling: cautionary tales from aggression and psychopathy. *Assessment*, *13*(3), 328–341.

215. Lyttelton, O. C., Karama, S., Ad-Dab'bagh, Y., Zatorre, R. J., Carbonell, F., Worsley, K., & Evans, A. C. (2009). Positional and surface area asymmetry of the human cerebral cortex. *NeuroImage*, *46*(4), 895–903.

216. MacDonald, D., Kabani, N., Avis, D., & Evans, A. C. (2000). Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *NeuroImage*, *12*(3), 340–356.

217. Madan, C. R., & Kensinger, E. A. (2017). Test–retest reliability of brain morphology estimates. *Brain Informatics*, *4*(2), 107–121.

218. Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: a systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review*, *34*(1), 73–86.

219. Mahendiran, T., Dupuis, A., Crosbie, J., Georgiades, S., Kelley, E., Liu, X., Nicolson, R., Schachar, R., Anagnostou, E., & Brian, J. (2019). Sex Differences in Social Adaptive Function in Autism Spectrum Disorder and Attention-Deficit Hyperactivity Disorder. *Frontiers in Psychiatry / Frontiers Research Foundation*, *10*, 607.

220. Maintz, J. B., & Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis*, *2*(1), 1–36.

221. Mak-Fan, K. M., Taylor, M. J., Roberts, W., & Lerch, J. P. (2012). Measures of Cortical Grey Matter Structure and Development in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders,42*(3),419–427. https://doi.org/10.1007/s10803-011-1261-6

222. Mandy, W., Murin, M., & Skuse, D. (2015). *The Cognitive Profile in Autism Spectrum Disorders*,*180*, 34–45.

223. Marchetto, M. C., Belinson, H., Tian, Y., Freitas, B. C., Fu, C., Vadodaria, K., Beltrao-Braga, P., Trujillo, C. A., Mendes, A. P. D., Padmanabhan, K., Nunez, Y., Ou, J., Ghosh, H., Wright, R., Brennand, K., Pierce, K., Eichenfield, L., Pramparo, T., Eyler, L., … Muotri, A. R. (2017). Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. *Molecular Psychiatry*, *22*(6), 820–835.

224. Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., … Dosenbach, N. U. F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*, 654-660.

225. Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An Overview of Autism Spectrum Disorder, Heterogeneity and Treatment Options. *Neuroscience Bulletin*, *33*(2), 183–193.

226. Matson, J. L., & Smith, K. R. M. (2008). Current status of intensive behavioral interventions for young children with autism and PDD-NOS. *Research in Autism Spectrum Disorders*,*2*(1), 60–74. https://doi.org/10.1016/j.rasd.2007.03.003

227. May, T., Adesina, I., McGillivray, J., & Rinehart, N. J. (2019). Sex differences in neurodevelopmental disorders. In *Current Opinion in Neurology,32*(4),622–626. https://doi.org/10.1097/wco.0000000000000714

228. Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., … Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1412), 1293–1322.

229. McCauley, J. B., Pickles, A., Huerta, M., & Lord, C. (2020). Defining Positive Outcomes in More and Less Cognitively Able Autistic Adults. *Autism Research,13*(9), 1548–1560. https://doi.org/10.1002/aur.2359

230. McGovern, C. W., & Sigman, M. (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry*, *46*(4), 401–408. https://doi.org/10.1111/j.1469-7610.2004.00361.x

231. McIntosh, A. R., & Lobaugh, N. J. (2004). Partial least squares analysis of neuroimaging data: applications and advances. *NeuroImage*, *23 Suppl 1*, S250–S263.

232. McIntosh, A. R., & Mišić, B. (2013). Multivariate statistical analyses for neuroimaging data. *Annual Review of Psychology*, *64*, 499–525.

233. McKavanagh, R., Buckley, E., & Chance, S. A. (2015). Wider minicolumns in autism: a neural basis for altered processing? *Brain: A Journal of Neurology*, *138*(Pt 7), 2034–2045.

234. McPartland, J. C. (2016). Considerations in biomarker development for neurodevelopmental disorders. *Current Opinion in Neurology*, *29*(2), 118–122.

235. McRobbie, D. W., Moore, E. A., Graves, M. J., & Prince, M. R. (n.d.). *MRI from Picture to Proton*. Retrieved July 23, 2022, from https://www.academia.edu/download/33333282/2003_McRobbie_From_picture_to_proton.pdf

236. Mei, T., Llera, A., Floris, D. L., Forde, N. J., Tillmann, J., Durston, S., Moessnang, C., Banaschewski, T., Holt, R. J., Baron-Cohen, S., Rausch, A., Loth, E., Dell'Acqua, F., Charman, T., Murphy, D. G. M., Ecker, C., Beckmann, C. F., Buitelaar, J. K., & EU-AIMS LEAP group. (2020). Gray matter covariations and core symptoms of autism: the EU-AIMS Longitudinal European Autism Project. *Molecular Autism*, *11*(1), 86.

237. Menary, K., Collins, P. F., Porter, J. N., Muetzel, R., Olson, E. A., Kumar, V., Steinbach, M., Lim, K. O., & Luciana, M. (2013). Associations between cortical thickness and general intelligence in children, adolescents and young adults. *Intelligence*, *41*(5), 597–606.

238. Meng, X., Iraji, A., Fu, Z., Kochunov, P., Belger, A., Ford, J., McEwen, S., Mathalon, D. H., Mueller, B. A., Pearlson, G., Potkin, S. G., Preda, A., Turner, J., van Erp, T., Sui, J., & Calhoun, V. D. (2022). Multimodel order independent component analysis: A data-driven method for evaluating brain functional network connectivity within and between multiple spatial scales. *Brain Connectivity*, *12*(7), 617–628.

239. Mensen, V. T., Wierenga, L. M., van Dijk, S., Rijks, Y., Oranje, B., Mandl, R. C. W., & Durston, S. (2017). Development of cortical thickness and surface area in autism spectrum disorder. *NeuroImage. Clinical*, *13*, 215–222.

240. Mieres, A. C., Kirby, R. S., Armstrong, K. H., Murphy, T. K., & Grossman, L. (2012). Autism spectrum disorder: an emerging opportunity for physical therapy. *Pediatric*

Physical Therapy: The Official Publication of the Section on Pediatrics of the American Physical Therapy Association, *24*(1), 31–37.

241. Mišić, B., Doesburg, S. M., Fatima, Z., Vidal, J., Vakorin, V. A., Taylor, M. J., & McIntosh, A. R. (2015). Coordinated Information Generation and Mental Flexibility: Large-Scale Network Disruption in Children with Autism. *Cerebral Cortex*,*25*(9), 2815–2827. https://doi.org/10.1093/cercor/bhu082
242. Mitteroecker, P., Gunz, P., Neubauer, S., & Müller, G. (2012). How to Explore Morphological Integration in Human Evolution and Development? *Evolutionary Biology*,*39*(4), 536–553. https://doi.org/10.1007/s11692-012-9178-3
243. Molnár, Z., & Rockland, K. S. (2020). Chapter 5 - Cortical columns. In J. Rubenstein, P. Rakic, B. Chen, & K. Y. Kwan (Eds.), *Neural Circuit and Cognitive Development (Second Edition)* (pp. 103–126). Academic Press.
244. Moradi, E., Khundrakpam, B., Lewis, J. D., Evans, A. C., & Tohka, J. (2017). *Predicting symptom severity in autism spectrum disorder based on cortical thickness measures in agglomerative data*. https://doi.org/10.1101/039180
245. Morgan, B. R., Ibrahim, G. M., Vogan, V. M., Leung, R. C., Lee, W., & Taylor, M. J. (2019). Characterization of Autism Spectrum Disorder across the Age Span by Intrinsic Network Patterns. *Brain Topography*,32(3), 461–471. https://doi.org/10.1007/s10548-019-00697-w
246. Morgan, J. T., Chana, G., Pardo, C. A., Achim, C., Semendeferi, K., Buckwalter, J., Courchesne, E., & Everall, I. P. (2010). Microglial activation and increased microglial density observed in the dorsolateral prefrontal cortex in autism. *Biological Psychiatry*, *68*(4), 368–376.
247. Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain: A Journal of Neurology*, *120 (Pt 4)*, 701–722.
248. Mutlu, A. K., Schneider, M., Debbané, M., Badoud, D., Eliez, S., & Schaer, M. (2013). Sex differences in thickness, and folding developments throughout the cortex. *NeuroImage*, *82*, 200–207.
249. Nakamura, K., & Fisher, E. (2009). Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. *NeuroImage*, *44*(3), 769–776.

250. Nakua, H., Hawco, C., Forde, N. J., Jacobs, G. R., Joseph, M., Voineskos, A. N., Wheeler, A. L., Lai, M.-C., Szatmari, P., Kelley, E., Liu, X., Georgiades, S., Nicolson, R., Schachar, R., Crosbie, J., Anagnostou, E., Lerch, J. P., Arnold, P. D., & Ameis, S. H. (2022). Cortico-amygdalar connectivity and externalizing/internalizing behavior in children with neurodevelopmental disorders. *Brain Structure & Function*, *227*(6), 1963–1979.

251. Narr, K. L., Woods, R. P., Thompson, P. M., Szeszko, P., Robinson, D., Dimtcheva, T., Gurbani, M., Toga, A. W., & Bilder, R. M. (2007). Relationships between IQ and regional cortical gray matter thickness in healthy adults. *Cerebral Cortex* , *17*(9), 2163–2171.

252. Nordahl, C. W., Dierker, D., Mostafavi, I., Schumann, C. M., Rivera, S. M., Amaral, D. G., & Van Essen, D. C. (2007). Cortical folding abnormalities in autism revealed by surface-based morphometry. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(43), 11725–11735.

253. O'Toole, A. J., Jiang, F., Abdi, H., Pénard, N., Dunlop, J. P., & Parent, M. A. (2007). Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. *Journal of Cognitive Neuroscience*, *19*(11), 1735–1752.

254. Ofner, M., Coles, A., Decou, M. L., Do, M., Bienek, A., Snider, J., & Ugnat, A. (2018). *Autism spectrum disorder among children and youth in Canada 2018*. Public Health Agency of Canada Ottawa, ON.

255. Ozonoff, S., Iosif, A.-M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., Rogers, S. J., Rozga, A., Sangha, S., Sigman, M., Steinfeld, M. B., & Young, G. S. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(3), 256–266.e1–e2.

256. Pagnozzi, A. M., Conti, E., Calderoni, S., Fripp, J., & Rose, S. E. (2018). A systematic review of structural MRI biomarkers in autism spectrum disorder: A machine learning perspective. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, *71*, 68–82.

257. Pai, A., Shetty, R., & Chowdhury, Y. S. (2021). Magnetic Resonance Imaging Physics. In *StatPearls.* StatPearls Publishing.

258. Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M. J., Grant, M. D., Franz, C. E., Xian, H., Tsuang, M., Fischl, B., Seidman, L., Dale, A., & Kremen, W. S. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cerebral Cortex*, *19*(11), 2728–2735.

259. Papez, J. W. (1929). Book Review. *The New England Journal of Medicine*, *200*(10), 516–516.

260. Patel, M. (2022). *Analysis methods for the study of multimodal MRI and individual variability in brain structure and function*. https://escholarship.mcgill.ca/concern/theses/qf85nh307

261. Patel, R., Mackay, C. E., Jansen, M. G., Devenyi, G. A., O'Donoghue, M. C., Kivimäki, M., Singh-Manoux, A., Zsoldos, E., Ebmeier, K. P., Chakravarty, M. M., & Suri, S. (2022). Inter- and intra-individual variation in brain structural-cognition relationships in aging. *NeuroImage*, *257*, 119254.

262. Patel, R., Steele, C. J., Chen, A. G. X., Patel, S., Devenyi, G. A., Germann, J., Tardif, C. L., & Chakravarty, M. M. (2020). Investigating microstructural variation in the human hippocampus using non-negative matrix factorization. *NeuroImage*, *207*, 116348.

263. Peters, A., & Sethares, C. (1996). Myelinated axons and the pyramidal cell modules in monkey primary visual cortex. *The Journal of Comparative Neurology*, *365*(2), 232–255.

264. Pienaar, R., Fischl, B., Caviness, V., Makris, N., & Grant, P. E. (2008). A methodology for analyzing curvature in the developing brain from preterm to adult. *International Journal of Imaging Systems and Technology*, *18*(1), 42–68.

265. Piven, J., Berthier, M. L., Starkstein, S. E., Nehme, E., Pearlson, G., & Folstein, S. (1990). Magnetic resonance imaging evidence for a defect of cerebral cortical development in autism. *The American Journal of Psychiatry*, *147*(6), 734–739.

266. Pobbe, R. L. H., Pearson, B. L., Defensor, E. B., Bolivar, V. J., Young, W. S., 3rd, Lee, H.-J., Blanchard, D. C., & Blanchard, R. J. (2012). Oxytocin receptor knockout mice display deficits in the expression of autism-related behaviors. *Hormones and Behavior*, *61*(3), 436–444.

267. Poline, J.-B., & Brett, M. (2012). The general linear model and fMRI: does love last forever? *NeuroImage*, *62*(2), 871–880.

268. Pomponio, R., Erus, G., Habes, M., Doshi, J., Srinivasan, D., Mamourian, E., Bashyam, V., Nasrallah, I. M., Satterthwaite, T. D., Fan, Y., Launer, L. J., Masters, C. L., Maruff, P., Zhuo, C., Völzke, H., Johnson, S. C., Fripp, J., Koutsouleris, N., Wolf, D. H., … Davatzikos, C. (2020). Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *NeuroImage*, *208*, 116450.

269. Pontious, A., Kowalczyk, T., Englund, C., & Hevner, R. F. (2008). Role of intermediate progenitor cells in cerebral cortex development. *Developmental Neuroscience*, *30*(1-3), 24–32.

270. Posar, A., & Visconti, P. (2019). Long-term outcome of autism spectrum disorder. *Turk Pediatri Arsivi*, *54*(4), 207–212.

271. Postema, M. C., van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Filho, G. B., Calderoni, S., Calvo, R., Daly, E., Deruelle, C., Di Martino, A., Dinstein, I., Duran, F. L. S., Durston, S., Ecker, C., Ehrlich, S., Fair, D., Fedor, J., … Francks, C. (2019). Altered structural brain asymmetry in autism spectrum disorder in a study of 54 datasets. *Nature Communications*, *10*(1), 4958.

272. Pua, E. P. K., Thomson, P., Yang, J. Y.-M., Craig, J. M., Ball, G., & Seal, M. (2021). Individual Differences in Intrinsic Brain Networks Predict Symptom Severity in Autism Spectrum Disorders. *Cerebral Cortex*, *31*(1), 681–693.

273. Radua, J., Vieta, E., Shinohara, R., Kochunov, P., Quidé, Y., Green, M. J., Weickert, C. S., Weickert, T., Bruggemann, J., Kircher, T., Nenadić, I., Cairns, M. J., Seal, M., Schall, U., Henskens, F., Fullerton, J. M., Mowry, B., Pantelis, C., Lenroot, R., … ENIGMA Consortium collaborators. (2020). Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *NeuroImage*, *218*, 116956.

274. Rakic, P. (1988). Specification of cerebral cortical areas. *Science*, *241*(4862), 170–176.

275. Rakic, P. (1995). A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends in Neurosciences*, *18*(9), 383–388.

276. Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nature Reviews. Neuroscience*, *10*(10), 724–735.

277. Rao, V. S., Raman, V., & Mysore, A. V. (2015). Issues related to obtaining intelligence quotient-matched controls in autism research. *Indian Journal of Psychological Medicine*, *37*(2), 149–153.

278. Raznahan, A., Greenstein, D., Lee, N. R., Clasen, L. S., & Giedd, J. N. (2012). Prenatal growth in humans and postnatal brain maturation into late adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(28), 11366–11371.

279. Raznahan, A., Lerch, J. P., Lee, N., Greenstein, D., Wallace, G. L., Stockman, M., Clasen, L., Shaw, P. W., & Giedd, J. N. (2011). Patterns of coordinated anatomical change in human cortical development: a longitudinal neuroimaging study of maturational coupling. *Neuron*, *72*(5), 873–884.

280. Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G. L., Greenstein, D., Clasen, L., Gogtay, N., & Giedd, J. N. (2011). How does your cortex grow? *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(19), 7174–7177.

281. Remer, J., Croteau-Chonka, E., Dean, D. C., D'Arpino, S., Dirks, H., Whiley, D., & Deoni, S. C. L. (2017). Quantifying cortical development in typically developing toddlers and young children, 1–6 years of age. *NeuroImage*, *153*, 246–261.

282. Rice, D., & Barone, S., Jr. (2000). Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. *Environmental Health Perspectives*, *108 Suppl 3*, 511–533.

283. Richman, D. P., Stewart, R. M., Hutchinson, J. W., & Caviness, V. S., Jr. (1975). Mechanical model of brain convolutional development. *Science*, *189*(4196), 18–21.

284. Robbins, S. (2004). Tuning and comparing spatial normalization methods. *Medical Image Analysis,8(*3), 311–323). https://doi.org/10.1016/j.media.2004.06.009

285. Robert, C., Patel, R., Blostein, N., Steele, C. J., & Chakravarty, M. M. (2022). Analyses of microstructural variation in the human striatum using non-negative matrix factorization. *NeuroImage*, *246*, 118744.

286. Rogge, N., & Janssen, J. (2019). The Economic Costs of Autism Spectrum Disorder: A Literature Review. *Journal of Autism and Developmental Disorders*, *49*(7), 2873–2900.

287. Romero-Garcia, R., Warrier, V., Bullmore, E. T., Baron-Cohen, S., & Bethlehem, R. A. I. (2019). *Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism*. https://doi.org/10.1101/208223

288. Romero-Garcia, R., Whitaker, K. J., Váša, F., Seidlitz, J., Shinn, M., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., NSPN Consortium, Bullmore, E. T., & Vértes, P. E.

(2018). Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex. *NeuroImage*, *171*, 256–267.

289. Ronan, L., & Fletcher, P. C. (2015). From genes to folds: a review of cortical gyrification theory. *Brain Structure & Function*, *220*(5), 2475–2483.
290. Rossignol, D. A., & Frye, R. E. (2012). Mitochondrial dysfunction in autism spectrum disorders: a systematic review and meta-analysis. *Molecular Psychiatry*, *17*(3), 290–314.
291. Ruigrok, A. N. V., & Lai, M.-C. (2020). Sex/gender differences in neurology and psychiatry: Autism. *Handbook of Clinical Neurology*, *175*, 283–297.
292. Rylaarsdam, L., & Guemez-Gamboa, A. (2019). Genetic Causes and Modifiers of Autism Spectrum Disorder. *Frontiers in Cellular Neuroscience*, *13*, 385.
293. Ryland, H. K., Hysing, M., Posserud, M.-B., Gillberg, C., & Lundervold, A. J. (2014). Autistic features in school age children: IQ and gender effects in a population-based cohort. *Research in Autism Spectrum Disorders*, *8*(3), 266–274.
294. Safar, K., Vandewouw, M. M., Pang, E. W., de Villa, K., Crosbie, J., Schachar, R., Iaboni, A., Georgiades, S., Nicolson, R., Kelley, E., Ayub, M., Lerch, J. P., Anagnostou, E., & Taylor, M. J. (2022). Shared and Distinct Patterns of Functional Connectivity to Emotional Faces in Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder Children. *Frontiers in Psychology*, *13*, 826527.
295. Safar, K., Vandewouw, M. M., & Taylor, M. J. (2021). Atypical development of emotional face processing networks in autism spectrum disorder from childhood through to adulthood. *Developmental Cognitive Neuroscience,51,* 101003. https://doi.org/10.1016/j.dcn.2021.101003
296. Safar, K., Yuk, V., Wong, S. M., Leung, R. C., Anagnostou, E., & Taylor, M. J. (2020). Emotional face processing in autism spectrum disorder: Effects in gamma connectivity. *Biological Psychology*, *149*, 107774.
297. Sanfelici, R., Ruef, A., Antonucci, L. A., Penzel, N., Sotiras, A., Dong, M. S., Urquijo-Castro, M., Wenzel, J., Kambeitz-Ilankovic, L., Hettwer, M. D., Ruhrmann, S., Chisholm, K., Riecher-Rössler, A., Falkai, P., Pantelis, C., Salokangas, R. K. R., Lencer, R., Bertolino, A., Kambeitz, J., … PRONIA Consortium. (2022). Novel Gyrification Networks Reveal Links with Psychiatric Risk Factors in Early Illness. *Cerebral Cortex, 32*(8), 1625–1636.

298. Santos, S., Ferreira, H., Martins, J., Gonçalves, J., & Castelo-Branco, M. (2022). Male sex bias in early and late onset neurodevelopmental disorders: Shared aspects and differences in Autism Spectrum Disorder, Attention Deficit/hyperactivity Disorder, and Schizophrenia. *Neuroscience and Biobehavioral Reviews*, *135*, 104577.

299. Sato, W., Kochiyama, T., Uono, S., Yoshimura, S., Kubota, Y., Sawada, R., Sakihama, M., & Toichi, M. (2017). Reduced Gray Matter Volume in the Social Brain Network in Adults with Autism Spectrum Disorder. In *Frontiers in Human Neuroscience* (Vol. 11). https://doi.org/10.3389/fnhum.2017.00395

300. Schaer, M., Cuadra, M. B., Tamarit, L., Lazeyras, F., Eliez, S., & Thiran, J.-P. (2008). A surface-based approach to quantify local cortical gyrification. *IEEE Transactions on Medical Imaging*, *27*(2), 161–170.

301. Schaer, M., Kochalka, J., Padmanabhan, A., Supekar, K., & Menon, V. (2015). Sex differences in cortical volume and gyrification in autism. *Molecular Autism*, *6*, 42.

302. Schaer, M., Ottet, M.-C., Scariati, E., Dukes, D., Franchini, M., Eliez, S., & Glaser, B. (2013). Decreased frontal gyrification correlates with altered connectivity in children with autism. *Frontiers in Human Neuroscience*, *7*, 750.

303. Scheel, C., Rotarska-Jagiela, A., Schilbach, L., Lehnhardt, F. G., Krug, B., Vogeley, K., & Tepest, R. (2011). Imaging derived cortical thickness reduction in high-functioning autism: key regions and temporal slope. *NeuroImage*, *58*(2), 391–400.

304. Schumann, C. M., Bloss, C. S., Barnes, C. C., Wideman, G. M., Carper, R. A., Akshoomoff, N., Pierce, K., Hagler, D., Schork, N., Lord, C., & Courchesne, E. (2010). Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(12), 4419–4427.

305. Segovia, F., Holt, R., Spencer, M., GÃ³rriz, J. M., RamÃ-rez, J., Puntonet, C. G., Phillips, C., Chura, L., Baron-Cohen, S., & Suckling, J. (2014). Identifying endophenotypes of autism: a multivariate approach. *Frontiers in Computational Neuroscience,8*. https://doi.org/10.3389/fncom.2014.00060

306. Seidlitz, J., Váša, F., Shinn, M., Romero-Garcia, R., Whitaker, K. J., Vértes, P. E., Wagstyl, K., Kirkpatrick Reardon, P., Clasen, L., Liu, S., Messinger, A., Leopold, D. A., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., NSPN Consortium, Raznahan, A., & Bullmore, E. T. (2018). Morphometric Similarity Networks Detect Microscale Cortical

Organization and Predict Inter-Individual Cognitive Variation. *Neuron*, *97*(1), 231–247.e7.

307. Sekhon, J. S. (2011). Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R. *Journal of Statistical Software*, *42*(1), 1–52.
308. Sha, Z., van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Bernhardt, B., Bolte, S., Busatto, G. F., Calderoni, S., Calvo, R., Daly, E., Deruelle, C., Duan, M., Duran, F. L. S., Durston, S., Ecker, C., Ehrlich, S., Fair, D., … Francks, C. (2022). Subtly altered topological asymmetry of brain structural covariance networks in autism spectrum disorder across 43 datasets from the ENIGMA consortium. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-022-01452-7
309. Shan, X., Uddin, L. Q., Xiao, J., He, C., Ling, Z., Li, L., Huang, X., Chen, H., & Duan, X. (2022). Mapping the Heterogeneous Brain Structural Phenotype of Autism Spectrum Disorder Using the Normative Model. *Biological Psychiatry*, *91*(11), 967–976.
310. Sharda, M., Khundrakpam, B. S., Evans, A. C., & Singh, N. C. (2016). Disruption of structural covariance networks for language in autism is modulated by verbal ability. In *Brain Structure and Function*, *221*(2), 1017–1032. https://doi.org/10.1007/s00429-014-0953-z
311. Shattuck, D. W., & Leahy, R. M. (2001). Automated graph-based analysis and correction of cortical volume topology. *IEEE Transactions on Medical Imaging*, *20*(11), 1167–1177.
312. Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., Lounds, J., & Lord, C. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *37*(9), 1735–1747.
313. Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., Evans, A., Rapoport, J., & Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. *Nature*, *440*(7084), 676–679.
314. Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., Rapoport, J. L., Giedd, J. N., & Wise, S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(14), 3586–3594.

315. Shen, M. D., Kim, S. H., McKinstry, R. C., Gu, H., Hazlett, H. C., Nordahl, C. W., Emerson, R. W., Shaw, D., Elison, J. T., Swanson, M. R., Fonov, V. S., Gerig, G., Dager, S. R., Botteron, K. N., Paterson, S., Schultz, R. T., Evans, A. C., Estes, A. M., Zwaigenbaum, L., … Infant Brain Imaging Study Network. (2017). Increased Extra-axial Cerebrospinal Fluid in High-Risk Infants Who Later Develop Autism. *Biological Psychiatry*, *82*(3), 186–193.

316. Silbereis, J. C., Pochareddy, S., Zhu, Y., Li, M., & Sestan, N. (2016). The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. *Neuron*, *89*(2), 248–268.

317. Simon, H. D., Ding, C., He, X., & Jin, R. (2005). *On the Equivalence of Nonnegative Matrix Factorization and K-means- Spectral Clustering*. Lawrence Berkeley National Laboratory.

318. Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, *17*(1), 87–97.

319. Smith, E., Thurm, A., Greenstein, D., Farmer, C., Swedo, S., Giedd, J., & Raznahan, A. (2016). Cortical thickness change in autism during early childhood. *Human Brain Mapping*,*37*(7), 2616–2629. https://doi.org/10.1002/hbm.23195

320. Song, G., Han, J., Zhao, Y., Wang, Z., & Du, H. (2017). A Review on Medical Image Registration as an Optimization Problem. *Current Medical Imaging Reviews*, *13*(3), 274–283.

321. Sotiras, A., Resnick, S. M., & Davatzikos, C. (2015). Finding imaging patterns of structural covariance via Non-Negative Matrix Factorization. *NeuroImage*, *108*, 1–16.

322. Sotiras, A., Toledo, J. B., Gur, R. E., Gur, R. C., Satterthwaite, T. D., & Davatzikos, C. (2017). Patterns of coordinated cortical remodeling during adolescence and their associations with functional specialization and evolutionary expansion. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(13), 3527–3532.

323. Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(38), 8223–8231.

324. Stoner, R., Chow, M. L., Boyle, M. P., Sunkin, S. M., Mouton, P. R., Roy, S., Wynshaw-Boris, A., Colamarino, S. A., Lein, E. S., & Courchesne, E. (2014). Patches of disorganization in the neocortex of children with autism. *The New England Journal of Medicine*, *370*(13), 1209–1219.

325. Stoyanov, D., Kandilarova, S., Aryutova, K., Paunova, R., Todeva-Radneva, A., Latypova, A., & Kherif, F. (2020). Multivariate Analysis of Structural and Functional Neuroimaging Can Inform Psychiatric Differential Diagnosis. *Diagnostics (Basel, Switzerland)*, *11*(1). https://doi.org/10.3390/diagnostics11010019

326. Strike, L. T., Hansell, N. K., Couvy-Duchesne, B., Thompson, P. M., de Zubicaray, G. I., McMahon, K. L., & Wright, M. J. (2018). Genetic Complexity of Cortical Structure: Differences in Genetic and Environmental Factors Influencing Cortical Surface Area and Thickness. *Cerebral Cortex*, *29*(3), 952–962.

327. Stuart, E. A., King, G., Imai, K., & Ho, D. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*. https://dash.harvard.edu/handle/1/11130519

328. Sun, T., & Hevner, R. F. (2014). Growth and folding of the mammalian cerebral cortex: from molecules to malformations. *Nature Reviews. Neuroscience*, *15*(4), 217–232

329. Supekar, K., Iyer, T., & Menon, V. (2017). The influence of sex and age on prevalence rates of comorbid conditions in autism. *Autism Research: Official Journal of the International Society for Autism Research*, *10*(5), 778–789.

330. Sydnor, V. J., Larsen, B., Bassett, D. S., Alexander-Bloch, A., Fair, D. A., Liston, C., Mackey, A. P., Milham, M. P., Pines, A., Roalf, D. R., Seidlitz, J., Xu, T., Raznahan, A., & Satterthwaite, T. D. (2021). Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron*, *109*(18), 2820–2846.

331. Tallinen, T., Chung, J. Y., Biggins, J. S., & Mahadevan, L. (2014). Gyrification from constrained cortical expansion. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(35), 12667–12672.

332. Tang, L., Mostafa, S., Liao, B., & Wu, F.-X. (2019). A network clustering based feature selection strategy for classifying autism spectrum disorder. *BMC Medical Genomics*, *12*(Suppl 7), 153.

333. Tartaglia, N. R., Wilson, R., Miller, J. S., Rafalko, J., Cordeiro, L., Davis, S., Hessl, D., & Ross, J. (2017). Autism Spectrum Disorder in Males with Sex Chromosome Aneuploidy: XXY/Klinefelter Syndrome, XYY, and XXYY. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *38*(3), 197–207.

334. Tau, G. Z., & Peterson, B. S. (2010). Normal development of brain circuits. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 147–168.

335. Taylor, J. L., & Seltzer, M. M. (2010). Changes in the Autism Behavioral Phenotype During the Transition to Adulthood. *Journal of Autism and Developmental Disorders. 40*(12), 1431–1446). https://doi.org/10.1007/s10803-010-1005-z

336. Tohka, J., Zijdenbos, A., & Evans, A. (2004). Fast and robust parameter estimation for statistical partial volume models in brain MRI. *NeuroImage*, *23*(1), 84–97.

337. Toma, C., Hervás, A., Balmaña, N., Salgado, M., Maristany, M., Vilella, E., Aguilera, F., Orejuela, C., Cuscó, I., Gallastegui, F., Pérez-Jurado, L. A., Caballero-Andaluz, R., de Diego-Otero, Y., Guzmán-Alvarez, G., Ramos-Quiroga, J. A., Ribasés, M., Bayés, M., & Cormand, B. (2013). Neurotransmitter systems and neurotrophic factors in autism: association study of 37 genes suggests involvement of DDC. *The World Journal of Biological Psychiatry,14*(7), 516–527. https://doi.org/10.3109/15622975.2011.602719

338. Tordjman, S., Cohen, D., Anderson, G. M., Botbol, M., Canitano, R., Coulon, N., & Roubertoux, P. L. (2018). Repint of "Reframing autism as a behavioral syndrome and not a specific mental disorder: Implications of genetic and phenotypic heterogeneity." *Neuroscience & Biobehavioral Reviews*, *89*, 132–150. https://doi.org/10.1016/j.neubiorev.2018.01.014

339. Toro, R., & Burnod, Y. (2005). A morphogenetic model for the development of cortical convolutions. *Cerebral Cortex*, *15*(12), 1900–1913.

340. Toro, R., Perron, M., Pike, B., Richer, L., Veillette, S., Pausova, Z., & Paus, T. (2008). Brain size and folding of the human cerebral cortex. *Cerebral Cortex*, *18*(10), 2352–2357.

341. Tunç, B., Yankowitz, L. D., Parker, D., Alappatt, J. A., Pandey, J., Schultz, R. T., & Verma, R. (2019). Deviation from normative brain development is associated with symptom severity in autism spectrum disorder. *Molecular Autism*, *10*, 46.

342. Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320.

343. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289.

344. Van Essen, D. C. (1997). A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature*, *385*(6614), 313–318.

345. van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G. F., Calderoni, S., Daly, E., Deruelle, C., Di Martino, A., Dinstein, I., Duran, F. L. S., Durston, S., Ecker, C., Fair, D., Fedor, J., Fitzgerald, J., Freitag, C. M., Gallagher, L., … Buitelaar, J. K. (2018). Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: Results from the ENIGMA ASD working group. *The American Journal of Psychiatry*, *175*(4), 359–369.

346. van Wijngaarden-Cremers, P. J. M., van Eeten, E., Groen, W. B., Van Deurzen, P. A., Oosterling, I. J., & Van der Gaag, R. J. (2014). Gender and age differences in the core triad of impairments in autism spectrum disorders: a systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, *44*(3), 627–635.

347. Vandekar, S. N., Shinohara, R. T., Raznahan, A., Roalf, D. R., Ross, M., DeLeo, N., Ruparel, K., Verma, R., Wolf, D. H., Gur, R. C., Gur, R. E., & Satterthwaite, T. D. (2015). Topologically dissociable patterns of development of the human cerebral cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(2), 599–609.

348. Vargason, T., Grivas, G., Hollowood-Jones, K. L., & Hahn, J. (2020). Towards a Multivariate Biomarker-Based Diagnosis of Autism Spectrum Disorder: Review and Discussion of Recent Advancements. *Seminars in Pediatric Neurology*, *34*, 100803.

349. Varikuti, D. P., Genon, S., Sotiras, A., Schwender, H., Hoffstaedter, F., Patil, K. R., Jockwitz, C., Caspers, S., Moebus, S., Amunts, K., Davatzikos, C., & Eickhoff, S. B. (2018). Evaluation of non-negative matrix factorization of grey matter in age prediction. *NeuroImage,173*, 394–410. https://doi.org/10.1016/j.neuroimage.2018.03.007

350. Vismara, L. A., & Rogers, S. J. (2010). Behavioral treatments in autism spectrum disorder: what do we know? *Annual Review of Clinical Psychology*, *6*, 447–468.

351. Vogan, V. M., Morgan, B. R., Smith, M. L., & Taylor, M. J. (2019). Functional changes during visuo-spatial working memory in autism spectrum disorder: 2-year longitudinal functional magnetic resonance imaging study. *Autism: The International Journal of Research and Practice*, *23*(3), 639–652.

352. Vovk, U., Pernus, F., & Likar, B. (2007). A Review of Methods for Correction of Intensity Inhomogeneity in MRI. *IEEE Transactions on Medical Imaging*, *26*(3), 405–421.

353. Wagstyl, K., & Lerch, J. P. (2018). Cortical Thickness. In *Neuromethods* (pp. 35–49). https://doi.org/10.1007/978-1-4939-7647-8_3

354. Wallace, G. L., Dankner, N., Kenworthy, L., Giedd, J. N., & Martin, A. (2010). Age-related temporal and parietal cortical thinning in autism spectrum disorders. *Brain: A Journal of Neurology*, *133*(Pt 12), 3745–3754.

355. Wallace, G. L., Eisenberg, I. W., Robustelli, B., Dankner, N., Kenworthy, L., Giedd, J. N., & Martin, A. (2015). Longitudinal Cortical Development During Adolescence and Young Adulthood in Autism Spectrum Disorder: Increased Cortical Thinning but Comparable Surface Area Changes. *Journal of the American Academy of Child & Adolescent Psychiatry,54*(6), 464–469. https://doi.org/10.1016/j.jaac.2015.03.007

356. Wallace, G. L., Robustelli, B., Dankner, N., Kenworthy, L., Giedd, J. N., & Martin, A. (2013). Increased gyrification, but comparable surface area in adolescents with autism spectrum disorders. *Brain: A Journal of Neurology*, *136*(Pt 6), 1956–1967.

357. Wang, B., Mezlini, A. M., Demir, F., Fiume, M., Tu, Z., Brudno, M., Haibe-Kains, B., & Goldenberg, A. (2014). Similarity network fusion for aggregating data types on a genomic scale. *Nature Methods*, *11*(3), 333–337.

358. Wang, F., Lian, C., Wu, Z., Zhang, H., Li, T., Meng, Y., Wang, L., Lin, W., Shen, D., & Li, G. (2019). Developmental topography of cortical thickness during infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(32), 15855–15860.

359. Wang, M., Huang, T.-Z., Fang, J., Calhoun, V. D., & Wang, Y.-P. (2020). Integration of Imaging (epi)Genomics Data for the Study of Schizophrenia Using Group Sparse Joint Nonnegative Matrix Factorization. *IEEE/ACM Transactions on Computational*

Biology and Bioinformatics,17(5), 1671–1681. https://doi.org/10.1109/tcbb.2019.2899568

360. Wang, Y., Hu, D., Wu, Z., Wang, L., Huang, W., & Li, G. (2022). Developmental abnormalities of structural covariance networks of cortical thickness and surface area in autistic infants within the first 2 years. *Cerebral Cortex*, *32*(17), 3786–3798.
361. Waye, M. M. Y., & Cheng, H. Y. (2018). Genetics and epigenetics of autism: A Review. *Psychiatry and Clinical Neurosciences*, *72*(4), 228–244.
362. Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. *APA PsycTests*. https://doi.org/10.1037/t15171-000
363. Wechsler, D. (2003). Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV). *APA PsycTests*.
364. Wechsler, D. (2005). *Wechsler individual achievement test – Second edition (WIAT-II)*. San Antonio, TX: The Psychological Corporation
365. Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II)*. Pearson.
366. Wechsler,D. (2012). Wechsler abbreviated scale of intelligence. *APA PsycTeststst*. https://doi.org/10.1037/t15170-000
367. Wechsler, D. J. S. A. P. C. (2014). *Wechsler Intelligence Scale for Children–Fifth Edition (WISC-V)*. Bloomington, MN: Pearson.
368. Weiss, L. A., Arking, D. E., Gene Discovery Project of Johns Hopkins & the Autism Consortium, Daly, M. J., & Chakravarti, A. (2009). A genome-wide linkage and association scan reveals novel loci for autism. *Nature*, *461*(7265), 802–808.
369. Werling, D. M., & Geschwind, D. H. (2013). Sex differences in autism spectrum disorders. *Current Opinion in Neurology*, *26*(2), 146–153.
370. White, T., Su, S., Schmidt, M., Kao, C.-Y., & Sapiro, G. (2010). The development of gyrification in childhood and adolescence. *Brain and Cognition*, *72*(1), 36–45.
371. Wolff, N., Stroth, S., Kamp-Becker, I., Roepke, S., & Roessner, V. (2022). Autism Spectrum Disorder and IQ – A Complex Interplay. *Frontiers in Psychiatry*,*13*. https://doi.org/10.3389/fpsyt.2022.856084
372. Won, J. H., Youn, J., & Park, H. (2022). Enhanced neuroimaging genetics using multi-view non-negative matrix factorization with sparsity and prior knowledge. *Medical Image Analysis,77*, 102378. https://doi.org/10.1016/j.media.2022.102378

373. Wong, E., Palande, S., Wang, B., Zielinski, B., Anderson, J., & Thomas Fletcher, P. (2016). Kernel partial least squares regression for relating functional brain network topology to clinical measures of behavior. *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. https://doi.org/10.1109/isbi.2016.7493506

374. Woodman, A. C., Smith, L. E., Greenberg, J. S., & Mailick, M. R. (2015). Change in autism symptoms and maladaptive behaviors in adolescence and adulthood: the role of positive family processes. *Journal of Autism and Developmental Disorders*, *45*(1), 111–126.

375. Worsham, W., Dalton, S., & Bilder, D. A. (2021). The Prenatal Hormone Milieu in Autism Spectrum Disorder. *Frontiers in Psychiatry*,*12*. https://doi.org/10.3389/fpsyt.2021.655438

376. Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language*,*72*, 37–48). https://doi.org/10.1016/j.jml.2013.12.003

377. Yang, D. Y.-J., Beam, D., Pelphrey, K. A., Abdullahi, S., & Jou, R. J. (2016). Cortical morphological markers in children with autism: a structural magnetic resonance imaging study of thickness, area, volume, and gyrification. *Molecular Autism*, *7*, 11.

378. Yang, G., Pan, F., & Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature*, *462*(7275), 920–924.

379. Yang, Z., & Oja, E. (2010). Linear and Nonlinear Projective Nonnegative Matrix Factorization. *IEEE Transactions on Neural Networks*,*21*(5), 734–749. https://doi.org/10.1109/tnn.2010.2041361

380. Yee, Y., Fernandes, D. J., French, L., Ellegood, J., Cahill, L. S., Vousden, D. A., Spencer Noakes, L., Scholz, J., van Eede, M. C., Nieman, B. J., Sled, J. G., & Lerch, J. P. (2018). Structural covariance of brain region volumes is associated with both structural connectivity and transcriptomic similarity. *NeuroImage*, *179*, 357–372.

381. York, R. (2012). Residualization is not the answer: Rethinking how to address multicollinearity. *Social Science Research*, *41*(6), 1379–1386.

382. Yu, Q., Ouyang, A., Chalak, L., Jeon, T., Chia, J., Mishra, V., Sivarajan, M., Jackson, G., Rollins, N., Liu, S., & Huang, H. (2016). Structural Development of Human Fetal and Preterm Brain Cortical Plate Based on Population-Averaged Templates. *Cerebral Cortex*, *26*(11), 4381–4391.

383. Yuan, Z., & Oja, E. (2005). Projective Nonnegative Matrix Factorization for Image Compression and Feature Extraction. *Image Analysis*, 333–342.

384. Yuk, V., Urbain, C., Anagnostou, E., & Taylor, M. J. (2020). Frontoparietal Network Connectivity During an N-Back Task in Adults With Autism Spectrum Disorder. *Frontiers in Psychiatry / Frontiers Research Foundation*, *11*, 551808.

385. Zabihi, M., Oldehinkel, M., Wolfers, T., Frouin, V., Goyard, D., Loth, E., Charman, T., Tillmann, J., Banaschewski, T., Dumas, G., Holt, R., Baron-Cohen, S., Durston, S., Bölte, S., Murphy, D., Ecker, C., Buitelaar, J. K., Beckmann, C. F., & Marquand, A. F. (2019). Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using Normative Models. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *4*(6), 567–578.

386. Zeighami, Y., Fereshtehnejad, S.-M., Dadar, M., Louis Collins, D., Postuma, R. B., Mišić, B., & Dagher, A. (2019). A clinical-anatomical signature of Parkinson's disease identified with partial least squares and magnetic resonance imaging. *NeuroImage, 190*, 69–78. https://doi.org/10.1016/j.neuroimage.2017.12.050

387. Zhang, S., Liu, C.-C., Li, W., Shen, H., Laird, P. W., & Zhou, X. J. (2012). Discovery of multi-dimensional modules by integrative analysis of cancer genomic data. *Nucleic Acids Research,40*(19), 9379–9391. https://doi.org/10.1093/nar/gks725

388. Zhang, X., Feng, Y., Chen, W., Li, X., Faria, A. V., Feng, Q., & Mori, S. (2019). Linear Registration of Brain MRI Using Knowledge-Based Multiple Intermediator Libraries. *Frontiers in Neuroscience*, *13*, 909.

389. Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57

390. Zhou, H., Schaefer, G., & Shi, C. (2009). Fuzzy C-Means Techniques for Medical Image Segmentation. In Y. Jin & L. Wang (Eds.), *Fuzzy Systems in Bioinformatics and Computational Biology* (pp. 257–271). Springer Berlin Heidelberg.

391. Zhou, T., Kang, J., Cong, F., & Li, D. X. (2020a). Early childhood developmental functional connectivity of autistic brains with non-negative matrix factorization. *NeuroImage. Clinical*, *26*, 102251.

392. Zhou, T., Kang, J., Cong, F., & Li, X. (2020b). Stability-driven non-negative matrix factorization-based approach for extracting dynamic network from resting-state EEG. *Neurocomputing*, *389*, 123–131.

393. Zhu, H., Ibrahim, J. G., Tang, N., Rowe, D. B., Hao, X., Bansal, R., & Peterson, B. S. (2007). A statistical analysis of brain morphology using wild bootstrapping. *IEEE Transactions on Medical Imaging*, *26*(7), 954–966.

394. Zielinski, B. A., Anderson, J. S., Froehlich, A. L., Prigge, M. B. D., Nielsen, J. A., Cooperrider, J. R., Cariello, A. N., Thomas Fletcher, P., Alexander, A. L., Lange, N., Bigler, E. D., & Lainhart, J. E. (2012). scMRI Reveals Large-Scale Brain Network Abnormalities in Autism. *PLoS ONE*,*7*(11), e49172. https://doi.org/10.1371/journal.pone.0049172

395. Zielinski, B. A., Gennatas, E. D., Zhou, J., & Seeley, W. W. (2010). Network-level structural covariance in the developing brain. *Proceedings of the National Academy of Sciences*, *107*(42), 18191–18196.

396. Zielinski, B. A., Prigge, M. B. D., Nielsen, J. A., Froehlich, A. L., Abildskov, T. J., Anderson, J. S., Thomas Fletcher, P., Zygmunt, K. M., Travers, B. G., Lange, N., Alexander, A. L., Bigler, E. D., & Lainhart, J. E. (2014). Longitudinal changes in cortical thickness in autism and typical development. *Brain*,*137*(6),1799–1812. https://doi.org/10.1093/brain/awu083

397. Zijdenbos, A., Forghani, R., & Evans, A. (1998). Automatic quantification of MS lesions in 3D MRI brain data sets: Validation of INSECT. *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, 439–448.

398. Zilles, K., Armstrong, E., Schleicher, A., & Kretschmann, H. J. (1988). The human pattern of gyrification in the cerebral cortex. *Anatomy and Embryology*, *179*(2), 173–179.

399. Zilles, K., Schleicher, A., Langemann, C., Amunts, K., Morosan, P., Palomero-Gallagher, N., Schormann, T., Mohlberg, H., Bürgel, U., Steinmetz, H., Schlaug, G., & Roland, P. E. (1997). Quantitative analysis of sulci in the human cerebral cortex: development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture. *Human Brain Mapping*, *5*(4), 218–221.

400. Zwaigenbaum, L., Brian, J. A., & Ip, A. (2019). Early detection for autism spectrum disorder in young children. *Paediatrics & Child Health*, *24*(7), 424–443.

Appendix

Supplementary Table 1. Statistical differences between included and excluded subjects

Table S1. Statistical differences in demographic variables between subjects included and excluded after quality control (QC). Total before QC (n=1097); after QC (n=762). ASD = Autism Spectrum Disorder. NT = Neurotypical. FIQ = Fullscale IQ.

	Included	**Excluded**	**Test Statistic**
ASD	262	190	χ^2 (df =1) = 47.90 p<0.001
NT	500	145	
Male	428	242	χ^2 (df = 1) = 25.28 p<0.001
Female	334	93	
Median Age	16.42	11.00	*U*=72566, p<0.001
Median FIQ	111.00	108.00	*U*=70932, p<0.001

Supplementary Table 2. Demographic information of unmatched subjects

Table S2. Demographic information for unmatched subjects used for vertex-wise univariate analyses and CovBat validation analyses (n=668). ASD = Autism Spectrum Disorder. NT = Neurotypical. FIQ = Fullscale IQ.

Before Matching	**POND** N=137	**SickKids** N=320	**UK AIMS** N=211	**Total** N=668
Female_ASD	18	19	46	83
Male_ASD	46	60	54	160
Female_NT	36	117	45	198
Male_NT	37	124	66	227
ASD: NT	64:73	79:241	100:111	243:425
Age(years) Range [median]	4.37-21.96 [13.78]	4-49 [13]	18-52 [25.76]	4-52 [16.29]
FIQ Range [median]	44-142 [106]	69-149 [111]	73-137 [117]	44-149 [112]

Supplementary Table 3. Demographic information of matched subjects

Table S3. Demographic information for matched subjects used for multivariate analyses (n=486). ASD = Autism Spectrum Disorder. NT = Neurotypical. FIQ = Fullscale IQ.

After Matching	**POND** N=108	**SickKids** N=196	**UK AIMS** N=182	**Total** N=486
Female_ASD	18	19	46	83
Male_ASD	46	60	54	160
Female_NT	13	36	34	83
Male_NT	31	81	48	160
ASD: NT	64:44	79:117	100:82	243:243
Age(years) Range [median]	4.47-21.9[13.58]	4-49[12]	18-52[25]	4-52[18.42]
FIQ Range [median]	44-142[105]	69-144[111]	73-137[116]	44-144[111.5]

Supplementary Table 4. Additional information for subjects used in LV1 qualitative characterization

Table S4. Additional clinical and cognitive measures for matched subjects included in qualitative characterization of LV1 (Figure 4.11). NA = Not Available.

Assessments	**UK AIMS** N=182	**POND** N=108	**Total** N=290
ADOS_G_Total Range [median]	2-22 [10] N=97 ASD	-	2-22 [10] N=97 ASD
Verbal IQ Range [median]	71-144 [114] N=182	43-134 [101] N=104 NA=4	43-144 [109] N=286
Perceptual IQ Range [median]	67-138 [115] N=182	46-160 [106] N=104 NA=4	46-160 [112] N=286
Social Communication Questionnaire Total Score Range [median]	-	0-34 [8] N=113 NA=3	0-34 [8] N=113 NA=3
Repetitive Behaviour Scale Total Score Range [median]	-	0-85 [9] N=124 NA=4	0-85 [9] N=124 NA=4
Short Sensory Profile Total Score Range [median]	-	75-190 [151] N=93 NA=155	75-190 [151] N=93 NA=155

Supplementary Figure 1. Vertex-wise general linear model results

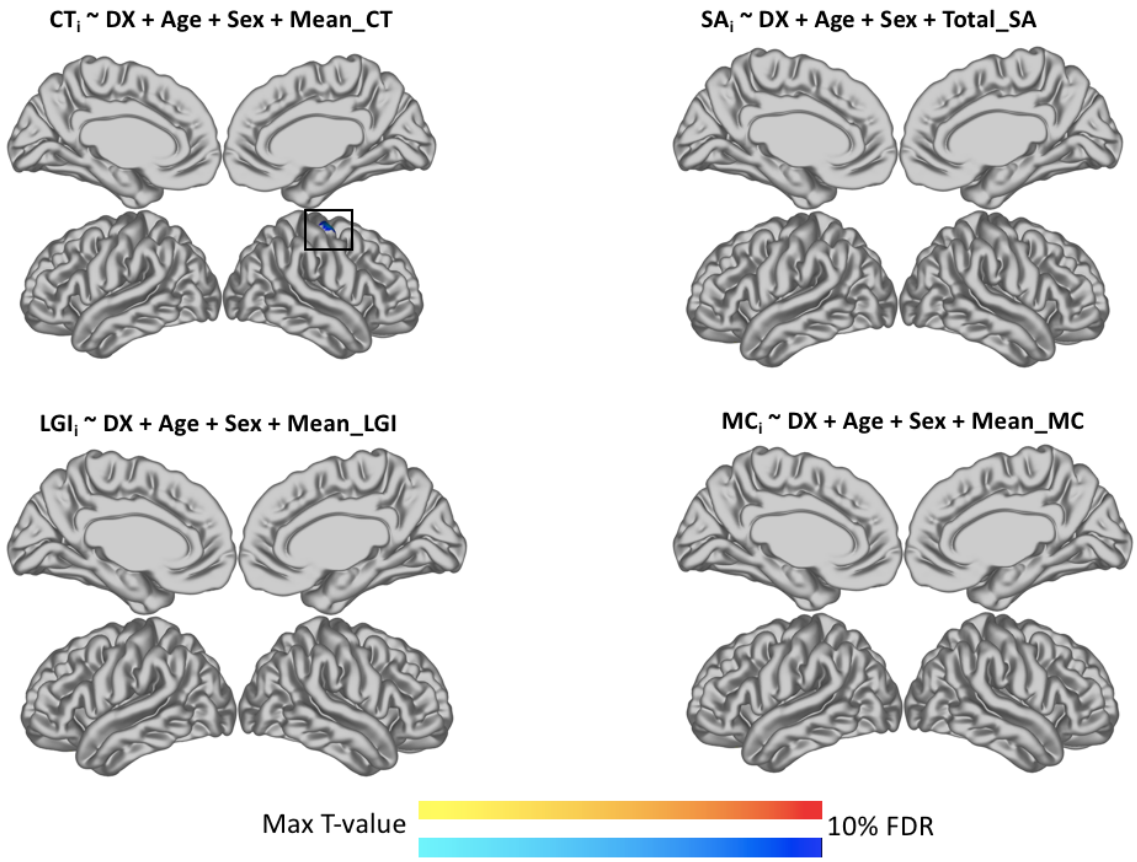


Figure S1. Results of vertex-wise linear models for subjects who passed QC and have complete diagnosis, age, sex, and FIQ data (n=668), projected onto population surface averages. Corresponding statistical models are shown above each brain map. CovBat-corrected values were used as input. Multiple comparison correction was performed with FDR. T-values of the 'diagnosis' predictor term are shown for vertices that remained significant at the 10% FDR threshold. A significant CT decrease was found in the right precentral gyrus of ASD subjects relative to NT controls. No other models yielded significant diagnostic differences at this threshold. DX = Diagnosis. ASD = Autism Spectrum Disorder. FDR = False Discovery Rate. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature.

Supplementary Figure 2. Vertex-wise standard deviation of raw measures

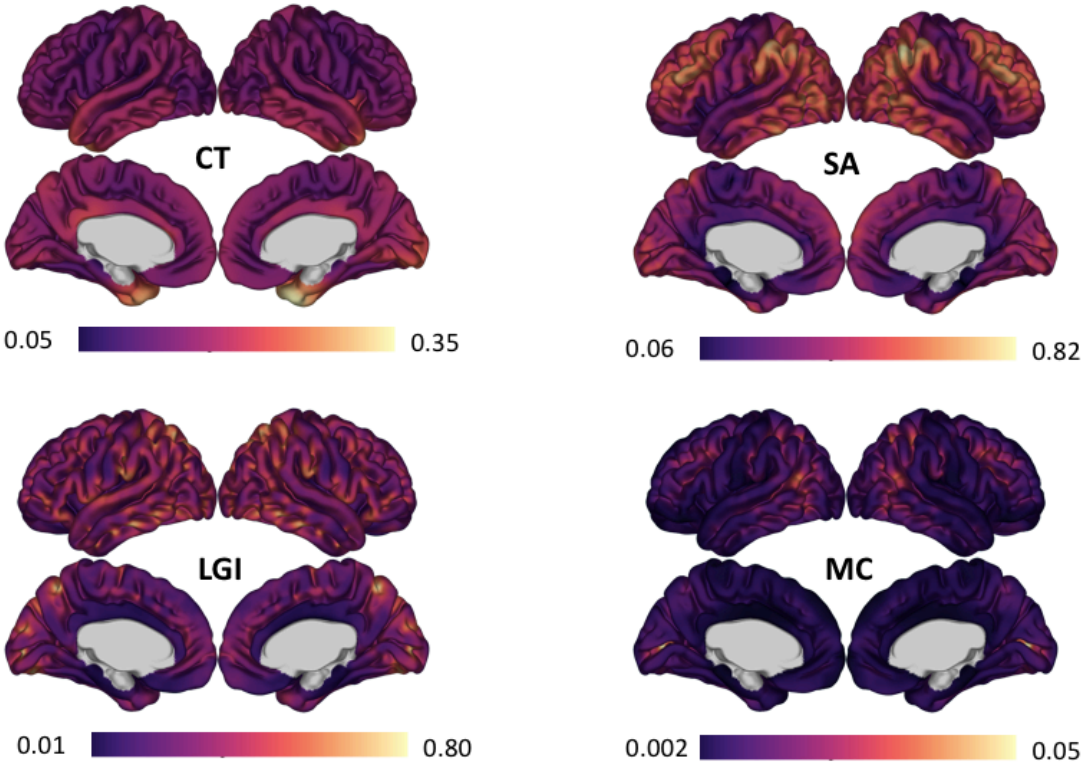


Figure S2. Standard deviation of raw vertex-wise cortical metrics for n=668 subjects. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature.

Supplementary Figure 3. Correlation of NMF subject weights with global brain measures: Non-GMV-residualized input measures

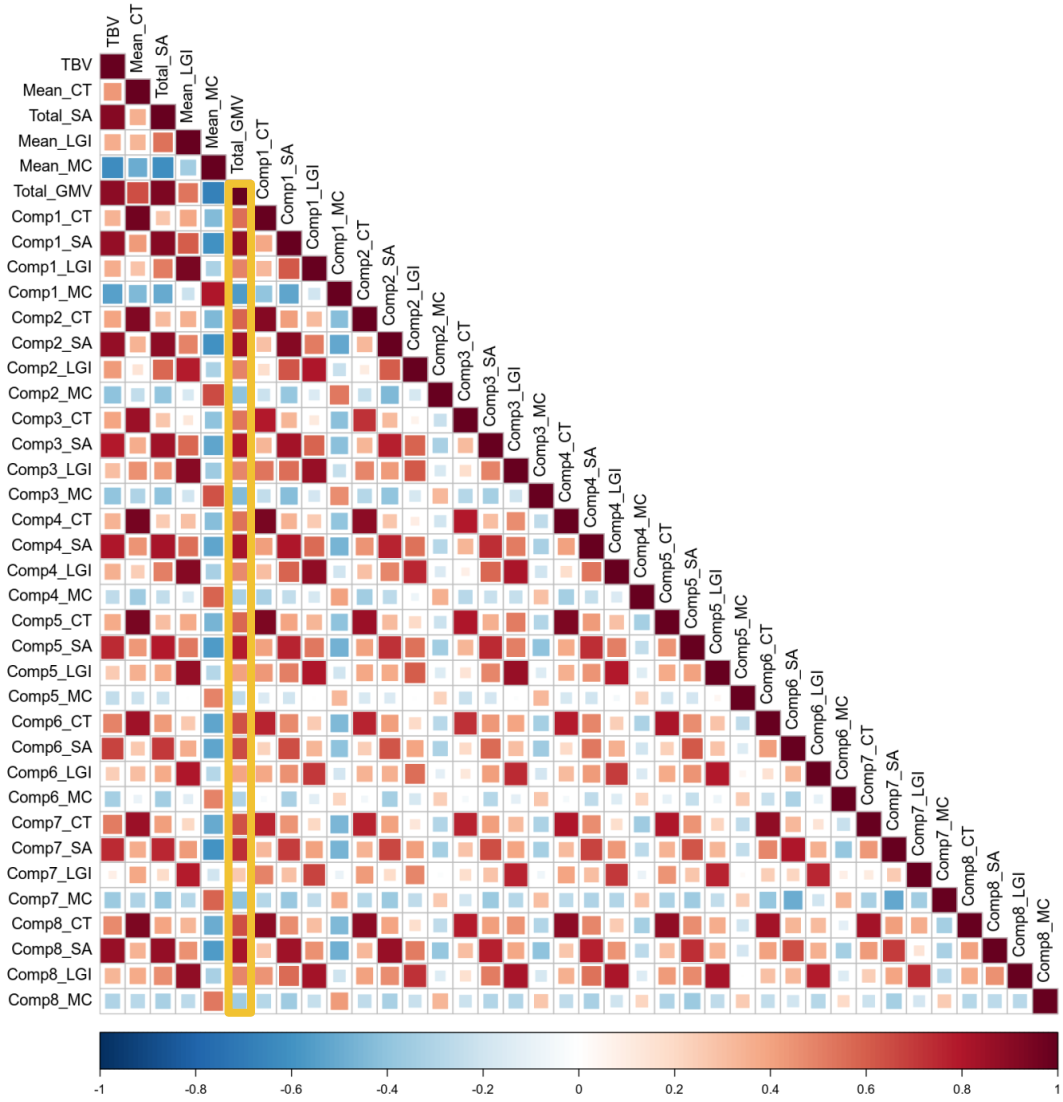
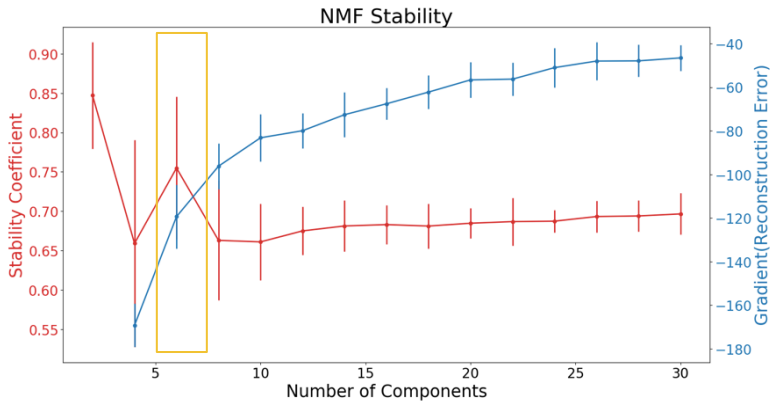


Figure S3. NMF was performed on CovBat-corrected data of unmatched subjects, without residualizing any global measures. An eight component decomposition was chosen following a stability analysis. H matrix component-by-metric weights are correlated with global measures (TBV, mean CT, total SA, mean LGI, mean MC, GMV, TBV). GMV displayed the strongest correlation with NMF H weights across components.

Supplementary Figure 4. Multivariate results for unmatched subjects

A.



B. i) **ii)**

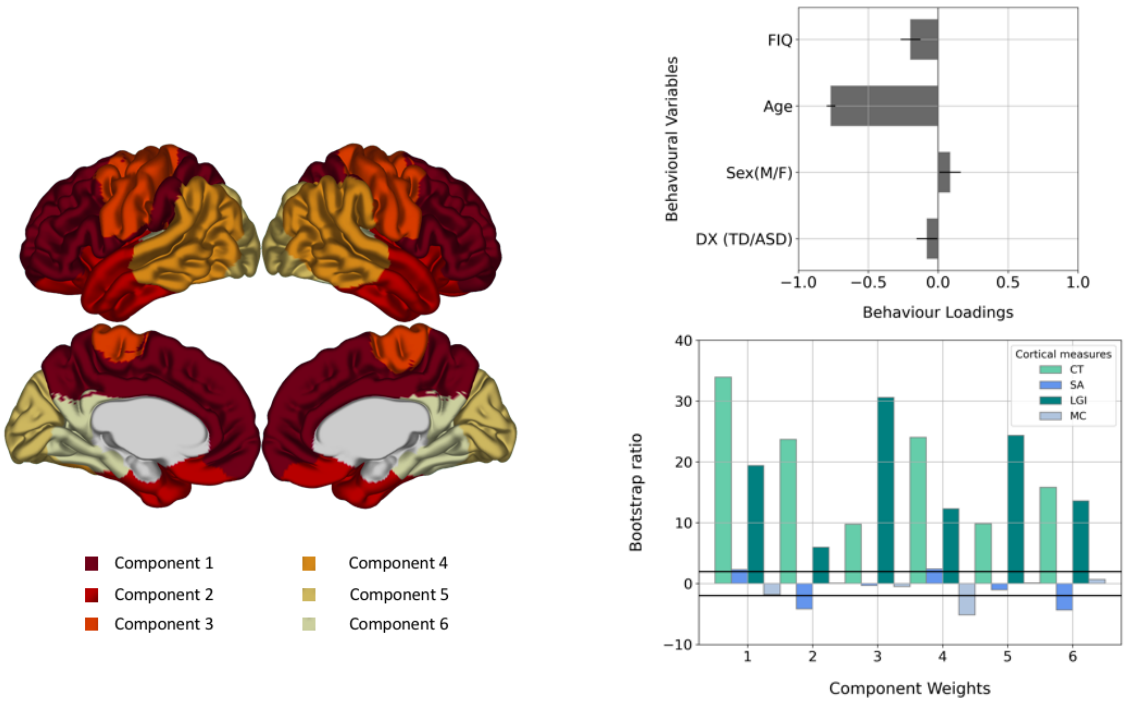
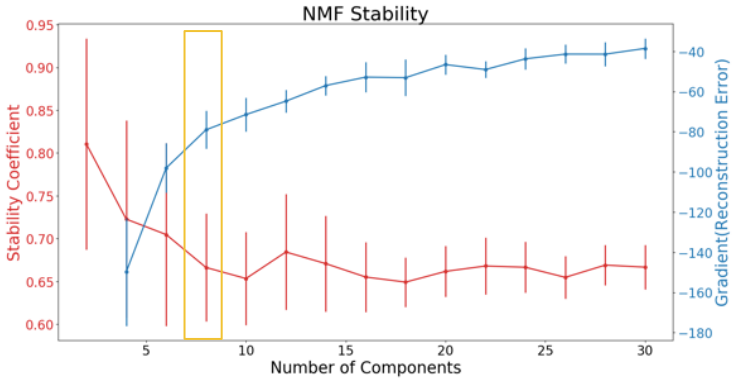


Figure S4. Results of Non-negative Matrix Factorization (NMF) and Partial Least Squares Correlation (PLSC) analyses using vertex-wise input matrices of CovBat-corrected and GMV-residualized data for n=668 subjects who passed QC and had available diagnosis, age, sex, and FIQ data. **A.** Results of split-half stability analysis revealed that a six component decomposition achieved optimal balance of generalizability and accuracy. **B. i)** Spatial map of six-component NMF decomposition, projected onto a population surface average. Each vertex of the W matrix was assigned to the highest weighted component in a winner-take-all fashion. **ii)**. PLSC results for latent variable 1 (LV) explain 91.4% covariance ($p<0.0001$). Barplots (top) depict behavioural variables contributing to LV1 as correlation coefficients, with significant contributions shown in dark grey. Barplots (bottom) show bootstrap ratios (BSRs) of cortical measures for each component (x-axis), such that a BSR of 2.58 is used to threshold significance. M= Male. F= Female. DX = Diagnosis. TD = Typically Developing. ASD = Autism Spectrum Disorder. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature.

Supplementary Figure 5. Multivariate results for unmatched subjects with non-GMV-residualized data

A.



B. i) **ii)**

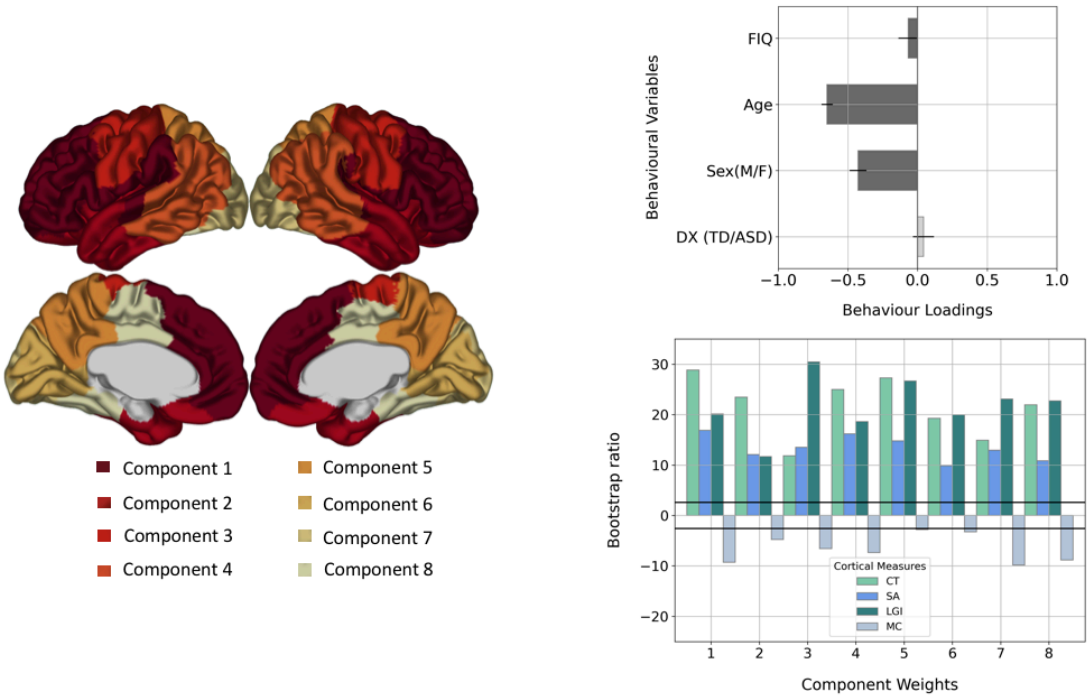
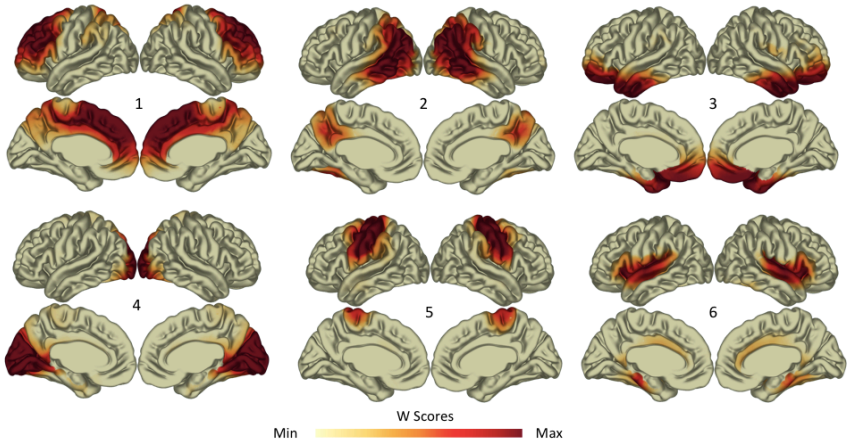


Figure S5. Results of Non-negative Matrix Factorization (NMF) and Partial Least Squares Correlation (PLSC) analyses using vertex-wise input matrices of CovBat-corrected and non-GMV-residualized data for n=668 subjects who passed QC and had available diagnosis, age, sex, and FIQ data. **A.** Results of split-half stability analysis revealed that an eight component decomposition achieved optimal balance of generalizability and accuracy. **B. i)** Spatial map of eight component NMF decomposition, projected onto a population surface average. Each vertex of the W matrix was assigned to the highest weighted component in a winner-take-all fashion. **ii)**. PLSC results for latent variable 1 (LV) explaining 85.6% covariance ($p<0.0001$). Barplots (top) depict behavioural variables contributing to LV1 as correlation coefficients, with significant contributions shown in dark grey. Barplots (bottom) show bootstrap ratios (BSRs) of cortical measures for each component (x-axis), such that a BSR of 2.58 is used to threshold significance. M= Male. F= Female. DX = Diagnosis. TD = Typically Developing. ASD = Autism Spectrum Disorder. CT = Cortical Thickness. SA = Surface Area.LGI = Local Gyrification Index. MC = Mean Curvature.

Supplementary Figure 6. Six component NMF decomposition on matched, GMV-residualized data

A.



B.

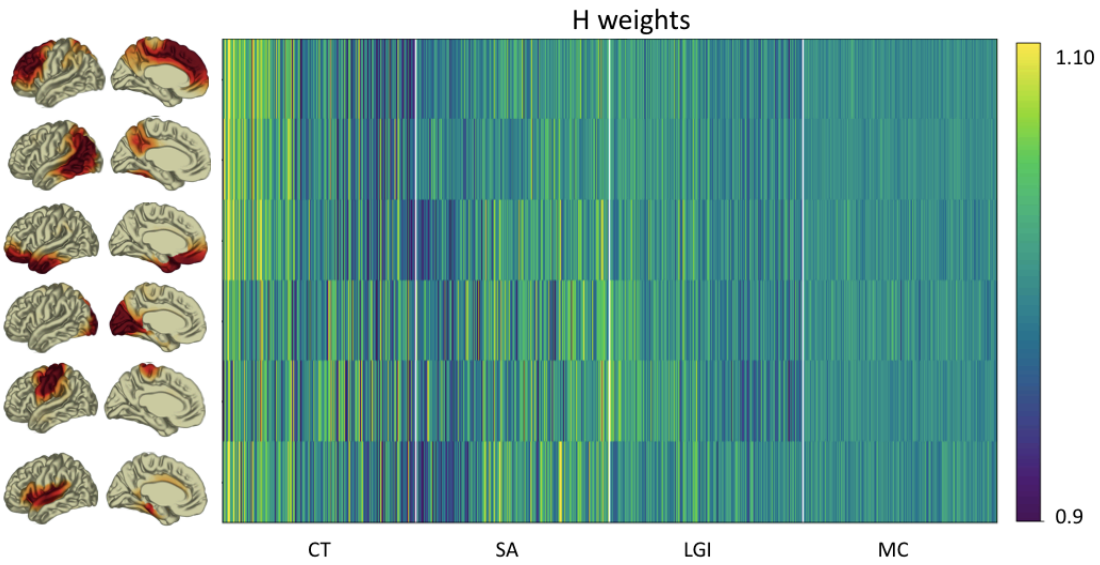


Figure S6. A. Spatial maps for a six component decomposition showing covariance patterns of the four cortical measures overlaid on population surface averages. **B.** Brighter colours indicate a higher participant-by-metric loading onto a given component's vertex, suggesting greater magnitude of that metric within the subject's covariance pattern. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature.

Supplementary Figure 7. Six Component PLSC Results for LV1

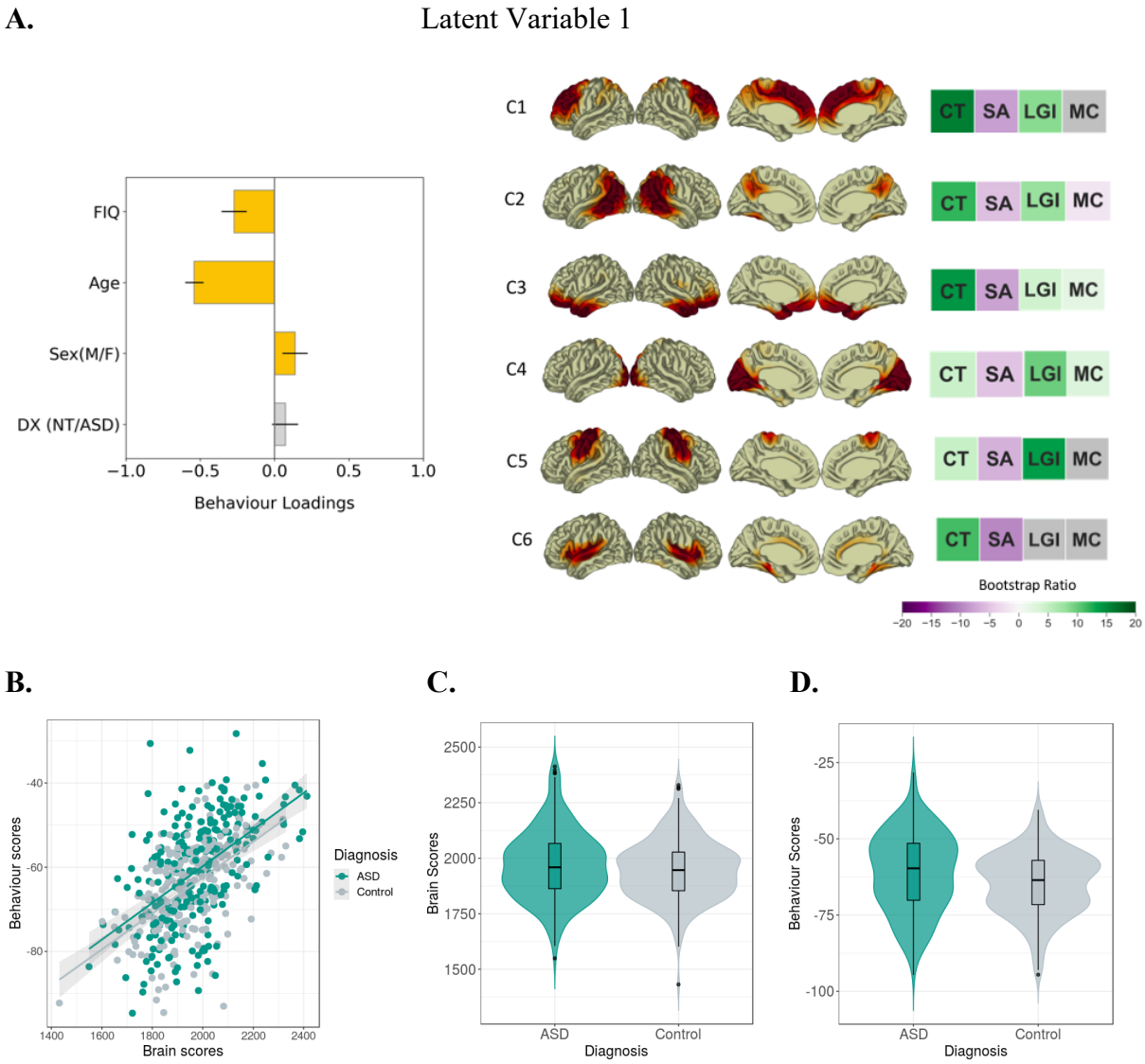


Figure S7. Partial least squares correlation (PLSC) results showing significant latent variables (LV) that capture the relationship between subjects' demographic characteristics (barplots) and NMF covariance weightings (brain heat maps). **A**. Results of LV1. Barplots show each behavioural measure's contribution to LV, expressed as a correlation coefficient. Error bars display bootstrapped 95% confidence intervals (CI). Orange coloured bars are indicative of measures that contribute significantly to the LV. Brain maps on the right depict the six component NMF solution, with adjacent heat maps showing bootstrap ratios (BSR) of the NMF component-by-metric loadings. A BSR of 1.96 (corresponding to a 95% CI) is used to threshold significance. Coloured squares (non-grey) indicate cortical measures within each covariance pattern that contribute significantly to the LV. Positive BSR are displayed in green and negative in purple. **B.** Correlation of subjects' brain and behaviour coloured by diagnostic group ($r=0.54$, $p<0.0001$). **C.** Brain scores grouped by diagnosis. **D.** Behaviour scores grouped by diagnosis. LV = Latent Variable. FIQ = Fullscale IQ. M = Male. F = Female. DX = Diagnosis. NT = Neurotypical. ASD = Autism Spectrum Disorder. C = Component. CT = Cortical Thickness. SA = Surface Area. LGI = Local Gyrification Index. MC = Mean Curvature.

Supplementary Table 5. 16 Component PLSC Bootstrap Ratios

Table S5. Bootstrap ratios (BSR) are shown for the dominant metric within each component. Positive BSR values indicate increase, whereas negative BSR values indicate decrease of a given metric within the covariance pattern captured by the component.

Component	**Anatomical Descriptor**	**Dominant Metric**	**BSR value**
1	Dorsal sensorimotor	LGI	14.63
2	Posterior frontal	CT	13.16
3	Orbitofrontal	CT	16.86
4	Ventral sensorimotor & superior temporal gyrus	CT	8.48
5	Anterior temporal	CT	11.33
6	Posterior temporal & inferior parietal	CT	12.33
7	Pars triangularis/opercularis, and lateral sulcus	CT	13.37
8	Cuneus	LGI	8.94
9	Dorsolateral prefrontal cortex & medial frontal gyrus	CT	18.65
10	Superior parietal	LGI	11.81
11	Ventral occipital	LGI	8.86
12	Anterior cingulate	CT	12.26
13	Parieto-occipital fissure	LGI	10.11
14	Precuneus	CT	13.11
15	Lingual gyrus	LGI	5.77
16	Parahippocampal gyrus	SA	-10.84