Characterizing Subcortical Morphometric Differences and Heterogeneity in Autism Spectrum Disorder

David MacDonald

Integrated Program in Neuroscience

McGill University, Montreal

January, 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science (M.Sc.)

Table of Contents

Abstract

Autism Spectrum Disorder (ASD) presents with significant heterogeneity, which obscures our understanding of the neuroanatomical correlates of ASD. This may be why neuroimaging studies of the subcortical structures of the thalamus, globus pallidus and striatum have produced inconsistent and contradictory results. These structures are critical mediators of functions known to be affected in ASD, including sensory gating and motor function.

We examined both volumetric and fine-grained localized shape differences in ASD using a large (n=3145), cross-sectional dataset of T1-weighted structural MRI scans from 33 sites, including both males and females, while investigating three potentially important sources of neuroanatomical heterogeneity: sex, age, and intelligence quotient (IQ). To minimize the impact of site effects, all scans underwent rigorous motion and segmentation quality control and identical processing and segmentation using MAGeTBrain. All data was processed in site-wise batches. Linear models, including sex, age, IQ, and total brain volume as covariates, were computed per-site and then combined across sites using random-effects meta-analysis. To assess the importance of age, sex, and IQ, nested models were compared using the Akaike Information Criterion (AIC).

We observed no differences in thalamic, pallidal, or striatal volume in ASD ($p > .26$, | Hedges' g^*| < 0.05 for all structures). Including age, but not sex, improved the fit for both the pallidum and striatum, but not for the thalamus. Repeating the analysis using best-fitting models did not alter the outcome, nor did conducting IQ- and age-centered analyses (age: $p > 0.29$ for all structures, IQ: $p > 0.21$), or sex-stratified analysis ($p > 0.22$). However, we did find a variety of localized shape differences in all three structures which were not apparent in the volumetric analysis. Also, age was an important variable across more than half of most surfaces. Age-

centered shape analysis indicated a variety of age-dependent regional differences. We found no correlations between symptom severity and neuroanatomical alterations in ASD, though this portion of the study was significantly underpowered compared to other analyses.

Overall, our findings help confirm that the neurodevelopment of the striatum, globus pallidus and thalamus are affected in ASD, in a subtle location-dependent manner that is not reflected in overall structure volumes, and that is highly non-uniform across the lifespan.

Résumé

Les troubles du spectre de l'autisme (TSA) présentent une hétérogénéité significantive, laquelle voile notre compréhension des corrélats neuroanatomiques des TSA. Ceci pourrait expliquer pourquoi les études de neuroimagerie des structures sous-corticales du thalamus, globus pallidus et le striatum produisent des résultats inconsistants et contradictoires. Ces structures sont des médiatrices critiques de fonctions connus pour être affectées dans les TSA, incluant le déclenchement sensoriel et la fonction motrice.

Nous avons examinés les différences de volumétrie et de forme localisée à grain fin dans les TSA en utilisant une large base de données (n=3145) transversale comportant des scans structurels IRM de séquences pondérés T1 provenant de 33 sites, incluant des hommes et des femmes, tout en enquêtant sur trois potentiellement importantes sources d'hétérogénétié neuroanatomique: le sexe, l'âge et le quotient intellectuel (QI).

Afin de réduire l'impact de l'effet de site autant que possible, tous les scans ont été soumis a un rigoureux contrôle de qualité du mouvement et de la segmentation et soumis à un processus et une segmentation identique en utilisant MAGeTBrain. Toutes les données ont été transformées en lots par site. Des modèles linéaires, incluant le sex, l'âge, le QI et le volume cébébral total comme covariables, ont été caclulés par site et ensuite combinés à travers les sites selon une méta-analyse à effets aléatoires. Afin d'évaluer l'importance de l'âge, du sexe et du QI, des modèles imbriqués ont été comparés selon le Critère d'Information d'Akaike (CIA).

Nous n'avons observés aucunes différences dans les volumes thalamique, pallidal ou striatal dans les TSA ($p > .26$, |Hedges' $g^*| < 0.05$ pour toutes les structures). En incluant l'âge, mais non le sex, une amélioration de la correspondance a été observée pour le pallidum et le striatum, mais non pour le thalamus. Une répétition de l'analyse en utilisant les modèles les

mieux adaptés n’a pas changé la résultante, ni les analyses conduitent sur les mesures centrées sur le QI et l’âge (âge: $p > 0.29$ pour toutes les structures, QI: $p > 0.21$), ou les analyses stratifiées selon le sexe ($p > 0.22$). Toutefois, nous avons trouvé une variété de différences de forme localisée dans les trois structures, qui n’étaient pas apparentes dans l’analyse volumétrique. De plus, l’âge s’avère être une variable importante pour plus que la moitié de la surface pour la plupart des structures. Les analyses de forme centrées sur l’âge ont indiqués une variété de différences régionales dépendantes de l’âge. Nous n’avons pas trouvé de corrélation entre la sévérité des symptômes et les altérations neuroanatomiques dans les TSA, bien qu’une partie de cette étude manque significativement de puissance statistique comparée aux autres analyses.

Dans l’ensemble, nos résultats aident à confirmer que le neurodéveloppement du thalamus, globus pallidus et du striatum sont impactés d'une manière subtile dépendant de l'emplacement dans les TSA. Par ailleurs, ces variations ne se reflètent pas dans les volumes globaux de la structure et ne sont pas vraiment pas uniformes tout au long de la durée de vie.

Acknowledgements

I have to begin by expressing my sincere gratitude to my supervisor, Dr. Mallar Chakravarty, for taking a chance on me as a graduate student, for inviting me into what is a truly remarkable research organization in the CoBrA Lab, and for his unfailing support and patience through no end of unforeseen challenges, including an international pandemic.

I am grateful to all of my present and former colleagues at the CoBrA Lab. Special thanks are due to Saashi Bedford, who got me started in the lab, and who paved the way for this project by assembling the data and developing and refining many of the techniques we used. Thanks to Gabe Devenyi, who played a key role in developing many of the tools used in this project and for his unfailing support throughout. To Rai, who helped ease my transition into the lab, and who was always eager to talk about neuroscience. To Christina, who spent endless hours staring at segmented brains to help with quality control. To Aurelie, Lani, Vanessa, and Justine, who were always supportive, helpful, and friendly, and who made it fun to come into the lab even when it was dark and snowy. Thanks are due also to Louis Theroux at the Douglas Mental Health University Institute, whose deft handling of any and all bureaucratic challenges that came up allowed me to focus on the work, and whose good humour and encyclopedic knowledge of history always made lunch time more interesting.

I would also like to thank my committee members, Drs. Boris Bernhardt, Sarah Lippé, and Marie Brossard-Racine for their support, their insight, and their always helpful feedback. I am also grateful to the National Science and Engineering Research Council (NSERC) and the Fonds de recherche du Québec - Santé (FRQS), who provided funding throughout my Master's project.

And of course, to my friends and family who supported me through this work: my brothers AJ and Dan, my parents Norm and Sherry, and all of my friends who were there for me through both the ups and downs that come with taking on a project like this.

Contribution of Authors

This thesis represents work that is original, unpublished, independent work completed by the author, David MacDonald. Data processing and analysis, except where indicated below, were completed by the author, as has been all writing and figure generation. Preliminary results from this project were presented at the Québec Bio-Imaging Network Scientific Day in June, 2022, and at the Organization for Human Brain Mapping annual conference in Glasgow, UK, also in June, 2022.

Saashi Bedford assembled the dataset from public and non-public sources, performed pre-processing, led the quality control process for motion, performed the CIVET processing from which Total Brain Volumes were derived, and used the meta-analytic technique in her own study at the cortical level. Dr. Gabriel Devenyi maintained the preprocessing and MAGeT Brain pipelines, and developed the brain masking functionality we used to improve segmentation accuracy, as well as helping in troubleshooting and problem solving, particularly with the RMINC statistical package. Stephanie Tullo developed and refined the MAGeT Brain atlases used in this project. Christina Kazazian and Mallar Chakravarty assisted with segmentation quality control. My supervisor, Mallar Chakravarty, designed and implemented the original MAGeT Brain algorithm, and he conceptualized the initial project and provided invaluable guidance throughout.

The Autism Brain Imaging Data Exchange (ABIDE; http://fcon_1000.projects.nitrc.org/indi/abide/) provided a large proportion of the data used in this project through their two open-access releases. The following collaborators also generously contributed data to this project: at the National Institute of Mental Health (Bethesda, Maryland, USA) - Dr. Armin Raznahan, Dr. Elizabeth Smith, Dr. Audrey Thurm and Dr. Rhoshel Lenroot;

at the Hospital for Sick Children (Toronto, ON, Canada) - Dr. Evdokia Anagnostou, Dr. Jason P. Lerch and Dr. Margot J. Taylor; at the University of Cambridge (Cambridge, UK) - Dr. Meng-Chuan Lai, Dr. Lindsay R. Chura, Dr. Simon Baron-Cohen, Dr. Dorothea L. Floris, Dr. Rosemary J. Holt, Dr. Michael V. Lombardo and Dr. Amber N. V. Ruigrok; and at King's College London (London, UK) - Dr. Michael C. Craig, Dr. Christine Ecker and Dr. Declan G. M. Murphy. The data sets from the University of Cambridge and from King's College London were collected for the UK Medical Research Council Autism Imaging Multisite Study (UK MRC-AIMS) and the Cambridge Family Study of Autism.

List of Abbreviations

ABIDE	Autism Brain Imaging Data Exchange
ACA	Anterior Cingulate Cortex
ADOS	Autism Diagnostic Observation Schedule
ADOS-CSS	Autism Diagnostic Observation Schedule - Calibrated Severity Score
AIC	Akaike Information Criterion
ANTs	Advanced Normalization Tools
ASD	Autism Spectrum Disorder
BEaST	Brain Extraction based on nonlocal Segmentation Technique
CAM	Cambridge Family Study of Autism
Cambridge	University of Cambridge - UK-AIMS Site
DLPFC	Dorsolateral Prefrontal Cortex
DX	Diagnostic status (ASD or typically developing)
ENIGMA	Enhancing Neuro Imaging Genetics through Meta-Analysis Consortium
FDR	False Discovery Rate
FEF	Frontal Eye Fields
FIQ	Full-scale Intelligence Quotient
GABA	γ-Aminobutyric Acid
IoP	Institute of Psychiatry, King's College, London - UK-AIMS Site
IP	Institut Pasteur - ABIDE site
IQ	Intelligence Quotient
KKI	Kennedy Krieger Institute - ABIDE site
LGN	Lateral Geniculate Nucleus of the thalamus
LOFC	Lateral Orbitofrontal Cortex
MAGeT	Multiple Automatically Generated Templates
MAX_MUN	Ludwig Maximilians University Munich - ABIDE Site
MOFC	Medial Orbitofrontal Cortex
MRI	Magnetic Resonance Imaging
NIMH	National Institute of Mental Health
NYU	New York University - ABIDE site
OHSU	Oregon Health and Science University - ABIDE Site
Q1K	Quebec 1000 Families Initiative
QC	Quality Control
SDSU	San Diego State University - ABIDE Site
TBV	Total Brain Volume
TD	Typically Developing
TORONTO	SickKids / University of Toronto
UK-AIMS	UK Medical Research Council Autism Imaging Multicentre Study
UM	University of Michigan - ABIDE Site

List of Figures

List of Tables

1. Introduction and Statement of Problem

In the first English-language clinical report of Autism Spectrum Disorder (ASD), Leo Kanner (1943) suggested that the disorder might be something "inborn". Theories of the cause, mechanism, and nature of ASD have shifted greatly in the intervening decades, including an unfortunate period in the mid 20th century, during which a psychoanalytic view was ascendant that placed the blame on inadequate parenting (Boucher, 2008). While the exact etiology has remained elusive, what has become clearer is that ASD is a biologically-based, highly heritable neurodevelopmental disorder with a significant genetic component (Chaste & Leboyer, 2012).

Estimates of the prevalence of ASD vary widely, with a worldwide median of about 62 / 10 000, though there are reports of increasing prevalence (Elsabbagh et al., 2012; Levy et al., 2009). Among children and youth in Canada, the prevalence is 1/66 (Ofner et al., 2018). As a spectrum disorder, symptom profiles are quite heterogeneous by definition, as is the degree of impairment an individual experiences. Some individuals pursue advanced education, marriage, children, and a career; in some cases diagnosis may not even come until adulthood. For other individuals, impairment may be severe, and life-long support may be required even with basic functions of self-care (Boucher, 2008).

There are, thus far, no biological criteria to identify ASD, and diagnosis relies entirely on the observation of behaviour. There is no single, definitive symptom; rather, a constellation of symptoms from two categories must be present for an ASD diagnosis. First, social interaction and communication must be impaired, in at least the contexts of nonverbal communication, maintaining and understanding relationships, and social-emotional reciprocity. Second, there must be "restricted, repetitive patterns of behaviour, interests, or activities," which must include at least two of the following: stereotyped movements or speech; inflexible, routinized or

ritualized patterns of behaviour and an aversion to change; intense, fixated interest in a restricted set of topics; and/or sensory-perceptual hyper- or hypo-acuity and interest. These symptoms must be persistent, present early in development, and not be better explained by either global developmental delay or intellectual developmental disorder, and they must cause significant impairment in functioning. Specifiers, such as “with/without accompanying language impairment,” or “with/without accompanying intellectual impairment” may be added to the diagnosis to encode common features that are not diagnostically necessary, but that have a significant impact on functional outcomes (American Psychiatric Association, 2022). Language development is generally atypical, though whether that means that language function is impaired or even enhanced depends on the individual (Lai et al., 2014). While there are many common co-morbidities, intellectual disability, anxiety, and the dysregulation of sleep, immune function, and/or gastrointestinal function are particularly common (Lai et al., 2014).

The presentation varies greatly by age, by sex, and by overall intellectual function (Lai et al., 2014). With appropriate primary care screening, ASD can be diagnosed within the first three years of life, though signs are often present even in the first year; nevertheless, in many cases, ASD diagnoses do not occur until children reach school age and teachers call attention to symptoms (C. P. Johnson et al., 2007). This is unfortunate, because the success of the available treatment options depends on the intervention beginning as early as possible. Though several forms of treatment exist, they generally involve behavioural interventions (Lai et al., 2014; Lord et al., 2020), and can be extremely expensive due to their time-intensive, one-on-one nature.

Studying the neurobiology of ASD has proven difficult for a number of reasons indicated above: there is a great deal of heterogeneity in symptom profiles and co-morbidities, the etiology is poorly understood, and there are no reliable biomarkers that indicate the presence of ASD.

ASD presentation varies greatly over the lifespan for reasons that are poorly understood, but that probably encompass both neurodevelopmental changes as well as learned behaviours (for example, compensatory strategies). It varies greatly by sex, and while it is prevalent in males at a rate three to four times higher than in females, females with ASD are more likely than males to be more severely affected (Boucher, 2008; C. P. Johnson et al., 2007; Lai et al., 2014). Even common comorbidities occur in different combinations in different individuals, and some, such as intellectual disability, can greatly affect the presentation. This heterogeneity at the behavioural level presumably indicates heterogeneous biology. In this case, we would expect to find a lack of agreement in the literature about the exact nature of the neurobiological alterations that are present in ASD, and, as outlined in Chapter 2, this is in fact the case. In this study, we attempt to address this by examining the neurobiology of ASD at the subcortical level, while accounting for three potential sources of neurobiological heterogeneity: sex, age, and full-scale intelligence quotient (FIQ), since variations along each of these dimensions has been associated with variations in both ASD symptom profiles and neurodevelopment.

2. Background Information and Relevant Literature

The subcortex is relatively understudied in comparison to the cortex, both in general (Alkemade et al., 2013) and in relation to ASD in particular. In both subcortical structures and the cerebral cortex, however, significant biological heterogeneity in ASD has made the task of studying its neurobiology more difficult. This is apparent in the literature reviewed below, which at times contains divergent and contradictory reports about the nature of the neuroanatomical differences associated with ASD.

As a neurodevelopmental disorder, ASD is characterized by alterations in normal trajectories of brain development. These alterations may be inconsistent across individuals. While the focus of the present study is on the subcortical neuroanatomy of ASD and not its genetics, neuroanatomical development is ultimately, largely controlled by genetics and so it may be helpful briefly to review some of the literature that explores these phenomena.

2.1 Genetic and Neurobiological Heterogeneity in ASD

Just as we would expect to find heterogeneous biology underlying the heterogeneous clinical presentations in ASD, we would expect to find wide variation in the genetic backgrounds on which that biology develops over the lifespan. Indeed this is the case, and while there has been a great deal of discussion about potential environmental risk factors for ASD (Gaugler et al., 2014), twin studies provided evidence of a strong genetic component as early as 50 years ago (Folstein & Rutter, 1977). However, no single genetic cause has been found. Loci across the genome, on nearly every chromosome, have been identified as contributing potential genetic risk factors of ASD (Müller, 2007), which suggests that ASD is polygenic in nature (Boucher, 2008).

Reviewing this state of affairs, Courchesne et al. (2019) proposed a set of 72 “high confidence,” relatively highly penetrant ASD genes. Over 90% of these genes are highly expressed during prenatal development, with 68% peaking prenatally and downregulated after birth, while the expression of the remaining 32% peaks during the first few postnatal years. Most such genes are pleiotropic, with expression in many different regions, and with functions focused on basic organizational processes of the central nervous system: proliferation, migration, neurite outgrowth, and synaptogenesis (Courchesne et al., 2019). What is striking is not only that these high confidence ASD genes play a role in a well-defined set of neurodevelopmental functions, but also that there are so many of them, converging on the same pathways. This lends support to the polygenic hypothesis, while also hinting at a wide potential degree of genetic heterogeneity in ASD: there are many possible combinations of mutations, many of them relatively common (Y. Zhang et al., 2021), in this set of genes that all converge on the same pathways, and may influence neurodevelopment in subtly different ways depending on the specific combination. In this sense, ASD is a diffuse disorder, in which many (or even all) brain systems are affected (Müller, 2007).

At the macroscopic level, abnormalities in both grey (Brieber et al., 2007) and white matter (Ameis & Catani, 2015) have been widely reported. There also appear to be differences in microstructure at the boundary of grey and white matter (Olafson et al., 2021). Early overgrowth of the whole brain, followed by a relative slowing of brain size growth, was observed over 20 years ago (Courchesne et al., 2001). Just after birth, neonates who will go on to be diagnosed with ASD show no difference (or a small decrease) in brain volume when compared to their typically developing peers; however, by about age 3, autistic brains are about 10% larger than expected (Courchesne et al., 2005; Redcay & Courchesne, 2005). The relative overgrowth

normalizes by early adulthood, likely because the brains of typically developing children have been undergoing steady growth and have “caught up” in size to those of their autistic peers (Redcay & Courchesne, 2005). As with many topics in the neurobiology of ASD, some controversy surrounds even this oft-reported finding (Raznahan et al., 2013). While this may reflect phenotypic heterogeneity, it is also possible that sampling bias, including lack of representation of females and across the ASD severity spectrum, or other possibly related factors such as in-scanner motion, may contribute to conflicting findings (Bedford et al., 2019; Ducharme et al., 2016).

As noted above, the cortex has been studied with greater intensity than subcortical structures in ASD. As widely studied as the cortex is in ASD, however, even here there is disagreement about the specific nature of the differences, which may, in part be due to study methodologies that have not always been able to account for the significant neurobiological heterogeneity observed in ASD (Bedford et al., 2019). There are grounds to consider a number of subcortical structures as fruitful territory to explore in seeking an understanding of the neurobiology of ASD. Here we focus on three structures that play a central role in many of the functions and behaviours that are affected in ASD: the thalamus, the striatum and the globus pallidus.

2.2 The Thalamus

The thalamus has long been understood to serve as a key relay of sensory information in its journey from the periphery to the brain (Kandel et al., 2013; Steriade & Llinás, 1988). This makes it an interesting target of study in research on ASD, a disorder that often includes as a symptom sensory hypo- or hyper-acuity. There is another reason to be interested in the thalamus,

however: it also mediates and modulates a great deal of cortico-cortical communication (Murray Sherman & Guillery, 2013). Aside from olfactory information, all sensory signals are channeled through the thalamus before reaching the cortex. Nevertheless, the thalamus is not a simple relay. In fact, only about 10% of the inputs to the thalamus are sensory (Sherman, 2005). At the most basic level, the thalamus can act on the sensory signals that it relays by either passing those signals on to the forebrain, or, as in the case of sleep, which is also often affected in ASD, the thalamus can suppress forebrain sensory input (Steriade & Llinás, 1988). However, in addition to acting as a direct relay, the thalamus performs two other distinct functions.

First, in its role as sensory relay, the thalamus modulates the signals that it passes to the cortex, based on input from the cortex and brainstem (Sherman, 2005). For example, the lateral geniculate nucleus (LGN) receives retinotopic projections from the retina, and relays the information received to the primary visual cortex. However, only about 7% of the input to the relay cells of the LGN comes from the retina; the rest is composed mainly of afferents whose sources are more or less evenly divided between local interneurons, layer six of the primary visual cortex, and from parabrachial nucleus in the brainstem (Sherman, 2005). Anatomical and electrophysiological evidence indicates strongly that retinal afferents act to drive LGN relay cells, while the other 93% of input fibers act to modulate relay cell activity (reviewed in Sherman, 2005; and Sherman & Guillery, 2011).

This pattern repeats itself throughout the various nuclei of the thalamus, even in areas where there is no peripheral input. These nuclei are driven instead by cortical input arriving in axons from cortical layer five (Sherman & Guillery, 2011). Otherwise the structure and function of these nuclei seems to be analogous to those of the sensory nuclei: they receive modulatory input via feedback connections from layer six of the cortex, and send output to other areas of the

cerebral cortex in a generally topographical manner. These higher order nuclei are thought to provide another channel of cortico-cortical communication, by mediating and modulating connections between cortical areas (Sherman & Guillery, 2011). Alterations in thalamic structure and function can therefore affect not only the transmission of sensory information to the cortex, but communication between areas of the cortex as well. Behavioural data confirms this: focal strokes affecting the thalamus can affect a wide range of behaviours and cognitions, including mood, memory, attention, and language use (Carrera & Bogousslavsky, 2006). Disruptions in thalamocortical connectivity in ASD have been observed using both anatomical and functional connectivity measures (Nair et al., 2013).

2.3 The Basal Ganglia

Another set of structures plays a similarly central role in functions and behaviours that are affected in ASD: the basal ganglia. The principal structures in the basal ganglia are the striatum (composed of caudate nucleus and putamen) and the globus pallidus, and these are the structures we propose to examine. However, other smaller structures are considered to be part of the basal ganglia, including the substantia nigra and the subthalamic nucleus. The striatum and globus pallidus are central in a parallel set of processing loops that run from the cerebral cortex, through the striatum, then the globus pallidus, and finally through the thalamus and back to cerebral cortex (Alexander & Crutcher, 1990). These loops are directly involved in behaviours that are affected in ASD, as discussed below.

Four main loops have been described: the motor loop, the oculomotor loop, the prefrontal/associative loop(s), and the limbic loop, named both for their cortical regions of origin, as well as for their function (Alexander & Crutcher, 1990). With some simplification, the

motor loop originates in motor areas of the frontal cortex (primary motor cortex, premotor cortex and supplementary motor area), projects topographically to the putamen, which then projects to the globus pallidus (both internal and external segments). The globus pallidus projects to ventral portions of the thalamus, which then closes the loop through projections back to the motor areas of the cortex (Alexander et al., 1991). The structure of all other loops is similar. The oculomotor loop proceeds from frontal eye fields (FEF) to the caudate, through the globus pallidus to ventral and medial thalamus, then back to FEF. The associative loop runs from dorsolateral prefrontal cortex (DLPFC) and lateral orbitofrontal cortex (LOFC), via anterior caudate, globus pallidus, and ventral and medial thalamus, back to the DLPFC and LOFC. The limbic loop runs from anterior cingulate (ACA) and medial orbitofrontal cortex (MOFC), through the ventral striatum, globus pallidus, ventral and medial thalamus, and back to ACA and MOFC (Alexander et al., 1991; Wichmann & DeLong, 2013). These loops have a stereotyped structure, each consisting of "direct", permissive pathways and "indirect", inhibitory pathways (Smith et al., 1998). In addition to these loops operating in parallel with each other, within each loop, projections are topographical in nature, are organized into segregated, parallel subcircuits, each of which may underlie a different aspect of behaviour (Alexander et al., 1991; Wichmann & DeLong, 2013). We can refine this highly simplified picture of these anatomical and functional loops a little more by noting that, while they terminate only in the locations indicated, they receive some supplemental input from widely distributed cortical regions from all lobes of the brain, as well as feedback connections from the thalamus (Wichmann & DeLong, 2013).

These loops do not act as simple relays. The signals in the striatum, for example, are highly modulated, by cholinergic and GABAergic interneurons and by dopaminergic, likely reward- and reinforcement-related neurons from other basal ganglia and midbrain structures

(Wichmann & DeLong, 2013). While the literature describing the function of these loops is vast, for the purposes of this study it may be sufficient to note that the evidence suggests a role in reinforcement learning, as well as selecting, initiating, executing, and evaluating behaviours, including movements and cognitive behaviours (Wichmann & DeLong, 2013).

Behaviours mediated by these basal ganglia circuits are known to be affected in ASD. For example, some verbal behaviours, as well as aspects of empathic behaviour and behaving in concert with social norms are mediated by the prefrontal loops, while the limbic loop has a role in motivated behaviour (Wichmann & DeLong, 2013). The oculomotor loop is responsible for regulating saccades, via the caudate nucleus and substantia nigra projections to the superior colliculus (Wichmann & DeLong, 2013). Saccadic eye movements are known to be affected in ASD, though it remains unclear to what extent these effects originate in the basal ganglia (B. P. Johnson et al., 2012; Schmitt et al., 2014).

2.4 The Thalamus and Basal Ganglia in ASD

There is ample evidence that both the structure and function of the striatum and globus pallidus are affected in ASD. For example, Langen et al. (2007) found larger caudate nucleus volumes, in a small sample of medication naive children, adolescents and adults. With a larger sample, the same group confirmed these findings and noted an altered developmental trajectory as well: caudate volumes continued to increase in children with ASD, when controlling for total brain volume, whereas in typically developing children this volume decreased over childhood and adolescence (Langen et al., 2009). Increased caudate volume was also found to be related to repetitive behaviours in individuals with ASD (Langen et al., 2014). Similar results have been found for the globus pallidus in adults and children (Turner et al., 2016).

Given the central role of the thalamus, striatum, and globus pallidus in behaviours known to be affected in ASD, and the evidence of structural and functional differences in ASD, it is not surprising that the neuroanatomy of these structures in ASD has been studied before, though they remain understudied in comparison to the cortex. Unfortunately, there has been very little agreement on the exact nature of the differences. For the thalamus, for example, there are reports that, compared to typically developing controls, the volume of the thalamus in individuals with ASD is larger (Lin et al., 2015), smaller (McAlonan et al., 2008; Sussman et al., 2016; Tsatsanis et al., 2003), and not significantly different (Estes et al., 2011; Haar et al., 2016; Schuetze et al., 2016; Turner et al., 2016; W. Zhang et al., 2018). Similarly, the striatum has been reported to be larger (Hollander et al., 2005; 2007; Turner et al., 2016), smaller (McAlonan et al., 2008; van Rooij et al., 2018), and not significantly different (Estes et al., 2011; Haar et al., 2016; Schuetze et al., 2016; Sussman et al., 2016; W. Zhang et al., 2018). The findings vary as well for the globus pallidus (Estes et al., 2011; Haar et al., 2016; McAlonan et al., 2008; Schuetze et al., 2016; Sussman et al., 2016; Turner et al., 2016; van Rooij et al., 2018; W. Zhang et al., 2018).

2.5 Potential Sources of Neuroanatomical Heterogeneity in ASD

Noting a similar phenomenon at the cortical level, Bedford et al. (2019) proposed three potential sources of this lack of concordance: heterogeneity within the ASD samples that had not been parsed, issues with data quality, and the use of gross volumetric measures that may obscure subtle, localized differences in structure area or shape. They parsed the heterogeneity by accounting for age, sex, and intellectual functioning as measured by IQ, adopted a rigorous quality-control protocol, and used a more sensitive vertex-based technique to uncover effects that varied over the surface of the brain.

There is evidence, beyond the study mentioned above, that sex, age, and IQ contribute meaningfully to neuroanatomical differences in ASD. For example, in a study of high-functioning children and adolescents on the ASD spectrum, Langen et al. (2009) found both a main effect of ASD on caudate volumes, and an age-by-diagnosis interaction, suggesting that the caudate developed along a different trajectory in ASD. The sample did not include adults, and, while it did include females, the male:female ratio was skewed (~10:1). Zhang et al. (2018) found an altered trajectory for the putamen volume, but only in females. Schuetze et al. (Schuetze et al., 2016) reported focal age-by-diagnosis interactions in the shape of the right globus pallidus and striatum in males with ASD, as well as a significant effect of IQ on striatal and thalamic morphology.

Data quality has been shown to be a significant factor affecting the outcome of MRI-based neuroanatomical studies. Rigorous, manual quality control protocols to verify both scan quality (particularly motion) and downstream processing such as segmentation improves the sensitivity of various cortical measures (Bedford et al., 2019). In another study in which data samples with and without strong quality control were matched for power, cortical thickness trajectories detected in the data were highly dependent on data quality, with trajectories shifting towards higher-order (quadratic, cubic) trajectories in lower-quality data (Ducharme et al., 2016). That said, the effect of motion is much better understood on cortical measures than on subcortical measures.

3. Rationale for the Study, Hypotheses, and Specific Aims

The overarching goal of this project was to try to resolve the disagreements about the nature of subcortical anatomical differences in ASD, using spatially-sensitive techniques that account for variation due to age, sex, and IQ, in a rigorously quality-controlled, large, highly-powered multisite dataset. In doing so, we hoped to parse some of the heterogeneity that may be interfering with finding clear results, and to reduce the effects of spurious, inter-site differences as much as possible.

Based on previous literature, we hypothesized that we would be able to find localized, regional differences in thalamic and basal ganglia morphometry that do not show up in global, volumetric comparisons. Furthermore, we expected to find that sex, age, and/or IQ would be significant contributors to heterogeneity in these measures.

Specific Aim 1: Determine whether the thalamus, striatum, and globus pallidus differ in volume in ASD, and whether sex, age, and IQ are important factors in explaining any such differences.

Specific Aim 2: Determine whether localized differences in morphometry are apparent in these same structures in ASD, that are not detected by the volumetric method above. Evaluate the role of sex, age, and IQ in any such differences.

4. Methods

4.1 Dataset

Analysis was performed on a large, multisite dataset (n = 3145) of T1-weighted MRI scans of the head, consisting of a combination of open-source and closed-source data (Bedford et al., 2019; Olafson et al., 2021). The demographic composition of this dataset is detailed in Tables 1 and 2, but in brief, it contains both male and female participants, both including individuals with ASD and developing typically, ranging in age from 2 to 65. Datasets include both releases of the multi-site ABIDE dataset (Di Martino et al., 2014, 2017), as well as data from the National Institute of Mental Health (USA), the Hospital for Sick Children (Canada), the Cambridge Family Study of Autism (UK), and the UK Medical Research Council Autism Imaging Multicentre Study (UK), comprising a total of 32 sites. Sites with fewer than 3 females in the ASD group after quality control were excluded from the analysis.

4.2 Visual Inspection and Preprocessing

Raw scans were inspected visually and rated by two independent raters, as described by Bedford et al. (2019). Scans with significant motion or other artefacts were processed but were excluded from analysis. All scans, regardless of quality, were preprocessed. Preprocessing was performed in site-wise batches, using the minc-bpipe-library pipeline (https://github.com/CoBrALab/minc-bpipe-library). This pipeline applies bias field correction, to remove artefactual, low-frequency inhomogeneity in image intensities, using the iterative N4 algorithm (Tustison et al., 2010). Next, brain masks were generated and used to extract the brains from surrounding tissue (skull, fat, etc.), using the Brain Extraction based on nonlocal Segmentation Technique (BEaST; Eskildsen et al., 2012). Total brain volume (TBV) was

estimated using CIVET 1.1.12, by summing estimates of grey matter volume, white matter volume, and cerebrospinal fluid volume (Bedford et al., 2019; J. P. Lerch & Evans, 2005).

Adjustments were made to the preprocessing pipeline for some sites to maximize the amounts of usable data; the site-based, meta-analytic nature of the statistical analysis ensures that we did not introduce new site-based confounds by doing this. Scans from ABIDE I and II releases from New York University (NYU) were processed with blood vessel masking activated in CIVET. The UK-AIMS scans were subjected to an intensity standardization relative to the MNI ICBM 152 average brain template (Collins et al., 1994) using minc_nyul. Scans from both ABIDE releases from the Kennedy Krieger Institute (KKI) were pre-processed without brain masking or extraction. Scans from the ABIDE II release from the Institut Pasteur (IP) without bias field correction.

4.3 Segmentation and Volume Computation

Segmentation of the thalamus, striatum and globus pallidus was performed using the Multiple Automatically Generated Templates (MAGeT Brain) algorithm (Chakravarty et al., 2013; Pipitone et al., 2014), in batches by site. MAGeT Brain makes use of five atlases in which voxels of the structures of interest are labeled. These labels were derived from serial histological data, fit to a single high-resolution MRI template (Chakravarty et al., 2006), then warped to five new high-resolution MRI atlas templates (Tullo et al., 2018). The 5 label sets were then propagated to a set of 21 template scans, drawn from the data in each batch, giving 105 label sets for each structure. These labels are then propagated to each scan in the batch, which produces 105 candidate labels for each structure and for each scan in the dataset. These are merged into a final set of labels for each structure and each scan through a voxel-wise majority vote. Label

propagation from one image to another is performed at each step using the Advanced Normalization Tools (ANTs) library (Avants et al., 2014), by first performing an affine registration, then a non-linear registration between the two volumes, and then using the generated transforms to warp the labels into the space of the second volume. The volume of each structure, for each participant, was computed using the number of voxels in the label for that structure and the voxel volume.

While other software exists to label and compute the volume of subcortical structures, including FreeSurfer (https://surfer.nmr.mgh.harvard.edu), and FSL-FIRST (https://fsl.fmrib.ox.ac.uk), MAGeT Brain has been shown to compute labels for these structures that more closely resemble the gold standard of manually generated labels created by expert raters (Makowski et al., 2018). All scans at each site were segmented, however only those scans that passed the motion quality control procedure described above were included in the downstream analyses.

All labels were inspected visually and rated for segmentation quality by one or more expert raters, and inaccurately labelled structures were excluded from downstream analysis.[1] Only individual structures that failed were excluded, meaning that successful segmentations from the same participant were included in the analysis. For this reason, the sample size varies somewhat between structures, as detailed in Section 5: Results.

To improve segmentation quality after the initial run (see Section 5: Results), two modifications were made to this workflow. First, to reduce over-segmentation of the striatum, which was a common failure, atlases with manually corrected striatal segmentations were used. Second, segmentation was performed on left and right hemispheres separately, and a subcortical

1 Saashi Bedford, Mallar Chakravarty, Christina Kazazian, Emily Olafson, and the author all contributed to segmentation quality control at various stages. All segmentation quality control for the final segmentations was completed by the author.

mask was applied to remove the cortex from the registration process. Only the results of this improved image-processing workflow are presented.

4.4 Morphometry

To capture shape differences, two morphometric measures - surface area and displacement - were computed by MAGeT Brain at each of many vertices across the surfaces of each structure, as described by Tullo et al. (2018) and other manuscripts from our group (Chakravarty et al., 2015; Raznahan et al., 2014; Shaw et al., 2016) and summarized here. For each structure, a model was created from all five atlas templates. These were registered to each other, first linearly and then non-linearly using ANTs (Avants et al., 2014), then averaged together. Using this average model atlas, surface meshes covering the individual structures (left and right striatum, thalamus, and globus pallidus) were computed using the marching cubes algorithm and manually smoothed. The surfaces were then re-meshed with vertex spacing corresponding to voxel size, resulting in model object surfaces with ~13 000 vertices per striatum, ~6 500 vertices per thalamus, and ~3 000 vertices per globus pallidus (Tullo et al., 2018).

The concatenation of the non-linear portion of the transformations used to generate the model, along with those used during registration steps outlined in the previous section, produces a deformation field that describes the shape of each subject structure in the dataset, relative to the model for that structure. The displacement of each vertex is determined by computing the dot product of the deformation at that point with the surface normal, resulting in a measure of expansion (positive displacement) or contraction (negative displacement) relative to the model

surface (Bussy et al., 2021; Makowski et al., 2018; Raznahan et al., 2014; Schuetze et al., 2016; Tullo et al., 2019).

To compute surface area at each vertex, the transforms described above are applied to the model mesh, bringing it into the subject space, and then merged using the median vertex position. The surface is tessellated using a Voronoi parcellation, by connecting the midpoints of the line segments joining each vertex to its neighbours in the mesh, creating a polygon around each vertex. The surface area at a given vertex is computed as the area of this polygon (J. P. Lerch et al., 2008; Shaw et al., 2014).

4.5 Statistical Analysis

The three metrics outlined above - structure-wise volume, vertex-wise displacement, and vertex-wise surface area - were computed for all scans in the dataset, however all data from scans which failed motion QC, or for structures with failed segmentation, were excluded from the analysis described here. However, because these measures were computed and retained, an extension to this project is possible in which we could evaluate the effect of rigorous quality control on the study outcome. This has been shown to be relevant at the cortical level (Bedford et al., 2019; Ducharme et al., 2016).

4.5.1 Statistical Analysis - Volumetric

4.5.1.1 Volumetric Case-Control Analysis

The volumes of the left and right thalamus, globus pallidus, and striatum were modeled using linear regression. In all models, diagnostic status (DX, either ASD or TD) was the

predictor, and total brain volume (TBV) was included as a covariate. In the initial model, sex and age were also included as covariates.

To reduce the effect of non-biological variation between sites due to factors such as the use of different scanner hardware, software, and scanning protocols, models for each structure were computed for each site. Hedges' *g**, which is an unbiased effect size indicator based on the more familiar Cohen's *d*, was computed for the effect of diagnosis in each model, and these effect sizes were combined across all sites using random-effects meta-analysis. Random-effects meta-analysis was chosen because it does not rely on the assumption that the effect sizes being modeled are the same for all sites (Borenstein et al., 2010).

The basic model *(1)* below was used for case-control volumetric analyses, where "struct" represents the left or right striatum, thalamus, or globus pallidus. This was then refit as model *(2)* with the addition of FIQ, for the subset of data for which FIQ data was available

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \beta_2 TBV + \beta_3 Age + \beta_4 Sex + \varepsilon_{struct} \quad (1)$$

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \beta_2 TBV + \beta_3 Age + \beta_4 Sex + \beta_5 FIQ + \varepsilon_{struct} \quad (2)$$

4.5.1.2 Volumetric Heterogeneity-Focused Analysis: Sex, Age, and FIQ

To determine whether sex, age, and FIQ are important contributors of volumetric heterogeneity (beyond the effect of total brain volume), a partially nested series of models was fit as described above, including the global model (3) containing all terms, models with sex and age terms removed (4, 5), and a model with only TBV as a covariate (6), and a model with no covariates (7). These were re-fit including FIQ for those participants for whom FIQ data was available.

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \beta_2 TBV + \beta_3 Age + \beta_4 Sex + \beta_5 Age*Diagnosis + \beta_6 Sex*Diagnosis + \varepsilon_{struct} \quad (3)$$

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \beta_2 TBV + \beta_3 Age + \beta_5 Age*Diagnosis + \varepsilon_{struct} \quad (4)$$

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \beta_2 TBV + \beta_4 Sex + \beta_6 Sex*Diagnosis + \varepsilon_{struct} \quad (5)$$

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \beta_2 TBV + \varepsilon_{struct} \quad (6)$$

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \varepsilon_{struct} \quad (7)$$

The relative fit of these models was estimated by computing the Akaike Information Criterion (AIC) for each (Akaike, 1974). This is a measure derived from information theory that provides relative measures of model fit that can be compared across nested models. Akaike weights were computed from the AIC for each model, which accounts for both goodness of fit and site size, and indicates the strength of evidence in favour of each of the models in a set (Burnham & Anderson, 2002), and weighted according to site size. Because some datasets did not include FIQ information, the analysis was first performed without including FIQ to maximize statistical power, then repeated on the smaller data set with FIQ included. Correction for multiple comparisons was applied after performing the meta-analysis, using the False Discovery Rate (FDR) method of Benjamini and Hochberg (1995).

4.5.2 Statistical Analysis - Morphometric

4.5.2.1 Morphometric Case-Control Analysis

Statistical analysis of the morphometric data was analogous to that of the volumetric data, except that it was performed at every vertex in each structure. Note that a random-effects meta-analysis across sites was performed for each model, at each vertex. As with the volumetric analysis, models without FIQ were run to maximize statistical power, then the models controlling for FIQ were refit with the subset of data that includes FIQ. Due to the very large

number of linear models and meta-analyses, FDR was used for multiple comparisons correction. FDR correction is performed across all p-values generated across all vertices, for all structures, for a given model. Effect size measures for DX were then thresholded at FDR < .05, and plotted on the model surfaces. The basic models were analogous to those used in the volumetric study, as shown below. SA_i and $Disp_i$ represent the surface area and displacement values, respectively, at vertex "i".

$$SA_i = \beta_{0(i)} + \beta_{1(i)}Diagnosis + \beta_{2(i)}TBV + \beta_{3(i)}Age + \beta_{4(i)}Sex + \varepsilon_i \quad (8)$$

$$Disp_i = \beta_{0(i)} + \beta_{1(i)}Diagnosis + \beta_{2(i)}TBV + \beta_{3(i)}Age + \beta_{4(i)}Sex + \varepsilon_i \quad (9)$$

4.5.2.2 Morphometric Heterogeneity-Focused Analysis: Sex, Age, and FIQ

Model selection was performed somewhat differently than in the volumetric analysis. Vertex-wise models with and without the term of interest were fit by site, and compared in pairs. For example, to evaluate the importance of age in models of surface area, the following two models were compared at each vertex:

$$SA_i = \beta_{0(i)} + \beta_{1(i)}Diagnosis + \beta_{2(i)}TBV + \beta_{3(i)}Age + \beta_{4(i)}Sex + \varepsilon_i \quad (10)$$

$$SA_i = \beta_{0(i)} + \beta_{1(i)}Diagnosis + \beta_{2(i)}TBV + \beta_{4(i)}Sex + \varepsilon_i \quad (11)$$

To do this, AIC was calculated for each model, for each site, and the model with the lowest AIC was selected. These were combined across sites using a winner-take-all approach, weighted by site size, resulting in vertex-wise maps indicating where on the surface of the structure the inclusion of age / sex improved fit. The analysis was then repeated including FIQ.

4.6 Follow-up Analyses: Sex-Stratification, and Age/FIQ Centering

Where age or FIQ were found to be important explanatory variables, an age- or FIQ-centered analysis was conducted to evaluate the interaction between ASD diagnosis and age/FIQ. This was done by repeating the case-control analysis described above, with an age- or FIQ-by-diagnosis interaction term included in the models and computing the model with age centered at 2-year intervals, or FIQ centered at 10-point intervals. This provided an indication of the effect of ASD on structure volume, vertex-wise surface area, and vertex-wise displacement at each age interval, without sacrificing statistical power by stratifying the dataset. FDR correction at 5% was done across all vertices, all structures, all age or FIQ intervals, and both measures. The models used are given below, where X_j represents the age in years. For FIQ, analogous models were used centering on FIQ and controlling for age and sex.

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \beta_2 TBV + \beta_3(Age\text{-}X_j) + \beta_4(Age\text{-}X_j)*Diagnosis + \beta_5 Sex + \varepsilon_{struct} \quad (12)$$

$$SA_i = \beta_{0(i)} + \beta_{1(i)} Diagnosis + \beta_{2(i)} TBV + \beta_{3(i)}(Age\text{-}X_j) + \beta_{4(i)}(Age - X_j) + \beta_{5(i)} Sex + \varepsilon_i \quad (13)$$

$$Disp_i = \beta_{0(i)} + \beta_{1(i)} Diagnosis + \beta_{2(i)} TBV + \beta_{3(i)}(Age\text{-}X_j) + \beta_{4(i)}(Age - X_j) + \beta_{5(i)} Sex + \varepsilon_i \quad (14)$$

4.7 Confirmatory Analyses: Linear Mixed Effects and ComBat Harmonization

For confirmation of our results and for homology with previous multi-site ASD literature, we repeated the case-control analyses using two different techniques that account for intersite differences. First, a mega-analysis was conducted using linear mixed effect models (Harrison et al., 2018), with site included as a random factor. Linear mixed models were computed using the

RMINC package in R (J. Lerch et al., n.d.), and were specified as follows, using a random intercept term for site:

$$V_{struct} = \beta_0 + \beta_1 Diagnosis + \beta_2 TBV + \beta_3 Age + \beta_4 Sex + \beta_5 FIQ + (1|site) + \varepsilon_{struct} \quad (15)$$

$$SA_i = \beta_{0(i)} + \beta_{1(i)} Diagnosis + \beta_{2(i)} TBV + \beta_{3(i)} Age + \beta_{4(i)} Sex + \beta_{5(i)} FIQ + (1|site) + \varepsilon_i \quad (16)$$

$$Disp_i = \beta_{0(i)} + \beta_{1(i)} Diagnosis + \beta_{2(i)} TBV + \beta_{3(i)} Age + \beta_{4(i)} Sex + \beta_{5(i)} FIQ + (1|site) + \varepsilon_i \quad (17)$$

A second mega-analysis was conducted by first harmonizing volume, surface area, and displacement data across sites, while preserving variation due to age, sex, IQ, and total brain volume, then computing linear regressions on the entire dataset. Harmonization was completed using ComBat, a technique first used in genetics research that uses an empirical Bayes algorithm to remove variation due to batch effects while retaining variation due to biological factors (Fortin et al., 2017, 2018; W. E. Johnson et al., 2007). ComBat harmonization was performed with the neuroCombat library for R (Fortin et al., 2017, 2018), prior to multiple linear regression using the basic case-control comparison models described above. Structure volumes were harmonized with respect to site, using the default options, including using parametric priors as well as empirical Bayes-based shrinkage of location and scale parameters across features. For volumetric data, the six structure volumes (left and right globus pallidus, striatum and thalamus) comprised the features. For vertex-wise data, the surface area or displacement at each vertex comprised the set of features. TBV, Sex, and Age were included as biological covariates, to protect against the removal of variance explained by these variables.

4.8 Symptom Severity

We assessed the relationship between our subcortical morphometry measures and ASD symptom severity, as measured by the Autism Diagnostic Observation Schedule Calibrated Severity Score (ADOS-CSS). This measure was only available for *n*=239 participants with ASD after motion QC (192 male, 47 female), across five sites (KKI, NYU, OHSU, SDSU, TORONTO, UM). The site-wise models used were as follows:

$$V_{struct} = \beta_0 + \beta_1 CSS + \beta_2 TBV + \beta_3 Age + \beta_4 Sex + \varepsilon_{struct} \quad (18)$$

$$SA_i = \beta_{0(i)} + \beta_{1(i)} CSS + \beta_{2(i)} TBV + \beta_{3(i)} Age + \beta_{4(i)} Sex + \varepsilon_i \quad (19)$$

$$Disp_i = \beta_{0(i)} + \beta_{1(i)} CSS + \beta_{2(i)} TBV + \beta_{3(i)} Age + \beta_{4(i)} Sex + \varepsilon_i \quad (20)$$

The semi-partial correlation between structure-wise volume, vertex-wise surface area, and vertex-wise displacement and ADOS-CSS, while controlling for TBV, age and sex (Aloe & Becker, 2012) was then computed from β_1 at each site. These were pooled across sites using random-effects meta-analysis for each structure and vertex. FDR correction for multiple comparisons was done across all vertices, all structures, and both measures.

5. Results

5.1 Number of sites and individuals after quality control

Of the 3145 scans over 32 separate sites in the complete dataset, 1118 participants across 20 sites were excluded because the sites did not have 4 or more females per group after quality control, and were therefore too small for statistical modelling. The NIMH dataset (130 participants) was also excluded from analysis because of segmentation failure, perhaps due to the very young ages (range 1-9 years) of the participants, though it was included in previous studies that characterized this dataset at the cortical level and did not rely on subcortical segmentations (Bedford et al., 2019; Olafson et al., 2021). This left 1897 participants across 11 sites. Of these, 1331 (342 M-ASD, 497 M-TD, 125 F-ASD, 367 F-TD) scans passed motion QC. All of the other excluded sites were in the ABIDE datasets. For details, refer to Tables 1 and 2.

The number of participants varied depending on the specific subcortical structure being examined, as segmentation quality control was performed on a per-structure basis. The most common failures were the over-segmentation of the caudate nucleus into the third ventricle, and the under-segmentation of the anterior caudate nucleus. These failures were reduced by using atlases with manually corrected striatal segmentations, by running left and right brains separately, and by using a subcortical mask. Quality control results are summarized in Table 3 (original MAGeT Workflow) and Table 4 (corrected atlases, separate left/right runs, and subcortical mask). Table 4 also shows the number of participants with FIQ data.

5.2 No volume differences in ASD

We did not observe any significant volume differences in ASD in the thalamus, striatum, or globus pallidus, when controlling for TBV, age and sex (Figure 1). This was true both when

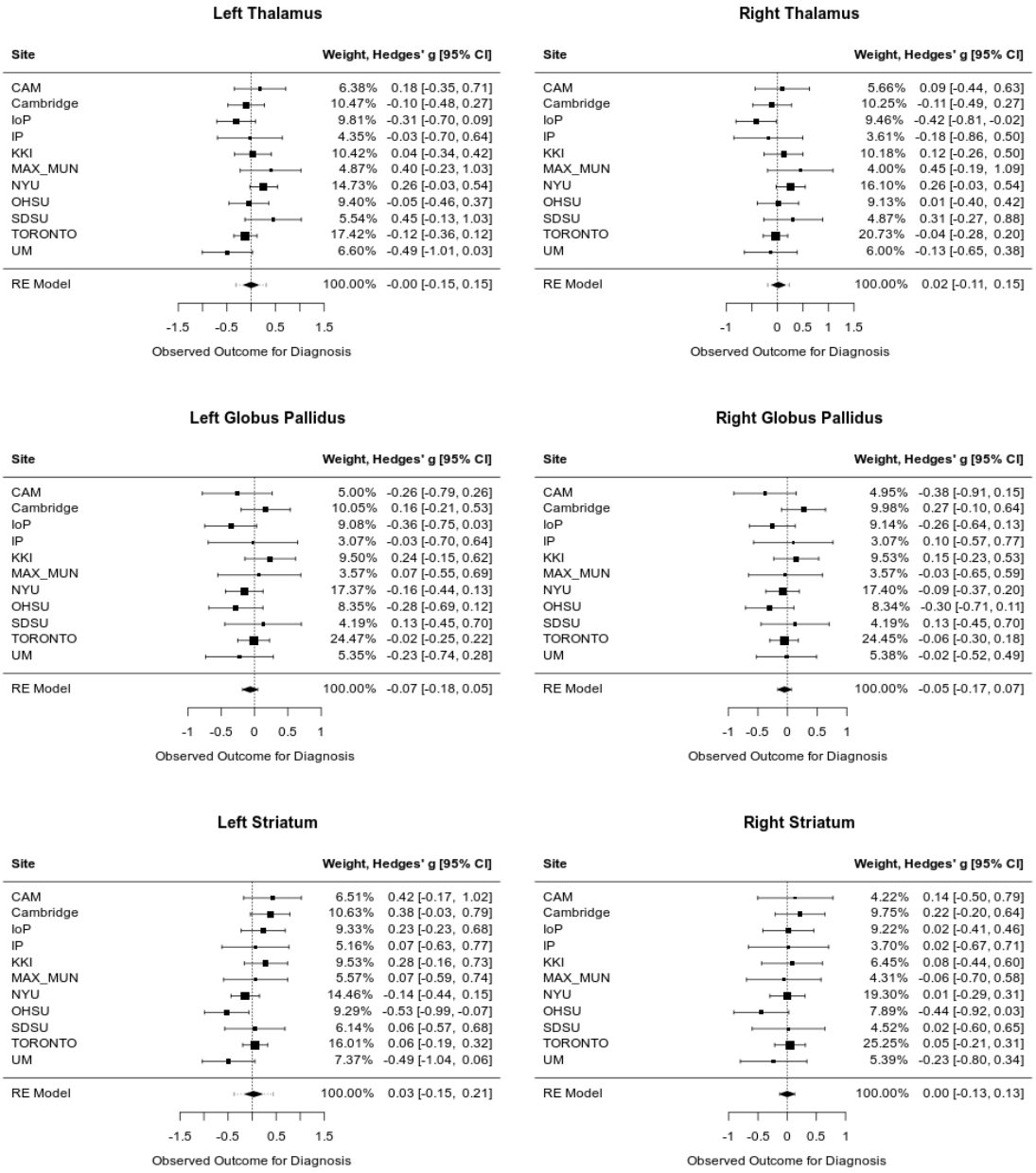

Figure 1: Forest plots showing results of random-effects meta-analysis across sites for all structures. Hedges *g** effect sizes are reported for main effect of diagnosis in model structure_volume ~ diagnosis + total_brain_volume + age + sex. No main effect of diagnosis was observed. Columns are: site name, forest plot, site weight, Hedges *g** estimate and 95% confidence intervals. Site codes are: CAM - Cambridge Family Study of Autism; Cambridge - UKAIMS Cambridge site; IoP - UKAIMS Institute of Psychiatry site; IP - ABIDE Institut Pasteur; KKI - ABIDE Kennedy Krieger Institute; MAX_MUN - ABIDE Ludwig Maximilians University Munich; OHSU - ABIDE Oregon Health and Science University; SDSU - ABIDE San Diego State University; TORONTO - SickKids / University of Toronto; UM - ABIDE University of Michigan

effect sizes were pooled across sites using random-effects meta-analysis, as well as within most sites, with two exceptions: reduced left striatal volume in ASD in the ABIDE OHSU site, and reduced right thalamic volume in ASD in the UKAIMS Institute of Psychiatry site. There were several nearly significant differences, including reduced right striatal volume at OHSU and left thalamic volume at IoP, increased left striatal volume at UKAIMS Cambridge, decreased left striatal volume at ABIDE UM, decreased left pallidal volume at IoP, increased left and right thalamic volume at ABIDE NYU, and decreased left thalamic volume at ABIDE UM. The results were similar when the analysis was repeated including FIQ as a covariate: we did not observe any significant volume differences in ASD in any of the six structures.

5.3 Role of age, sex, and IQ in volumetric models

Model selection using site size-weighted Akaike weights indicated that age, but not sex, improved model fit for both left and right striatum (Akaike weights 0.46 left, 0.51 right). There was somewhat weaker evidence that both age and sex improved model fit for both left and right thalamus (Akaike weight 0.42 left and right), and weak evidence that model fit was best when neither age nor sex were included for both left and right globus pallidus (Akaike weight 0.35 left, 0.36 right). Akaike weights for all candidate models are shown in Figure 2.

We did not observe any effect of ASD diagnosis on subcortical volumes in any of the structures when following up with sex-stratified, age-centered, or FIQ-centered analyses.

5.4 Localized effects alterations of surface area and shape in ASD

Although no volumetric effects of diagnosis were found, localized differences in both surface area and shape (relative displacement) were found in all structures following vertex-wise

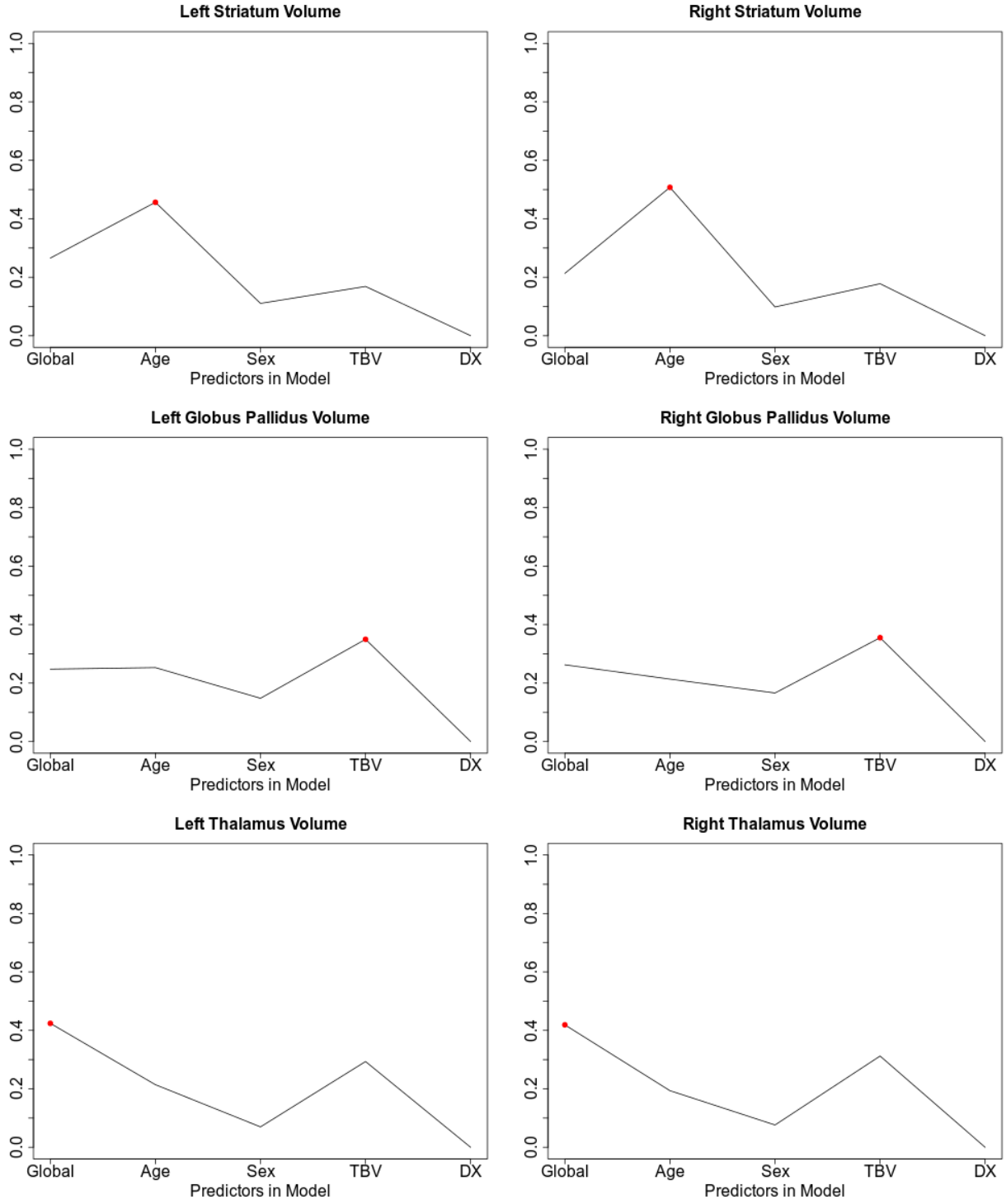

Figure 2: Site-size weighted average of Akaike weights showing evidence for each of five candidate models for left and right striatum, globus pallidus, and thalamus. Red dots indicate maximum value across candidate models. Models are: DX (volume ~ diagnosis), TBV (volume ~ diagnosis + TBV), Sex (volume ~ diagnosis + TBV + Sex + DX*Sex), Age (volume ~ diagnosis + TBV + Age + diagnosis*Age), Global (volume ~ diagnosis + TBV + Age + diagnosis*Age + Sex + diagnosis*Sex)

analysis (FDR < .05), as shown in Figure 3.

In the striatum, with the exception of a small, bilateral region of areal expansion on the anterior caudate surface and a very small region of areal contraction on the right dorsal caudate surface, all surface area and displacement effects were limited to the putamen. These mostly comprised a region of areal contraction on the central portion of the right dorsal putamen, and small, mainly anterior regions of areal expansion, as well as a left, anterior region of inward displacement (decreased convexity), near the nucleus accumbens, and two patches of outward displacement on the left antero-ventral and postero-dorsal putamen.

In the thalamus, the largest region of altered surface area was a bilateral region of increased surface area on the ventral posterior surface, approximately corresponding with the ventral surface of the pulvinar. There was also a fairly large region of positive displacement, approximately corresponding with the more lateral surface of the pulvinar and the ventral posterolateral nucleus.

No bilateral effects were observed in the globus pallidus. Surface area effects were mainly a patch of areal expansion on the left lateral surface, a patch of areal contraction on the right posterior medial surface. Displacement effects included a region of positive displacement on the central dorsal medial surface of the left pallidum, and a region of negative displacement on the left anterior ventral surface.

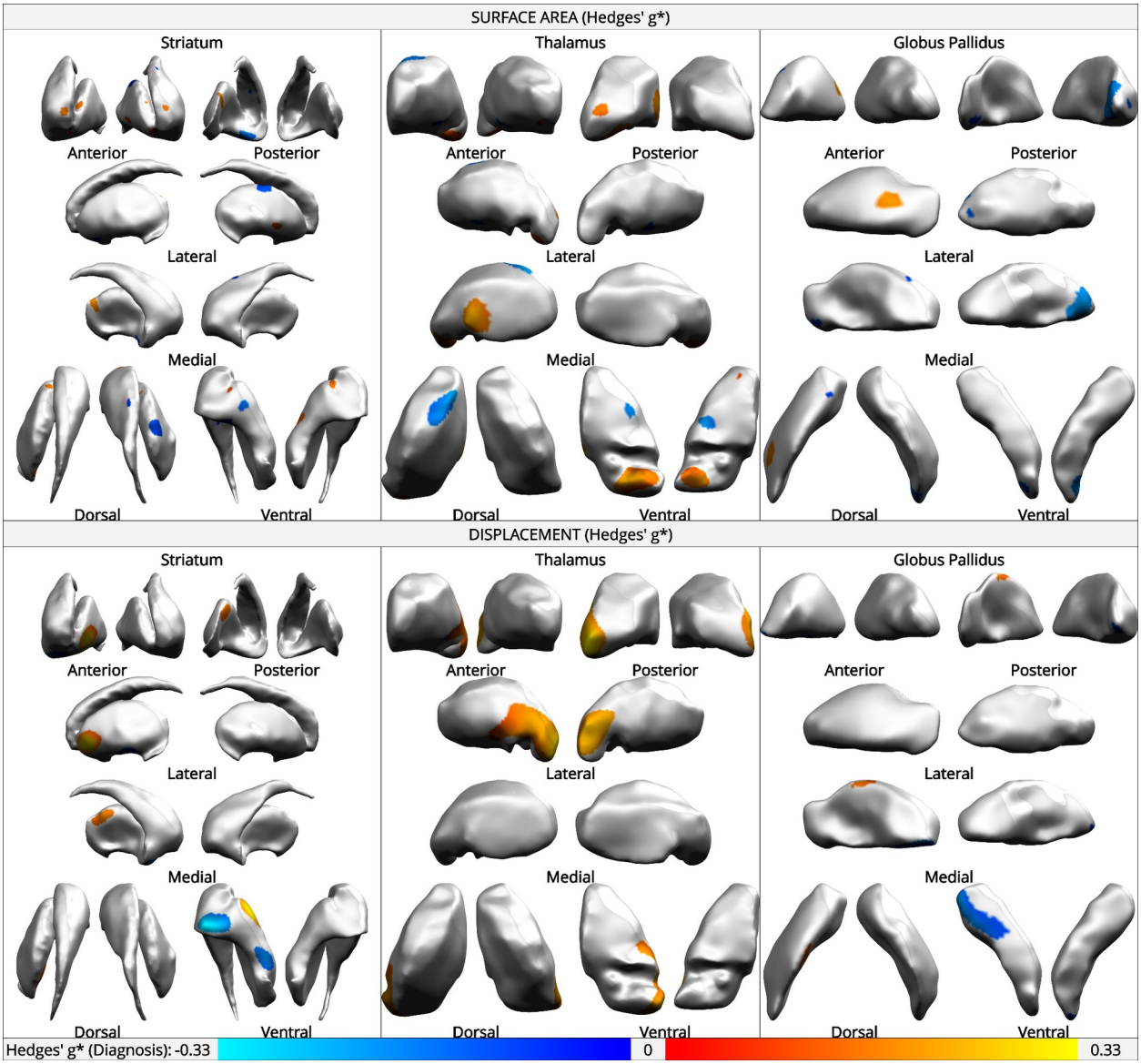

Figure 3: Hedges' *g** main effect of ASD diagnosis on vertex-wise surface area and displacement in the left and right striatum, thalamus, and globus pallidus, when controlling for total structure volume, age, and sex. Warm colours indicate positive effects, cool colours indicate negative effects (*g** range -0.3 to 0.3). Surface area is the Voronoi area surrounding a vertex; displacement represents relative convexity (positive) or concavity (negative).

5.5 Role of age, sex, and FIQ in morphometric models

Vertex-wise model selection analyses indicated that age contributes to the variation in surface-based measures, particularly for displacement, over the surface of most of the globus pallidus and lateral thalamus, as well as much of the striatum (Figure 4). Including FIQ does not improve fit when modeling surface area across most structures, with the exception of a small region on the anterior dorsal caudate. There are, however, regions on the surface of all three structures where FIQ improves the fit of models of displacement (Figure 5). The influence of sex in surface-area models was limited to a relatively small proportion of overall area, with few contiguous regions. Sex was important over somewhat larger regions in all three structures in models of displacement, but these regions still accounted for less than half of overall surface area (Figures 6, 7).

Because of the large proportion of vertices for which age was found to be an important explanatory variable, followup age-centered analyses were performed. A representative example is shown in Figure 8, for age-centered thalamic displacement, centered at intervals of five years. This shows large patches of relative increased convexity in ASD in the dorso- and medio-lateral right thalamus, as well as a region of the rostro-ventral thalamus roughly corresponding to the pulvinar, but only in childhood. These effects fade by adolescence. In adulthood, relatively little of the thalamus shows any shape effects of ASD, though a small region of decreased convexity in ASD on the left medial wall appears to be relatively stable throughout adulthood, while there is a patch of relative increased concavity appearing from middle age on, near the left pulvinar region.

The remainder of the age-centered results are shown in Supplementary Figures SF-1 through SF-5.

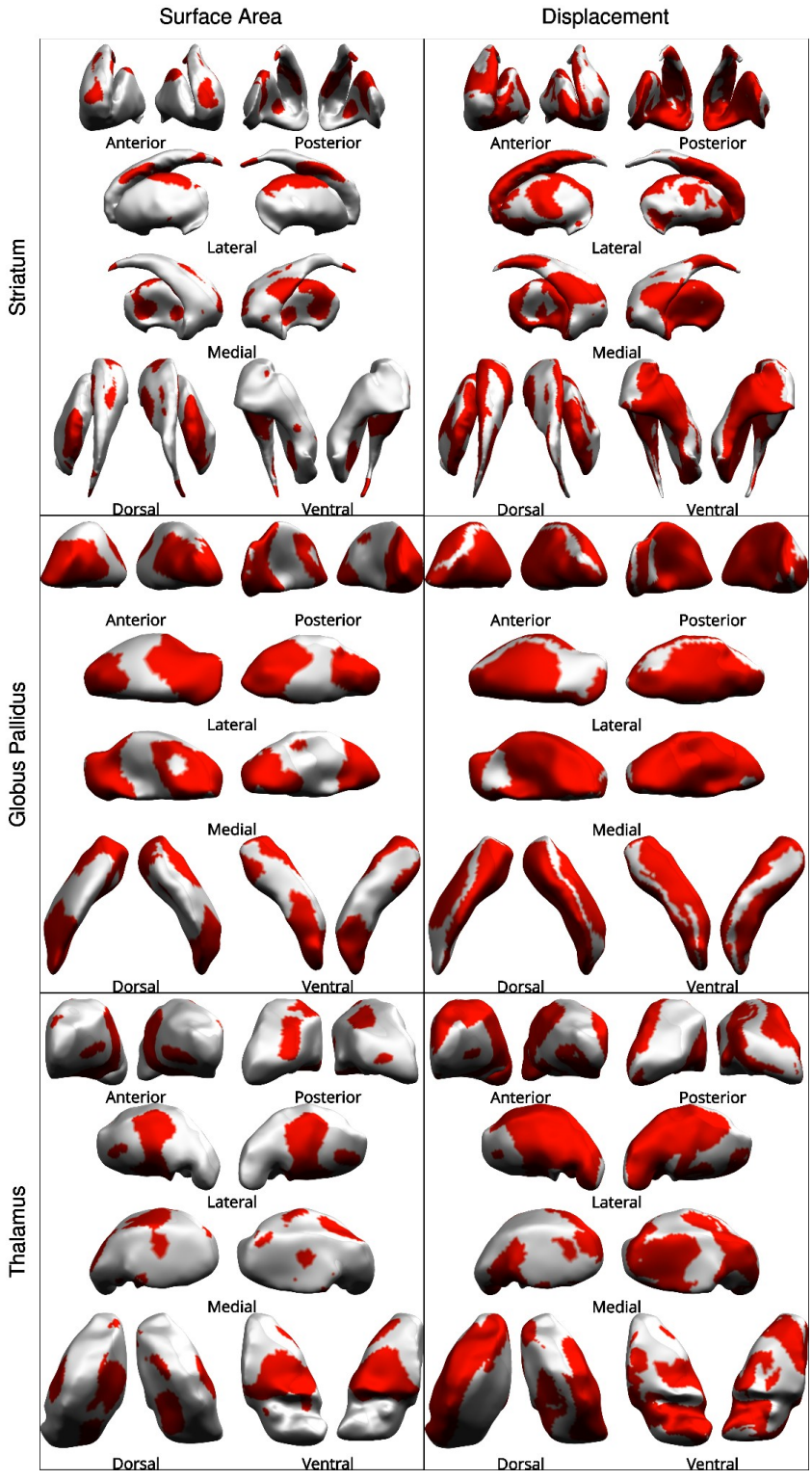

Figure 4: Age maps: Regions on the surface of left and right striatum, thalamus, and globus pallidus in which the inclusion of age and age-by-diagnosis terms improved model fit, for linear models of surface area and displacement on diagnosis, when controlling for sex. Red indicates improved fit with age terms.

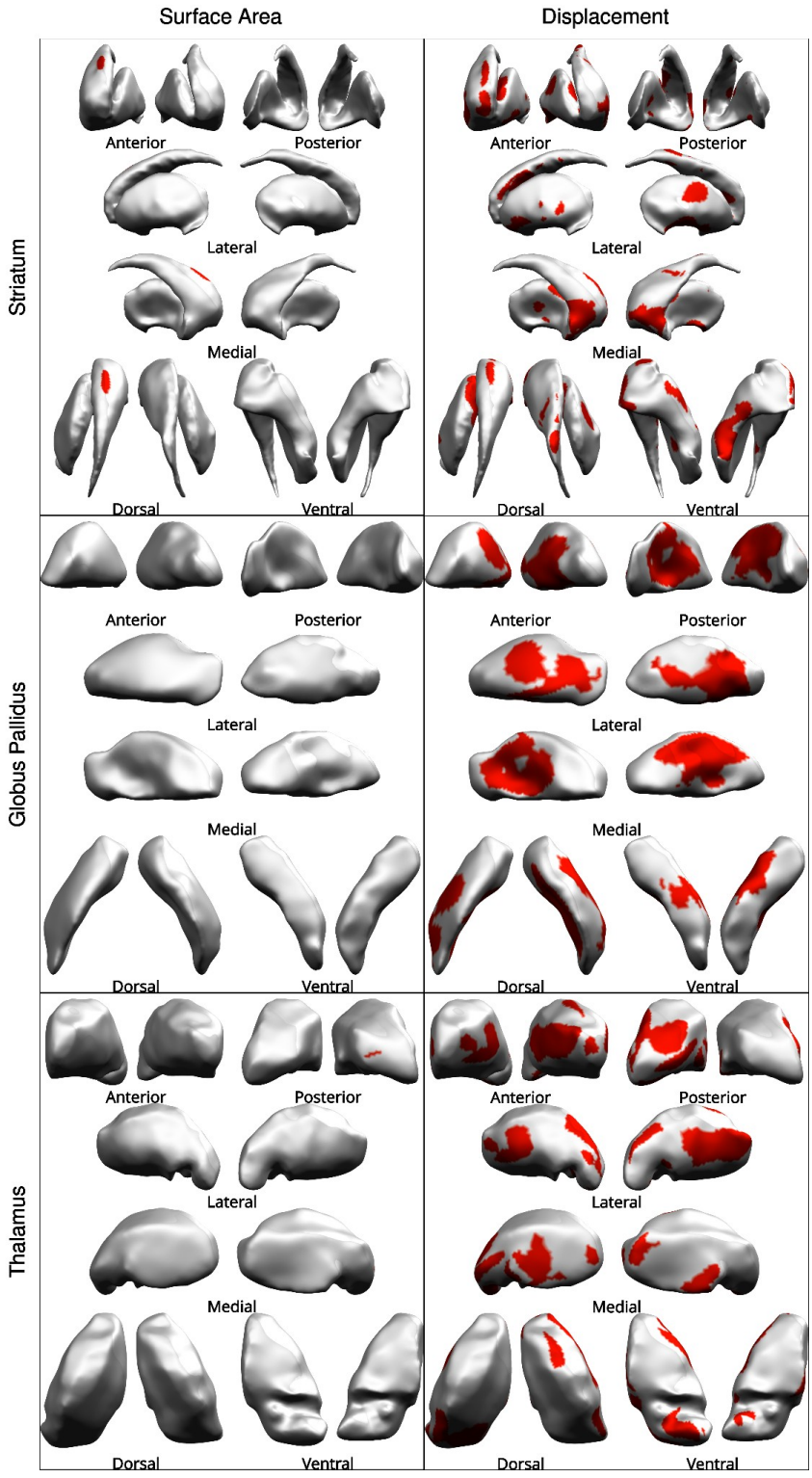

Figure 5: FIQ maps: Regions on the surface of left and right striatum, thalamus, and globus pallidus in which the inclusion of FIQ and FIQ-by-diagnosis terms improved model fit, for linear models of surface area and displacement on diagnosis, controlling for age and sex. Red indicates improved fit with FIQ.

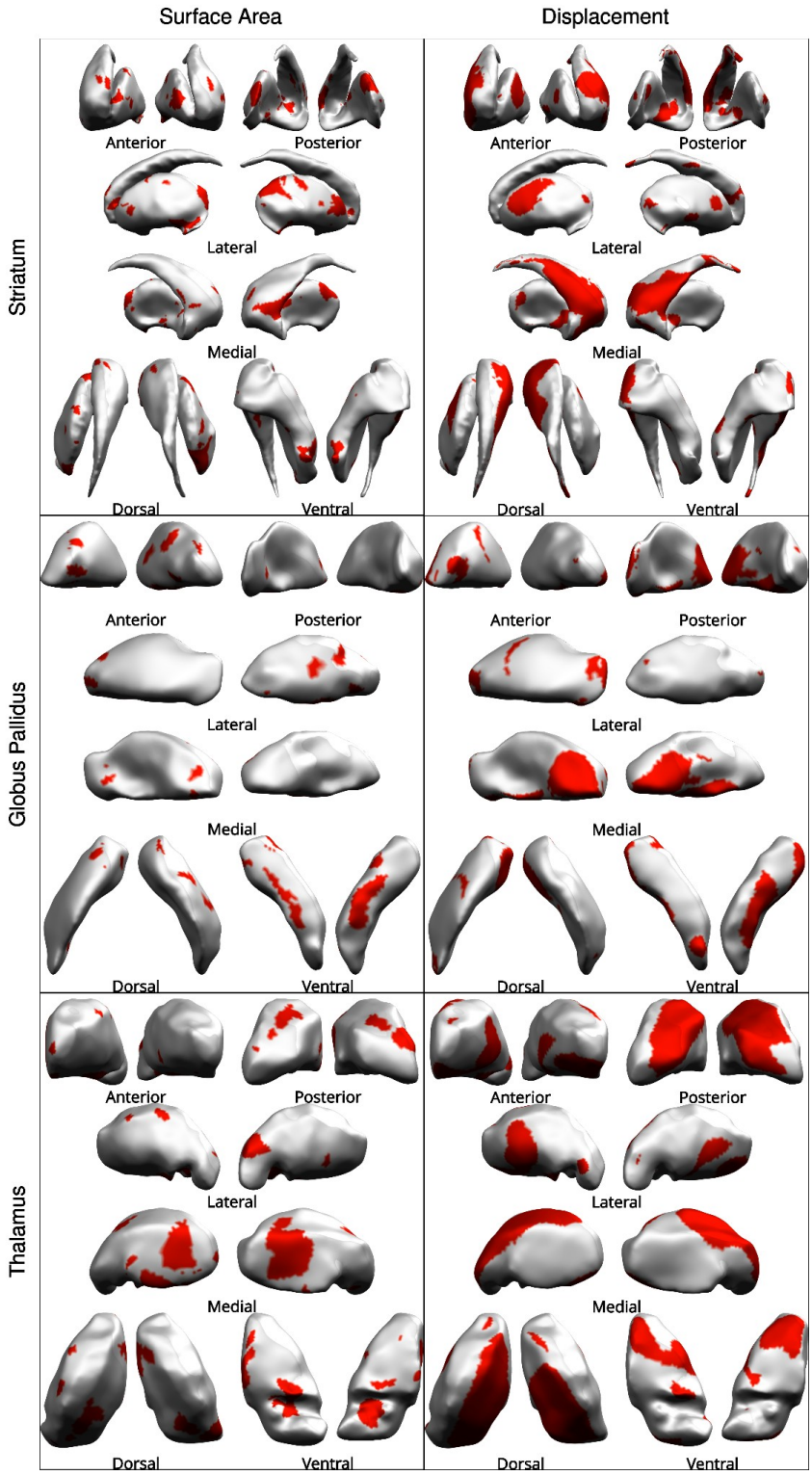

Figure 6: Sex maps: Regions on the surface of left and right striatum, thalamus, and globus pallidus in which the inclusion of sex and sex-by-diagnosis terms improved model fit, for linear models of surface area and displacement on diagnosis, when controlling for age. Red indicates improved fit with sex terms.

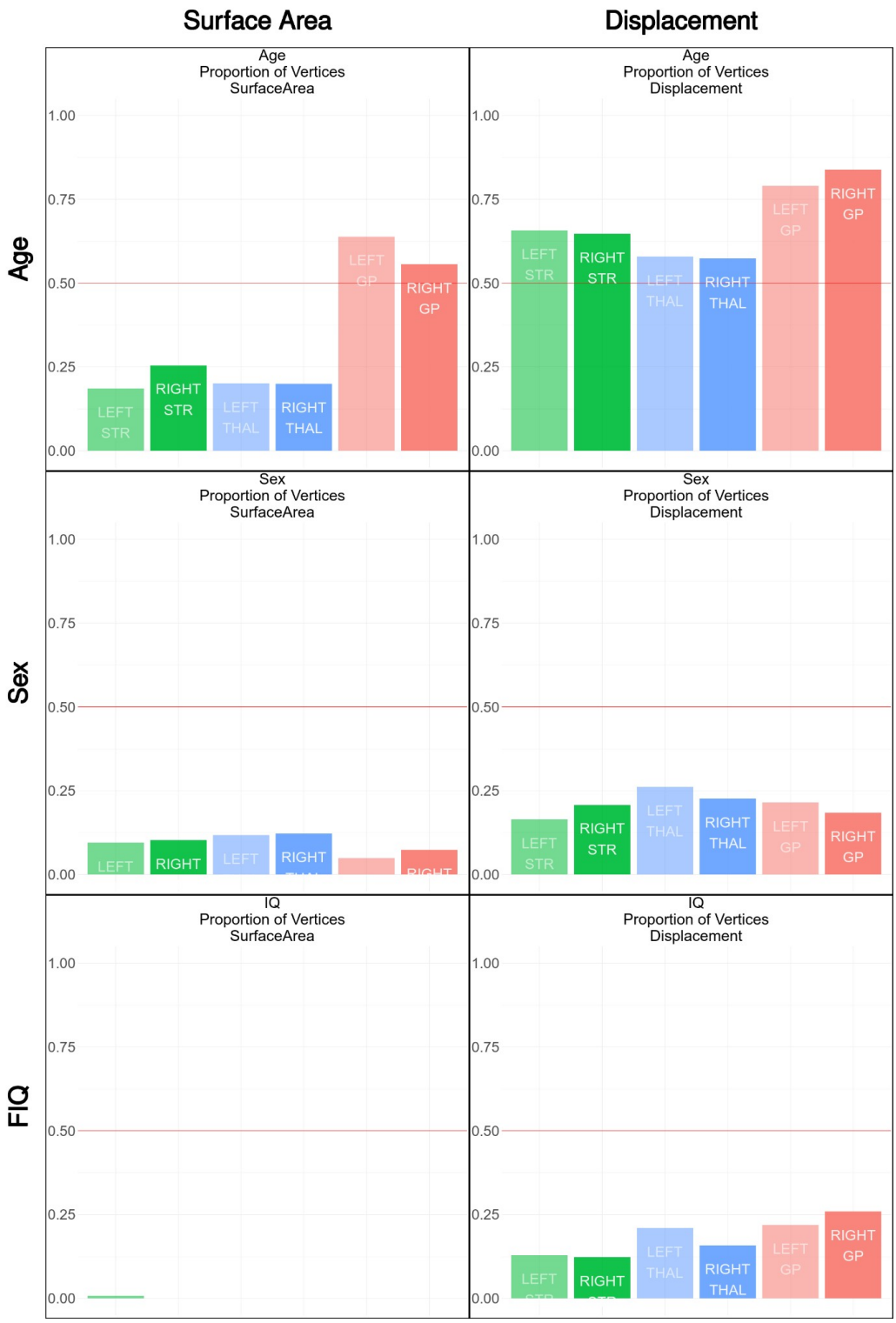

Figure 7: Proportion of vertices for which sex, age, and FIQ improved model fit across left (light) and right (dark) striatum (green), thalamus (blue), and globus pallidus (red). Linear models of main effect of ASD diagnosis on vertex-wise surface area (left column) and displacement (right column). Red line indicates 50%.

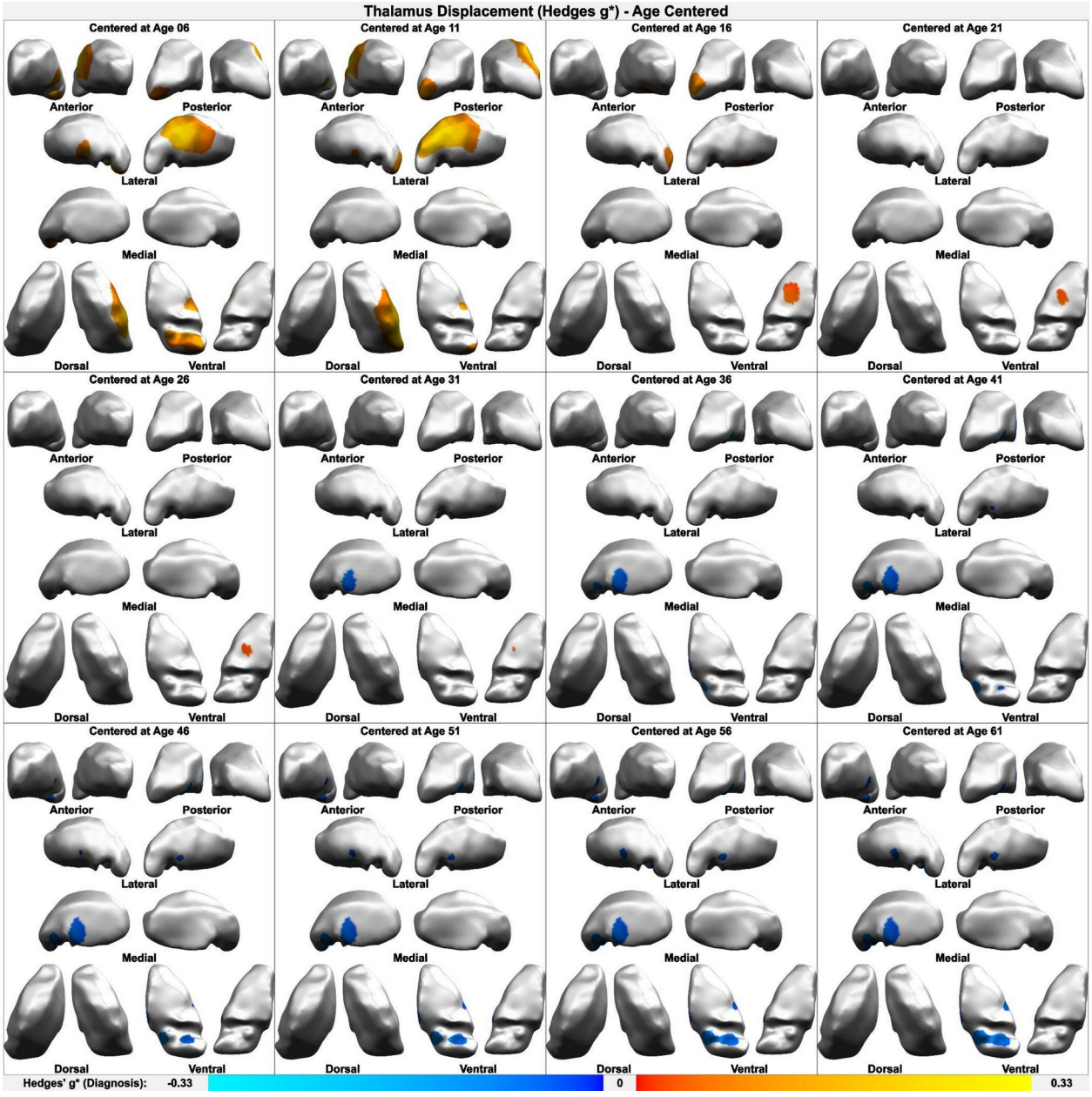

Figure 8: Age-centered analysis, centered on five-year intervals from ages 6-61. Hedges' *g** main effect of ASD diagnosis on vertex-wise displacement in the left and right thalamus, when controlling for total structure volume and sex. Warm colours indicate positive effects, cool colours indicate negative effects (*g** range -0.33 to 0.33). Displacement represents relative convexity (positive) or concavity (negative).

5.6 Linear mixed model and ComBat-harmonized mega-analyses confirm results, are less sensitive

A linear mixed model mega-analysis, including site as a random effect, indicated only two small regions of significant effects of ASD diagnosis after multiple comparisons correction, which roughly comprise a subset of the regions of significant effects detected using meta-analysis as described above. These were a region of reduced convexity around the left pulvinar, and a region of areal contraction around the right posterior pole of the globus pallidus. These results are shown in Figure 9.

Modeling the same data using linear models, after harmonizing across sites using ComBat, while preserving variation due to age and sex, as well as due to age, sex, and FIQ, reproduced the same general patterns, however none of the effects survived correction for multiple comparisons (Figures 10, 11).

5.7 No association between symptom severity and subcortical volumes or morphometry

No association was found between scores on the ADOS-CSS and subcortical volumes, vertex-wise surface area, or vertex-wise displacement ($p > .05$ for all structures). Only six sites (ABIDE-KKI, ABIDE-NYU, ABIDE-OHSU, ABIDE-SDSU, ABIDE-UM, Toronto) reported ADOS-CSS scores, so this analysis was performed with n=241 after motion QC and site exclusion.

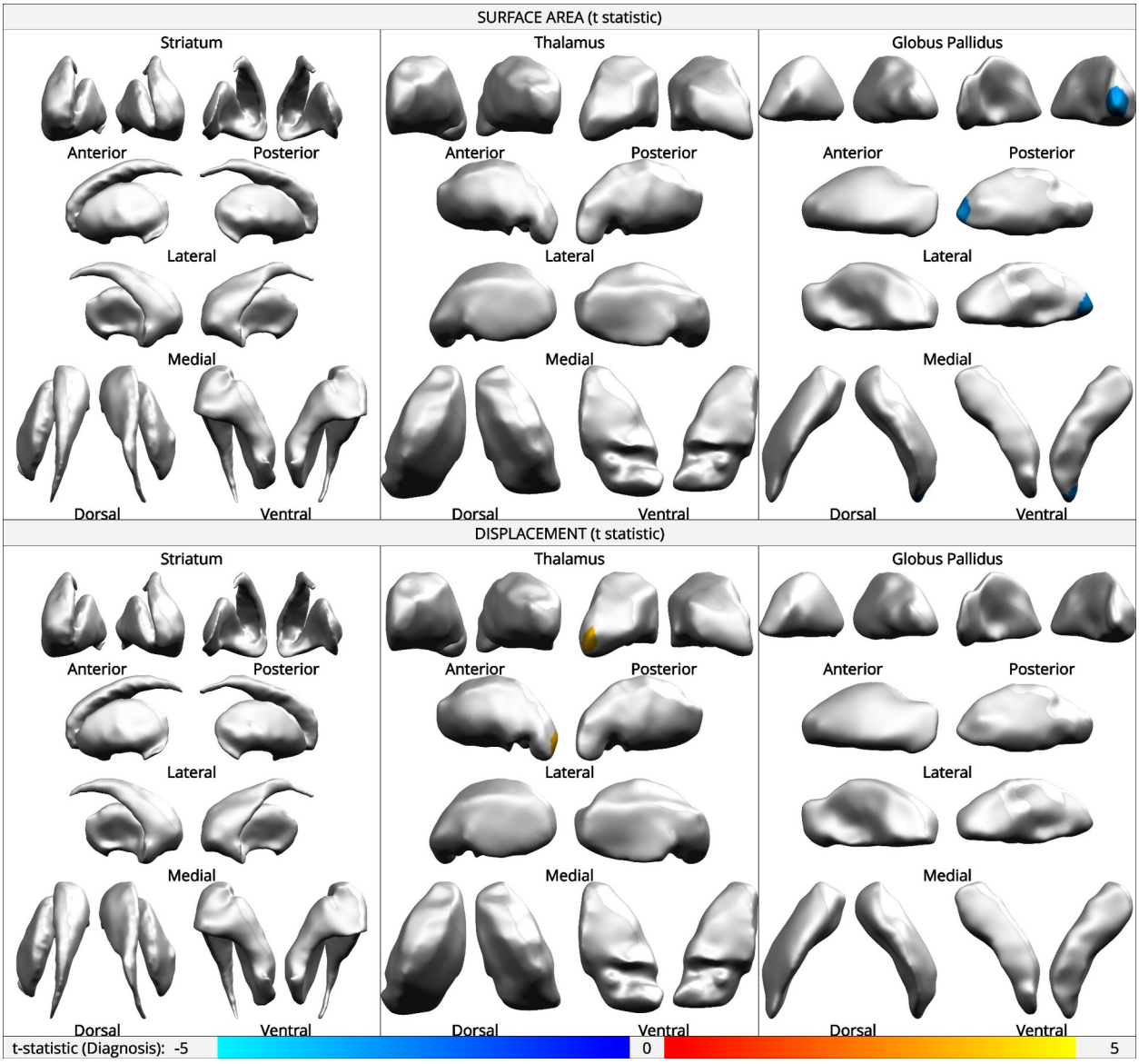

Figure 9: Linear-mixed model mega-analysis. Main effect of ASD diagnosis on vertex-wise surface area and displacement in the left and right striatum, thalamus, and globus pallidus, when controlling for total structure volume, age, and sex. Warm colours indicate positive effects, cool colours indicate negative effects (t-value range -5 to 5). Surface area is the Voronoi area surrounding a vertex; displacement represents relative convexity (positive) or concavity (negative). Thresholded at FDR < .05.

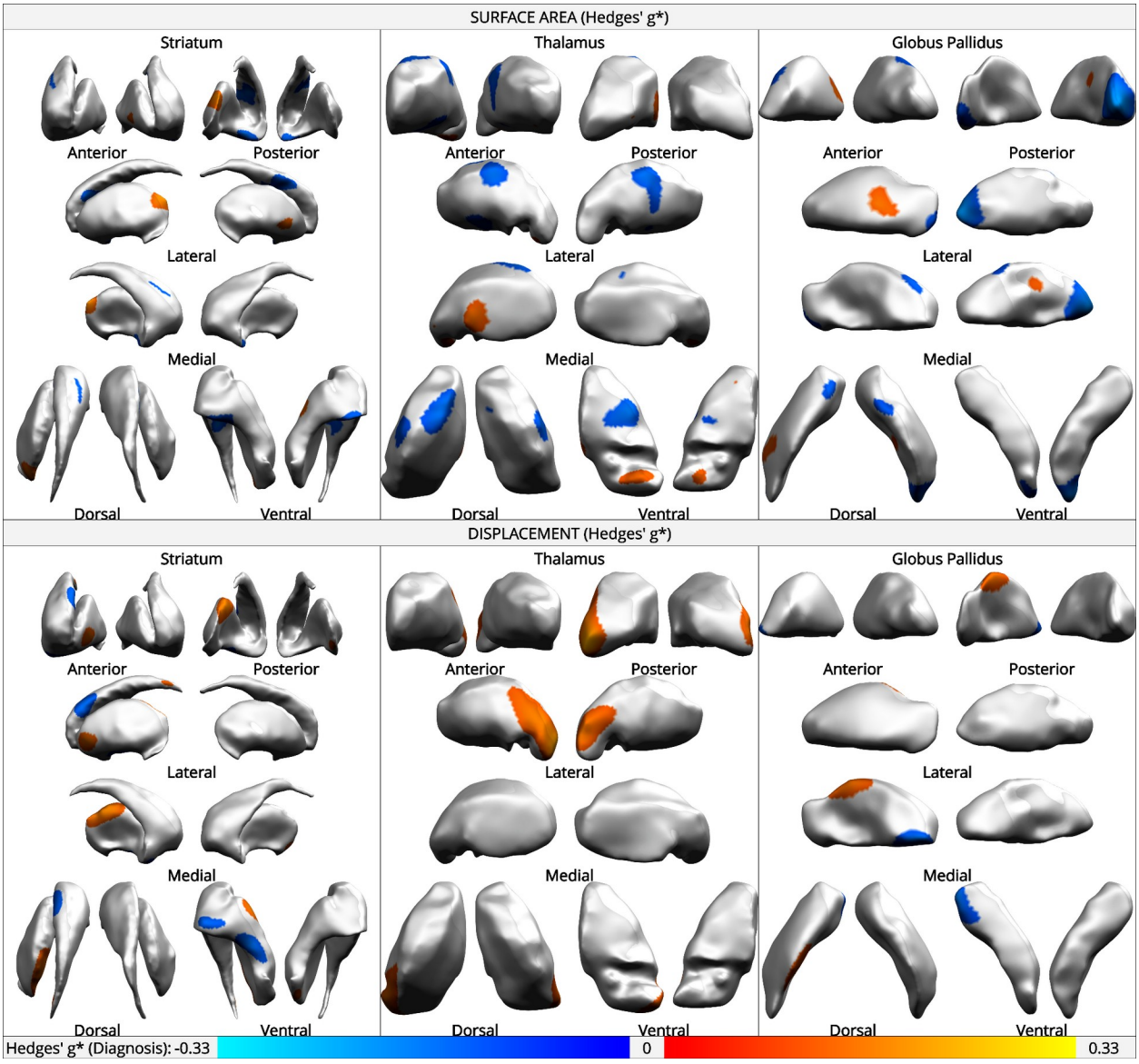

Figure 10: Uncorrected for multiple comparisons, ComBat-harmonized mega-analysis without controlling for FIQ. Main effect of ASD diagnosis, on vertex-wise surface area and displacement in the left and right striatum, thalamus, and globus pallidus, when controlling for total structure volume, age, and sex. Warm colours indicate positive effects, cool colours indicate negative effects (Hedges' g* range -0.3 to 0.3). Surface area is the Voronoi area surrounding a vertex; displacement represents relative convexity (positive) or concavity (negative). Thresholded at $p < .05$. No effects survived FDR multiple comparisons correction.

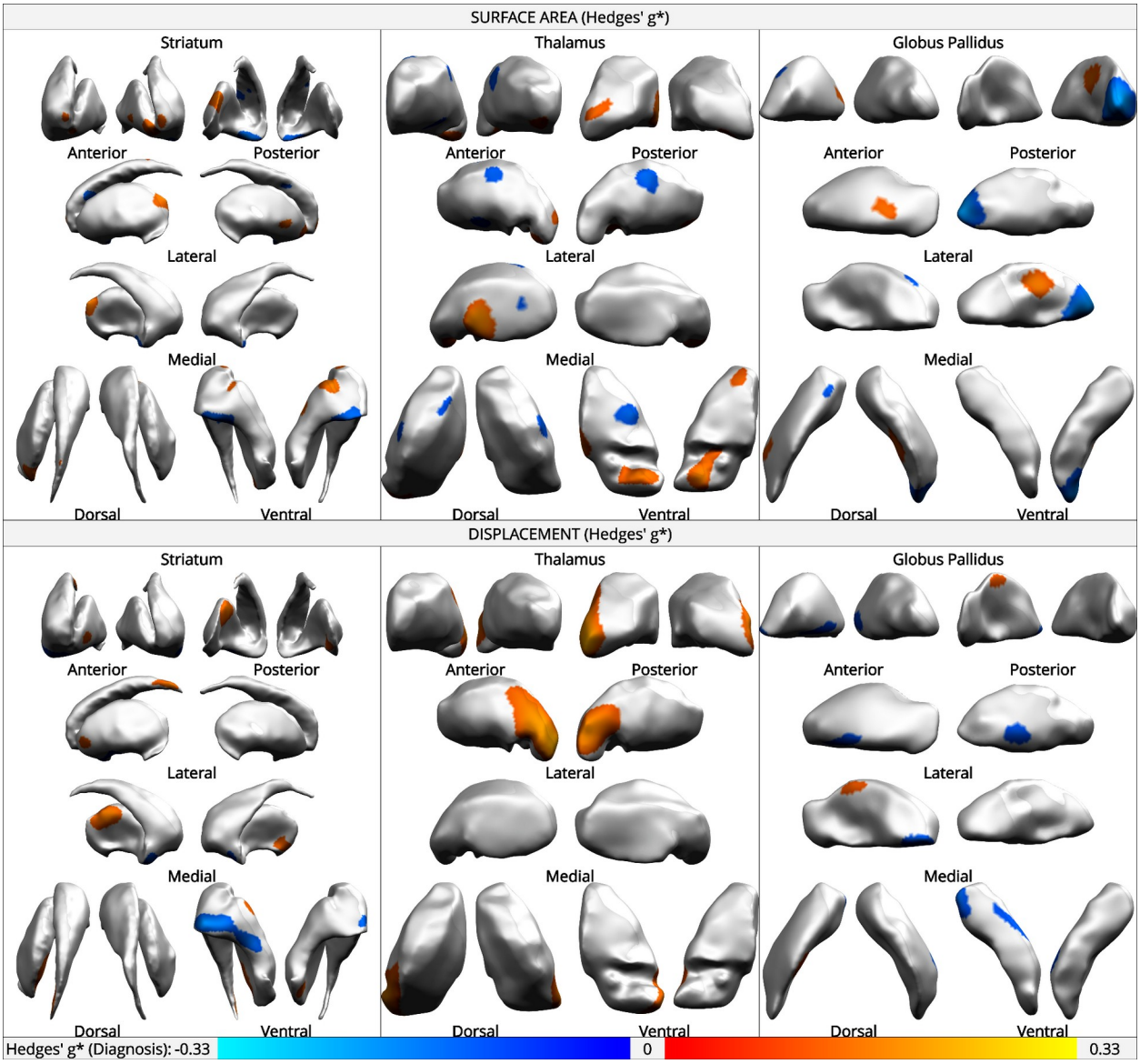

Figure 11: Uncorrected for multiple comparisons, ComBat-harmonized mega-analysis, controlling for FIQ. Main effect of ASD diagnosis, on vertex-wise surface area and displacement in the left and right striatum, thalamus, and globus pallidus, when controlling for total structure volume, age, sex, and FIQ. Warm colours indicate positive effects, cool colours indicate negative effects (Hedges' *g** range -0.3 to 0.3). Surface area is the Voronoi area surrounding a vertex; displacement represents relative convexity (positive) or concavity (negative). Thresholded at $p < .05$. No effects survived FDR multiple comparisons correction.

6. Discussion

In this study we examined volumetric and morphometric features of the striatum, globus pallidus, and thalamus in a large, multi-site, cross-sectional MRI dataset containing both males and females, with individuals from 5 to 65 years of age. We found no volumetric differences between the ASD and TD groups in any of the structures, but did find several regions of altered surface area and convexity in all three structures. Furthermore, age was an important explanatory variable across more than 50% of all surfaces when considering convexity/concavity, and across more than 50% of the globus pallidus when considering surface area. Sex and FIQ were found to be important explanatory variables across 10-25% of vertices, when considering displacement, and across fewer than 10% of vertices when considering surface area. No association was found between symptom severity and of the volumetric or morphometric measures. Overall, our findings help confirm that the neurodevelopment of the striatum, globus pallidus and thalamus are affected in ASD, in subtle ways that are not consistent across space or time.

Our findings suggest that it is important to take age into account when looking for neuroanatomical differences associated with ASD, because it contributes to neuroanatomical heterogeneity. While we did not find volumetric differences in ASD in any of the structures across the lifespan, we did find a complex pattern of spatially localized group differences that were highly dependent upon age. Normative studies have shown that, in typically developing individuals, the thalamus, striatum, and globus pallidus do not undergo uniform growth, but rather show complex patterns of localized growth and contraction through childhood, adolescence, and adulthood (Raznahan et al., 2014; Tullo et al., 2019). Any effects that are detected may therefore be highly dependent on the age of participants in the sample, and this may, in part, explain why the literature on this question is so discordant. It may be that localized

expansion and contraction in different regions has a cancellation effect, such that overall volume changes may be very slight and may depend on the ages of the participants in any given sample. Also, these growth patterns differ between structures, and change over the lifetime. The volume of the striatum, for instance, tends to peak during childhood or adolescence and then decrease monotonically throughout adulthood, whereas thalamic volume can remain stable for decades (Dima et al., 2022; Tullo et al., 2019; Wierenga et al., 2014), though the timing of peak volume attainment can differ between the sexes (Raznahan et al., 2014).

Overall, although we found no group differences in volumes at any point in the lifespan, we did find that localized surface area and displacement group differences were more pronounced in childhood and adolescence, and attenuated and occupying different regions in adulthood. In the thalamus, the overall pattern was a shift from relatively more convex shape in ASD in childhood, to relatively less convex/more concave shape in ASD in later adulthood, as well as regions of relative areal contraction in ASD in childhood, particularly in the left anterior thalamus, which showed the reverse patterns beginning in late middle age (Figure SF-3). In the globus pallidus, the effects were quite asymmetrical, with regions of surface area expansion on the lateral face of the left globus pallidus in childhood, while the dominant pattern on the right was of relative areal contraction in ASD at the anterior and posterior poles during the same time period. Through adulthood, the right globus pallidus shows no diagnostic effects, while there is a region of relative areal expansion on the left medial wall. Nevertheless, none of these effects were sufficiently large to register as volumetric differences in ASD in any of the structures studied here.

While the under-representation of females in ASD studies is improving, the low prevalence of females with ASD has led to the disorder being understudied in females, despite

evidence that sex modulates the effect of ASD diagnosis on neuroanatomy (Bedford et al., 2019; Lai et al., 2017). Consequently, there is very little evidence regarding sexual dimorphism in ASD in the structures studied here. One recent large study performed by the ENIGMA consortium found no sex-by-diagnosis interaction effect on the volumes of any of these structures (van Rooij et al., 2018), which is consistent with our results. The picture is less clear when looking at localized morphometry: including sex in the model did improve fit over between 5% and 10% of vertices when evaluating relative differences in surface area, and between 15% and 25% of vertices for relative displacement differences (see Figure 7). In particular, the medial caudate bilaterally, and much of the dorsolateral thalamus, when considering the effect of ASD on localized displacement, may exhibit qualitatively sexually dimorphic effects (Figure 6), despite the lack of evidence for a sex-by-diagnosis interaction in their volumes.

In addition to sex, FIQ is often unexamined in studies of ASD, despite evidence that differences in FIQ are associated with subcortical neuroanatomical differences. Various studies have found correlations between FIQ and regional volumes of both cortex and subcortical structures (Burgaleta et al., 2014; Colom et al., 2013; Grazioplene et al., 2015). We did not evaluate this directly in the structures under study, but rather asked whether accounting for FIQ improves our modelling of the effect of ASD on their volume and morphometry. We did not find any evidence that this is the case, which is consistent with several recent reports (Schuetze et al., 2016; van Rooij et al., 2018).

We used the ADOS-2 CSS as a measure of symptom severity, to evaluate the relationship with our measures of subcortical volumes and morphometry. Unfortunately, there was no consistent measure available in all of the datasets: a variety of ADOS versions and modules were used. This resulted in a relatively underpowered study, with data from a maximum of n=241

participants. Our finding of no relationship in any of the structures is at odds with the findings of other groups, which have found associations between restricted, repetitive behaviours and pallidal surface area (Schuetze et al., 2016), as well as growth rates of the caudate (Langen et al., 2014). However, the fact that our findings diverge from those in the literature may be due to the relatively low statistical power in this portion of the study, for which the available dataset comprised less than a fifth of data available for the other analyses reported here.

A number of limitations of this study should be considered. First, the data used was compiled from many smaller datasets, each collected for other purposes. There is therefore no harmonization between datasets in terms of sample characteristics, inclusion/exclusion criteria, MRI hardware, software, or acquisition parameters, or behavioural measures captured. To allow for meaningful statistical analysis, we used a meta-analytic technique that has been used successfully before (Bedford et al., 2019; Olafson et al., 2021), and followed up with two other commonly used data combination techniques: linear mixed models and ComBat harmonization. A related issue is the paucity of behavioural measures that were available across multiple sites, and the lack of other potentially relevant demographic information, such as socioeconomic status. There is some evidence that accounting for such behavioural and demographic heterogeneity improves the sensitivity of surface-based morphological measures, at least in small samples (Qiu et al., 2010). Large datasets that include consistent behavioural measures, such as the Quebec 1000 Families cohort now underway (q1k.ca), will make possible large studies that account for behavioural heterogeneity in a more comprehensive way.

The cross-sectional nature of the data also limits to some degree the scope of interpretation of these results. While some datasets did include longitudinal scans, for the majority of participants only a single timepoint was available. This makes it difficult to draw any

direct conclusions about how neurodevelopmental trajectories may be altered in ASD. And, though the dataset includes participants from ages 5 to 65, it is heavily weighted towards the younger end of the spectrum, with relatively sparse representation in middle age and beyond. Finally, after quality control and removal of sites with too few females to allow for statistical analysis, we were able to retain between 31% and 40% of participants, depending on the structure. This is an unfortunately high rate of data attrition, however, given that including relatively poor quality scans can introduce artifactual effects (Bedford et al., 2019; Ducharme et al., 2016), it was a necessary step. That said, while the effect of motion on cortical measures is well documented, particularly at distal regions such as the orbitofrontal cortex and temporal poles, the effect on the subcortical morphometric measures used here are currently less well understood and, considering their central location in the brain, may not be as drastic.

7. Conclusion

The subcortex remains understudied in ASD in relation to the cortical sheet, and yet subcortical structures like the basal ganglia and thalamus play a central role in a variety of behaviours that are common in ASD. Even when cortical systems are a target of study, the thalamus and basal ganglia are relevant because both mediate and modulate a large amount of cortico-cortical communication. Unfortunately, the results of neuroimaging studies of both cortical and subcortical structures in ASD have been discordant, and it has been difficult to find a consistent narrative. This may be due to biological heterogeneity that reflects the underlying genetic heterogeneity in ASD.

This study attempts to parse potential sources of this heterogeneity by examining the role of sex, age, and IQ in neuroanatomical alterations in the thalamus, striatum, and globus pallidus in ASD, using a combination of techniques from statistics and information theory. It does this using a large dataset that includes both males and females and that covers the lifespan from early childhood to the seventh decade. This dataset provides the power required to find subtle neuroanatomical differences, but presents certain technical challenges in combining data between sites while minimizing the effect of non-biological, inter-site differences. We overcame these challenges by using a prospective meta-analytic technique, allowing us first to model the effect of ASD within each site, and then combine the effects across sites.

This meta-analytic technique was used to examine subcortical volumes, then was extended and used in a series of surface-based, vertex-wise analyses that have been shown to be sensitive to localized alterations of shape that cannot be detected using global volumetric measurements. To our knowledge, this is the first time that this combination of techniques has been used to investigate subtle alterations in subcortical structure in ASD. We found no

volumetric differences in ASD in any of the structures. However, we found a number of localized surface area and shape differences in all three structures, which were not apparent in the volumetric analysis. Also, age was an important variable across more than half of most surfaces. Follow-up age-centered analysis indicated a variety of age-dependent regional differences. We found no correlations between symptom severity and neuroanatomical alterations in ASD, though this portion of the study was significantly underpowered compared to other analyses.

Overall, our findings help confirm that the neurodevelopment of the striatum, globus pallidus and thalamus are affected in ASD, in a subtle location-dependent manner that is not reflected in overall structure volumes, and that is highly non-uniform across the lifespan.

Tables

Table 1: Demographics (Included Sites)

	ASD		Typically Developing		
	Male	**Female**	**Male**	**Female**	**Age Range**
ABIDE - IP	14	8	12	22	6-46
ABIDE - KKI	58	19	123	65	8-12
ABIDE - MAX MUN	21	3	29	4	7-58
ABIDE - NYU	135	19	107	28	5-39
ABIDE - OHSU	43	7	42	29	7-15
ABIDE - SDSU	39	8	39	8	7-18
ABIDE - UM	58	10	59	18	8-28
CAM	39	17	20	20	12-18
Cambridge	29	32	32	33	18-49
IoP	43	22	41	21	18-52
TORONTO	106	25	196	194	4-65
Subtotal	**585**	**170**	**700**	**442**	
	755		**1142**		
Grand Total	**1897**				

Demographic data for all sites that were included in statistical analysis. Site codes are: CAM - Cambridge Family Study of Autism; Cambridge - UKAIMS Cambridge site; IoP - UKAIMS Institute of Psychiatry site; IP - ABIDE Institut Pasteur; KKI - ABIDE Kennedy Krieger Institute; MAX_MUN - ABIDE Ludwig Maximilians University Munich; OHSU - ABIDE Oregon Health and Science University; SDSU - ABIDE San Diego State University; TORONTO - SickKids / University of Toronto; UM - ABIDE University of Michigan.

Table 2: Demographics (Excluded Sites)

	ASD		Typically Developing		
	Male	**Female**	**Male**	**Female**	**Age Range**
ABIDE - BNI	29	0	29	0	18-64
ABIDE - CALTECH	15	4	15	4	17-56
ABIDE - CMU	11	3	10	3	19-40
ABIDE - EMC	22	5	22	5	6-10
ABIDE - ETH	13	0	24	0	13-30
ABIDE - GU	43	8	28	27	8-13
ABIDE - IU	16	4	15	5	17-54
ABIDE - KUL	28	0	0	0	18-35
ABIDE - LEUVEN	26	3	30	5	12-32
ABIDE - OLIN	17	3	14	2	10-24
ABIDE - ONRC	20	4	20	15	18-31
ABIDE - PITT	26	4	23	4	9-35
ABIDE - SBL	15	0	15	0	20-64
ABIDE - STANFORD	16	4	16	4	7-12
ABIDE - TCD	21	0	21	0	10-20
ABIDE - TRINITY	24	0	25	0	12-25
ABIDE - UCD	14	4	10	4	12-17
ABIDE - UCLA	63	7	50	11	7-17
ABIDE - USM	73	2	56	3	8-50
ABIDE - YALE	20	8	20	8	7-17
NIMH	68	17	29	16	1-9
Subtotal	**580**	**80**	**472**	**116**	
	660		**588**		
Grand Total	**1248**				

Demographic information for sites that were excluded from statistical analysis, due either to insufficient females in the ASD or control groups after quality control, or failed segmentation. Site codes are: BNI - Barrow Neurological Institute; CALTECH - California Institute of Technology; CMU - Carnegie Mellon University; EMC - Erasmus University Medical Center Rotterdam; ETH - ETH Zürich; GU - Georgetown University; IU - Indiana University; KUL - Katholieke Universiteit Leuven; LEUVEN - University of Leuven; OLIN - Olin, Institute of Living at Hartford Hospital; ONRC - Olin Neuropsychiatry Research Center, Institute of Living at Hartford Hospital; PITT - University of Pittsburgh School of Medicine; SBL - Social Brain Lab, BCN NIC UMC Groningen and Netherlands Institute for Neurosciences; STANFORD - Stanford University; TCD - Trinity Center for Health Sciences (Release 2); TRINITY - Trinity Center for Health Sciences (Release 1); UCD - University of California, Davis; UCLA - University of California, Los Angeles; USM - University of Utah School of Medicine; YALE - Yale Child Study Center; NIMH - National Institute of Mental Health

Table 3: Quality Control Results, Original Workflow

	Total	Motion Passed	L Str Passed	R Str Passed	L GP Passed	R GP Passed	L Thal Passed	R Thal Passed
ABIDE - IP	56	39	26	31	38	38	38	38
ABIDE - KKI	265	156	119	117	156	156	156	156
ABIDE - MAX MUN	57	41	28	34	41	41	37	39
ABIDE - NYU	289	195	148	157	195	195	190	188
ABIDE - OHSU	121	98	77	76	98	98	96	95
ABIDE - SDSU	94	48	39	41	47	47	47	47
ABIDE - UM	145	66	44	46	66	66	62	63
CFSA	96	57	39	35	57	57	53	54
UKAIMS - Cambridge	126	117	82	81	114	114	109	110
UKAIMS - IoP	127	110	68	77	105	105	99	97
TORONTO	521	404	254	270	402	387	366	374
Total	1897	1331	924	965	1319	1304	1253	1261

Segmentation results for original MAGeT Brain workflow, showing the number of participants per site, including total number, number that passed motion QC, and number that passed segmentation QC for each structure. Only sites that were not excluded are shown.

Table 4: Quality Control Results, Improved Workflow

	Total	Motion Passed	L Str Passed	R Str Passed	L GP Passed	R GP Passed	L Thal Passed	R Thal Passed
ABIDE - IP	56	39	35	37	38	38	38	37
ABIDE - KKI	265	156	118	101	156	156	155	155
ABIDE - MAX MUN	57	41	37	39	41	41	41	40
ABIDE - NYU	289	195	177	170	195	195	192	190
ABIDE - OHSU	121	98	79	76	98	98	95	97
ABIDE - SDSU	94	48	40	40	47	47	47	47
ABIDE - UM	145	66	56	53	66	66	65	65
CFSA	96	57	44	38	57	57	55	55
UKAIMS - Cambridge	126	117	94	86	114	114	109	109
UKAIMS - IoP	127	110	74	82	103	103	99	100
TORONTO	521	404	336	323	403	403	397	397
Total	1897	1331	1090	1045	1318	1318	1293	1292
w/FIQ		1220	1002	953	1216	1216	1194	1192

Segmentation results for improved MAGeT Brain workflow (left-right split runs and subcortical masking), showing the number of participants per site, including total number, number that passed motion QC, and number that passed segmentation QC for each structure. Only sites that were not excluded are shown. Last row indicates the number of participants with FIQ data available. Overall segmentation accuracy improved compared to the initial workflow (Table 3), particularly in the striatum.

Supplementary Figures

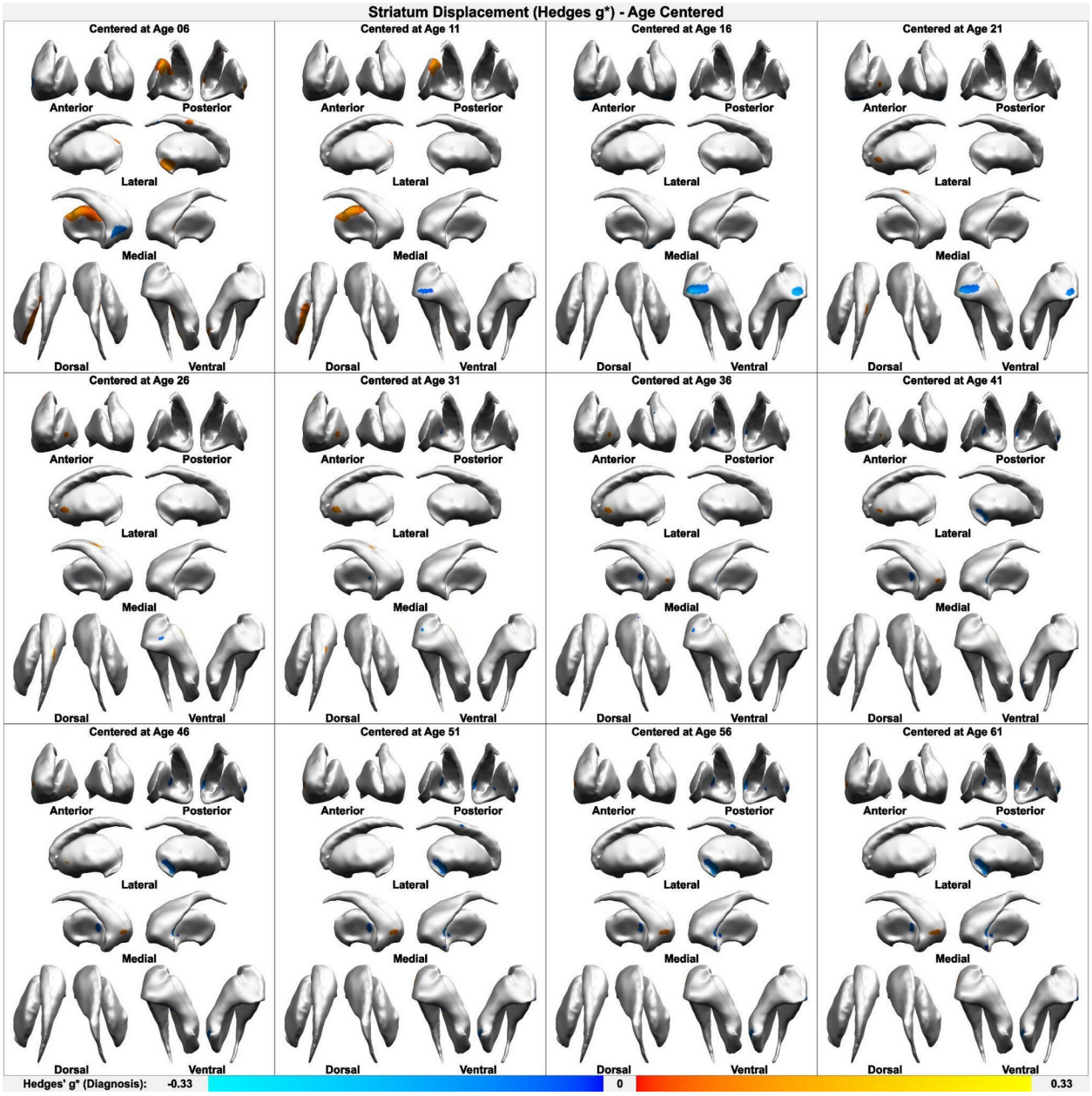

Figure SF-1: Age-centered analysis, centered on five-year intervals from ages 6-61. Hedges' *g** main effect of ASD diagnosis on vertex-wise displacement in the left and right striatum, when controlling for total structure volume and sex. Warm colours indicate positive effects, cool colours indicate negative effects (*g** range -0.33 to 0.33). Displacement represents relative convexity (positive) or concavity (negative).

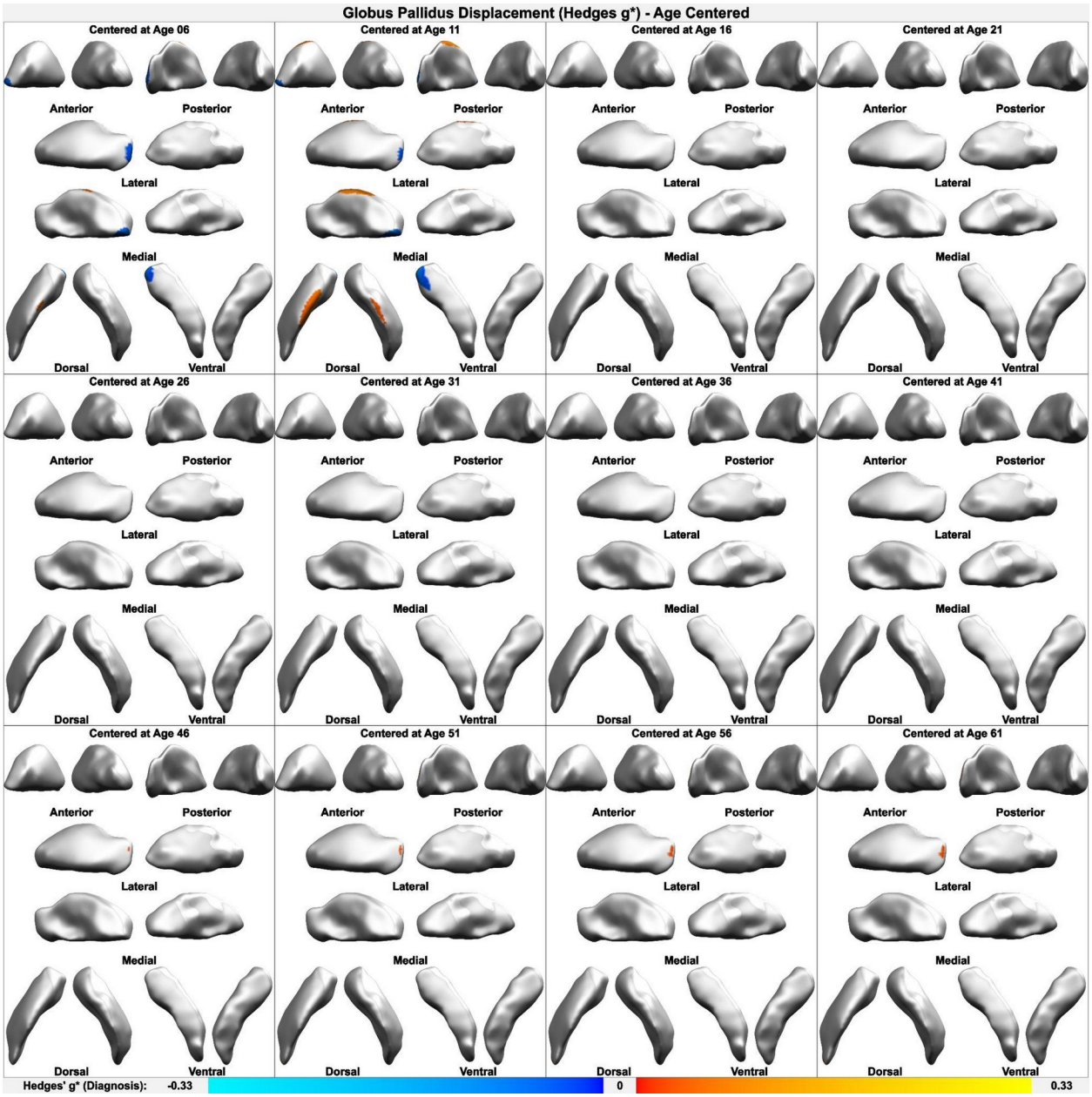

Figure SF-2: Age-centered analysis, centered on five-year intervals from ages 6-61. Hedges' *g** main effect of ASD diagnosis on vertex-wise displacement in the left and right globus pallidus, when controlling for total structure volume and sex. Warm colours indicate positive effects, cool colours indicate negative effects (*g** range -0.33 to 0.33). Displacement represents relative convexity (positive) or concavity (negative).

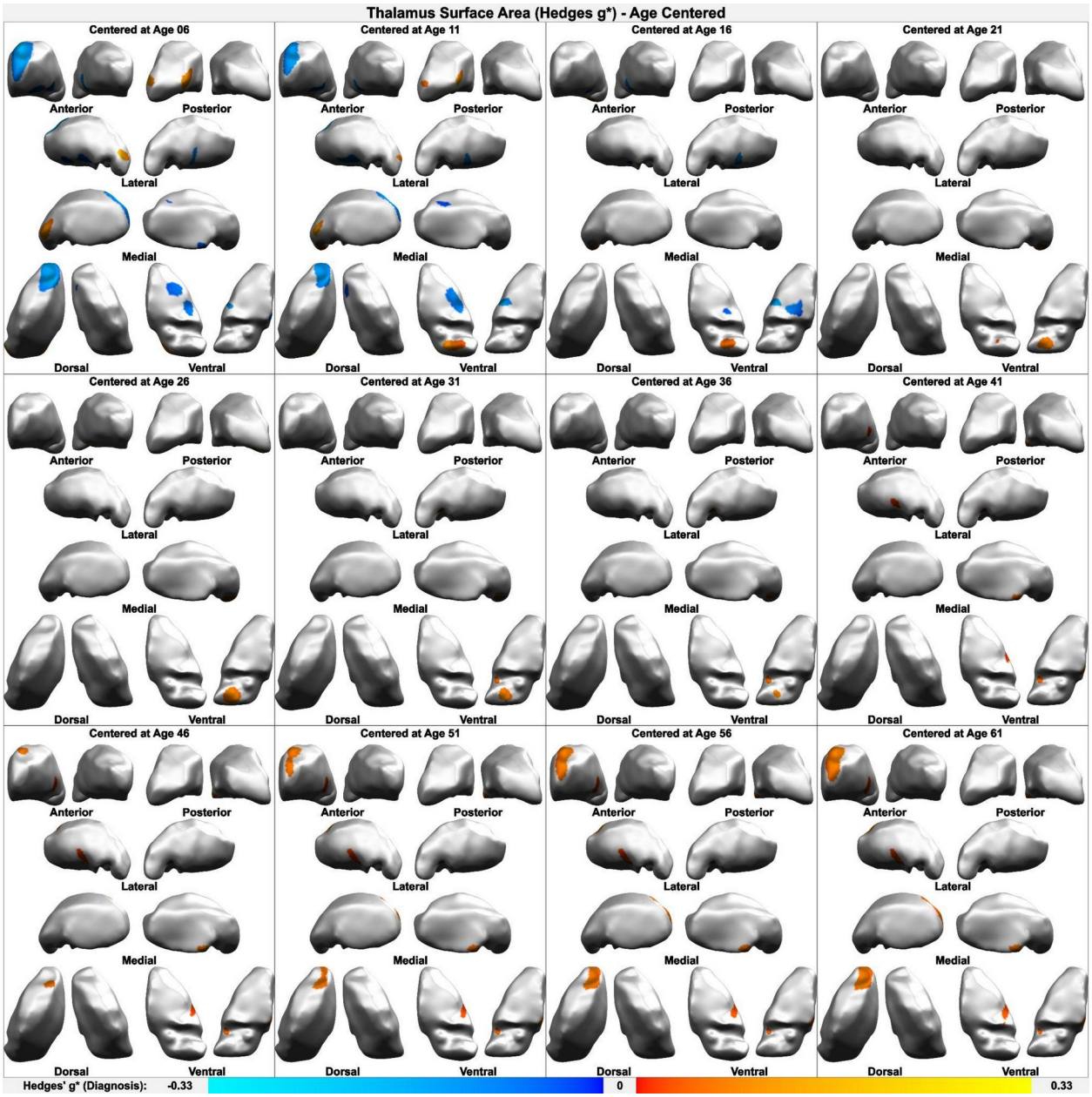

Figure SF-3: Age-centered analysis, centered on five-year intervals from ages 6-61. Hedges' *g** main effect of ASD diagnosis on vertex-wise surface area in the left and right thalamus, when controlling for total structure volume and sex. Warm colours indicate positive effects, cool colours indicate negative effects (*g** range -0.33 to 0.33). Surface area is the Voronoi area surrounding a vertex.

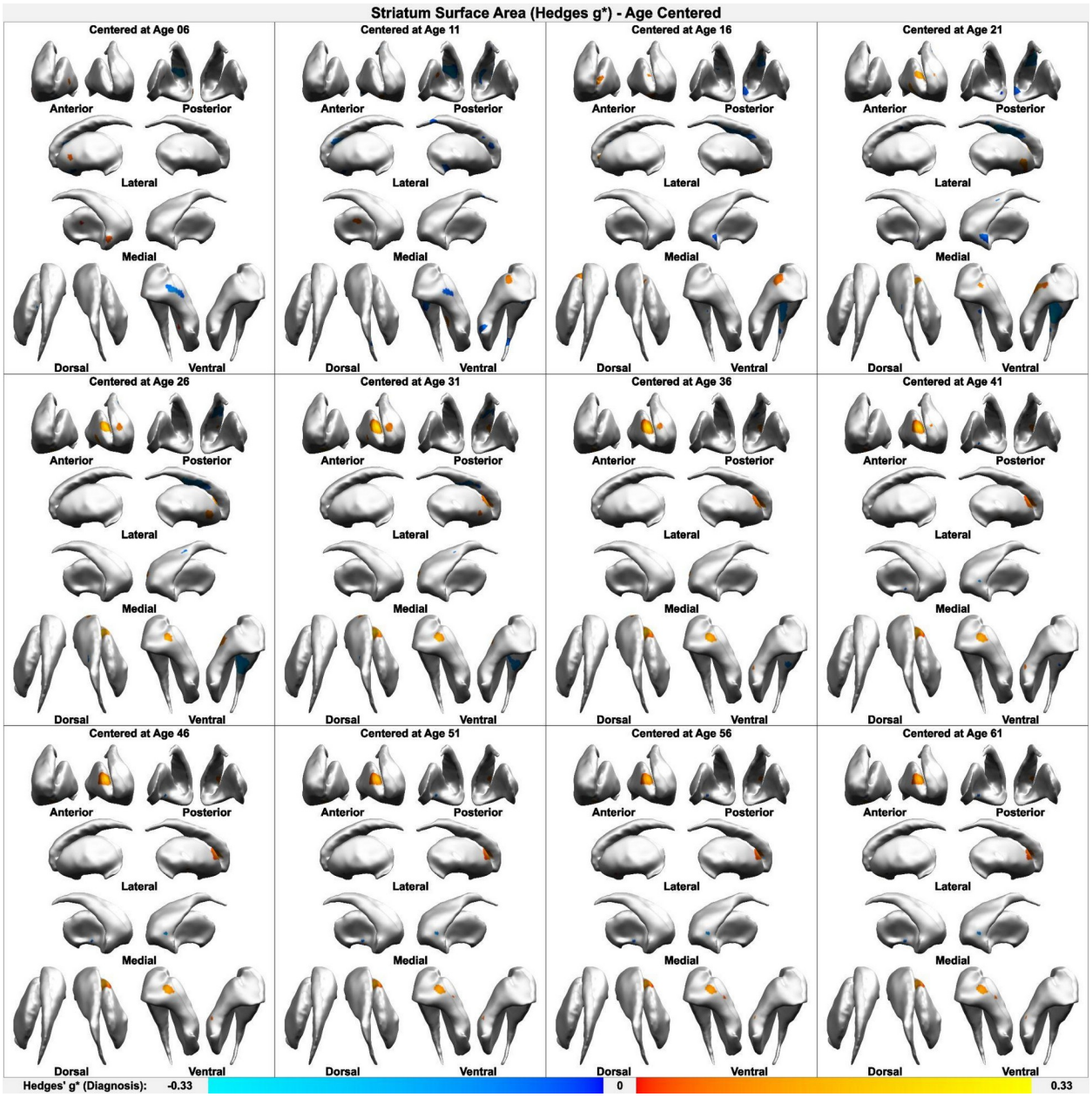

Figure SF-4: Age-centered analysis, centered on five-year intervals from ages 6-61. Hedges' *g** main effect of ASD diagnosis on vertex-wise surface area in the left and right striatum, when controlling for total structure volume and sex. Warm colours indicate positive effects, cool colours indicate negative effects (*g** range -0.33 to 0.33). Surface area is the Voronoi area surrounding a vertex.

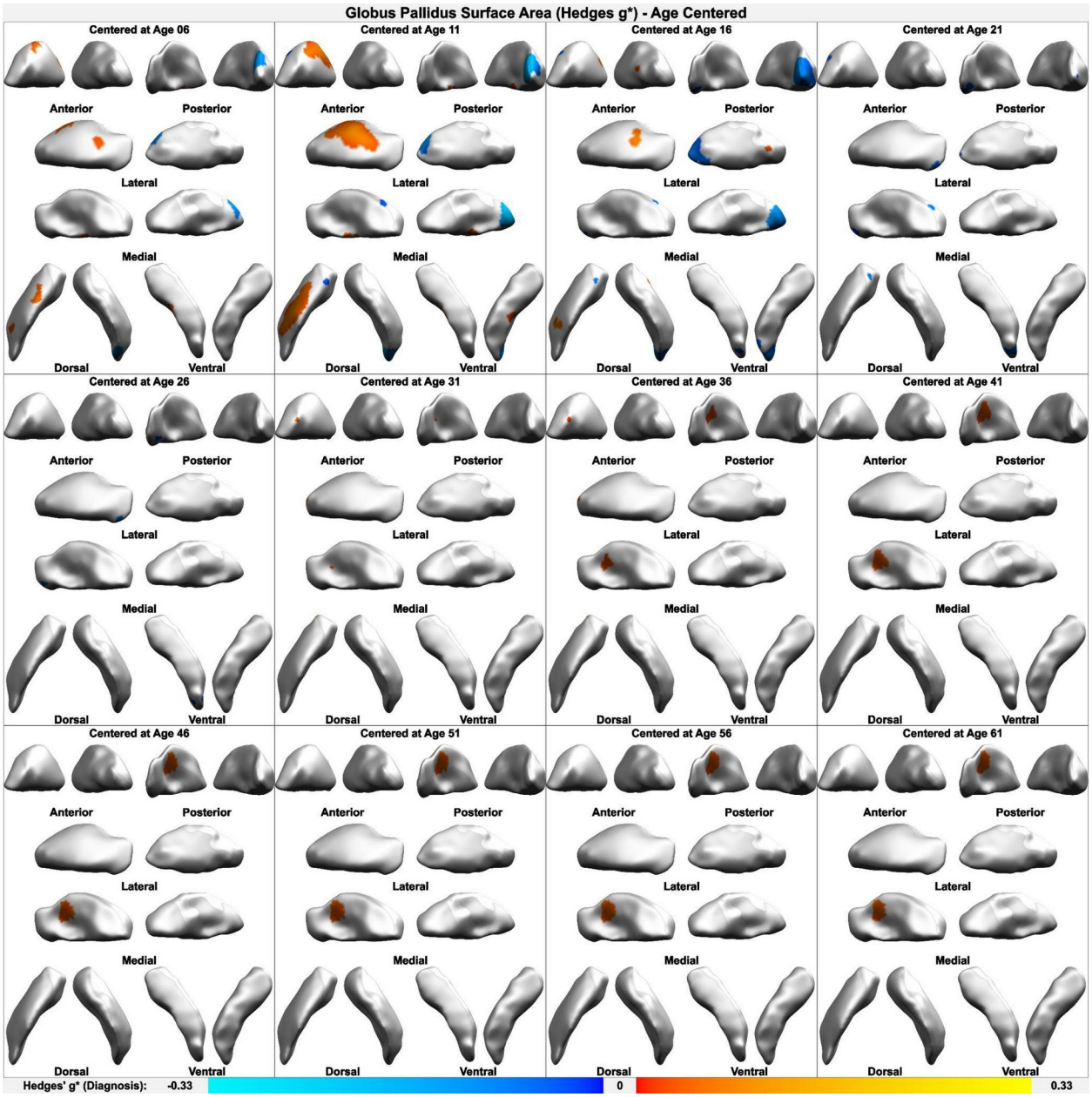

Figure SF-5: Age-centered analysis, centered on five-year intervals from ages 6-61. Hedges' *g** main effect of ASD diagnosis on vertex-wise surface area in the left and right globus pallidus, when controlling for total structure volume and sex. Warm colours indicate positive effects, cool colours indicate negative effects (*g** range -0.33 to 0.33). Surface area is the Voronoi area surrounding a vertex.

References

Akaike, H. (1974). A New Look at the Statistical Model Identification. In *Springer Series in Statistics* (pp. 215–222). https://doi.org/10.1007/978-1-4612-1694-0_16

Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neurosciences*, *13*(7), 266–271.

Alexander, G. E., Crutcher, M. D., & DeLong, M. R. (1991). Chapter 6 Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. In H. B. M. Uylings, C. G. Van Eden, J. P. C. De Bruin, M. A. Corner, & M. G. P. Feenstra (Eds.), *Progress in Brain Research* (Vol. 85, pp. 119–146). Elsevier.

Alkemade, A., Keuken, M. C., & Forstmann, B. U. (2013). A perspective on terra incognita: uncovering the neuroanatomy of the human subcortex. *Frontiers in Neuroanatomy*, *7*, 40.

Aloe, A. M., & Becker, B. J. (2012). An Effect Size for Regression Predictors in Meta-Analysis. *Journal of Educational and Behavioral Statistics: A Quarterly Publication Sponsored by the American Educational Research Association and the American Statistical Association*, *37*(2), 278–297.

Ameis, S. H., & Catani, M. (2015). Altered white matter connectivity as a neural substrate for social impairment in Autism Spectrum Disorder. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *62*, 158–181.

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5-TR*. American Psychiatric Association Publishing.

Avants, B. B., Tustison, N. J., Stauffer, M., Song, G., Wu, B., & Gee, J. C. (2014). The Insight ToolKit image registration framework. *Frontiers in Neuroinformatics*, *8*, 44.

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. (2019). Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and

cortical morphometry in autism spectrum disorder. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-019-0420-6

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.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111.

Boucher, J. (2008). *The Autistic Spectrum: Characteristics, Causes and Practical Issues*. SAGE.

Brieber, S., Neufang, S., Bruning, N., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., Fink, G. R., & Konrad, K. (2007). Structural brain abnormalities in adolescents with autism spectrum disorder and patients with attention deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *48*(12), 1251–1258.

Burgaleta, M., MacDonald, P. A., Martínez, K., Román, F. J., Álvarez-Linera, J., Ramos González, A., Karama, S., & Colom, R. (2014). Subcortical regional morphology correlates with fluid and spatial intelligence. *Human Brain Mapping*, *35*(5), 1957–1968.

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.). Springer-Verlag.

Bussy, A., Patel, R., Plitman, E., Tullo, S., Salaciak, A., Bedford, S. A., Farzin, S., Béland, M.-L., Valiquette, V., Kazazian, C., Tardif, C. L., Devenyi, G. A., & Chakravarty, M. M. (2021). Hippocampal shape across the healthy lifespan and its relationship with cognition. *Neurobiology of Aging*, *106*, 153–168.

Carrera, E., & Bogousslavsky, J. (2006). The thalamus and behavior: effects of anatomically distinct strokes. *Neurology*, *66*(12), 1817–1823.

Chakravarty, M. M., Bertrand, G., Hodge, C. P., Sadikot, A. F., & Collins, D. L. (2006). The creation of a brain atlas for image guided neurosurgery using serial histological data. *NeuroImage*, *30*(2), 359–376.

Chakravarty, M. M., Rapoport, J. L., Giedd, J. N., Raznahan, A., Shaw, P., Collins, D. L., Lerch, J. P., &

Gogtay, N. (2015). Striatal shape abnormalities as novel neurodevelopmental endophenotypes in schizophrenia: a longitudinal study. *Human Brain Mapping*, *36*(4), 1458–1469.

Chakravarty, M. M., Steadman, P., van Eede, M. C., Calcott, R. D., Gu, V., Shaw, P., Raznahan, A., Collins, D. L., & Lerch, J. P. (2013). Performing label-fusion-based segmentation using multiple automatically generated templates: MAGeT Brain: Label Fusion Segmentation Using Automatically Generated Templates. *Human Brain Mapping*, *34*(10), 2635–2654.

Chaste, P., & Leboyer, M. (2012). Autism risk factors: genes, environment, and gene-environment interactions. *Dialogues in Clinical Neuroscience*, *14*(3), 281–292.

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.

Colom, R., Burgaleta, M., Román, F. J., Karama, S., Alvarez-Linera, J., Abad, F. J., Martínez, K., Quiroga, M. Á., & Haier, R. J. (2013). Neuroanatomic overlap between intelligence and cognitive factors: morphometry methods provide support for the key role of the frontal lobes. *NeuroImage*, *72*, 143–152.

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.

Courchesne, E., Pramparo, T., Gazestani, V. H., Lombardo, M. V., Pierce, K., & Lewis, N. E. (2019). The ASD Living Biology: from cell proliferation to clinical phenotype. *Molecular Psychiatry*, *24*(1), 88–107.

Courchesne, E., Redcay, E., Morgan, J. T., & Kennedy, D. P. (2005). Autism at the beginning: microstructural and growth abnormalities underlying the cognitive and behavioral phenotype of autism. *Development and Psychopathology*, *17*(3), 577–597.

Dima, D., Modabbernia, A., Papachristou, E., Doucet, G. E., Agartz, I., Aghajani, M., Akudjedu, T. N.,

Albajes-Eizagirre, A., Alnaes, D., Alpert, K. I., Andersson, M., Andreasen, N. C., Andreassen, O. A., Asherson, P., Banaschewski, T., Bargallo, N., Baumeister, S., Baur-Streubel, R., Bertolino, A., … Karolinska Schizophrenia Project (KaSP). (2022). Subcortical volumes across the lifespan: Data from 18,605 healthy individuals aged 3-90 years. *Human Brain Mapping*, *43*(1), 452–469.

Di Martino, A., O'Connor, D., Chen, B., Alaerts, K., Anderson, J. S., Assaf, M., Balsters, J. H., Baxter, L., Beggiato, A., Bernaerts, S., Blanken, L. M. E., Bookheimer, S. Y., Braden, B. B., Byrge, L., Castellanos, F. X., Dapretto, M., Delorme, R., Fair, D. A., Fishman, I., … Milham, M. P. (2017). Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Scientific Data*, *4*, 170010.

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S. Y., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D. A., Gallagher, L., Kennedy, D. P., Keown, C. L., Keysers, C., … Milham, M. P. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, *19*(6), 659–667.

Ducharme, S., Albaugh, M. D., Nguyen, T.-V., Hudziak, J. J., Mateos-Pérez, J. M., Labbe, A., Evans, A. C., Karama, S., & Brain Development Cooperative Group. (2016). Trajectories of cortical thickness maturation in normal brain development--The importance of quality control procedures. *NeuroImage*, *125*, 267–279.

Elsabbagh, M., Divan, G., Koh, Y.-J., Kim, Y. S., Kauchali, S., Marcín, C., Montiel-Nava, C., Patel, V., Paula, C. S., Wang, C., Yasamy, M. T., & Fombonne, E. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research: Official Journal of the International Society for Autism Research*, *5*(3), 160–179.

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.

Estes, A., Shaw, D. W. W., Sparks, B. F., Friedman, S., Giedd, J. N., Dawson, G., Bryan, M., & Dager, S.

R. (2011). Basal ganglia morphometry and repetitive behavior in young children with autism spectrum disorder. *Autism Research: Official Journal of the International Society for Autism Research*, *4*(3), 212–220.

Folstein, S., & Rutter, M. (1977). Infantile autism: a genetic study of 21 twin pairs. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *18*(4), 297–321.

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.

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.

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., Mahajan, M., Manaa, D., Pawitan, Y., Reichert, J., Ripke, S., Sandin, S., Sklar, P., Svantesson, O., Reichenberg, A., Hultman, C. M., Devlin, B., Roeder, K., & Buxbaum, J. D. (2014). Most genetic risk for autism resides with common variation. *Nature Genetics*, *46*(8), 881–885.

Grazioplene, R. G., G Ryman, S., Gray, J. R., Rustichini, A., Jung, R. E., & DeYoung, C. G. (2015). Subcortical intelligence: caudate volume predicts IQ in healthy adults. *Human Brain Mapping*, *36*(4), 1407–1416.

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

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794.

Hollander, E., Anagnostou, E., Chaplin, W., Esposito, K., Haznedar, M. M., Licalzi, E., Wasserman, S., Soorya, L., & Buchsbaum, M. (2005). Striatal volume on magnetic resonance imaging and repetitive

behaviors in autism. *Biological Psychiatry*, *58*(3), 226–232.

Johnson, B. P., Rinehart, N. J., Papadopoulos, N., Tonge, B., Millist, L., White, O., & Fielding, J. (2012). A closer look at visually guided saccades in autism and Asperger's disorder. *Frontiers in Integrative Neuroscience*, *6*, 99.

Johnson, C. P., Myers, S. M., & American Academy of Pediatrics Council on Children With Disabilities. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1183–1215.

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* , *8*(1), 118–127.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (Eds.). (2013). *Principles of Neural Science, 5th ed*. McGraw-Hill.

Kanner, L. (1943). Autistic disturbances of affective contact. *The Nervous Child*, *2*(3), 217–250.

Lai, M.-C., Lerch, J. P., Floris, D. L., Ruigrok, A. N. V., Pohl, A., Lombardo, M. V., & Baron-Cohen, S. (2017). Imaging sex/gender and autism in the brain: Etiological implications. *Journal of Neuroscience Research*, *95*(1-2), 380–397.

Lai, M.-C., Lombardo, M. V., & Baron-Cohen, S. (2014). Autism. *The Lancet*, *383*(9920), 896–910.

Langen, M., Bos, D., Noordermeer, S. D. S., Nederveen, H., van Engeland, H., & Durston, S. (2014). Changes in the development of striatum are involved in repetitive behavior in autism. *Biological Psychiatry*, *76*(5), 405–411.

Langen, M., Durston, S., Staal, W. G., Palmen, S. J. M. C., & van Engeland, H. (2007). Caudate nucleus is enlarged in high-functioning medication-naive subjects with autism. *Biological Psychiatry*, *62*(3), 262–266.

Langen, M., Schnack, H. G., Nederveen, H., Bos, D., Lahuis, B. E., de Jonge, M. V., van Engeland, H., & Durston, S. (2009). Changes in the developmental trajectories of striatum in autism. *Biological Psychiatry*, *66*(4), 327–333.

Lerch, J., Hammill, C., Nikelski, J., van Eede, M., Cassel, D., Yee, Y., Devenyi, G., & Fernandes, D.

(n.d.). *RMINC*. Github. Retrieved September 27, 2022, from https://github.com/Mouse-Imaging-Centre/RMINC

Lerch, J. P., Carroll, J. B., Spring, S., Bertram, L. N., Schwab, C., Hayden, M. R., & Henkelman, R. M. (2008). Automated deformation analysis in the YAC128 Huntington disease mouse model. *NeuroImage*, *39*(1), 32–39.

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

Levy, S. E., Mandell, D. S., & Schultz, R. T. (2009). Autism. *The Lancet*, *374*(9701), 1627–1638.

Lin, H.-Y., Ni, H.-C., Lai, M.-C., Tseng, W.-Y. I., & Gau, S. S.-F. (2015). Regional brain volume differences between males with and without autism spectrum disorder are highly age-dependent. *Molecular Autism*, *6*, 29.

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E. J. H., Jones, R. M., Pickles, A., State, M. W., Taylor, J. L., & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature Reviews. Disease Primers*, *6*(1), 5.

Makowski, C., Béland, S., Kostopoulos, P., Bhagwat, N., Devenyi, G. A., Malla, A. K., Joober, R., Lepage, M., & Chakravarty, M. M. (2018). Evaluating accuracy of striatal, pallidal, and thalamic segmentation methods: Comparing automated approaches to manual delineation. *NeuroImage*, *170*, 182–198.

McAlonan, G. M., Suckling, J., Wong, N., Cheung, V., Lienenkaemper, N., Cheung, C., & Chua, S. E. (2008). Distinct patterns of grey matter abnormality in high-functioning autism and Asperger's syndrome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *49*(12), 1287–1295.

Müller, R.-A. (2007). The study of autism as a distributed disorder. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(1), 85–95.

Murray Sherman, S., & Guillery, R. W. (2013). *Functional Connections of Cortical Areas: A New View from the Thalamus*. MIT Press.

Nair, A., Treiber, J. M., Shukla, D. K., Shih, P., & Müller, R.-A. (2013). Impaired thalamocortical

connectivity in autism spectrum disorder: a study of functional and anatomical connectivity. *Brain: A Journal of Neurology*, *136*(Pt 6), 1942–1955.

Ofner, M., Coles, A., Decou, M. L., Do, M. T., Bienek, A., Snider, J., & Ugnat, A.-M. (2018). *Autism Spectrum Disorder among Children and Youth in Canada 2018*. Public Health Agency of Canada. https://www.canada.ca/en/public-health/services/publications/diseases-conditions/autism-spectrum-disorder-children-youth-canada-2018.html

Olafson, E., Bedford, S. A., Devenyi, G. A., Patel, R., Tullo, S., Park, M. T. M., Parent, O., 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., … Chakravarty, M. M. (2021). Examining the Boundary Sharpness Coefficient as an Index of Cortical Microstructure in Autism Spectrum Disorder. *Cerebral Cortex* , *31*(7), 3338–3352.

Pipitone, J., Park, M. T. M., Winterburn, J., Lett, T. A., Lerch, J. P., Pruessner, J. C., Lepage, M., Voineskos, A. N., Chakravarty, M. M., & Alzheimer's Disease Neuroimaging Initiative. (2014). Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. *NeuroImage*, *101*, 494–512.

Qiu, A., Adler, M., Crocetti, D., Miller, M. I., & Mostofsky, S. H. (2010). Basal ganglia shapes predict social, communication, and motor dysfunctions in boys with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(6), 539–551, 551.e1–e4.

Raznahan, A., Shaw, P. W., Lerch, J. P., Clasen, L. S., Greenstein, D., Berman, R., Pipitone, J., Chakravarty, M. M., & Giedd, J. N. (2014). Longitudinal four-dimensional mapping of subcortical anatomy in human development. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(4), 1592–1597.

Raznahan, A., Wallace, G. L., Antezana, L., Greenstein, D., Lenroot, R., Thurm, A., Gozzi, M., Spence, S., Martin, A., Swedo, S. E., & Giedd, J. N. (2013). Compared to what? Early brain overgrowth in autism and the perils of population norms. *Biological Psychiatry*, *74*(8), 563–575.

Redcay, E., & Courchesne, E. (2005). When is the brain enlarged in autism? A meta-analysis of all brain

size reports. *Biological Psychiatry*, *58*(1), 1–9.

Schmitt, L. M., Cook, E. H., Sweeney, J. A., & Mosconi, M. W. (2014). Saccadic eye movement abnormalities in autism spectrum disorder indicate dysfunctions in cerebellum and brainstem. *Molecular Autism*, *5*(1), 47.

Schuetze, M., Park, M. T. M., Cho, I. Y., MacMaster, F. P., Chakravarty, M. M., & Bray, S. L. (2016). Morphological Alterations in the Thalamus, Striatum, and Pallidum in Autism Spectrum Disorder. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *41*(11), 2627–2637.

Shaw, P., De Rossi, P., Watson, B., Wharton, A., Greenstein, D., Raznahan, A., Sharp, W., Lerch, J. P., & Chakravarty, M. M. (2014). Mapping the development of the basal ganglia in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*(7), 780–789.e11.

Shaw, P., Weingart, D., Bonner, T., Watson, B., Park, M. T. M., Sharp, W., Lerch, J. P., & Chakravarty, M. M. (2016). Defining the neuroanatomic basis of motor coordination in children and its relationship with symptoms of attention-deficit/hyperactivity disorder. *Psychological Medicine*, *46*(11), 2363–2373.

Sherman, S. M. (2005). Thalamic relays and cortical functioning. *Progress in Brain Research*, *149*, 107–126.

Sherman, S. M., & Guillery, R. W. (2011). Distinct functions for direct and transthalamic corticocortical connections. *Journal of Neurophysiology*, *106*(3), 1068–1077.

Smith, Y., Bevan, M. D., Shink, E., & Bolam, J. P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience*, *86*(2), 353–387.

Steriade, M., & Llinás, R. R. (1988). The functional states of the thalamus and the associated neuronal interplay. *Physiological Reviews*, *68*(3), 649–742.

Sussman, D., Leung, R. C., Chakravarty, M. M., Lerch, J. P., & Taylor, M. J. (2016). The developing human brain: age-related changes in cortical, subcortical, and cerebellar anatomy. *Brain and*

Behavior, *6*(4), e00457.

Tsatsanis, K. D., Rourke, B. P., Klin, A., Volkmar, F. R., Cicchetti, D., & Schultz, R. T. (2003). Reduced thalamic volume in high-functioning individuals with autism. *Biological Psychiatry*, *53*(2), 121–129.

Tullo, S., Devenyi, G. A., Patel, R., Park, M. T. M., Louis Collins, D., & Mallar Chakravarty, M. (2018). *Warping an atlas derived from serial histology to 5 high-resolution MRIs*. Figshare. https://doi.org/10.6084/m9.figshare.c.4052768

Tullo, S., Patel, R., Devenyi, G. A., Salaciak, A., Bedford, S. A., Farzin, S., Wlodarski, N., Tardif, C. L., PREVENT-AD Research Group, Breitner, J. C. S., & Chakravarty, M. M. (2019). MR-based age-related effects on the striatum, globus pallidus, and thalamus in healthy individuals across the adult lifespan. *Human Brain Mapping*, *40*(18), 5269–5288.

Turner, A. H., Greenspan, K. S., & van Erp, T. G. M. (2016). Pallidum and lateral ventricle volume enlargement in autism spectrum disorder. *Psychiatry Research. Neuroimaging*, *252*, 40–45.

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.

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.

Wichmann, T., & DeLong, M. R. (2013). The Basal Ganglia. In E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of Neural Science, 5th ed.* (pp. 982–998). McGraw-Hill.

Wierenga, L., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B., & Durston, S. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage*, *96*, 67–72.

Zhang, W., Groen, W., Mennes, M., Greven, C., Buitelaar, J., & Rommelse, N. (2018). Revisiting subcortical brain volume correlates of autism in the ABIDE dataset: effects of age and sex. *Psychological Medicine*, *48*(4), 654–668.

Zhang, Y., Liu, X., Guo, R., Xu, W., Guo, Q., Hao, C., Ni, X., & Li, W. (2021). Biological implications of genetic variations in autism spectrum disorders from genomics studies. *Bioscience Reports*, *41*(7). https://doi.org/10.1042/BSR20210593