NeuroImage: Clinical 32 (2021) 102888

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Dissociating regional gray matter density and gray matter volume in autism spectrum condition

Lisa D. Yankowitz [a,b,*], Benjamin E. Yerys [a,c], John D. Herrington [a,c], Juhi Pandey [a,c], Robert T. Schultz [a,c,d]

[a] Center for Autism Research, Children's Hospital of Philadelphia, 2716 South St, Philadelphia, PA 19104, United States
[b] Department of Psychology, University of Pennsylvania, 425 S. University Ave, Philadelphia, PA 19104, United States
[c] Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd Philadelphia, PA 19105, United States
[d] Department of Pediatrics Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd Philadelphia, PA 19105, United States

A B S T R A C T

Background: Despite decades of research, there is continued uncertainty regarding whether autism is associated with a specific profile of gray matter (GM) structure. This inconsistency may stem from the widespread use of voxel-based morphometry (VBM) methods that combine indices of GM density and GM volume. If GM density or volume, but not both, prove different in autism, the traditional VBM approach of combining the two indices may obscure the difference. The present study measures GM density and volume separately to examine whether autism is associated with alterations in GM volume, density, or both.

Methods: Differences in MRI-based GM density and volume were examined in 6–25 year-olds with a diagnosis of autism spectrum disorder (n = 213, 80.8% male, IQ 47–154) and a typically developing (TD) sample (n = 190, 71.6% male, IQ 67–155). High-resolution T1-weighted anatomical images were collected on a single MRI scanner. Regional density and volume were estimated via whole-brain parcellation comprised of 1625 parcels. Parcel-wise analyses were conducted using generalized additive models while controlling the false discovery rate (FDR, $q < 0.05$). Volume differences in the 68-region Desikan-Killiany atlas derived from Freesurfer were also examined, to establish the generalizability of findings across methods.

Results: No density differences were observed between the autistic and TD groups, either in individual parcels or whole brain mean density. Increased volume was observed in autism compared to the TD group when controlling for Age, Sex, and IQ, both at the level of the whole brain (total volume) and in 45 parcels (2.8% of total parcels). Parcels with greater volume included inferior, middle, and superior temporal gyrus, inferior and superior frontal gyrus, precuneus, and fusiform gyrus. Converging evidence from a standard Freesurfer pipeline also identified greater volume in a number of overlapping regions.

Limitations: The method for determining "density" is not a direct measure of neuronal density, and this study cannot reveal underlying cellular differences. While this study represents possibly the largest single-site sample of its kind, it is underpowered to detect very small differences.

Conclusions: These results provide compelling evidence that autism is associated with regional GM volumetric differences, which are more prominent than density differences. This underscores the importance of examining volume and density separately, and suggests that direct measures of volume (e.g. region-of-interest or tensor-based morphometry approaches) may be more sensitive to autism-relevant differences in neuroanatomy than concentration/density-based approaches.

1. Introduction

The search for neuroanatomical differences associated with autism spectrum condition has produced a wide range of disparate and often contradictory findings. It is widely recognized that inconsistent results in this literature relate to small sample sizes (Sacco et al., 2007), substantial clinical and etiological heterogeneity (Jeste & Geschwind, 2014), difficulty adequately accounting for individual differences with known relationships to brain structure, including age, sex, and IQ (Bedford et al., 2020), and variability in neuroimaging methods (Carp, 2012). Though the field is approaching agreement about neural systems involved in the neurobiology of autism through a combination of functional and structural imaging modalities (e.g., see Ecker, 2017), there is not a precise understanding of regional alterations in gray matter (GM) structure in autism.

At least 9 meta-analyses of GM differences in autism have been published within the past 10 years (Carlisi et al., 2017; Cauda et al., 2011; DeRamus & Kana, 2014; Duerden et al., 2012; Liu et al., 2017; Lukito et al., 2020; Nickl-Jockschat et al., 2012; Via et al., 2011; Yang et al., 2016). The fact that there have been so many meta-analyses on this topic is itself evidence that it is difficult to synthesize and summarize findings about GM in autism. There are fairly consistent reports of enlargement in inferior, middle, or superior temporal cortex (Carlisi et al., 2017; Cauda et al., 2011; DeRamus & Kana, 2014; Duerden et al.,

* Corresponding author.
E-mail address: lisayank@sas.upenn.edu (L.D. Yankowitz).

https://doi.org/10.1016/j.nicl.2021.102888
Received 14 August 2021; Received in revised form 18 October 2021; Accepted 16 November 2021
Available online 19 November 2021

2012; Liu et al., 2017; Lukito et al., 2020; Yang et al., 2016), middle or superior frontal cortex (DeRamus & Kana, 2014; Duerden et al., 2012; Liu et al., 2017; Lukito et al., 2020; Via et al., 2011), and precuneus or posterior cingulate cortex (Cauda et al., 2011; Duerden et al., 2012; Lukito et al., 2020; Nickl-Jockschat et al., 2012), as well as reductions in hippocampus (DeRamus & Kana, 2014; Duerden et al., 2012; Lukito et al., 2020; Nickl-Jockschat et al., 2012; Via et al., 2011) and amygdala (Cauda et al., 2011; Duerden et al., 2012; Nickl-Jockschat et al., 2012). However, even these broad findings are not reported by every meta-analysis, and the precise location – and even hemisphere – of findings varies. These findings also occur in the context of structures which have alternatively been reported as significantly increased or decreased in GM across different meta-analyses, including insula, anterior cingulate cortex, and the cerebellum. Furthermore, there is controversy over the developmental course of GM structural changes (Courchesne, 2004; Yankowitz et al., 2020), with some studies indicating that structural changes are only apparent at specific ages (e.g., Duerden et al., 2012). The lack of a precise understanding of gray matter (GM) alterations in autism diminishes its potential to increase understanding of the etiology of autism or serve as a clinically useful biomarker.

While the importance of increasing sample sizes and understanding heterogeneity are increasingly recognized in autism research, less attention has been paid to key MRI methodological differences that can have a major impact on study findings. Many studies use voxel-based morphometry, a technique that quantifies differences in GM structure at the voxel level (Ashburner & Friston, 2000). At a high level, voxel-based morphometry essentially quantifies how much gray matter is present in an area of the brain, and statistically examines whether this value varies with an important characteristic, such as diagnosis. Voxel based morphometry typically consists of 1) segmentation of T1 images resulting in probability maps of GM at each voxel, 2) spatial normalization to a group template, 3) modulation (scaling) of probability maps to preserve local volumes, 4) spatial smoothing, and 5) statistical analysis (for example, voxelwise t-tests for group differences in modulated segmentation values). Researchers may or may not include step 3 – modulation – and the decision of whether to report modulated or unmodulated values is controversial (Eckert et al., 2006; Radua et al., 2014). An intuitive justification for using modulation is that performing spatial normalization to a group template – necessary for conducting analyses in the same region across individuals – artificially expands some regions and shrinks others, resulting in the need to correct (modulate) the probability values to preserve original volumes. Analysis of unmodulated values are said to reflect regional concentration differences, whereas modulated values are thought to reflect volume differences (Ashburner & Friston, 2000; Good et al., 2001).

In autism research, these modulated values are inconsistently reported as reflecting GM *volume* (Rojas et al., 2006), GM *concentration* (Boddaert et al., 2004), or GM *density* (Mei et al., 2020). Mathematically, modulated GM maps represent a multiplication of two sources of variation: concentration/density differences and volume differences. This multiplicative effect may result in decreased sensitivity to detect differences in either of these sources of variance (Eckert et al., 2006; Radua et al., 2014). While most autism researchers are interested in identifying specific brain regions that differ in GM, the ability to create distinct GM volume and GM density measures at this regional scale has been difficult.

Recent work that has successfully dissociated GM density and volume on the regional scale has also demonstrated that these GM properties vary independently with age and sex. In a sample of 1,189 8–23 year-old typically developing individuals, GM density was found to increase with age, while GM volume decreased across adolescence (Gennatas et al., 2017). Additionally, females had greater GM density but smaller GM volume than males. Finally, when these measures were multiplied by each other (in rough imitation of modulated VBM), the resulting quality ('GM mass') largely resembled volume, but showed sometimes weaker and occasionally opposite relationships with age and sex as volume did. Given that density and volume vary in opposing directions with both age and sex, it is possible that combining these measures obscures true differences in either metric, and this problem may contribute to inconsistent GM findings in autism. While the issue of multiplicative error is problematic in any research context, the problem may be particularly pronounced in autism research, where it is common to have imbalanced sex ratios between groups and narrow age bands that vary between studies.

The goal of this study is to use the analytic pipeline developed by Gennatas et al, (2017) to investigate whether differences in GM density, volume, or both are apparent in autism as compared to a typically developing (TD) comparison group. Following Gennatas et al., we use the terminology of "density" and "volume." Crucially, density is not a measure of neuronal density, which cannot be obtained from conventional structural MR images. Here, *density* is used to refer to the concentration of GM within a region, which is derived from probability value from the probabilistic segmentation, which in turn reflects histology and partial volume effects. *Volume* intuitively refers to absolute volume. The biological interpretation of density is less straightforward than volume. The probabilistic segmentation assigns a probability of tissue class membership based on signal intensity. Signal intensity in a T1 image is influenced by a number of factors, including histological properties, partial volume effects, and other influences. Regarding histology, work comparing T1 images to post-mortem staining indicates that T1 images largely reflect a mix of cytoarchitecture and myeloarchitecture, meaning that signal intensity (and therefore density) reflects a mix of neuronal density and myelination (Eickhoff et al., 2005). Regarding partial volume, voxels in the brain may fall entirely within GM, entirely within white matter, or span the gray-white boundary. A voxel that is entirely comprised of GM is likely to be assigned a very high posterior probability for the tissue class label of GM, whereas a voxel entirely within white matter would receive a very low probability value. A voxel that spans the gray-white boundary ('partial volume') would receive an intermediate density value, reflecting the fact that the concentration or density of GM within that voxel is lower than a voxel entirely comprised of GM. T1 could also be affected by other sources of noise or unmeasured variance, for example hydration levels, which affect GM volume (Nakamura et al., 2014), as well as tissue perfusion, cholesterol levels, unmeasured head motion, steroid levels, time of day, exercise, and mental activities (Weinberger and Radulescu, 2020). Overall, the GM density measure represents how confident an algorithm is that a voxel is GM, and this confidence is driven primarily by the concentration of GM within a voxel and the histology of the voxel.

The present study had two major aims: (1) to replicate the Age and Sex effects of GM density and volume reported in (Gennatas et al., 2017) in our typically developing sample, (2) to separately examine the contributions of GM density and volume in autism. The purpose of performing the replication of (Gennatas et al., 2017) was two-fold. First, because this is a novel analytic technique, this replication served as a check that the method was performing as expected. Second, direct replication is often crucially lacking in neuroscience (Button et al., 2013). We sought to combat this by providing an independent replication attempt of the novel findings presented by Gennatas et al. (although our sample is smaller, it is very similar in age range). We addressed the second aim (GM density and volume differences in autism) using one of the largest single-site, single-scanner cohorts collected in autism research to date. By leveraging the power of this large, methodologically homogenous and well-characterized sample, we address an important question about GM in autism - the contributions of GM volume and density to regional structural differences. Finally, as a reliability check, we examined our aims in the context of two software pipelines that provide volumetric information - the Advanced Normalization Toolkit (ANTs, our primary analysis pipeline; Avants et al., 2011) and Freesurfer (another widely used structural pipeline; http://surfer.nmr.mgh.harvard.edu/).

2. Material and methods

2.1. Sample

Data from this study came from several neuroimaging projects conducted at the Center for Autism Research at the Children's Hospital of Philadelphia – all using the same structural imaging sequence (Yankowitz et al., 2020). For individuals on the autism spectrum, final autism spectrum disorder diagnosis was made by expert clinical judgment using DSM-IV-TR criteria using results from the Autism Diagnostic Observation Schedule (Lord et al., 2000) and the Autism Diagnostic Interview-Revised (Rutter et al., 2005). All diagnostic subcategories (autism,

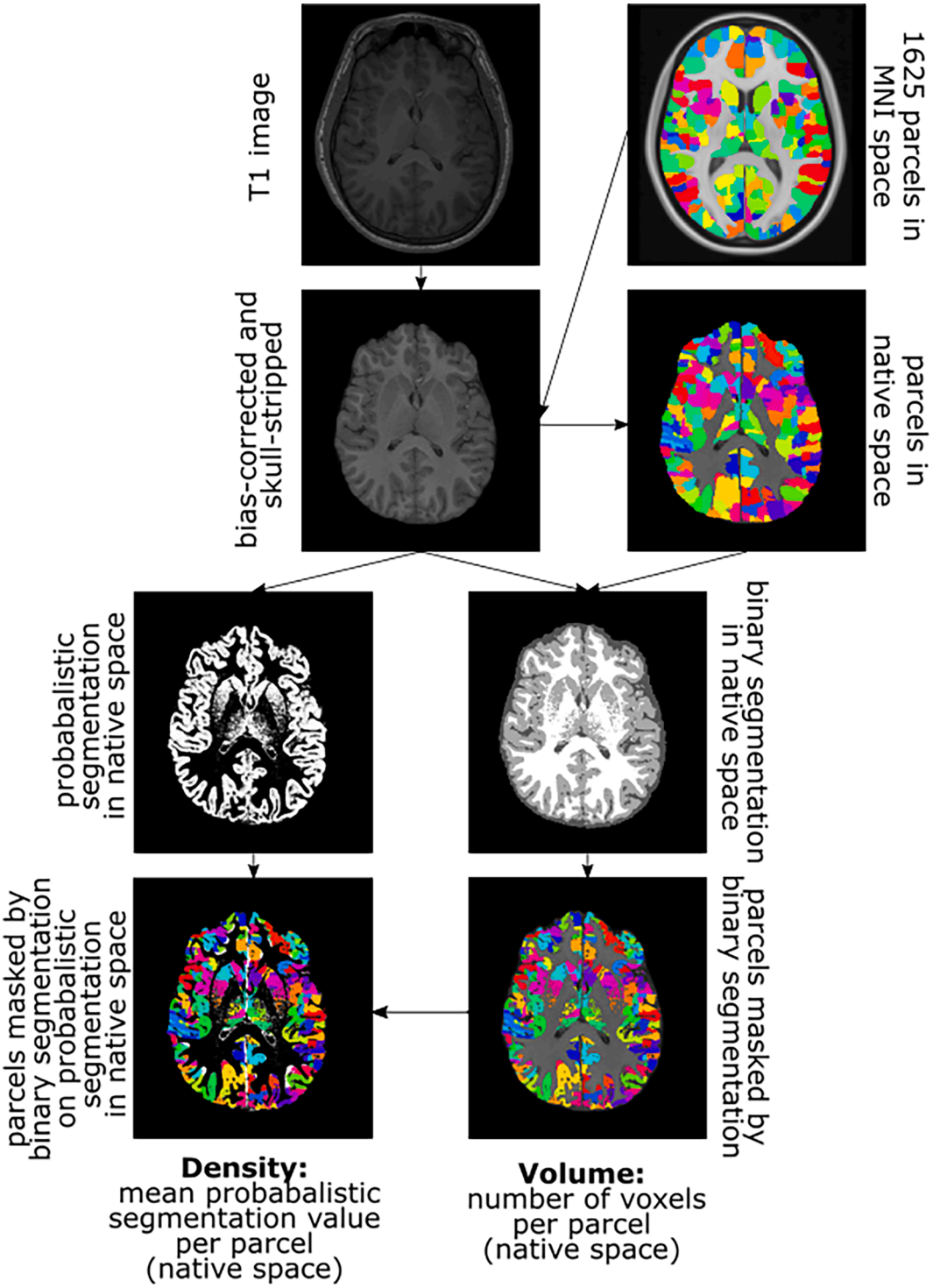

Fig. 1. Processing pipeline for neuroimaging data. T1-weighted images were bias-corrected and skull-stripped. The parcellation described in Gennatas et al., 2017 (1625 parcels) was warped from MNI space to subject-native space. Each subject's own binary (or deterministic) gray matter segmentation was applied as a mask, resulting in the final parcels. The size of each (warped) parcel represents gray matter volume within that parcel. Each skull-stripped image was also subjected to probabilistic (or "soft") segmentation. The mean of these soft segmentation values was extracted from each parcel, masked by the binary segmentation, as gray matter density.

Asperger's, PDD-NOS) were pooled into a single "autism" group in this study, in keeping with DSM-5. Typically developing (TD) participants were excluded if parents reported that their child had any known traumatic brain injury, genetic, medical, or neurological condition which affects cognitive functioning; extreme premature birth; any first- or second degree relative with autism spectrum disorder; or if the child had significant reported symptoms of affective disorders, anxiety, or attention-deficit/hyperactivity disorder. Participants ranged from 6 to 25 years of age. Maternal educational attainment (measured on a 6-point scale), a proxy for socio-economic status (which is known to correlate with brain structure; Noble et al., 2015), did not significantly differ between groups (Fisher's Exact Test, $p = 0.225$).

2.2. Image acquisition

Anatomical images were acquired on a research-dedicated Siemens 3 T wide-bore Magnetom Verio TIM scanner with a 12-channel head coil and a Siemens MPRAGE sequence (voxel size = 0.9 × 0.41 × 0.41 mm, TR = 1900, TE = 2.54, flip angle = 9).

2.3. Image Preprocessing and quality assurance

Fig. 1 provides an overview of the processing pipeline. We aimed to replicate, as directly as possible, the processing and analytic pipeline described in Gennatas et al. (2017). The following multi-step processing and quality control procedure was applied to the sample reported in Yankowitz et al., 2020 (n = 456), in order to apply more stringent criteria to support the regional analyses presented in this paper, as well as to emulate the quality assurance (QA) pipeline used by (Gennatas et al., 2017). All images were rated for motion quality on a 1–6 scale by a reliable rater (Intraclass correlation ICC(2,1) = 0.95 for a subset of images double-coded with an expert rater), with 1 indicating no visible motion artifacts and 6 indicating significant widespread quality issues. Images were N3 bias corrected with Advanced Normalization Tools (ANTs, Tustison et al., 2010) and brain extracted with LABEL (Shi et al., 2012). Brain extractions were visually inspected, and manually edited with ITK-SNAP (Yushkevich et al., 2006) if imperfections were identified, ensuring no cortex or cerebellum was excluded. Images with ratings of 5 or 6 were excluded (n = 4). Images with ratings of 3 or 4 were reviewed for inclusion by LDY and/or BEY, with final decisions made by consensus, resulting in 42 additional exclusions. The 410 remaining images were subjected to the automated QA procedure used by Gennatas et al. (2017), which attempts to identify outliers by examining correlations of T1 images and density images with other subjects, as described next. Each image was normalized to MNI space using nonlinear (diffeomorphic) warping using ANTs. Normalized images were converted to 1D vectors. Each of these was correlated with the remaining participants' MNI-normalized vectors, and the mean of each participant's correlations was taken. Participants with a mean correlation of > 2 SD from the mean were excluded (n = 7). In practice, this resulted in the exclusion of images with skull stripping that resulted in excess non-brain matter left in the neck area. After QA, our sample consisted of 403 participants (213 autism, 190 TD). Demographic and clinical information for these participants is shown in Table 1.

3-class (GM, white matter, and CSF) segmentation was performed using Atropos (Avants et al., 2011). Briefly, Atropos was initialized with no template priors, and performed K-means clustering on the first iteration to provide an initial segmentation. This output was updated with two additional iterations, with the segmentation output of the previous iteration used as the initialization to optimize the posterior probabilities of tissue class labels. This results in two outputs: the deterministic ("binary" or "hard") class labels for GM, white matter, and CSF, and the probabilistic (or "soft") segmentation values.

Table 1
Participant characteristics.

	TD (N = 190)	Autism (N = 213)	P-value
Sex			
Male	136 (71.6%)	172 (80.8%)	0.0406
Female	54 (28.4%)	41 (19.2%)	
Age			
Mean (SD)	13.4 (4.23)	13.1 (3.58)	0.49
Median [Min, Max]	12.4 [6.35, 25.6]	12.6 [6.36, 25.9]	
IQ			
Mean (SD)	114 (15.9)	101 (20.2)	<0.001
Median [Min, Max]	114 [67.0, 155]	101 [47.0, 154]	
Motion Rating			
Mean (SD)	2.66 (0.496)	2.69 (0.528)	0.535
Median [Min, Max]	3.00 [1.00, 4.00]	3.00 [1.00, 4.00]	
ADOS-CSS		N = 210	
Mean (SD)		6.88 (2.26)	
Median [Min Max]		7.00 [1.00, 10.0]	
Maternal Educational Attainment			
Some high school	15 (7.9%)	13 (6.1%)	0.225
High school	0 (0%)	0 (0%)	
Some college / 2 year degree / Associate's degree	22 (11.6%)	37 (17.4%)	
4 year college degree, Bachelor's degree	46 (24.2%)	66 (31.0%)	
Master's degree / Graduate degree-NOS	41 (21.6%)	36 (16.9%)	
Doctoral degree	8 (4.2%)	6 (2.8%)	
Not Reported	58 (30.5%)	55 (25.8%)	

Note. Differences were assessed by t-tests (Age, IQ, Motion Rating, ADOS-CSS), chi-squared test (Sex), or Fisher's exact test (Maternal Educational Attainment). Motion was rated on a 1–6 scale with 1 indicating no visible motion artifacts and 6 indicating significant widespread quality issues.

2.4. Parcellation warping and data extraction

For each subject, the GMD1625 parcellation (Gennatas et al., 2017) was warped to subject native space by applying the inverse of the subject-to-MNI transformation. This high-resolution parcellation includes 1625 parcels, which represent GM density "peaks" (i.e., cortical gyri and subcortical nuclei) derived from an Age- and Sex-matched sample of 240 individuals from the Philadelphia Neurodevelopmental Cohort (Gennatas et al., 2017). These parcellations in native space were masked by the subject's binary GM mask, as derived from the deterministic segmentation, to ensure that only GM was included in the analysis of volume and density. In subject native space, two values were extracted for each parcel: GM volume and GM density. GM volume was the number of voxels in the parcel in subject space, multiplied by the voxel size to provide volume in cubic centimeters. Total GM volume was calculated by summing the volume of all parcels. Thus, the biological interpretation of volume is straightforwardly the size of the parcels – with the caveat that because of individual differences in morphology, the parcellation may not cover all voxels labelled as GM by deterministic segmentation, thereby slightly underestimating overall GM volume (see Limitations for further discussion of this point). The advantage of this approach is that it is able to estimate GM volume at smaller spatial scales than methods that more accurately capture the whole brain (i.e., 1625 parcels, as compared to the 68 regions of interest (ROIs) in the Desikan-Killiany atlas (Desikan et al., 2006)).

GM density was the mean of the probabilistic segmentation values across voxels within each parcel, masked by the binary GM segmentation (Gennatas et al., 2017). Extraction of these density data in subject-native space eliminates the need to account for warping of the data during spatial registration, and allows for regionally specific analysis without interpolation (Gennatas et al., 2017). Whole brain mean GM

density was calculated as the average of each parcel density, weighted by the size of the parcel in voxels in the GMD1625 atlas.

Labels for parcels were extracted from two atlases (MNI and Talairach Daemon) using the atlasquery function of FSL (Jenkinson et al., 2012), assigning the label with the highest probability averaged across voxels within the parcel. Functional network labels were assigned using a 7-network solution (Yeo et al., 2011) by assigning the modal network label assigned to voxels within the parcel.

As a reliability check on our method of estimating volume and to ensure that volume-based results were not specific to our processing pipeline, regional volumes were calculated through a separate pipeline in Freesurfer (http://surfer.nmr.mgh.harvard.edu/). Freesurfer is a widely used and freely available cortical reconstruction and volumetric segmentation tool. Additional Freesurfer quality control steps were taken as described elsewhere (Yankowitz et al., 2020). Volumes for 68 ROIs were calculated using the Desikan-Killiany atlas (Desikan et al., 2006),

2.5. Clinical measures

Participants' cognitive ability ("IQ") was assessed with one of four standard instruments: the General Cognitive Ability score of the Differential Abilities Scale, Second Edition (Elliot, 2007), the Full Scale IQ of the Wechsler Intelligence Scale for Children, Fourth Edition (Wechsler et al., 2003), and the Wechsler Abbreviated Scale of Intelligence, First or Second Edition (Wechsler, 1999, 2011). IQ in the autism group ($M = 101$, $SD = 20.2$) was significantly lower than the TD group ($M = 114$, $SD = 15.9$, $t = 7.0$, $p < 0.001$), and is included in all models examining Group or Severity as a covariate. Clinical severity was assessed with the Autism Diagnostic Observation Schedule Calibrated Severity Score (Gotham et al., 2009), which is derived from clinician ratings and is hereafter referred to as "Severity."

2.6. Statistical approach

Whole-brain and voxel-wise comparisons were conducted using Generalized Additive Models (GAMs). GAMs are similar to general linear models, with the additional feature that predictors are replaced by smooth functions of themselves (indicated by *s(predictor)* in the model specifications). By using penalized splines with smoothing parameters selected by restricted maximum likelihood, nonlinear effects are efficiently estimated.

First, we attempted to replicate the findings of (Gennatas et al., 2017) using our TD sample. Within the TD sample only, to examine the effect of Age, models were fit separately for each sex with the form:

{total volume, mean density} ~ s(Age) (form 1)

To examine the effects of Sex, models were of the form:

{total volume, mean density} ~ Sex + s(Age) (form 2)

To examine Age-by-Sex interactions, models were fit of the form:

{total volume, mean density} ~ Sex + s(Age) + s(Age*Sex) (form 3)

To examine the effect of autism Group, two models were fit for the whole brain: one for volume, using the sum of GM voxels across all parcels, and one for density, using the mean GM density of all parcels (weighted by parcel size):

{total volume, mean density} ~ Group + Sex + IQ + s(Age) (form 4)

Similarly, two models were fit for each individual parcel, one for volume and one for density. These models were corrected for multiple comparisons using FDR correction (Benjamini and Hochberg method (Benjamini & Hochberg, 1995)) separately for volume and density analyses, with $q < 0.05$ considered significant. To examine effects of autism symptom severity, models were fit using only the autism Group, with Severity replacing the Group term. Additionally, this model was fit for the volume of each Desikan-Killiany parcel.

To examine potential effects of head motion on findings, models were re-examined with motion ratings (visual ratings from 1 to 6, with details provided in section 2.3 Image Preprocessing and Quality Assurance) included as a covariate. Additionally, because the sex-ratio in the diagnostic Groups is particularly unbalanced in the oldest segment of our sample (i.e., there are no females over 20 in the autism group), all analyses of Group were re-run in only individuals under Age 20 as a sensitivity analysis. Finally, exploratory models testing the effect of Group separately by sex, and the effects of Age, Sex, and IQ in the autism group only are presented in Supplementary Material, Supplementary Figs. 1 and 2.

3. Results

3.1. Replication of previous Age and Sex effects

Our data were broadly consistent with prior reports of the effects of Age and Sex on GM. In the replication model for GM density (form 1), there was a significant effect of Sex ($p = 0.006$, with females having increased density relative to males, Fig. 2). Within both the males-only and females-only models of GM density (form 2), there was a significant effect of Age ($p_{male} = 0.042$, $p_{female} = 0.047$) with density increasing with Age in both sexes. There was not a significant Age-by-Sex interaction (form 3, $p = 0.9$).

In the replication model for GM volume (form 1), there was a significant effect of Sex ($p < 0.001$, with females having reduced volume relative to males). Within both the males-only and females-only models of GM volume (form 2), there was a significant effect of Age ($p_{male} < 0.001$, $p_{female} = < 0.001$), with volume decreasing with Age in both sexes. All of these findings are a direct replication of the main findings of (Gennatas et al., 2017), and support the overall conclusions of that paper: that GM density increases while volume decreases with Age, and that females show higher density but lower volume than males. There was not a significant Age-by-Sex interaction (form 3, $p = 0.13$).

3.2. Replication of previous Age and Sex findings within parcels

In a direct replication attempt of (Gennatas et al., 2017), models of forms 1, 2, and 3 were conducted for each parcel in the brain and corrected for multiple comparisons (FDR $q = 0.05$), to examine the effects of Age, Sex, and the interaction of Age and Sex on regional brain structure (Supplementary Material, Supplementary Tables 1, specific parcels listed in Supplementary Tables 6–9). There were significant Sex effects (form 2) on density in 71.7% of parcels and on volume in 67.6% of parcels. Gennatas et al. (2017) similarly report widespread Sex effects on both measures in all MNI structures, though they report a higher number of significant parcels throughout the brain.

There was a significant effect of Age on density (form 1) in 35.3% of parcels in males and 17.7% in females. For volume, the Age effect was significant for 30.5% of parcels in males and 38.6% in females. Notably, while modeling the effects of Age separately by sex theoretically gives a more precise estimate of the effect of Age for each sex, it also significantly reduces power (by splitting the sample roughly in half). Extracting the effects of Age from models in form 2 instead (that is, the effect of Age statistically controlling for Sex in the entire sample) results in a significant Age effect on density for 75.8% parcels, and on volume for 53.4% parcels. This pattern more closely mirrors the results reported by Gennatas et al., which found Age effects for nearly all parcels for density, but only 52% and 65% for volume (in males and females, respectively). In contrast to Gennatas et al.'s findings, there were no significant Sex-by-Age interactions predicting either density or volume

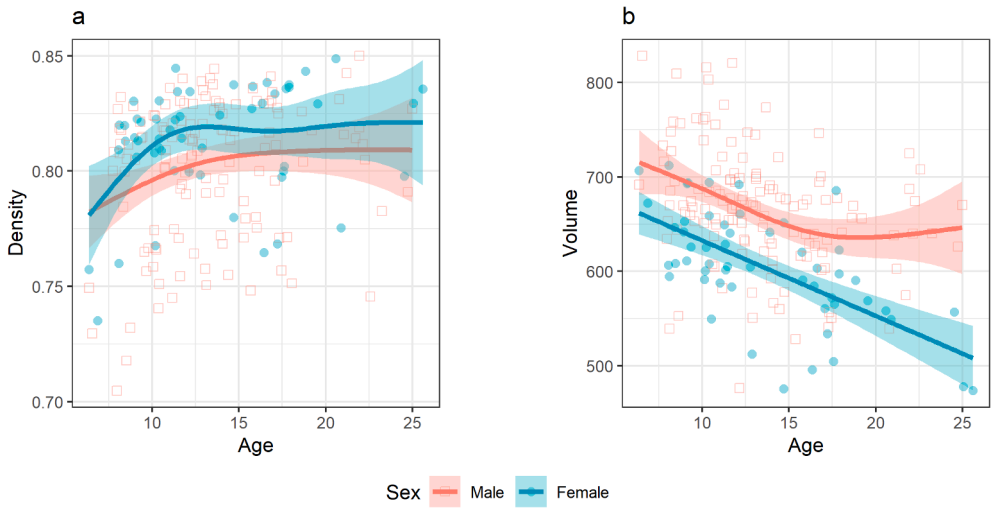

Fig. 2. Whole brain gray matter density and volume by Age and Sex in typical development. Whole brain mean gray matter density (a) increases with Age, and is higher in females than males. Whole brain mean gray matter volume (b) decreases with Age, and is higher in males than females.

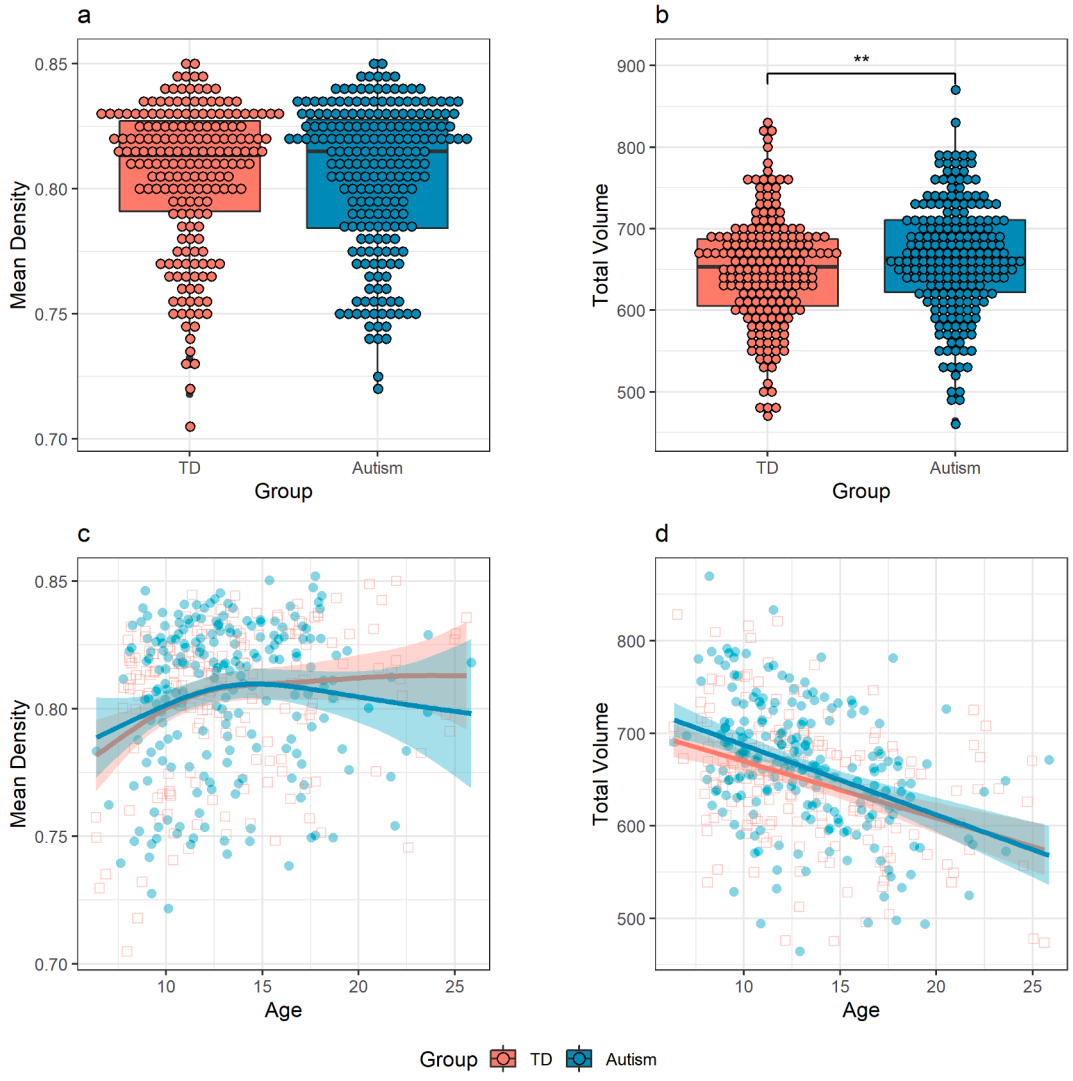

Fig. 3. Autism and whole brain gray matter density and volume. (a) There is no significant effect of Group on whole-brain gray matter density. (b) Autism is associated with larger gray matter volume. (c) Age is positively associated with mean density, with no significant Age-by-Group Interaction. (d) Age is negatively associated with total volume, with no significant Age-by-Group Interaction.

(form 3) in any parcels after correcting for multiple comparisons. Thus, we largely replicate the main findings reported in (Gennatas et al., 2017), namely widespread Sex effects on both density and volume, and Age effects on both measures which are more prominent in density, though we fail to replicate Age-by-Sex interaction effects, perhaps due to lower power. This replication provides evidence for the robustness of an analytic framework considering volume and density distinctly in analysis, and supports the use of this framework in examining differences related to autism.

3.3. Motion effects in the replication of previous Age and Sex effects

To confirm that Age- and Sex-effects in the TD sample were not driven by motion artifact, we re-ran analyses with motion ratings (visual ratings on a 1–6 scale) added as a covariate. When motion was added as a covariate in whole-brain replication models, Age remained a significant predictor of total volume in both sexes (form 1), Sex remained a significant predictor of volume and density (form 2), and the Sex-by-Age interaction remained a non-significant predictor of volume and density (form 3). The only notable changes were that in models of form 2, Age became a non-significant predictor of density with p-values just above the alpha level ($p_{male} = 0.071$, $p_{female} = 0.079$). Regional analyses were also conducted with motion as a covariate. TD results are shown in Supplementary Material, Supplementary Table 2. Similar to the whole brain results, the inclusion of motion ratings as a covariate resulted in reductions in the number of parcels showing significant effects of Age on density, particularly in the separate sex models. Findings were otherwise largely preserved, indicating that main findings are not the result of motion artifact.

3.4. Whole brain effect of Autism:

Our next set of analyses included effects of Group (autism versus TD). In the model predicting GM density (form 4), neither Group nor IQ were significant predictors ($p = 0.65$ and 0.30, respectively). There were significant effects of Sex ($p = 0.006$) and Age ($p < 0.001$). Female sex and older age was associated with increased GM density.

In the model predicting GM volume (form 4), there were significant effects of Group ($p = 0.009$), Sex ($p < 0.001$), Age ($p < 0.001$), and IQ ($p < 0.001$). Autism, male sex, younger age, and greater IQ were all associated with increased GM volume (Fig. 3). Because visual inspection of the data suggests a possible Age-by-Group interaction, a post-hoc test was conducted with an Age-by-Group interaction term added, which was not significant ($p = 0.25$). Additional post-hoc tests exploring Sex-by-Group and Age-by-Sex-by-Group on both volume and density revealed no significant interactions (see Supplementary Material). Within-sex effects of group are also shown in the Supplementary Material.

3.5. Regional effects of autism

Parallel analyses were conducted to examine region-specific differences in density and volume. Models (form 4) were fit for each parcel, testing the effect of Group controlling for Age, Sex, and IQ, and correcting for multiple comparisons. No parcels showed significant effects of Group on GM density. Conversely, there were 45 parcels with significantly increased GM volume in the autism group (Table 2 and Figs. 4 and 5, additional views in Supplementary Fig. 3). Of these, 26 fall in the temporal lobe (8.3% of parcels labelled as temporal lobe), 10 in the frontal lobe (1.9%), 5 in the parietal lobe (1.7%), 3 in the occipital lobe (1.6%), and 1 in the cerebellum (0.5%). With regard to functional networks, 19 parcels were labelled as Default (5.6% of parcels labelled as Default), 10 Somatomotor (5.6%), 6 Dorsal Attention (3.8%), 4 Visual (3.3%), 2 Frontoparietal (1.6%), 1 Limbic (0.7%), 1 Ventral Attention (0.7%), and 2 unlabeled (0.6%). To explore the general trend of the direction of volume differences among all parcels, even those not reaching statistical significance, a sign test was applied. This test indicated that a significantly greater number of parcels showed larger GM volume in autism compared to TD (number of parcels larger in autism = 1375, $p < 0.001$).

Table 2
Parcels showing significant volume effect of Group, controlling for Age, Sex, and IQ.

GMD1625 Parcel	Brain Region	Functional Network
84	Cerebellar Tonsil	*
183	Middle Temporal Gyrus	Default
185	Superior Temporal Gyrus	Limbic
221	Superior Temporal Gyrus	Default
249†	Inferior Temporal Gyrus	Default
261	*	Default
263	Superior Temporal Gyrus	Default
283	Fusiform Gyrus	Visual
298	Middle Temporal Gyrus	Default
351	Declive	Visual
358	Inferior Frontal Gyrus	Default
376	Middle Temporal Gyrus	Default
385	Declive	Visual
421	Inferior Frontal Gyrus	Default
423	Middle Temporal Gyrus	Default
425	Middle Temporal Gyrus	Default
427	Superior Temporal Gyrus	Default
473	Inferior Frontal Gyrus	Default
480	Middle Temporal Gyrus	Default
484	Sub-Gyral	Ventral Attention
544	Middle Frontal Gyrus	Default
563	Middle Temporal Gyrus	Default
564	Middle Temporal Gyrus	Default
595	Middle Occipital Gyrus	Visual
602	Sub-Gyral	Default
645†	Culmen	*
731	Superior Temporal Gyrus	Somatomotor
806	Superior Temporal Gyrus	Somatomotor
807	Superior Temporal Gyrus	Somatomotor
862	Superior Temporal Gyrus	Somatomotor
876†	Superior Temporal Gyrus	Somatomotor
878†	Superior Temporal Gyrus	Somatomotor
891†	Superior Temporal Gyrus	Somatomotor
981	Superior Temporal Gyrus	Somatomotor
1158	Precentral Gyrus	Dorsal Attention
1161	Inferior Frontal Gyrus	Dorsal Attention
1193	Precentral Gyrus	Dorsal Attention
1287	Postcentral Gyrus	Somatomotor
1308	Cingulate Gyrus	Frontoparietal
1318	Cingulate Gyrus	Frontoparietal
1435†	Precuneus	Dorsal Attention
1455	Superior Frontal Gyrus	Default
1502	Precuneus	Dorsal Attention
1549	Superior Parietal Lobule	Dorsal Attention
1608	Sub-Gyral	Somatomotor

Note. Brain region assignments come from the Talairach Daemon using the atlasquery function of FSL, and functional networks were assigned by the modal label for all the voxels within each parcel from the (Yeo et al., 2011) 7-Network Parcellation. * No label was assigned to the parcel. † Parcel is no longer significant when motion ratings are included as a covariate in the model.

3.6. Autism symptom Severity

Within the autism group, the effects of autism symptom severity (ADOS-CSS) were explored. These models were fit using models of form 4, except that they included only autistic participants, and Severity was entered in place of Group. At the whole brain level, Severity did not significantly predict either density ($p = 0.61$) or volume ($p = 0.73$). Similarly, at the regional level, no parcels showed significant effects of symptom Severity predicting either volume or density after multiple comparison correction.

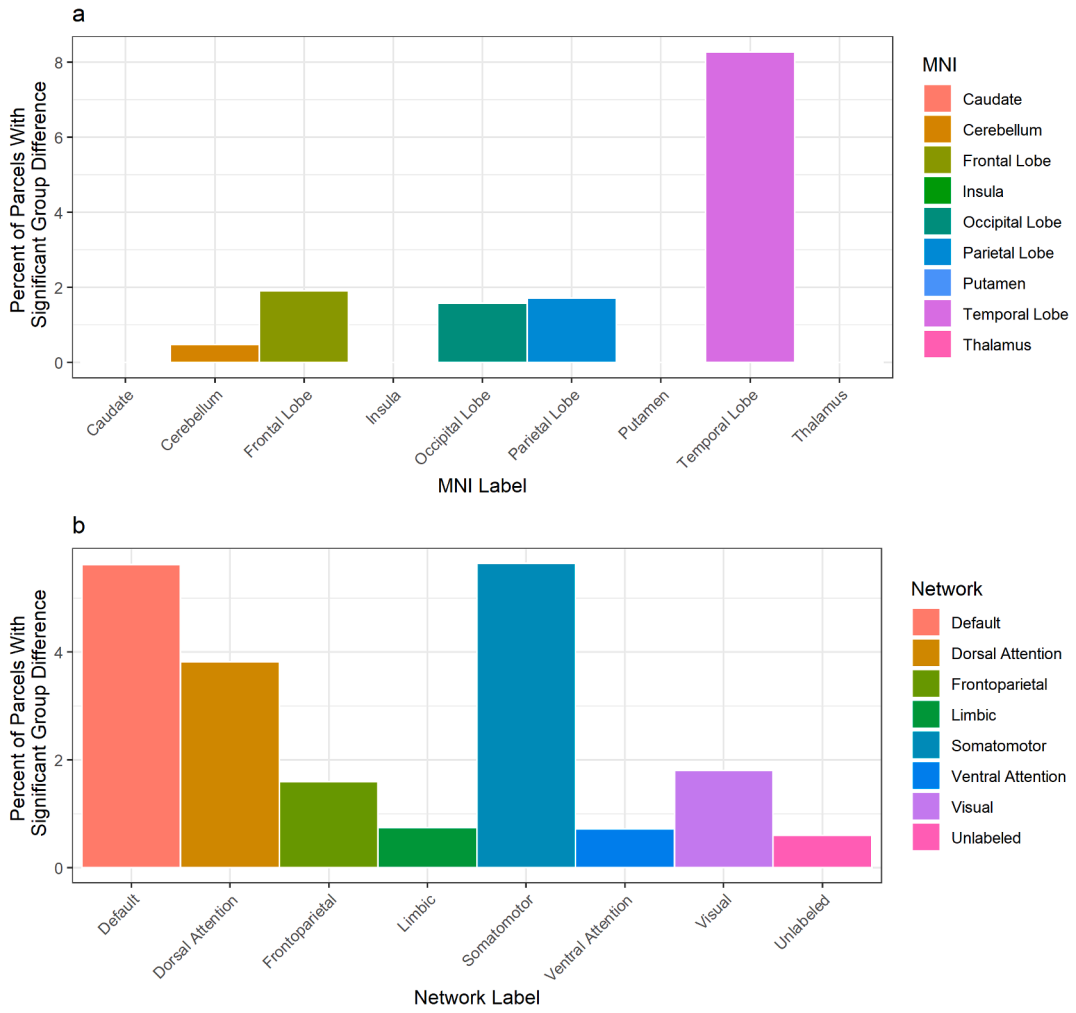

Fig. 4. Parcels showing a significant effect of larger volume in autism compared to TD, expressed as the percentage of significant parcels per MNI label (a) or per network as defined by the 7-network (Yeo et al., 2011) solution (b).

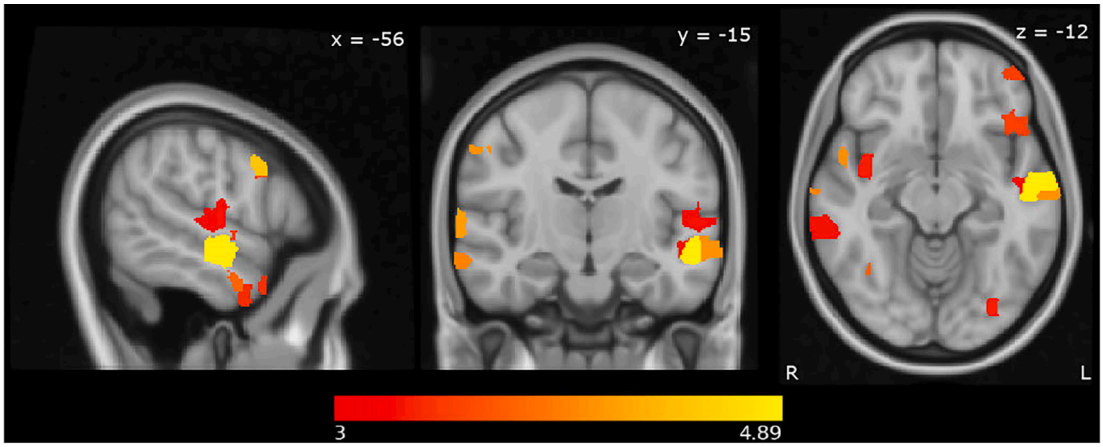

Fig. 5. Regional effects of autism on gray matter volume. Parcels showing significant effects of Group on gray matter volume, controlling for Age, Sex, and IQ, corrected for multiple comparisons ($q < 0.05$). The color of each parcel represents the estimate of the Group term from the GAM for that parcel, with more yellow values indicating larger effects. MNI coordinates shown. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.7. Motion

There was no difference in motion ratings between the autism ($M = 2.69$, $SD = 0.53$) and TD ($M = 2.66$, $SD = 0.50$) groups $t(400) = -0.62$, $p = 0.54$. Worse motion ratings were associated with younger age ($r = -0.29$, $p < 0.001$), reduced mean density ($r = -0.13$, $p = 0.008$), and increased GM volume ($r = 0.1$, $p = 0.036$). To explore the effect of motion on our findings, all results described above were re-run with

motion ratings included as a covariate. In whole brain models examining effects of autism, the effect of Group remained a significant predictor of volume and not density (form 4), and autism Severity remained a non-significant predictor within the autism group (form 4). In regional models examining the effect of Group including a motion covariate, there continued to be no significant differences in parcel-wise density; 39 parcels continued to show a significant effect of Group on volume (meaning that 6 parcels [249, 645, 876, 878, 891, 1435, see Table 2 for labels] lost significance when including motion in the model). Broadly, including motion as a covariate reduced the significance of the Age/GM density relationship, likely related to the correlation of motion and Age in this sample. Motion did not have substantial impact on Sex, Group, or autism Severity findings.

3.8. Age

Age-related sensitivity analyses were primarily focused on identifying whether Group differences were consistent across the age span of our sample (i.e., not driven by children/adolescents or adults alone). In the participants under age 20, as in the full sample, there was a significant effect of Group on total GM volume ($p = 0.0048$) but not mean GM density ($p = 0.44$) when controlling for Age, Sex, and IQ. In regional analyses, 0 parcels showed a significant effect of Group on GM density, while 67 showed a significant effect of Group on GM volume, again controlling for Age, Sex, and IQ. Of these, 43 overlapped with the parcels that were significant in the main analysis, and 24 newly emerged as significant. One parcel (794, in the insula) was significantly smaller in autism. All results are displayed in Supplementary Material, Supplementary Tables 3-4 and Supplementary Figs. 4-6. As in the whole-sample analysis, autism Severity was not a significant predictor of whole-brain or regional GM density or volume. Overall, the Age sensitivity analysis found similar results in the participants only under age 20 as the full sample, with additional parcels showing significant increases in volume in the autism group in the restricted sample.

3.9. Freesurfer volume measures

To establish that the primary findings of increased volume in autism are robust to differences in methodology and not specific to the processing pipeline, follow-up analyses were conducted examining regional volume as derived by Freesurfer. This approach yielded 23 significant regions, shown in Fig. 6 and Table 3.

Table 3
Regions of the Desikan-Killiany atlas showing greater volume for autism compared to no autism, controlling for Age, Sex, and IQ, surviving multiple comparison correction.

Hemisphere	Region	T-value	Q-value
R	Banks of the Superior Temporal Sulcus	2.66	0.031
L	Caudal Anterior Division of Cingulate Cortex	2.55	0.038
R	Cuneus	2.97	0.015
R	Inferior Temporal Gyrus	3.01	0.015
R	Insula	2.59	0.036
L	Lateral Orbitofrontal Cortex	2.97	0.015
R	Lateral Orbitofrontal Cortex	2.98	0.015
L	Middle Temporal Gyrus	3.73	0.002
R	Middle Temporal Gyrus	4.04	0.002
L	Postcentral Gyrus	2.65	0.031
R	Postcentral Gyrus	2.42	0.048
L	Posterior Cingulate	2.44	0.047
L	Precentral Gyrus	2.43	0.047
R	Precentral Gyrus	3.92	0.002
L	Rostral Anterior Cingulate	3.6	0.003
R	Superior Frontal Gyrus	2.69	0.031
L	Superior Parietal Cortex	3.24	0.01
R	Superior Parietal Cortex	4.76	0
L	Superior Temporal Gyrus	2.86	0.02
R	Superior Temporal Gyrus	3.87	0.002
L	Supramarginal Gyrus	3.61	0.003
R	Supramarginal Gyrus	3.07	0.015
L	Transverse Temporal Cortex	3.78	0.002

Note. No regions showed reduced volume in autism.

4. Discussion

This is the first study to measure GM density and GM volume separately in autistic youth to uncover which metric drives GM structural differences observed between autistic and TD youth. Our results show that at both the level of the whole brain and in specific brain regions, GM volume is significantly increased in autism. Notably, many of the regions showing increased GM volume in autism have been implicated in prior brain imaging studies of autism, including studies of structural differences, task-based functional differences, and intrinsic functional connectivity differences (Ecker, 2017). For example, enlargement was detected in inferior, middle, and superior, temporal gyri, as well as inferior and superior frontal gyri, all of which have been previously reported to show enhancement across *meta*-analyses of GM structure in autism (Carlisi et al., 2017; Cauda et al., 2011; DeRamus & Kana, 2014; Duerden et al., 2012; Liu et al., 2017; Lukito et al., 2020; Yang et al., 2016). While sensitivity analyses (i.e., including a covariate for motion or restricting the age of the sample) affected the exact number of significant parcels, the general pattern of enlargement remained consistent. These findings are supported by ROI-based analyses using a standard Freesurfer processing pipeline, which identified enlargement in similar regions.

In the context of significantly increased GM volume in autism, there were no differences detected between Groups in GM density. These results suggest that when previous studies of modulated VBM values have reported differences related to autism, these differences were likely driven more by volume than by density.

The findings of increased GM volume but not density in autism have implications for how future structural imaging studies of autism should be conducted. First, these findings suggest against using modulated SPM VBM values as an indicator of GM alterations in autism. While modulated VBM values are often reported as reflecting volume differences, these values incorporate the qualities described as *density* and *volume* in the present study. The utility of these modulated values has been previously criticized, as the multiplicative effects of combining these two properties appears to dilute effects that may be present in only one property (Eckert et al., 2006; Radua et al., 2014). The present study underscores that density and volume are separable properties which capture unique variance in autism, and thus support the recommendation not to combine them through modulation in VBM. Furthermore, these findings also indicate that volume, and not density, is the more sensitive measure for autism-related differences. Unmodulated VBM values, often described as concentration or density, are most similar to the metric reported as *density* in the current study. Thus, these findings hint that neither modulated nor unmodulated VBM values are optimal GM measures for sensitive and replicable autism research, relative to the direct measurement of regional GM volume. The reliance on these measures in prior VBM studies may contribute to inconsistencies of reports of gray matter structural differences in autism, as reporting on a weak signal (i.e. density/unmodulated VBM) or a signal rendered weaker or less reliable through multiplication (i.e. volume and density/modulated VBM) increases the likelihood of spurious or inconsistent results.

There are two main methods for measuring regional volume directly: region-of-interest approaches and tensor-based morphometry. A region-of-interest approach, such as the volumes reported by FreeSurfer, provides the volume of anatomically meaningful subunits of the brain. As demonstrated in this paper, the FreeSurfer ROI approach can identify areas of increased volume in autism. The FreeSurfer pipeline has the advantages of being free, popular, and easy to implement, and additionally breaks GM volume into its constituent parts: cortical thickness and surface area. Cortical thickness and surface area can have

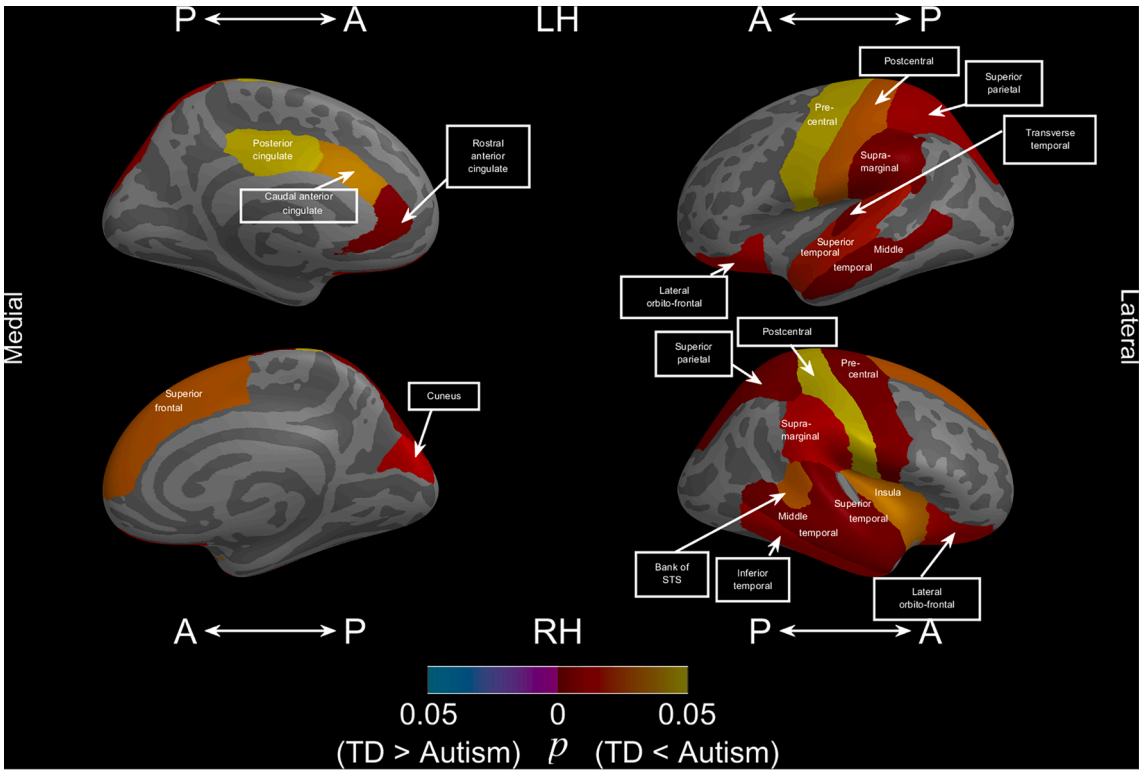

Fig. 6. Regions showing increased volume in autism relative to TD. Volume of Desikan-Killiany regions was calculated with Freesurfer. No regions were reduced in volume in autism.

independent associations with autism, though evidence remains mixed. For example, cortical thickness has been reported to be both increased (e.g., van Rooij et al., 2017) and decreased (e.g., Ecker et al., 2014) in autism, with potentially altered developmental trajectories (e.g., Zielinski et al., 2014). Increased surface area in autism has also been reported in some (Ohta et al., 2016) but not all (Raznahan et al., 2013) studies. Further study of cortical thickness and surface area can help elucidate differences in volume. The disadvantage of an ROI approach is the relatively large scale of the regions of interest (e.g., FreeSurfer commonly reports only 68 ROIs (Desikan et al., 2006)). Of course, a different set of ROIs can be used (e.g., the 1625 parcels used in this study), although doing so requires additional processing steps and renders comparison across studies challenging, to the degree that different researchers use different parcellations.

The second volumetric method, tensor-based morphometry (TBM, e.g., Hua et al., 2008), is a voxel-wise approach that measures volume directly. This method is roughly equivalent to directly analyzing the Jacobian determinant of the deformation matrix produced in VBM, rather than using this deformation to modulate the segmentation values. Though less common than SPM's VBM, TBM methods have been used in autism research (Brun et al., 2009), and publicly available software exists to implement these methods (e.g. RAVENS maps (Davatzikos et al., 2001) or Tensor Based Morphometry Analysis in BrainSuite, http://brainsuite.org/). Focus on volumetric region-of-interest or TBM approaches may provide more sensitivity and insight into the nature of GM alterations in autism.

The underlying mechanisms of GM volume alterations in autism cannot be determined by in vivo neuroimaging, but can be informed by other lines of research. Increased cortical surface area (Ohta et al., 2016) and increased cortical thickness (Ecker et al., 2014; Khundrakpam et al., 2017; Raznahan et al., 2013; Smith et al., 2016), have both previously been reported in autism. Cortical surface area and thickness, which multiply to produce cortical volume, are independently heritable with unique mechanistic underpinnings, implying that there are multiple underlying genetic influences on GM enlargement (Panizzon et al., 2009). An increased neuronal number has been found in post-mortem prefrontal cortex of autistic children (Courchesne et al., 2011). Increased dendritic spine density and decreased synaptic pruning has also been reported in post-mortem autistic brains (Tang et al., 2014). This has been linked with hyperactive mammalian target of rapamycin (mTOR) kinase, which has in turn been linked with a number of autism-linked genes, including t*SC1/TSC2*, *NF1*, and *PTEN* (Bourgeron, 2009). Additional evidence implicates genes regulating prenatal neural progenitor cells, including CHD8 (Sugathan et al., 2014), which has been linked to both autism and macrocephaly (Bernier et al., 2014), and WDFY3 (Orosco et al., 2014). Additionally, approaches examining structural covariance suggest that connectivity plays a role in morphological differences in autism (e.g., Mei et al., 2020; Zielinski et al., 2012)

We also broadly replicate the findings of (Gennatas et al., 2017) with regard to the differential effects of Age and Sex on GM density and volume. We find that while whole brain volume is larger in males than females, mean density is greater in females than males. We additionally confirm that while volume decreases during this age range, density increases. In addition to contributing to the literature on age and sex effects on brain volume, these results suggest that these study methods are sensitive and robust – and support their application to the autism data.

Across all effects (Age, Sex, and their interaction) and both outcome measures (density and volume), we find a lower percentage of significant parcels than Gennatas et al. (2017), and fail to detect significant Age-by-Sex interactions surviving multiple comparison correction. This is likely due to our smaller sample size (n = 190 vs. n = 1189) and reduced power. Notably, we nonetheless find evidence to support two of the main patterns of their regional results. Specifically, we find: 1) significant effects of Age on both density and volume, which are more widespread for density than volume, and 2) more widespread Sex effects than Age effects on volume. Thus, despite our smaller sample size, our results support the interpretation that density and volume develop differently by sex across adolescence.

It should be noted that although the term "GM density" is used here to describe the values from the probabilistic segmentation following Gennatas et al., 2017, this measure does not literally refer to the density of tissue. Neuronal density is not currently measurable by MRI or any in vivo measure, and this study cannot determine whether there are differences in neuronal count, size of neurons, myelination, or other molecular features that contribute to differences in tissue contrast (Weinberger and Radulescu, 2020). Rather, the aim of the study was to determine, of possible features which can be extracted from structural MRI, which are most sensitive to underlying differences associated with autism.

A limitation of this study is that when warped to individual subject space, the GM density parcellation does not cover all voxels labelled as GM by binary segmentation due to individual variability in morphometry. This issue is not unique to this method, and registration quality affects any regionally-specific measure of brain structure (Crum et al., 2003). Controlling for Age, sex, and IQ, the percentage of GM from the binary segmentation excluded by the parcellation is not significantly different between Groups (8.6% in TD vs. 8.8% in autism), although this percentage varies by Age and Sex (see Supplementary Material, Supplementary Figure 7 and Supplementary Table 5). No single value can capture all possible individual or group variation in brain structure. Our work indicates that volume is greater in a number of brain regions in autism, which we confirm using ROIs derived from Freesurfer, which does not suffer this issue.

While this study benefits from possibly the largest single-site, single-scanner sample of structural MRI of autistic and TD participants, it is not powered to detect small effect sizes. Thus, it is possible that GM density is altered in autistic individuals to a smaller degree than GM volume. This possibility could be explored in larger, multi-site datasets (e.g., ENIGMA (van Rooij et al., 2017) or ABIDE (Di Martino et al., 2017)). Independent replication of our autism findings in such samples would be important to demonstrate the robustness and generalizability of these results.

This case-control design was employed to address the question of whether GM density, volume, or both are associated with autism regionally when broken into constituent parts, but has limitations. The case-control design does not address the significant heterogeneity of autism, which certainly contributes to the lack of replicability in autism GM research (Jeste & Geschwind, 2014), and is better addressed by examining potential subgroups or subtypes. Additionally, though we examined the effects of Age statistically, other approaches including longitudinal designs and age normative modeling (e.g., Tunç et al., 2019) are better suited to assess the developmental course of structural differences in autism.

5. Conclusions

These findings indicate that GM volume, but not GM density, is related to autism at the whole brain level, and in regions with previously reported structural differences, including inferior, middle, and superior temporal gyrus, inferior and superior frontal gyrus, precuneus, and fusiform gyrus. These findings suggest research on brain structural differences in autism should focus on direct measures of morphometry, including volume, cortical thickness, and surface area to maximize the likelihood of identifying meaningful differences. These findings also confirm through replication that GM density and GM volume are dissociable at the regional level, and vary differentially with Age and Sex. Our replication of the dissociation of Age and Sex effects on density and volume, along with our new findings indicating an effect of autism on volume but not density, indicate the importance of considering these measures separately.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Institute of Health NIMH R01MH073084, RC1MH088791, R21MH098153, and R21MH092615, by NICHD 5U54HD086984, by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-132185, by the Pennsylvania Department of Health SAP #4100042728 and 4100047863, by the Robert Wood Johnson Foundation #66727, by Pfizer Inc. (no award number), and by Shire Development LLC (no award number).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2021.102888.

References

Ashburner, J., Friston, K.J., 2000. Voxel-Based Morphometry—The Methods. NeuroImage 11 (6), 805–821. https://doi.org/10.1006/nimg.2000.0582.

Avants, B.B., Tustison, N.J., Wu, J., Cook, P.A., Gee, J.C., 2011. An open source multivariate framework for n-tissue segmentation with evaluation on public data. Neuroinformatics 9 (4), 381–400. https://doi.org/10.1007/s12021-011-9109-y.

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., Ruigrok, A.N.V., Smith, E., Spencer, M.D., Suckling, J., Taylor, M.J., Thurm, A., Lai, M.-C., Chakravarty, M.M., 2020. Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. Mol. Psychiatry 25 (3), 614–628. 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. J. Royal Stat. Soc. Series B (Methodological) 57 (1), 289–300. JSTOR.

Bernier, R., Golzio, C., Xiong, B.o., Stessman, H., Coe, B., Penn, O., Witherspoon, K., Gerdts, J., Baker, C., Vulto-van Silfhout, A., Schuurs-Hoeijmakers, J., Fichera, M., Bosco, P., Buono, S., Alberti, A., Failla, P., Peeters, H., Steyaert, J., Vissers, L.L.M., Francescatto, L., Mefford, H., Rosenfeld, J., Bakken, T., O'Roak, B., Pawlus, M., Moon, R., Shendure, J., Amaral, D., Lein, E.d., Rankin, J., Romano, C., de Vries, B.A., Katsanis, N., Eichler, E., 2014. Disruptive CHD8 mutations define a subtype of autism early in development. Cell 158 (2), 263–276. https://doi.org/10.1016/j.cell.2014.06.017.

Boddaert, N., Chabane, N., Gervais, H., Good, C.D., Bourgeois, M., Plumet, M.-H., Barthélémy, C., Mouren, M.-C., Artiges, E., Samson, Y., Brunelle, F., Frackowiak, R.S.J., Zilbovicius, M., 2004. Superior temporal sulcus anatomical abnormalities in childhood autism: A voxel-based morphometry MRI study. NeuroImage 23 (1), 364–369. https://doi.org/10.1016/j.neuroimage.2004.06.016.

Bourgeron, T., 2009. A synaptic trek to autism. Curr. Opin. Neurobiol. 19 (2), 231–234. https://doi.org/10.1016/j.conb.2009.06.003.

Brun, C.C., Nicolson, R., Leporé, N., Chou, Y.-Y., Vidal, C.N., DeVito, T.J., Drost, D.J., Williamson, P.C., Rajakumar, N., Toga, A.W., Thompson, P.M., 2009. Mapping brain abnormalities in boys with autism. Hum. Brain Mapp. 30 (12), 3887–3900. https://doi.org/10.1002/hbm.20814.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafò, M.R., 2013. Power failure: Why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14 (5), 365–376. https://doi.org/10.1038/nrn3475.

Carlisi, C.O., Norman, L.J., Lukito, S.S., Radua, J., Mataix-Cols, D., Rubia, K., 2017. Comparative multimodal meta-analysis of structural and functional brain abnormalities in autism spectrum disorder and obsessive-compulsive disorder. Biol. Psychiatry 82 (2), 83–102. https://doi.org/10.1016/j.biopsych.2016.10.006.

Carp, J., 2012. The secret lives of experiments: Methods reporting in the fMRI literature. NeuroImage 63 (1), 289–300. https://doi.org/10.1016/j.neuroimage.2012.07.004.

Cauda, F., Geda, E., Sacco, K., D'Agata, F., Duca, S., Geminiani, G., Keller, R., 2011. Grey matter abnormality in autism spectrum disorder: An activation likelihood estimation meta-analysis study. J. Neurol. Neurosurg. Psychiatry 82 (12), 1304–1313. https://doi.org/10.1136/jnnp.2010.239111.

Courchesne, E., 2004. Brain development in autism: Early overgrowth followed by premature arrest of growth. Mental Retard. Dev. Disab. Res. Rev. 10 (2), 106–111. https://doi.org/10.1002/mrdd.20020.

Courchesne, E., Mouton, P.R., Calhoun, M.E., Semendeferi, K., Ahrens-Barbeau, C., Hallet, M.J., Barnes, C.C., Pierce, K., 2011. Neuron number and size in prefrontal cortex of children with autism. JAMA 306 (18), 2001–2010. https://doi.org/10.1001/jama.2011.1638.

Crum, W.R., Griffin, L.D., Hill, D.L.G., Hawkes, D.J., 2003. Zen and the art of medical image registration: Correspondence, homology, and quality. NeuroImage 20 (3), 1425–1437. https://doi.org/10.1016/j.neuroimage.2003.07.014.

Davatzikos, C., Genc, A., Xu, D., Resnick, S.M., 2001. Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. NeuroImage 14 (6), 1361–1369. https://doi.org/10.1006/nimg.2001.0937.

DeRamus, T.P., Kana, R.K., 2014. Anatomical likelihood estimation meta-analysis of grey and white matter anomalies in autism spectrum disorders. NeuroImage: Clin. 7, 525–536. https://doi.org/10.1016/j.nicl.2014.11.004.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31 (3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021.

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., Fitzgerald, J., Gallagher, L., Keehn, R.J.J., Kennedy, D.P., Lainhart, J.E., Luna, B., Mostofsky, S.H., Müller, R.-A., Nebel, M.B., Nigg, J.T., O'Hearn, K., Solomon, M., Toro, R., Vaidya, C.J., Wenderoth, N., White, T., Craddock, R.C., Lord, C., Leventhal, B., Milham, M.P., 2017. Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. Sci. Data 4 (1). https://doi.org/10.1038/sdata.2017.10.

Duerden, E.G., Mak-Fan, K.M., Taylor, M.J., Roberts, S.W., 2012. Regional differences in grey and white matter in children and adults with autism spectrum disorders: An activation likelihood estimate (ALE) meta-analysis. Autism Research 5 (1), 49–66. https://doi.org/10.1002/aur.2012.5.issue-110.1002/aur.235.

Ecker, C., 2017. The neuroanatomy of autism spectrum disorder: An overview of structural neuroimaging findings and their translatability to the clinical setting. Autism 21 (1), 18–28. https://doi.org/10.1177/1362361315627136.

Ecker, C., Shahidiani, A., Feng, Y., Daly, E., Murphy, C., D'Almeida, V., Deoni, S., Williams, S.C., Gillan, N., Gudbrandsen, M., Wichers, R., Andrews, D., Van Hemert, L., Murphy, D.G.M., 2014. The effect of age, diagnosis, and their interaction on vertex-based measures of cortical thickness and surface area in autism spectrum disorder. J. Neural Transm. 121 (9), 1157–1170. https://doi.org/10.1007/s00702-014-1207-1.

Eckert, M.A., Tenforde, A., Galaburda, A.M., Bellugi, U., Korenberg, J.R., Mills, D., Reiss, A.L., 2006. To modulate or not to modulate: Differing results in uniquely shaped Williams syndrome brains. NeuroImage 32 (3), 1001–1007. https://doi.org/10.1016/j.neuroimage.2006.05.014.

Eickhoff, S., Walters, N.B., Schleicher, A., Kril, J., Egan, G.F., Zilles, K., Watson, J.D.G., Amunts, K., 2005. High-resolution MRI reflects myeloarchitecture and cytoarchitecture of human cerebral cortex. Hum. Brain Mapp. 24 (3), 206–215. https://doi.org/10.1002/(ISSN)1097-019310.1002/hbm.v24:310.1002/hbm.20082.

Elliot, C., 2007. The Differential Abilities Scale, second ed. Harcourt Assessments Inc.

Gennatas, E.D., Avants, B.B., Wolf, D.H., Satterthwaite, T.D., Ruparel, K., Ciric, R., Hakonarson, H., Gur, R.E., Gur, R.C., 2017. Age-related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood. J. Neurosci. 37 (20), 5065–5073. https://doi.org/10.1523/JNEUROSCI.3550-16.2017.

Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N.A., Friston, K.J., Frackowiak, R.S. J., 2001. A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains. NeuroImage 14 (1), 21–36. https://doi.org/10.1006/nimg.2001.0786.

Gotham, K., Pickles, A., Lord, C., 2009. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. J. Autism Dev. Disord. 39 (5), 693–705. https://doi.org/10.1007/s10803-008-0674-3.

Hua, X., Leow, A.D., Parikshak, N., Lee, S., Chiang, M.-C., Toga, A.W., Jack, C.R., Weiner, M.W., Thompson, P.M., 2008. Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: An MRI study of 676 AD, MCI, and normal subjects. NeuroImage 43 (3), 458–469. https://doi.org/10.1016/j.neuroimage.2008.07.013.

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. NeuroImage 62 (2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015.

Jeste, S.S., Geschwind, D.H., 2014. Disentangling the heterogeneity of autism spectrum disorder through genetic findings. Nat. Rev. Neurol. 10 (2), 74–81. https://doi.org/10.1038/nrneurol.2013.278.

Khundrakpam, B.S., Lewis, J.D., Kostopoulos, P., Carbonell, F., Evans, A.C., 2017. Cortical thickness abnormalities in autism spectrum disorders through late childhood, adolescence, and adulthood: A large-scale MRI study. Cereb. Cortex 27 (3), 1721–1731. https://doi.org/10.1093/cercor/bhx038.

Liu, J., Yao, L., Zhang, W., Xiao, Y., Liu, L., Gao, X., Shah, C., Li, S., Tao, B., Gong, Q., Lui, S., 2017. Gray matter abnormalities in pediatric autism spectrum disorder: A meta-analysis with signed differential mapping. Eur. Child Adolesc. Psychiatry 26 (8), 933–945. https://doi.org/10.1007/s00787-017-0964-4.

Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Leventhal, B.L., DiLavore, P.C., Pickles, A., Rutter, M., 2000. The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. J. Autism Dev. Disord. 30 https://doi.org/10.1023/A:1005592401947.

Lukito, S., Norman, L., Carlisi, C., Radua, J., Hart, H., Simonoff, E., Rubia, K., 2020. Comparative meta-analyses of brain structural and functional abnormalities during cognitive control in attention-deficit/hyperactivity disorder and autism spectrum disorder. Psychol. Med. 50 (6), 894–919. https://doi.org/10.1017/S0033291720000574.

Mei, T., Llera, A., Floris, D.L., Forde, N.J., Tillmann, J., Durston, S., Moessnang, C., Banaschewski, T., Holt, R.J., Baron-Cohen, S., Rausch, A., Loth, E., Dell'Acqua, F., Charman, T., Murphy, D.G.M., Ecker, C., Beckmann, C.F., Buitelaar, J.K., the EU-AIMS LEAP group, 2020. Gray matter covariations and core symptoms of autism: The EU-AIMS Longitudinal European Autism Project. Molecular Autism 11 (1), 86. https://doi.org/10.1186/s13229-020-00389-4.

Nakamura, K., Brown, R.A., Araujo, D., Narayanan, S., Arnold, D.L., 2014. Correlation between brain volume change and T2 relaxation time induced by dehydration and rehydration: Implications for monitoring atrophy in clinical studies. NeuroImage: Clinical 6, 166–170. https://doi.org/10.1016/j.nicl.2014.08.014.

Nickl-Jockschat, T., Habel, U., Maria Michel, T., Manning, J., Laird, A.R., Fox, P.T., Schneider, F., Eickhoff, S.B., 2012. Brain structure anomalies in autism spectrum disorder—A meta-analysis of VBM studies using anatomic likelihood estimation. Hum. Brain Mapp. 33 (6), 1470–1489. https://doi.org/10.1002/hbm.v33.610.1002/hbm.21299.

Noble, K.G., Houston, S.M., Brito, N.H., Bartsch, H., Kan, E., Kuperman, J.M., Akshoomoff, N., Amaral, D.G., Bloss, C.S., Libiger, O., Schork, N.J., Murray, S.S., Casey, B.J., Chang, L., Ernst, T.M., Frazier, J.A., Gruen, J.R., Kennedy, D.N., Van Zijl, P., Mostofsky, S., Kaufmann, W.E., Kenet, T., Dale, A.M., Jernigan, T.L., Sowell, E.R., 2015. Family income, parental education and brain structure in children and adolescents. Nat. Neurosci. 18 (5), 773–778. https://doi.org/10.1038/nn.3983.

Ohta, H., Nordahl, C.W., Iosif, A.-M., Lee, A., Rogers, S., Amaral, D.G., 2016. Increased surface area, but not cortical thickness, in a subset of young boys with autism spectrum disorder: cortical thickness in autism spectrum disorder. Autism Res. 9 (2), 232–248. https://doi.org/10.1002/aur.2016.9.issue-210.1002/aur.1520.

Orosco, L.A., Ross, A.P., Cates, S.L., Scott, S.E., Wu, D., Sohn, J., Pleasure, D., Pleasure, S. J., Adamopoulos, I.E., Zarbalis, K.S., 2014. Loss of Wdfy3 in mice alters cerebral cortical neurogenesis reflecting aspects of the autism pathology. Nat. Commun. 5, 4692. https://doi.org/10.1038/ncomms5692.

Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Jernigan, T.L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M.J., Grant, M.D., Franz, C.E., Xian, H., Tsuang, M., Fischl, B., Seidman, L., Dale, A., Kremen, W.S., 2009. Distinct genetic influences on cortical surface area and cortical thickness. Cerebral Cortex (New York, NY) 19 (11), 2728–2735. https://doi.org/10.1093/cercor/bhp026.

Radua, J., Canales-Rodríguez, E.J., Pomarol-Clotet, E., Salvador, R., 2014. Validity of modulation and optimal settings for advanced voxel-based morphometry. NeuroImage 86, 81–90. https://doi.org/10.1016/j.neuroimage.2013.07.084.

Raznahan, A., Lenroot, R., Thurm, A., Gozzi, M., Hanley, A., Spence, S.J., Swedo, S.E., Giedd, J.N., 2013. Mapping cortical anatomy in preschool aged children with autism using surface-based morphometry. NeuroImage: Clinical 2, 111–119. https://doi.org/10.1016/j.nicl.2012.10.005.

Rojas, D.C., Peterson, E., Winterrowd, E., Reite, M.L., Rogers, S.J., Tregellas, J.R., 2006. Regional gray matter volumetric changes in autism associated with social and repetitive behavior symptoms. BMC Psychiatry 6 (1), 56. https://doi.org/10.1186/1471-244X-6-56.

Rutter, M., Le Couteur, A., Lord, C., & Faggioli, R. (2005). *ADI-R: Autism diagnostic interview—Revised: Manual*. OS, Organizzazioni speciali.

Sacco, R., Militerni, R., Frolli, A., Bravaccio, C., Gritti, A., Elia, M., Curatolo, P., Manzi, B., Trillo, S., Lenti, C., Saccani, M., Schneider, C., Melmed, R., Reichelt, K.-L., Pascucci, T., Puglisi-Allegra, S., Persico, A.M., 2007. Clinical, morphological, and biochemical correlates of head circumference in autism. Biol. Psychiatry 62 (9), 1038–1047. https://doi.org/10.1016/j.biopsych.2007.04.039.

Shi, F., Wang, L., Dai, Y., Gilmore, J.H., Lin, W., Shen, D., 2012. LABEL: Pediatric brain extraction using learning-based meta-algorithm. NeuroImage 62 (3), 1975–1986. https://doi.org/10.1016/j.neuroimage.2012.05.042.

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

Sugathan, A., Biagioli, M., Golzio, C., Erdin, S., Blumenthal, I., Manavalan, P., Ragavendran, A., Brand, H., Lucente, D., Miles, J., Sheridan, S.D., Stortchevoi, A., Kellis, M., Haggarty, S.J., Katsanis, N., Gusella, J.F., Talkowski, M.E., 2014. CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. PNAS 111 (42), E4468–E4477. https://doi.org/10.1073/pnas.1405266111.

Tang, G., Gudsnuk, K., Kuo, S.-H., Cotrina, M.L., Rosoklija, G., Sosunov, A., Sonders, M. S., Kanter, E., Castagna, C., Yamamoto, A., Yue, Z., Arancio, O., Peterson, B.S., Champagne, F., Dwork, A.J., Goldman, J., Sulzer, D., 2014. Loss of mTOR-Dependent Macroautophagy Causes Autistic-like Synaptic Pruning Deficits. Neuron 83 (5), 1131–1143. https://doi.org/10.1016/j.neuron.2014.07.040.

Tunç, B., Yankowitz, L.D., Parker, D., Alappatt, J.A., Pandey, J., Schultz, R.T., Verma, R., 2019. Deviation from normative brain development is associated with symptom severity in autism spectrum disorder. Molecular Autism 10 (1), 46. https://doi.org/10.1186/s13229-019-0301-5.

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 Trans. Med. Imaging 29 (6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908.

van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G.F., Calderoni, S., Daly, E., Deruelle, C., Di Martino, A., Dinstein, I., Duran, F.L.S., Durston, S., Ecker, C., Fair, D., Fedor, J., Fitzgerald, J., Freitag, C.M., Gallagher, L., Gori, I., Haar, S., Hoekstra, L., Jahanshad, N., Jalbrzikowski, M., Janssen, J., Lerch, J., Luna, B., Martinho, M.M., McGrath, J., Muratori, F., Murphy, C.M., Murphy, D.G.M., O'Hearn, K., Oranje, B., Parellada, M., Retico, A., Rosa, P., Rubia, K., Shook, D., Taylor, M., Thompson, P.M., Tosetti, M., Wallace, G.L., Zhou, F., Buitelaar, J.K., 2018. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: results From the ENIGMA ASD Working Group. Am. J. Psychiatry 175 (4), 359–369. https://doi.org/10.1176/appi.ajp.2017.17010100.

Via, E., Radua, J., Cardoner, N., Happé, F., Mataix-Cols, D., 2011. Meta-analysis of Gray matter abnormalities in autism spectrum disorder: should asperger disorder be

subsumed under a broader umbrella of autistic spectrum disorder? Arch. Gen. Psychiatry 68 (4), 409–418. https://doi.org/10.1001/archgenpsychiatry.2011.27.
Wechsler, D., 1999. Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation, Harcourt Brace & Company.
Wechsler, D., 2011. Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II). NCS Pearson.
Wechsler, D., Kaplan, E., Fein, D., Kramer, J., Morris, R., Delis, D., Maelender, A., 2003. Wechsler intelligence scale for children: Fourth edition (WISC-IV). Pearson.
Weinberger, D.R., Radulescu, E., 2020. Structural Magnetic Resonance Imaging All Over Again. JAMA Psychiatry 78 (1), 11. https://doi.org/10.1001/jamapsychiatry.2020.1941.
Yang, X., Si, T., Gong, Q., Qiu, L., Jia, Z., Zhou, M., Zhao, Y., Hu, X., Wu, M., Zhu, H., 2016. Brain gray matter alterations and associated demographic profiles in adults with autism spectrum disorder: A meta-analysis of voxel-based morphometry studies. Aust. N. Z. J. Psychiatry 50 (8), 741–753. https://doi.org/10.1177/0004867415623858.
Yankowitz, L.D., Herrington, J.D., Yerys, B.E., Pereira, J.A., Pandey, J., Schultz, R.T., 2020. Evidence against the "normalization" prediction of the early brain overgrowth hypothesis of autism. Molecular Autism 11 (1), 51. https://doi.org/10.1186/s13229-020-00353-2.
Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R. L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106 (3), 1125–1165. https://doi.org/10.1152/jn.00338.2011.
Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., 2006. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. NeuroImage 31 (3), 1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015.
Zielinski, B.A., Anderson, J.S., Froehlich, A.L., Prigge, M.B.D., Nielsen, J.A., Cooperrider, J.R., Cariello, A.N., Fletcher, P.T., Alexander, A.L., Lange, N., Bigler, E. D., Lainhart, J.E., Soriano-Mas, C., 2012. ScMRI Reveals Large-Scale Brain Network Abnormalities in Autism. PLoS ONE 7 (11), e49172. https://doi.org/10.1371/journal.pone.0049172.
Zielinski, B.A., Prigge, M.B.D., Nielsen, J.A., Froehlich, A.L., Abildskov, T.J., Anderson, J. S., Fletcher, P.T., Zygmunt, K.M., Travers, B.G., Lange, N., Alexander, A.L., Bigler, E. D., Lainhart, J.E., 2014. Longitudinal changes in cortical thickness in autism and typical development. *Brain: A*. J. Neurol. 137 (Pt 6), 1799–1812. https://doi.org/10.1093/brain/awu083.