Archival Report

Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using Normative Models

Mariam Zabihi, Marianne Oldehinkel, Thomas Wolfers, Vincent Frouin, David Goyard, Eva Loth, Tony Charman, Julian Tillmann, Tobias Banaschewski, Guillaume Dumas, Rosemary Holt, Simon Baron-Cohen, Sarah Durston, Sven Bölte, Declan Murphy, Christine Ecker, Jan K. Buitelaar, Christian F. Beckmann, and Andre F. Marquand

ABSTRACT

BACKGROUND: The neuroanatomical basis of autism spectrum disorder (ASD) has remained elusive, mostly owing to high biological and clinical heterogeneity among diagnosed individuals. Despite considerable effort toward understanding ASD using neuroimaging biomarkers, heterogeneity remains a barrier, partly because studies mostly employ case-control approaches, which assume that the clinical group is homogeneous.

METHODS: Here, we used an innovative normative modeling approach to parse biological heterogeneity in ASD. We aimed to dissect the neuroanatomy of ASD by mapping the deviations from a typical pattern of neuroanatomical development at the level of the individual and to show the necessity to look beyond the case-control paradigm to understand the neurobiology of ASD. We first estimated a vertexwise normative model of cortical thickness development using Gaussian process regression, then mapped the deviation of each participant from the typical pattern. For this, we employed a heterogeneous cross-sectional sample of 206 typically developing individuals (127 males) and 321 individuals with ASD (232 males) (6–31 years of age).

RESULTS: We found few case-control differences, but the ASD cohort showed highly individualized patterns of deviations in cortical thickness that were widespread across the brain. These deviations correlated with severity of repetitive behaviors and social communicative symptoms, although only repetitive behaviors survived corrections for multiple testing.

CONCLUSIONS: Our results 1) reinforce the notion that individuals with ASD show distinct, highly individualized trajectories of brain development and 2) show that by focusing on common effects (i.e., the "average ASD participant"), the case-control approach disguises considerable interindividual variation crucial for precision medicine.

Keywords: Autism, Cortical thickness, Gaussian process, Heterogeneity, Normative modeling, Outlier detection

https://doi.org/10.1016/j.bpsc.2018.11.013

Autism spectrum disorder (ASD) is a lifelong neurodevelopmental disorder diagnosed exclusively on the basis of symptomatology, period of onset, and impairment (i.e., impairments in social communication and interaction, alongside repetitive stereotyped behavior and sensory anomalies) (1). Autism is well recognized as being highly heterogeneous on multiple levels—for example, in terms of its clinical presentation and underlying neurobiology. Indeed, more than 100 genes (2) and many aspects of brain structure have been associated with ASD at the group level (3). Autism is also grounded in the process of brain maturation, and it is believed that alterations are evident throughout brain development (4,5). In particular, differences in cortical thickness (CT) have been reported across different studies and ages (6), which—together with differences in surface area (SA) (6–10)—underpin regional differences in brain volume in ASD

(10–13). However, the precise etiology of the disorder in terms of brain development and underlying mechanisms remain elusive.

The heterogeneity of ASD is a fundamental barrier to understanding the neurobiology of ASD and the development of interventions (14). Regional group-level differences have been reported across several neuroanatomical measures, including CT (8,10,15–22). However these findings show generally poor replication across studies (3,7,19,23,24) and small effect sizes (8,19). Heterogeneity is also evident in studies that have used classifiers to discriminate ASD participants from control subjects, which mostly show relatively low accuracy for predicting diagnosis, especially in large samples (19,25,26). An important reason for this is that most studies to date have employed a traditional case-control approach, which is based on the assumption that the clinical and control groups are

homogeneous entities (7,27). Thus, the case-control approach provides information about alterations at the group level or, in other words, in the "average ASD participant." However, different participants may have different symptom profiles and different etiological pathways, and resulting neurobiological changes may converge on the same symptoms. Therefore, to understand the neurobiology of ASD, it is important to understand the range of associated neurobiological variation, which may subsequently inform intervention at the level of the individual in the spirit of "precision medicine" (28). A common approach to study the biological heterogeneity underlying ASD is to find subtypes using clustering algorithms, mostly on the basis of symptoms or behavioral characteristics (29–34). This approach has been somewhat successful and is appropriate if the clinical cohort can be cleanly partitioned into a relatively small number of homogeneous subgroups on the basis of the chosen measures. However, it does not tackle heterogeneity within subgroups, and it may be the case that no clearly defined subgroups exist in the data. Moreover, subgroups derived from behavior or symptoms require extensive validation on external measures and still may not fully reflect the underlying biology (35,36).

Here, we apply a complementary normative modeling approach (36,37) to understand the biological heterogeneity of ASD. This shifts the focus away from group-level comparisons—which can detect consistent differences across groups of individuals (e.g., diagnoses or putative subtypes)—toward characterizing the degree of alteration in each individual, with reference to the typically developing (TD) brain. This allows us to detect and map neuroanatomical alterations at the level of the individual and has recently shown promise in understanding the biological variation of psychotic disorders (37). Normative modeling is analogous to the use of growth charts in pediatric medicine, which allow the development (e.g., in terms of height or weight) of each individual child to be measured against expected centiles of variation in the population. To achieve this, we first estimated a statistical model characterizing typical cortical development that accurately quantifies the variation within the population and across brain development. We then placed each individual ASD participant in relation to the typical distribution to identify alterations in individual cases with respect to the typical pattern of brain maturation. Our main goals were to 1) to map the neuroanatomical features by which each individual ASD participant differs from the expected TD pattern, across both different developmental stages and different levels of functioning, and thereby 2) demonstrate the value of normative modeling techniques for understanding the biological heterogeneity of ASD. For this, we employed data from a large international study (38) with harmonized data acquisition procedures and a design that naturally groups subjects according to different developmental stages. While normative modeling is suitable for many different aspects of brain structure or function, here we focused on CT, which is a sensitive and reliable measure of cortical morphology in ASD (6,8,39), although we also investigated SA. Ultimately, we hope this approach will yield a set of individualized neurobiological "fingerprints" facilitating a route toward precision medicine approaches in ASD (28).

METHODS AND MATERIALS

Participants

Full details on study design and clinical characteristics have been described previously (38). Briefly, we included all participants from the Longitudinal European Autism Project (40) cohort with a structural magnetic resonance imaging scan surviving quality control and the necessary clinical and demographic data. We included 206 TD individuals 7 to 31 years of age (127 males) (Table 1; Supplemental Table S1 and Supplemental Figure S1) and 321 individuals 6 to 31 years of age with ASD (232 males). There were no significant differences between the TD and ASD cohorts in age, but the IQ of ASD participants was lower than TD participants. Under the study design, each cohort was split into four subgroups according to age and level of intellectual ability (Table 1): 1) adults with ASD without intellectual disability (ID) and TD control subjects 18 to 30 years of age (IQ ≥70); 2) adolescents with ASD without ID and TD control subjects 12 to 17 years of age; 3) children with ASD without ID or TD control subjects 6 to 11 years of age; and 4) adolescents and adults with ASD and ID [i.e., full-scale IQ between 50 and 70 (1)] 12 to 30 years of age. Note that only TD participants were included in the estimation of the normative model.

TD participants were recruited via advertisement. Individuals with an existing ASD and/or mild ID diagnosis (according to DSM-5/ICD-10 criteria) were recruited from existing databases and clinic contacts across one of seven study sites: the Institute of Psychiatry, Psychology and

Table 1. Clinical Characteristics

Variable	ASD Cohort, $n = 321$ (89 Female)	TD Cohort, $n = 206$ (79 Female)	p Value
Age, Years	17.01 ± 5.79	17.14 ± 5.97	.93[a]
IQ			
Global IQ	100.89 ± 18.53, $n = 316$	108.22 ± 14.24	.00
Performance IQ	101.65 ± 20.14, $n = 316$	108.26 ± 15.72	.00
Verbal IQ	99.64 ± 18.53, $n = 313$	107.32 ± 16.13	.00
ADI-R	$n = 308$		
Social	16.20 ± 6.71		
Communication	13.11 ± 5.69		
Repetitive behavior	4.32 ± 2.69		
ADOS-2	$n = 258$		
Total	5.12 ± 2.77		
Social	5.78 ± 2.62		
Repetitive behavior	4.78 ± 2.76		
Schedule, n			
A: Adults	125	84	
B: Adolescents	112	70	
C: Children	64	52	
D: IQ <70	20	—	

Values are mean ± SD, except where noted.

ADI-R, Autism Diagnostic Interview-Revised; ADOS-2, Autism Diagnostic Observation Schedule, Second Edition; ASD, autism spectrum disorder; TD, typically developing.

[a]Not significant.

Dissecting Heterogeneity of ASD With Normative Modeling

Neuroscience, King's College London, London, United Kingdom; Autism Research Centre at the University of Cambridge, Cambridge, United Kingdom; Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands; University Medical Centre Utrecht, Utrecht, the Netherlands; Central Institute of Mental Health, Mannheim, Germany; and University Campus Bio-Medico, Rome, Italy. The combined information from the Autism Diagnostic Interview-Revised (ADI-R) (41) and Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) (42) were used to measure symptom severity (33). However, individuals with a clinical ASD diagnosis who did not reach conventional cutoffs on these instruments were not excluded. The ADI-R is a parent-reported measure of lifetime or past developmental window symptom severity, whereas the ADOS-2 is an expert rating of current symptoms. A standard set of exclusion criteria were applied and are provided in the Supplement. All subjects were scanned with a T1-weighted imaging protocol, and FreeSurfer (version 5.3; https://surfer. nmr.mgh.harvard.edu/) was used to estimate measures of regional CT and SA. See the Supplemental Methods for details.

Constructing a Normative Model of CT

An overview of the normative modeling approach is shown in Figure 1 and has been described previously (36). Briefly, Gaussian process regression (43) was used to estimate separate normative models of CT and SA at each vertex on the cortical surface (see Supplemental Methods for details). This normative model can be used to predict both the expected CT and the associated predictive uncertainty for each individual participant. The contours of predictive uncertainty can then be used to model centiles of variation within the cohort. This allows us to place each individual participant within the normative distribution, thereby quantifying the vertexwise deviation of CT from the healthy range across the brain.

To achieve this, we generated a developmental model of typical brain development by training a Gaussian process regression model on the TD cohort ($n = 206$) using age and gender as covariates (i.e., independent variables) to predict CT (i.e., dependent variable). In pediatric medicine, growth charts are normally estimated on the basis of a large population cohort (i.e., potentially including patients with various disorders based on the population prevalence). In our sample, the prevalence of ASD is much higher than in the population, so for simplicity and to avoid the normative model's being enriched for ASD, we estimated the normative model on the basis of the TD participants only. Moreover, while the amount of data we employ here is relatively small in comparison with population-based studies, our Bayesian statistical model provides a principled method to handle uncertainty and therefore automatically makes inferences more conservative as the number of data points decreases, although more data would allow more precise estimates. To assess generalization, we used 10-fold cross-validation before retraining the model using the whole dataset to make predictions on the ASD participants following standard practice in machine learning (Supplemental Methods). Importantly, all parameters were estimated using the training data using empirical Bayesian estimation (36), and the use of cross-validation ensures unbiased estimates for the
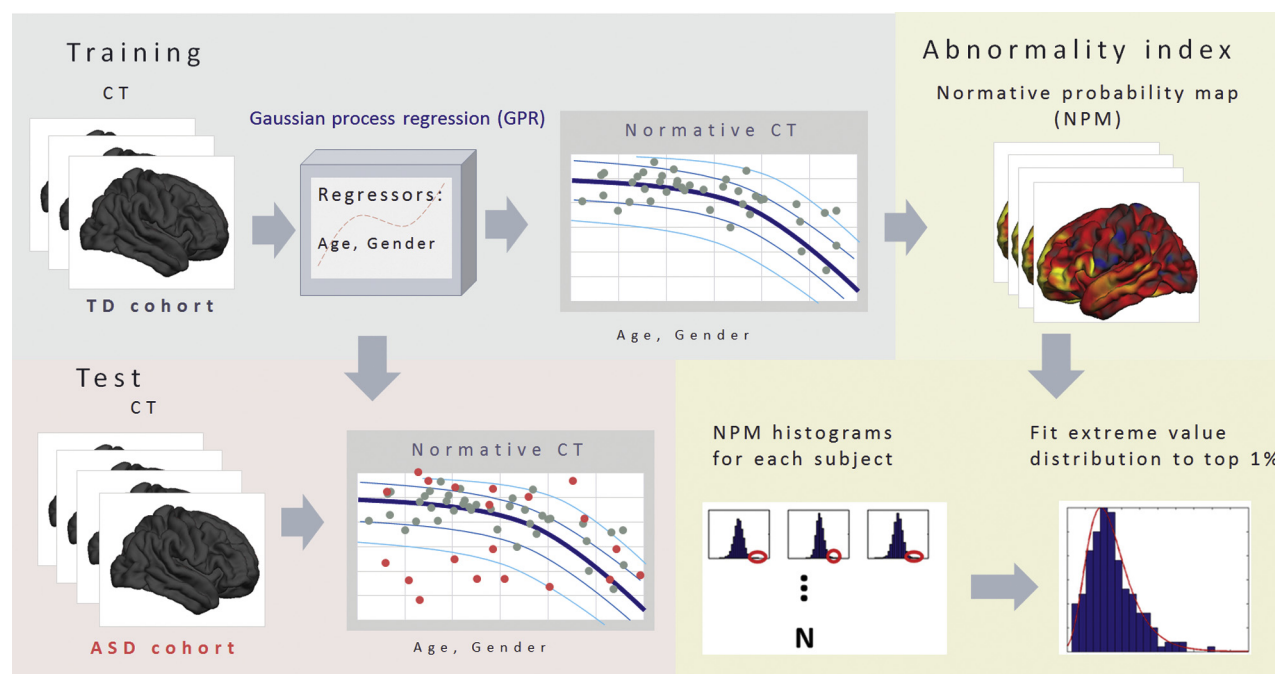

Figure 1. Methodological overview. First, a normative model was estimated from cortical thickness derived from typically developing (TD) subjects (gray dots). Then we used this model to predict cortical thickness (CT) in autism spectrum disorder (ASD) subjects (red dots). This allowed us to estimate normative probability maps, which show the regional deviations from the expected pattern in each subject. Finally, we generated a summary statistic quantifying the overall deviation for each subject by taking maximum deviation across brain using extreme value statistics.

TD cohort as well as for the ASD cohort. Therefore, deviations can be compared with one another.

Estimating Regional Deviations for Each Subject

To estimate a pattern of regional deviations from typical CT for each participant, we derived a normative probability map (NPM) that quantifies the deviation from the normative model for CT at each vertex. This was done by using the normative model to predict vertexwise estimates of CT for each individual participant, then estimating a subject-specific Z score (36) (Supplemental Methods). This provides a statistical estimate of how much each individual differs from the healthy pattern at each vertex. We thresholded the NPMs, correcting for multiple comparisons by controlling false discovery rate (FDR) at $p < .05$ within each participant, as in Marquand et al. (36).

To measure the spatial overlap of the individualized deviations across the cohort, we calculated an overlap map by counting the significant (FDR-corrected) vertices derived from the Z-score maps across all subject-level NPMs. The resulting summary maps indicate the spread of vertexwise deviations across the brain, separately for positive and negative deviations. This allowed us to identify a set of brain regions where participants had increased (positive deviation) or decreased (negative deviation) CT relative to the reference cohort.

To provide a simple comparison for these subject-level deviations, we also estimated a standard vertexwise general linear model to establish significant differences between groups including age as a covariate. We also investigated models including quadratic and cubic age terms (corrected using FDR at $p < .05$) and separate models for male and female subjects.

Constructing an Individual-Level Atypicality Score

A key benefit of normative modeling is a probabilistic interpretation of the deviations across all subjects. The NPMs therefore provide a multivariate measure of deviation from the normative range across all brain regions. This captures spatially distributed differences from the TD pattern. To better understand most important focal differences for each subject, we estimated a summary score for each participant capturing the individual's largest deviation from the typical pattern (which is potentially the most clinically relevant). This can be modeled using extreme value statistics (44) and is based on the notion that the expected maximum of any random variable converges to an extreme value distribution. Therefore, we estimated a maximum deviation for each subject by taking a trimmed mean of 1% of the top absolute deviations for each subject across all vertices and fit an extreme value distribution to these deviations.

Mapping Behavioral Associations

Last, to assess the clinical relevance of these deviations, we computed Spearman correlation coefficients between global and regional extreme deviation from the normative model and ADOS-2/ADI-R symptom severity scores ($p < .05$, FDR). The global measure (described above) provides an overall summary of the deviation for each individual, while the regional assessment helps to determine the functional correspondence of the deviations across individuals on a region-by-region

basis. The regional extreme deviation was computed as the trimmed mean of the 1% of top absolute deviations for each region after parcellating the cortex using the Desikan-Killiany atlas (45).

Checking for Potential Confounds

To investigate whether potential confounds could have influenced our findings, we estimated a separate normative model additionally including dummy regressors for IQ, site, and FreeSurfer Euler number (46). We also performed post hoc tests between the deviations from the normative model and potential confounding variables (IQ, comorbid symptoms, and surrogate measures of image quality) (see Supplemental Tables S2 and S3).

RESULTS

A Normative Model Quantifying the Decline of CT With Age

Figure 2 shows the developmental normative model of CT derived from the TD male cohort, thresholded to show vertices where the correlation between true and predicted labels was higher than predicted by chance ($p < .05$, FDR corrected) (see Supplemental Figure S3 for female cohort). The unthresholded map showing the correlation between true and predicted CT values is shown in Supplemental Figure S2 along with the root mean square error of the normative model across different vertices. In most regions, CT decreases consistently and approximately linearly with age. However, in some regions, CT followed a nonlinear (i.e., inverted U-shaped) trajectory with an early rise followed by a decline, e.g., in the inferior temporal and posterior frontal regions. This corresponds well with the known developmental trajectory of CT (47–51). The normative model for SA showed a similar, relatively global pattern of decline as for CT (not shown).

Widespread Deviations From the Normative Pattern of CT Among the ASD Cohort

Figure 3 shows the classical mass-univariate group difference (i.e., case-control) map between ASD and TD cohorts. This shows few significant differences between groups; only two small regions of increased CT in the superior frontal and parietal cortices survived FDR correction. There were also few significant differences when additionally including quadratic and cubic age terms and no differences in the age-by-diagnosis interaction. The separate models for male and female subjects also did not show any significant differences after FDR correction.

Figures 4 and 5 show a summary of the NPMs for the ASD and TD cohorts. Specifically, these figures show the number of participants in each group that deviate negatively (Figure 4) or positively (Figure 5) from the normative model at each vertex after intraindividual FDR correction. Importantly, and contrast to the general linear model, these deviations need not overlap between subjects. As expected, the TD cohort shows few significant deviations, indicating that the normative model provides a good fit for this cohort. Crucially, this fit was achieved under cross-validation and is therefore unbiased. Therefore, under the null hypothesis that ASD participants

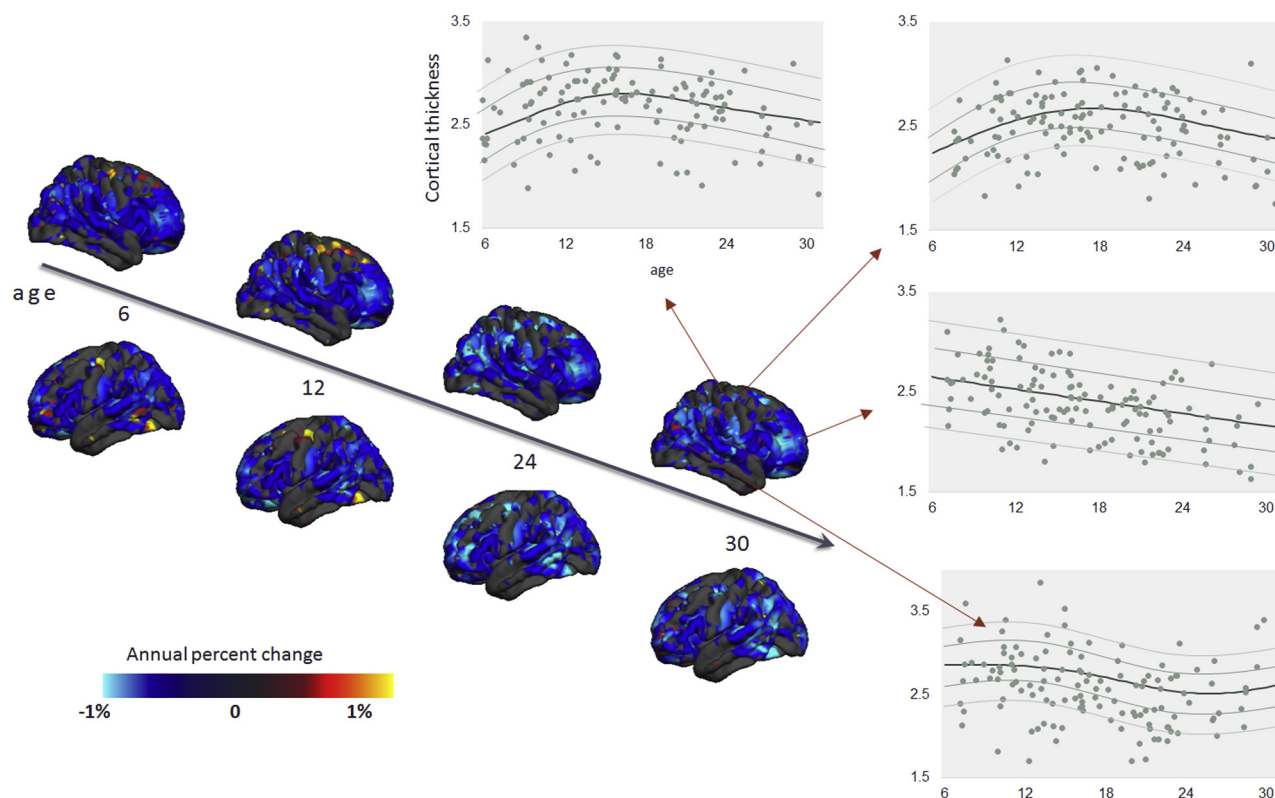Dissecting Heterogeneity of ASD With Normative Modeling

Figure 2. Normative model of developmental changes of cortical thickness across the developmental range in the typically developing male cohort (the model was estimated using both genders). Cortical thickness was predicted using a trained normative model across the age range of 6 to 31 years of age. The predicted cortical thickness map was thresholded so that only vertices that could accurately predict the true cortical thickness in the healthy cohort under cross-validation were retained (Pearson correlation, $p < .05$, false discovery rate). Blue and yellow vertices indicate reduced and increased cortical thickness, respectively. Moreover, the predicted cross-sectional developmental trajectories of cortical thickness in four randomly selected vertices are shown.

follow a similar trajectory of brain development to TD participants, there is no prior reason to expect that the fit will be better in TD than in ASD participants. In contrast, the total number of deviating vertices was noticeably higher in the ASD cohort and was widespread across the brain, suggesting that there are widespread and individualized deviations from the normative model in certain subsets of participants. When considering each age group separately, negative deviations were most prominent in children, whereas positive deviations were most prominent in adolescents and adults. The results were very similar for the models including IQ, scanning

site, and Euler number as covariates (Supplemental Figures S4 and S5), and a similar pattern of results was observed for SA, albeit with slight differences with respect to the pattern of deviations across brain regions (Supplemental Figures S6 and S7).

ASD Participants Deviate More Than TD Participants From the Normative Pattern of Development

Figure 6 shows the distribution of the most extreme deviations from the normative model across the brain. This shows that the
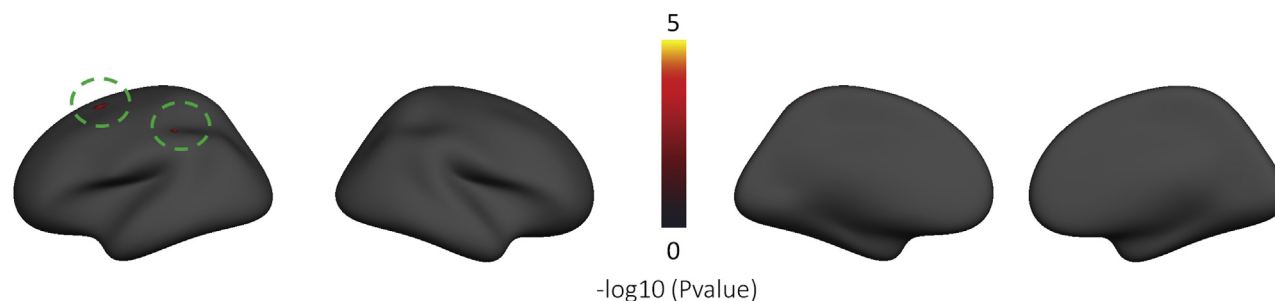

Figure 3. Vertexwise group differences between the autism spectrum disorder and typically developing cohorts after false discovery rate correction ($p < .05$). The green circles indicate the regions showing the vertexwise group difference. No vertices survived after false discovery rate correction in the vertexwise group differences map between autism spectrum disorder and typically developing female and male subjects.
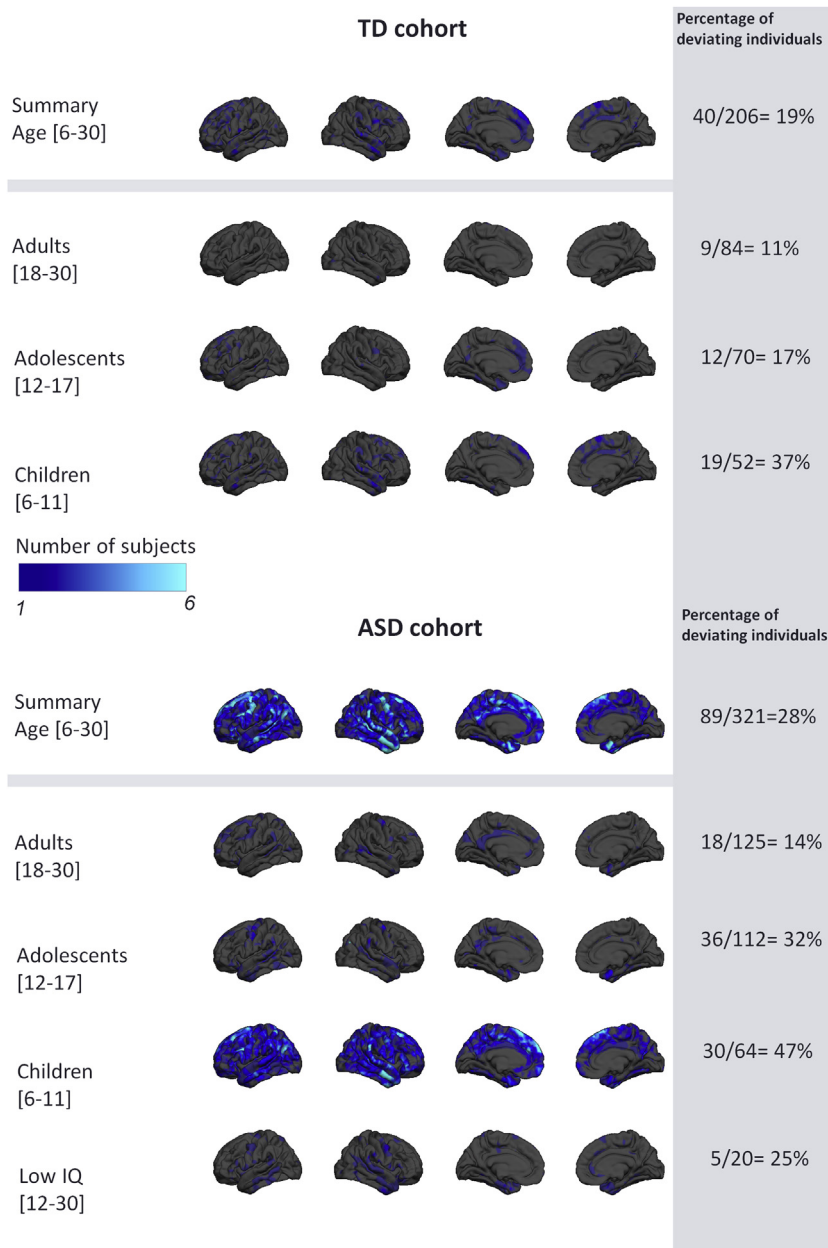
Figure 4. Overlap of vertexwise negative deviation across each cohort and schedule. This map shows the spatial distribution of individual subjects with significant deviations in each vertex after false discovery rate correction. The proportion of subjects contributing to each map is also shown (i.e., the proportion of subjects having deviations surviving false discovery rate correction). ASD, autism spectrum disorder; TD, typically developing.

maximum deviation across the brain is higher in the ASD cohort than the TD cohort and shows that the distribution of the ASD cohort is shifted toward the right, implying relatively more subjects with extreme deviations. Saliently, the top 15 deviating individuals belong to the ASD cohort, which is extremely unlikely to occur by chance ($p < .0005$, binomial test). The NPMs of these participants (Supplemental Figure S8) have highly individualized patterns of deviation not only with respect to brain regions, but also in sign, with some participants having positive deviations (i.e., greater CT) or negative deviations (reduced CT). These participants did not show a consistent pattern with respect to their symptom scores (Supplemental

Table S5), which underscores the degree of clinical and neurobiological heterogeneity within the ASD cohort. However, with regard to their demographic profile, subjects with predominantly positive deviations were adolescents or adults, while most subjects with negative deviations were children.

Association With Symptoms

Global deviations from the normative model were negatively associated with ADOS-2 repetitive behaviors ($\rho = -.21$, $p < .05$), and regional deviations were associated with symptoms in several brain regions (Figures 7 and 8). Associations were

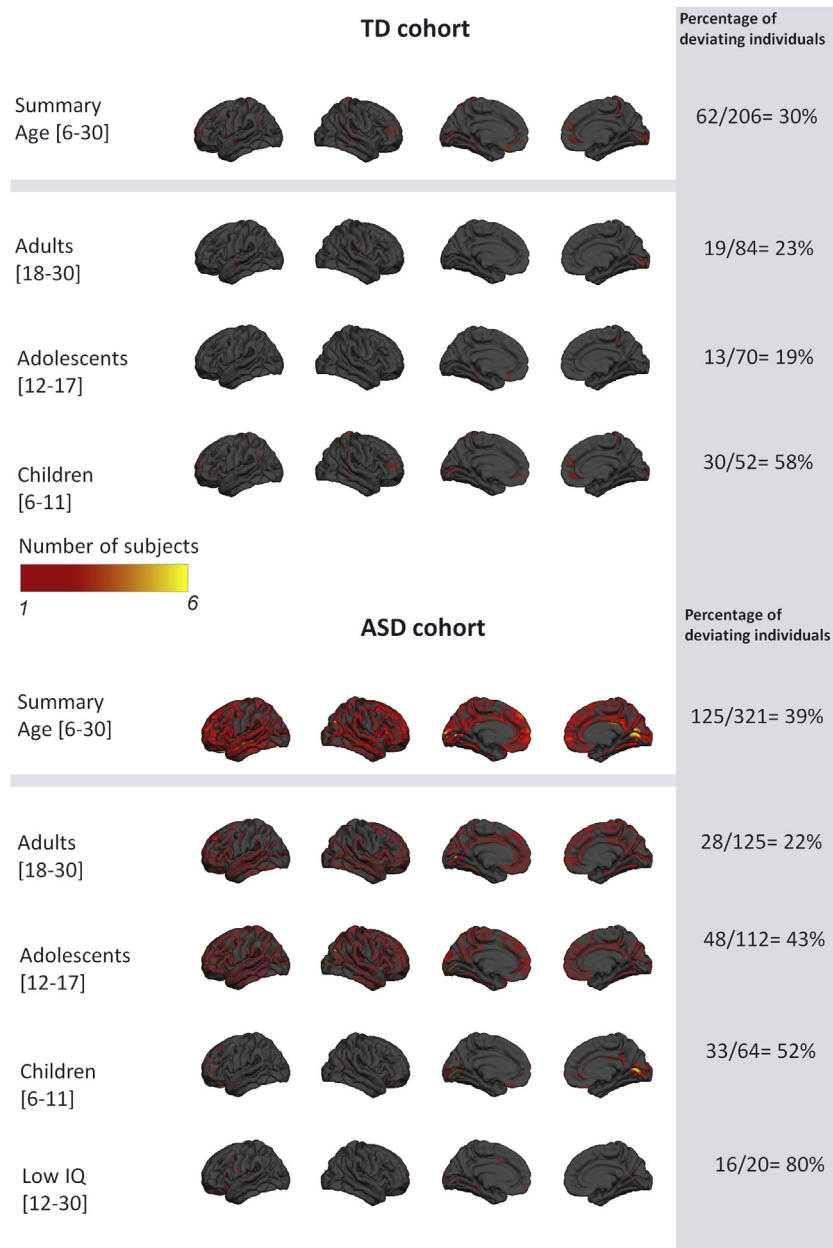Dissecting Heterogeneity of ASD With Normative Modeling

Figure 5. Overlap of vertexwise positive deviation across each cohort and schedule. See Figure 4 legend for further details. ASD, autism spectrum disorder; TD, typically developing.

found with symptom severity in the repetitive domain of the ADOS-2 or ADI-R in prefrontal regions in female subjects. In male subjects, a similar pattern was seen but did not survive multiple comparison correction, except for the superior frontal region in the ADI-R. Social interaction and communication scores also had nominally significant associations in female subjects, but these did not survive correction.

DISCUSSION

In this study, we aimed to dissect the heterogeneous neurobiology of ASD by mapping the deviation of each individual

participant from a normative model of CT development. In a large, heterogeneous cohort spanning a wide range of the ASD phenotype, we showed few significant group-level differences between ASD and TD cohorts in CT using a classical case-control analysis. In contrast, our normative modeling approach showed striking, widespread patterns of cortical atypicality at the level of individual ASD participant. These patterns were highly individualized across participants, distinct across different developmental stages, and associated with symptoms, especially repetitive behaviors. This supports the notion that a subset of ASD participants follow a different developmental trajectory than TD subjects, and that the
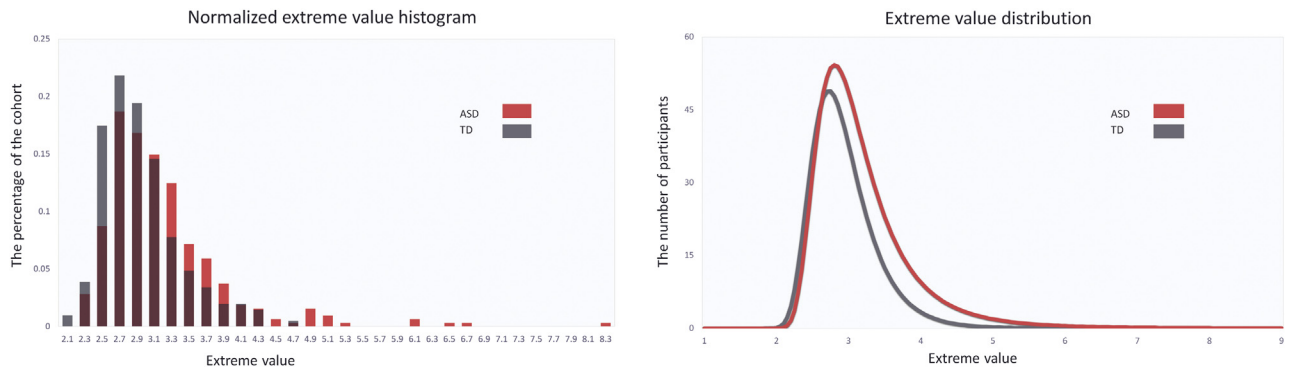
Figure 6. Extreme value histogram and distribution. ASD, autism spectrum disorder; TD, typically developing.

trajectory each ASD participant follows is highly individualized. From a methodological standpoint, our study shows that 1) it is necessary to look beyond the case-control paradigm to understand the heterogeneous neuroanatomy of ASD, 2)

normative modeling provides an alternative conceptual framework for understanding the heterogeneous neurobiology of ASD in terms of deviations from a typical pattern, and 3) focusing on an "average autistic individual" provides only a

Figure 7. Regional extreme value deviation correlation with autism spectrum disorder (ASD) symptoms for female subjects ($p < .05$) according to the Desikan-Killiany parcellation scheme. Blue and yellow regions indicate negative and positive association with autism spectrum disorder symptoms, respectively. Green circles indicate the regions that survived after false discovery rate correction. ADI, Autism Diagnostic Interview-Revised; ADOS II, Autism Diagnostic Observation Schedule, Second Edition; RRB, repetitive behavior.

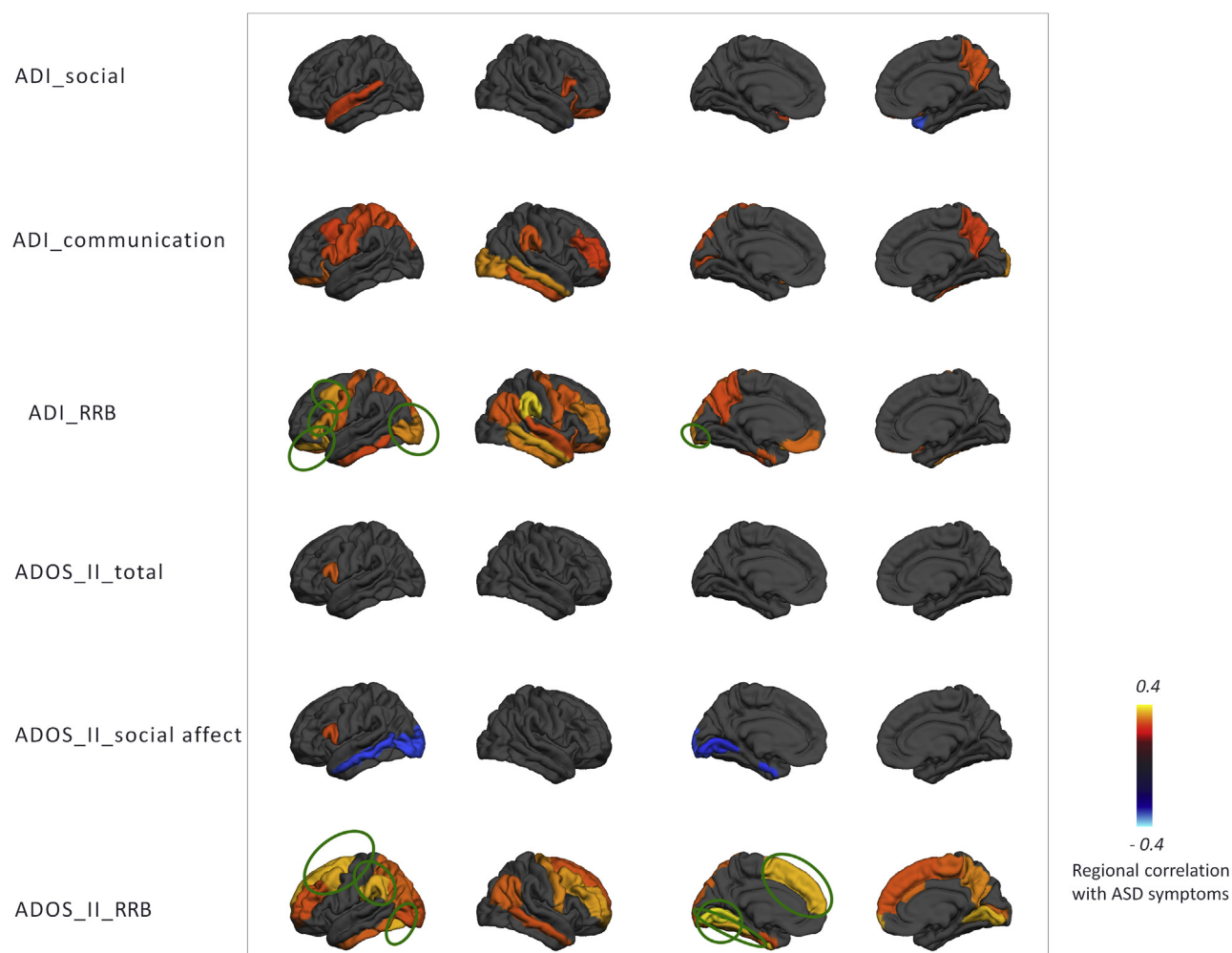Dissecting Heterogeneity of ASD With Normative Modeling

Figure 8. Regional extreme value deviation correlation with autism spectrum disorder (ASD) symptom for male subjects ($p < .05$) according to the Desikan-Killiany parcellation scheme. ADI, Autism Diagnostic Interview-Revised; ADOS II, Autism Diagnostic Observation Schedule, Second Edition; RRB, repetitive behavior.

partial reflection of the nature of the condition. In other words, the case-control approach focuses on common effects rather than interindividual variation. Capturing and capitalizing on such variation at the individual level is at the heart of precision medicine.

The normative model describes the variation in typical brain development showed a largely monotonic—and in some areas nonlinear—decrease of CT throughout development, consistent with previous neuroimaging studies (47–55). The fact that we observed widespread interindividual differences between ASD participants in terms of their deviations from the normative model explains why our classical case-control analysis revealed few significant differences and why several large previous neuroimaging studies have also only detected relatively modest group level effects (8,19). The heterogeneity underlying ASD is widely recognized (2,56–62); some studies have reported reductions in CT in ASD (15), whereas some studies have reported increases (16,63). Saliently, these inconsistencies remain evident even in large studies; for

example, a large study derived from the ENIGMA (Enhancing Neuro Imaging Genetics Through Meta Analysis) consortium demonstrated both regional increases and decreases in ASD at the group level that were consistent across development (8). Other studies—many derived from the ABIDE (Autism Brain Imaging Data Exchange) dataset (64)—have shown widespread increases in CT early in development that are attenuated later in development (19,20,48). Our results complement these studies because of our focus on studying individual variation within the ASD cohort. We show that 1) a subset of participants show decreased CT and SA in childhood while 2) other patients show regional increases in childhood in different areas (e.g., pericalcarine cortex), and 3) some participants show increased CT and SA in adolescence or adulthood. Crucially, however, these effects show minimal overlap across brain regions in different individuals. This is in line with another recent study applying normative modeling to ASD, which found effects in a subset of participants that were different from the main group effects (65). Thus, we consider that group-

level effects can be understood as the background on which individual variation is superimposed. The individualized deviations we report were mostly located in areas previously associated with ASD, such as the medial cortex including the cingulate and dorsomedial prefrontal regions, lateral prefrontal and parietal cortices, temporal cortices, and hippocampal formation (6,7,63,66,67). While some of these regions have been associated with social processing, the individual deviations in these regions were not associated with social interaction or communication symptoms at the group level. This could be for several reasons; for example, the anatomical patterns associated with these symptoms may be expressed in other measures of cortical anatomy [e.g., (68,69)] or in subcortical regions. Adults and adolescents had relatively fewer deviations, but these were positive (relatively increased CT and SA) and widespread across prefrontal and temporal cortices. Notably, we detected relatively few deviations in ASD with ID, which is important to exclude the possibility that these subjects were driving the effects described above. However, the ASD with ID group was relatively small ($n = 20$), so we do not draw strong conclusions about potential differences between ASD with and without ID.

The 15 subjects with the most atypical anatomy all had ASD, which is extremely unlikely to occur by chance. Moreover, these participants had individualized brain alterations and clinical characteristics. At the group level, the regional deviations we detected from the normative model were associated with the severity of lifetime and current autistic symptoms (ADI-R and ADOS-2, respectively), demonstrating that our model predictions may be clinically relevant. The deviation from the normative range was most informative about repetitive behavior symptom severity in that the strongest correlations were between CT in prefrontal regions with restricted repetitive behaviors, especially in female subjects and across both parental report via ADI-R and observer ratings of current symptoms via ADOS-2. These results broadly correspond with previous reports (6,70,71) and suggest that ASD may be more heterogeneous in male individuals, but we are cautious about this interpretation because we did not test it directly. Taken together, our results add weight to the importance of considering ASD in the context of a model of typical brain development and at the individual level (39,63,67).

Our findings should be considered in the light of several limitations. First, the trajectories of brain development were based on cross-sectional data and should be validated in a longitudinal cohort. Longitudinal follow-up data are currently being acquired and will be the subject of a future report. Moreover, while our sample size is similar to other neuroimaging studies of brain development [e.g., (72)], the model would yield more precise estimates with more data. Second, we registered all subjects to a standard adult template brain, as is standard in the field (10,63,67,73–75), which could cause bias. However, there were few deviations in the TD cohort, which makes this possibility unlikely. Third, our data do not permit strong inferences about the degree to which confounding variables may have influenced our findings. We found moderate associations between deviations from the normative models and a surrogate metric of image quality, but these were also associated with childhood ASD symptoms, comorbid attention-deficit/hyperactivity disorder symptoms, and IQ.

Moreover, our study design does not permit inferences about the direction of causality. For example, subjects with the most abnormal anatomy may also have the most impairment. Finally, we did not perform manual edits on the cortical surface reconstructions. While this eliminates one potential source of bias, the results need to be interpreted in the light of this, and it is possible that performing manual edits may improve the quality of the surface reconstructions in some cases.

In conclusion, we estimated a normative model of cortical development based on a large TD cohort and applied this model to a heterogeneous ASD cohort. Our results show that it is necessary to look beyond the case-control paradigm—which is limited to detecting group-level effects describing the "average ASD participant"—to understand the heterogeneous neurobiology of ASD. Normative modeling is well suited for this purpose, as it can chart the individualized deviation of each individual subject relative to the normative range, and hence provides an excellent tool for understanding the heterogeneity of psychiatric disorders.

ARTICLE INFORMATION

From the Department of Cognitive Neuroscience (MZ, MO, JKB, CFB, AFM), Department of Human Genetics (TW), Radboud University Medical Center; Donders Institute for Brain (MZ, MO, TW, JKB, CFB, AFM), Cognition and Behaviour, Radboud University; and Karakter Child and Adolescent Psychiatry University Centre (JKB), Nijmegen; and Department of Psychiatry, University Medical Centre (SD), Utrecht, the Netherlands; Department of Applied Psychology: Health, Development, Enhancement, and Intervention (JT), University of Vienna, Vienna, Austria; Neurospin (VF, DG), Institut des sciences du vivant Frédéric Joliot, CEA–Université Paris-Saclay,

Dissecting Heterogeneity of ASD With Normative Modeling

Gif-sur-Yvette; and Human Genetics and Cognitive Functions Unit (GD), Institut Pasteur, Paris, France; Department of Forensic and Neurodevelopmental Sciences (EL, DM, CE), Department of Psychology (TC, JT), Sackler Institute for Translational Neurodevelopment (DM), and Department of Neuroimaging (AFM), Institute of Psychiatry, Psychology and Neuroscience King's College London, London; Autism Research Centre (RH, SB-C), Department of Psychiatry, University of Cambridge, Cambridge; and Centre for Functional MRI of the Brain (CFB), University of Oxford, Oxford, United Kingdom; Department of Child and Adolescent Psychiatry and Psychotherapy (TB), Central Institute of Mental Health Mannheim, Mannheim; and Department of Child and Adolescent Psychiatry (CE), Psychosomatics and Psychotherapy, University Hospital Frankfurt am Main, Goethe University Frankfurt, Frankfurt, Germany; and the Center for Neurodevelopmental Disorders (SB), Division of Neuropsychiatry, Department of Women's and Children's Health; and Child and Adolescent Psychiatry (SB), Centre of Psychiatry Research, Stockholm County Council, Stockholm, Sweden.

Address correspondence to Mariam Zabihi, MSc., Radboudmc, Kapittelweg 29, 6525 EN Nijmegen, Gelderland, Netherlands; E-mail: m.zabihi@donders.ru.nl.

REFERENCES

1. American Psychiatric Association (2013): Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington, DC: American Psychiatric Press.
2. Betancur C (2011): Etiological heterogeneity in autism spectrum disorders: More than 100 genetic and genomic disorders and still counting. Brain Res 1380:42–77.
3. Ecker C (2016): The neuroanatomy of autism spectrum disorder: An overview of structural neuroimaging findings and their translatability to the clinical setting. Autism 21:18–28.
4. Walsh CA, Morrow EM, Rubenstein JLR (2008): Autism and brain development. Cell 135:396–400.
5. Schumann CM, Bloss CS, Barnes CC, Wideman GM, Carper RA, Akshoomoff N, et al. (2010): Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. J Neurosci 30:4419–4427.
6. Ecker C, Ginestet C, Feng Y, Johnston P, Lombardo MV, Lai M-C, et al. (2013): Brain surface anatomy in adults with autism: The relationship between surface area, cortical thickness, and autistic symptomsbrain surface anatomy in adults with autism. JAMA Psychiatry 70:59–70.
7. Hyde KL, Samson F, Evans AC, Mottron L (2010): Neuroanatomical differences in brain areas implicated in perceptual and other core features of autism revealed by cortical thickness analysis and voxel-based morphometry. Hum Brain Mapp 31:556–566.
8. van Rooij D, Anagnostou E, Arango C, Auzias G, Behrmann M, Busatto GF, et al. (2018): Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: Results from the ENIGMA ASD Working Group. Am J Psychiatry 175:359–369.
9. Ecker C, Bookheimer SY, Murphy DGM (2015): Neuroimaging in autism spectrum disorder: Brain structure and function across the lifespan. Lancet Neurol 14:1121–1134.
10. Hazlett HC, Poe M, Gerig G, Styner M, Chappell C, Smith RG, et al. (2011): Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. Arch Gen Psychiatry 68:467–476.
11. Piven J, Arndt S, Bailey J, Havercamp S, Andreasen NC, Palmer P (1995): An MRI study of brain size in autism. Am J Psychiatry 152:1145–1149.
12. Piven J, Arndt S, Bailey J, Andreasen N (1996): Regional brain enlargement in autism: A magnetic resonance imaging study. J Am Acad Child Adolesc Psychiatry 35:530–536.
13. Hardan AY, Minshew NJ, Mallikarjuhn M, Keshavan MS (2001): Brain volume in autism. J Child Neurol 16:421–424.
14. Lai MC, Lombardo MV, Chakrabarti B, Baron-Cohen S (2013): Subgrouping the autism "spectrum": Reflections on DSM-5. PLoS Biol 11:e1001544.
15. Hadjikhani N, Joseph RM, Snyder J, Tager-Flusberg H (2006): Anatomical differences in the mirror neuron system and social cognition network in autism. Cereb Cortex 16:1276–1282.
16. Mak-Fan KM, Taylor MJ, Roberts W, Lerch JP (2012): Measures of cortical grey matter structure and development in children with autism spectrum disorder. J Autism Dev Disord 42:419–427.
17. Wallace GL, Dankner N, Kenworthy L, Giedd JN, Martin A (2010): Age-related temporal and parietal cortical thinning in autism spectrum disorders. Brain 133:3745–3754.
18. Scheel C, Rotarska-Jagiela A, Schilbach L, Lehnhardt FG, Krug B, Vogeley K, Tepest R (2011): Imaging derived cortical thickness reduction in high-functioning autism: Key regions and temporal slope. Neuroimage 58:391–400.
19. Haar S, Berman S, Behrmann M, Dinstein I (2016): Anatomical abnormalities in autism? Cereb Cortex 26:1440–1452.
20. Khundrakpam BS, Lewis JD, Kostopoulos P, Carbonell F, Evans AC (2017): Cortical thickness abnormalities in autism spectrum disorders through late childhood, adolescence, and adulthood: A large-scale MRI study. Cereb Cortex 27:1721–1731.
21. Courchesne E, Karns CM, Davis HR, Ziccardi R, Carper RA, Tigue ZD, et al. (2011): Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. Neurology 57:245–254.
22. Ecker C, Shahidiani A, Feng Y, Daly E, Murphy C, D'Almeida V, et al. (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 121:1157–1170.
23. Ramaswami G, Geschwind DH (2018): Genetics of autism spectrum disorder. Handb Clin Neurol 147:321–329.
24. 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. Psychol Med 48:654–668.
25. Wolfers T, Buitelaar JK, Beckmann CF, Franke B, Marquand AF (2015): From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. Neurosci Biobehav Rev 57:328–349.
26. Sabuncu MR, Konukoglu E (2014): Clinical prediction from structural brain MRI scans: A large-scale empirical study. Neuroinformatics 13:31–46.
27. Damiano CR, Mazefsky CA, White SW, Dichter GS (2014): Future directions for research in autism spectrum disorders. J Clin Child Adolesc Psychol 43:828–843.
28. Insel TR, Cuthbert BN (2015): Brain disorders? Precisely. Science 348:499–500.
29. Lombardo MV, Pierce K, Eyler LT, Carter Barnes C, Ahrens-Barbeau C, Solso S, et al. (2015): Different functional neural substrates for good and poor language outcome in autism. Neuron 86:567–577.
30. Fountain C, Winter AS, Bearman PS (2012): Six developmental trajectories characterize children with autism. Pediatrics 129:e1112–e1120.
31. Fair DA, Bathula D, Nikolas MA, Nigg JT (2012): Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. Proc Natl Acad Sci U S A 109:6769–6774.
32. Costa Dias TG, Iyer SP, Carpenter SD, Cary RP, Wilson VB, Mitchel SH, et al. (2015): Characterizing heterogeneity in children with and without ADHD based on reward system connectivity. Dev Cogn Neurosci 11:155–174.
33. van Loo HM, de Jonge P, Romeijn J-W, Kessler RC, Schoevers RA (2012): Data-driven subtypes of major depressive disorder: A systematic review. BMC Med 10:156.
34. Bell MD, Corbera S, Johannesen JK, Fiszdon JM, Wexler BE (2013): Social cognitive impairments and negative symptoms in schizophrenia: Are there subtypes with distinct functional correlates? Schizophr Bull 39:186–196.
35. Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF (2016): Beyond lumping and splitting: A review of computational approaches for stratifying psychiatric disorders. Biol Psychiatry Cogn Neurosci Neuroimaging 1:433–447.
36. Marquand AF, Rezek I, Buitelaar J, Beckmann CF (2016): Understanding heterogeneity in clinical cohorts using normative models: Beyond case-control studies. Biol Psychiatry 80:552–561.

37. Wolfers T, Doan N, Kaufmann T, Alnæs D, Moberget T, Buitelaar JK, et al. (2018): Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. JAMA Psychiatry 75:1146–1155.

38. Loth E, Charman T, Mason L, Tillmann J, Jones EJH, Wooldridge C, et al. (2017): The EU-AIMS Longitudinal European Autism Project (LEAP): Design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. Mol Autism 8:24.

39. Anagnostou E, Taylor MJ (2011): Review of neuroimaging in autism spectrum disorders: What have we learned and where we go from here. Mol Autism 2:4.

40. Charman T, Loth E, Tillmann J, Crawley D, Wooldridge C, Goyard D, et al. (2017): The EU-AIMS Longitudinal European Autism Project (LEAP): Clinical characterisation. Mol Autism 8:27.

41. Rutter M (2003): Autism Diagnostic Interview, Revised. Torrance, CA: WPS.

42. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, Dilavore PC, et al. (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 30:205–223.

43. Rasmussen CE, Williams CKI (2006): Model selection and adaptation of hyperparameters. In: Gaussian Processes for Machine Learning. Cambridge, MA: MIT Press, 105–128.

44. Fisher RA, Tippett LHC (1928): Limiting forms of the frequency distribution of the largest or smallest member of a sample. Math Proc Cambridge Philos Soc 24:180–190.

45. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980.

46. Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9:179–184.

47. Walhovd KB, Fjell AM, Giedd J, Dale AM, Brown TT (2016): Through thick and thin: A need to reconcile contradictory results on trajectories in human cortical development. Cereb Cortex 27:1472–1481.

48. Zielinski BA, Prigge MBD, Nielsen JA, Froehlich AL, Abildskov TJ, Anderson JS, et al. (2014): Longitudinal changes in cortical thickness in autism and typical development. Brain 137:1799–1812.

49. Ducharme S, Albaugh MD, Nguyen T, Hudziak JJ, Mateos-Pérez JM, Labbe A, et al. (2016): NeuroImage Trajectories of cortical thickness maturation in normal brain development — The importance of quality control procedures. Neuroimage 125:267–279.

50. Tamnes CK, Herting MM, Goddings A-L, Meuwese R, Blakemore S-J, Dahl RE, et al. (2017): Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. J Neurosci 37:3402–3412.

51. Shaw P, Kabani NJ, Lerch JP, Eckstrand K, Lenroot R, Gogtay N, et al. (2008): Neurodevelopmental trajectories of the human cerebral cortex. J Neurosci 28:3586–3594.

52. Fjell AM, Grydeland H, Krogsrud SK, Amlien I, Rohani DA, Ferschmann L (2015): Development and aging of cortical thickness correspond to genetic organization patterns. Proc Natl Acad Sci U S A 112:15462–15467.

53. Fjell AM, Walhovd KB (2010): Structural brain changes in aging: Courses, causes and cognitive consequences. Rev Neurosci 21:187–221.

54. Mensen VT, Wierenga LM, van Dijk S, Rijks Y, Oranje B, Mandl RCW, Durston S (2016): Development of cortical thickness and surface area in autism spectrum disorder. Neuroimage Clin 13:215–222.

55. Thambisetty M, Wan J, Carass A, An Y, Prince JL, Resnick SM (2010): Longitudinal changes in cortical thickness associated with normal aging. Neuroimage 52:1215–1223.

56. Abrahams BS, Geschwind DH (2008): Advances in autism genetics: On the threshold of a new neurobiology. Nat Rev Genet 9:341–355.

57. Ecker C, Murphy D (2014): Neuroimaging in autism—from basic science to translational research. Nat Rev Neurol 10:82–91.

58. Geschwind DH, Levitt P (2007): Autism spectrum disorders: Developmental disconnection syndromes. Curr Opin Neurobiol 17:103–111.

59. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, et al. (2008): Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet 82:477–488.

60. Croen LA, Grether JK, Selvin S (2002): Descriptive epidemiology of autism in a California population: Who is at risk? J Autism Dev Disord 32:217–224.

61. Seltzer MM, Shattuck P, Abbeduto L, Greenberg JS (2004): Trajectory of development in adolescents and adults with autism. Ment Retard Dev Disabil Res Rev 10:234–247.

62. Ronald A, Happé F, Bolton P, Butcher LM, Price TS, Wheelwright S, et al. (2006): Genetic heterogeneity between the three components of the autism spectrum: A twin study. J Am Acad Child Adolesc Psychiatry 45:691–699.

63. Hardan AY, Muddasani S, Vemulapalli M, Keshavan MS, Minshew NJ (2006): An MRI study of increased cortical thickness in autism. Am J Psychiatry 163:1290–1292.

64. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, et al. (2014): The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry 19:659–667.

65. Bethlehem RAI, Seidlitz J, Romero-Garcia R, Lombardo MV (2018): Using normative age modelling to isolate subsets of individuals with autism expressing highly age-atypical cortical thickness features [published online ahead of print Jan 23]. bioRxiv.

66. Amaral DG, Schumann CM, Nordahl CW (2008): Neuroanatomy of autism. Trends Neurosci 31:137–145.

67. Jiao Y, Chen R, Ke X, Chu K, Lu Z, Herskovits E (2011): Predictive models of autism spectrum disorder based on brain regional cortical thickness. Neuroimage 50:589–599.

68. Ecker C, Andrews D, Dell'Acqua F, Daly E, Murphy C, Catani M, et al. (2016): Relationship between cortical gyrification, white matter connectivity, and autism spectrum disorder. Cereb Cortex 26:3297–3309.

69. Ecker C, Ronan L, Feng Y, Daly E, Murphy C, Ginestet CE, et al. (2013): Intrinsic gray-matter connectivity of the brain in adults with autism spectrum disorder. Proc Natl Acad Sci U S A 110:13222–13227.

70. Moradi E, Khundrakpam B, Lewis JD, Evans AC, Tohka J (2017): Predicting symptom severity in autism spectrum disorder based on cortical thickness measures in agglomerative data. Neuroimage 144:128–141.

71. Doyle-Thomas KAR, Duerden EG, Taylor MJ, Lerch JP, Soorya LV, Wang AT, et al. (2013): Effects of age and symptomatology on cortical thickness in autism spectrum disorders. Res Autism Spectr Disord 7:141–150.

72. Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, et al. (1999): Brain development during childhood and adolescence: A longitudinal MRI study. Nat Neurosci 2:861–863.

73. Hardan AY, Libove RA, Keshavan MS, Melhem NM, Minshew NJ (2009): A preliminary longitudinal magnetic resonance imaging study of brain volume and cortical thickness in autism. Biol Psychiatry 66:320–326.

74. McAlonan GM, Cheung V, Cheung C, Suckling J, Lam GY, Tai KS, et al. (2005): Mapping the brain in autism. A voxel-based MRI study of volumetric differences and intercorrelations in autism. Brain 128:268–276.

75. McAlonan GM, Suckling J, Wong N, Cheung V, Lienenkaemper N, Cheung C, Chua SE (2008): Distinct patterns of grey matter abnormality in high-functioning autism and Asperger's syndrome. J Child Psychol Psychiatry 49:1287–1295.