

Revisiting subcortical brain volume correlates of autism in the ABIDE dataset: effects of age and sex

W. Zhang¹, W. Groen², M. Mennes¹, C. Greven^{2,3}, J. Buitelaar^{2,3,4} and N. Rommelse^{2,4*}

¹Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands

²Karakter, Child and Adolescent Psychiatry University Center, Nijmegen, The Netherlands

³Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands

⁴Department of Psychiatry, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands

Background. Autism spectrum disorders (ASD) are characterized by substantial clinical, etiological and neurobiological heterogeneity. Despite this heterogeneity, previous imaging studies have highlighted the role of specific cortical and subcortical structures in ASD and have forwarded the notion of an ASD specific neuroanatomy in which abnormalities in brain structures are present that can be used for diagnostic classification approaches.

Method. A large ($N = 859$, 6–27 years, IQ 70–130) multi-center structural magnetic resonance imaging dataset was examined to specifically test ASD diagnostic effects regarding (sub)cortical volumes.

Results. Despite the large sample size, we found virtually no main effects of ASD diagnosis. Yet, several significant two- and three-way interaction effects of diagnosis by age by gender were found.

Conclusion. The neuroanatomy of ASD does not exist, but is highly age and gender dependent. Implications for approaches of stratification of ASD into more homogeneous subtypes are discussed.

Received 5 May 2017; Revised 21 June 2017; Accepted 21 June 2017; First published online 26 July 2017

Key words: ABIDE, ASD, brain structure, MRI, neuroimaging, volume.

The developmental nature of autism spectrum disorder (ASD) likely implies that different neural mechanisms may play a role during the unfolding of the disorder over age. The clearest example hereof is the abnormal overgrowth of the whole brain in ASD at early ages, followed by abnormal decelerated growth during puberty and possible degeneration in adulthood (Courchesne *et al.* 2011). A recent meta-analysis reported head circumference to be significantly larger in individuals with ASD compared with controls, with 822/5225 (15.7%) individuals with ASD displaying macrocephaly (Sacco *et al.* 2015). Further, structural brain imaging studies found relatively larger differences in brain volumes between individuals with ASD and controls in low functioning than in high functioning individuals with ASD (Sacco *et al.* 2015). Finally, the meta-analysis reported a significant interaction between age and total brain volume, resulting in larger head circumference and brain size during early childhood (Sacco *et al.* 2015). Age-specific anatomy may also be present for

subcortical structures in ASD, since a systematic review and meta-analysis reported a modifying effect of age on amygdala size and of age and IQ on the volume of cerebellar vermal lobules VI–VII in ASD (Stanfield *et al.* 2008; Sacco *et al.* 2015).

Similarly, sex is also a likely moderating factor in relating ASD to neuroanatomical correlates, with ASD being a male predominant disorder (Van Wijngaarden-Cremers *et al.* 2014). There are well known dissimilar patterns of brain development in healthy boys and girls, with large sex differences during development on the caudate nucleus, amygdala, and hippocampus and girls have faster maturing brains than boys (Giedd *et al.* 2012). A meta-analysis found regional sex differences in volume and tissue density of various brain structures, including the amygdala, hippocampus and insula, areas known to be implicated in sex-biased neuropsychiatric conditions (Ruigrok *et al.* 2014). However, most previous studies on the brain anatomy of ASD included no or only few female participants, making it unknown if male-reported anatomical differences in ASD also pertain to females with ASD.

Yet, despite these age and sex effects on brain anatomy potentially interacting with ASD diagnosis, attempts are still made to design one overarching

* Address for correspondence: N. Rommelse, Department of Psychiatry, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands. (Email: n.lambregts-rommelse@psy.umcn.nl)

diagnostic classification algorithm to aid in the diagnosis of ASD. The study by Ecker *et al.* (2012) is often cited as supporting the view that such a diagnostic classification algorithm is feasible, with a high level of accuracy when using structural brain imaging measures in advanced machine learning algorithms (Support Vector Machines). Since then, at least 20 papers have been published using magnetic resonance imaging (MRI)-based single subject prediction of ASD (see for extensive review Wolfers *et al.* 2015; Arbabshirani *et al.* 2017). With varying levels of classification accuracy, what stands out is that no overarching diagnostic classification algorithm can be derived that can differentiate all ASD patients from controls with acceptable sensitivity and specificity across studies. Importantly, in studies matching case-control subjects for age and sex, the classification algorithm varies substantially between participant pairs. These data therefore suggest that it is too simplistic to search for one common neurobiological substrate underlying ASD across development and sex. Instead, MRI markers may be used to stratify ASD into biologically more homogeneous subtypes (Marquand *et al.* 2016).

The current study was designed to advance existing work by examining the main and combined (interactive) effects of age and sex on (sub)cortical gray matter (GM) volumes in subjects with and without ASD. The Autism Brain Imaging Data Exchange (ABIDE) dataset was used for this purpose (Di Martino *et al.* 2014). The dataset includes 539 individuals with ASD and 573 age- and sex-matched controls, with standardized phenotyping and scanning approaches across sites (Di Martino *et al.* 2014). Previous reports using this dataset unfortunately did not explicitly analyze and/or report the combined effects of age and sex on case-control differences (Lai *et al.* 2013; Nielsen *et al.* 2013; Blackmon *et al.* 2015; Cerliani *et al.* 2015; Chen *et al.* 2015a, b; Di & Biswal, 2015; Jiang *et al.* 2015; Kucharsky Hiess *et al.* 2015; Schaer *et al.* 2015; Supekar & Menon, 2015; Valk *et al.* 2015; Venkataraman *et al.* 2015; Haar *et al.* 2016; Riddle *et al.* 2016). In the current study, we performed an analysis of whole brain GM and white matter (WM) volume and subcortical structure volumes in a large subsample of this dataset ($N = 859$; 6–27 years).

Material and methods

ABIDE dataset

For all analyses, we used the publicly available MRI datasets from the ABIDE program (Di Martino *et al.* 2014). ABIDE shares 1112 structural T1 MRI high-resolution images and phenotypic information on ASD, IQ, age, and sex. In total, the dataset comprises 539

individuals with ASD and 573 age-matched controls (7–64 years) from 17 different sites. Sites verified clinical DSM-IV diagnoses of ASD by combining clinical judgment and ADOS-G/ADI-R, clinical judgment only or ADOS-G/ADI-R only (see online Supplementary Fig. S6 and Di Martino *et al.* 2014). All contributions were based on studies approved by the local Institutional Review Boards, and data were fully anonymized. All data distributed were visually inspected before release (Di Martino *et al.* 2014).

Participants

We assessed the distribution, mean, standard deviation and range of age, sex, full-scale IQ, and number of participants per scanner site to survey group characteristics (see online Supplementary Figs S2–S5). The average age of the complete sample ($N = 1112$) was 17.05 years (s.d. = 8.04, range = 6.5–64.0). The age distribution was skewed, with 975/1112 (87.7%) being younger than 27.0 years. As the group with age 27.0 or older had a very broad distribution in age range and the regression analyses examining age effects were not robust in this group, the data of this subgroup were excluded from further analyses (see Fig. 1 flow chart). The average FIQ of the complete sample [available in 1084/1112 (97.5%)] was 108.3 (s.d. = 15.0, range = 41–148), showing a normal distribution. Participants with a FIQ <70, >130 or missing were excluded in a second step to approximate a homogeneous sample ($n = 76$, 7.8%). In a third step, 40/899 (4.4%) subjects were excluded from analyses because the manual visual inspection revealed geometric inaccuracies of segmentation. After applying these selection steps, 859 participants remained (77.2% of original sample), with 401 (46.7%) ASD and 458 (53.3%) control participants that did not differ in handedness (% right handed ASD = 87.3, % right handed controls = 90.0; $\chi^2 = 1.33$, $p = 0.51$), but did differ slightly in age [*mean* (M) age ASD = 15.1, M age controls = 15.6; $F = 4.54$, $p = 0.03$], sex (% males ASD = 87.6, % males controls = 82.6; $\chi^2 = 4.44$, $p = 0.04$), and FIQ (M FIQ ASD = 103.3, M FIQ controls = 110.2; $F = 68.2$, $p < 0.001$).

Structural image preprocessing and volumetric analysis

All images were skull stripped using the 3DSkullStrip algorithm in AFNI (Cox, 1996). FIRST (v5.0.0) was used to derive left and right hippocampus, amygdala, thalamus, globus pallidus, nucleus accumbens, caudate nucleus, and putamen volumes (Patenaude *et al.* 2011). FIRST is part of the FMRIB's Software Library (FSL) and performs both registration and segmentation. Registration entails an affine transformation with 12 degrees of freedom of the T1 structural MRI

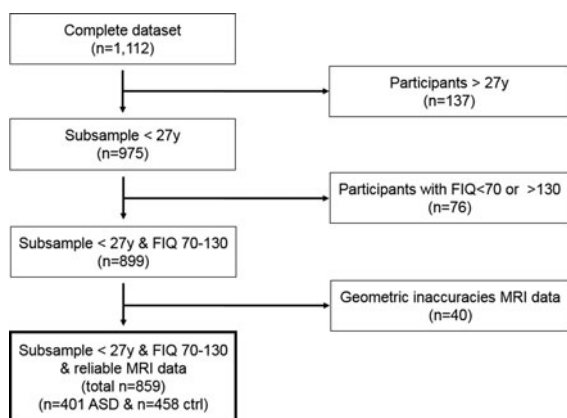


Fig. 1. Flow chart illustrating the number of participants remaining after applying selection criteria.

scan onto the 152MNI standard space. After subcortical registration, a sub-cortical mask is applied to segment the subcortical structures based on shape models and voxel intensities. The first author (WZ) visually inspected all registrations and segmentations, while blinded to diagnostic group. Registration or segmentation was deemed inaccurate (for example, insufficient contrast or poor image quality) in 40/899 subjects, thus yielding 859 segmented scans. FMRIB's Automated Segmentation Tool (FAST) was used to obtain WM and GM volumes (Zhang *et al.* 2001). FAST segments MR images into different tissue types, while also correcting for spatial intensity variations using a hidden Markov random field model and an associated Expectation–Maximization algorithm. The results of FIRST and FAST are robust and reliable and comparable with manual segmentation (Morey *et al.* 2009; Zhang *et al.* 2001). FAST is superior to other automated segmentation algorithms with respect to its volume detection accuracy (Eggert *et al.* 2012).

Data analyses

All analyses were carried out in SPSS version 24. Variables were normally distributed. Linear mixed models were used for the analyses with site as a random factor, allowing for the investigation of effects while correcting for non-independency of data collected per scanning site. The following main effects were examined: diagnosis (ASD, control), age (linear, quadratic), sex (male, female) and left-and-right hemisphere (as within subjects' measure; i.e. the left and right volume) on the brain volumetric measures (i.e. total GM, total WM, hippocampus, amygdala, thalamus, pallidus, nucleus accumbens, caudate nucleus, putamen). Age (linear and quadratic) was demeaned (i.e. centered) to reduce potential multicollinearity. All two, three and four-way interactions (diagnosis by sex by age/age² by right–left hemisphere) were also implemented in the

initial model. The model was simplified by eliminating the non-significant interaction effects, starting with the highest order effects. Main effects were never dropped from the model. All volumetric analyses were carried out with and without covarying for total brain volume (the sum of total GM volume and total WM volume) and FIQ to examine effects of these potential confounders. In order not to miss out on potentially important effects, a lenient alpha of 0.05 was considered significant. Cohen's *d* was used to indicate the effect size. We considered $d=0.2$ – 0.5 as a small, $d=0.5$ – 0.8 as a medium, and $d>0.8$ as a large effect size (Cohen, 1969). For *post hoc* analyses and visualization of results, age was stratified into three groups of equal age-range (child, adolescent, adult) yielding 407 participants in the 6–12 years group, 415 in the 13–19 years group and 153 in the 20–27 years group.

Results

Scanner site

Datasets were completely matched for sex, age and diagnosis within each scanner site; there were thus no confounding effects between scanner site and the main variables (sex, diagnosis). Small to moderate sized differences between scanner sites were present regarding age [$F(19, 839)=31.57, p<0.001, d=0.39$], FIQ [$F(19, 832)=3.38, p<0.001, d=0.13$], sex ($\chi^2=63.21, p<0.001, d=0.57$) and handedness ($\chi^2=88.43, p<0.001, d=0.69$) of the participants. Of note, diagnosis did not differ between scanner sites such that patients were balanced with age, sex, and IQ-matched controls across sites (see online Supplementary Figs S2–S4).

Main effects

Table 1 shows the data (mean, s.d.) of total GM and WM volume as well as of the right and left volumes of the hippocampus, caudate nucleus, putamen, amygdala, thalamus, globus pallidus, and nucleus accumbens, broken down by group (ASD *v.* controls). Table 2 summarizes the main results of the mixed linear models. There were no main effects of diagnosis on GM and WM, nor were there main effects of diagnosis on the subcortical structures except for the nucleus accumbens, with smaller nucleus accumbens volume in ASD (Table 2). The ASD main effect on the nucleus accumbens remained significant with total brain volume as covariate [$F(1, 843.6)=5.75, p=0.017, d=0.17$], but this ASD effect was lost when FIQ was included as a covariate [$F(1, 844.5)=0.78, p=0.38, d=0.06$].

We further detected a significant main effect of age on total GM (decreasing) and WM (increasing) and on all subcortical structures except the caudate nucleus

Table 1. Mean brain volumes (mm³) and standard deviations for ASD cases and controls

Brain structure	Volume											
	ASD											
	Total				Male				Female			
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Grey matter (GM)	613 525.4	76 689.6			620 137.7	72 356.1			567 107.0	90 016.3		
White matter (WM)	508 075.7	72 049.7			514 545.5	69 260.0			462 657.9	75 552.1		
	Left		Right		Left		Right		Left		Right	
Hippocampus	3691.7	850.5	3781.8	795.3	3734.0	852.4	3816.6	802.6	3394.4	781.9	3537.1	701.9
Amygdala	1014.0	309.7	984.2	335.2	1030.5	310.5	997.6	344.0	897.9	280.5	890.6	248.0
Thalamus	7894.3	1256.5	7746.0	1133.6	7951.3	1265.2	7797.3	1139.3	7494.7	1126.0	7385.8	1033.8
Globus pallidus	1689.7	359.2	1677.8	406.5	1709.6	363.6	1689.3	419.1	1550.5	294.5	1596.6	294.3
Nucleus accumbens	528.2	150.5	434.4	124.2	535.9	150.7	442.0	124.6	473.9	138.4	380.7	108.0
Caudate nucleus	3820.9	597.5	4042.8	663.0	3849.5	599.6	4078.4	659.2	3620.2	546.5	3793.3	641.7
Putamen	5301.4	950.1	5394.5	914.8	5366.8	925.5	5453.5	899.6	4842.7	1002.1	4980.6	922.7
	Controls											
	Total				Male				Female			
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
GM	604 451.0	70 650.9			615 020.2	68 489.6			555 987.5	59 574.8		
WM	502 774.3	59 640.6			513 184.5	57 121.5			455 040.1	46 399.5		
	Left		Right		Left		Right		Left		Right	
Hippocampus	3733.2	725.7	3783.7	697.2	3808.5	701.7	3859.7	669.6	3388.0	738.1	3435.4	718.8
Amygdala	1021.1	293.8	998.2	328.2	1040.3	297.5	1022.8	329.6	932.8	260.1	885.4	298.6
Thalamus	7879.2	1092.5	7751.8	990.6	7957.5	1084.2	7854.4	947.8	7520.1	1064.3	7281.3	1050.9
Globus pallidus	1668.4	333.9	1657.8	371.0	1687.1	947.8	1683.9	368.6	1582.7	241.0	1537.8	360.5
Nucleus accumbens	539.0	140.0	444.7	117.3	549.5	139.9	452.3	116.7	490.8	130.7	409.9	114.7
Caudate nucleus	3834.5	561.8	4043.1	585.9	3875.0	567.9	4103.6	584.3	3648.4	495.1	3766.0	511.4
Putamen	5360.2	728.5	5413.1	715.6	5464.6	680.8	5511.2	660.6	4881.7	752.3	4963.1	786.8

Table 2. Summary of linear mixed model results

Brain structure	Main effect				Interaction effect					
	Dx (<i>p/d</i>)	Age (<i>p/d</i>) Age ² (<i>p/d</i>)	Gender (<i>p/d</i>)	Hemisphere (<i>p/d</i>)	Dx × age/ age ²	Dx × gender	Dx × age/age ² × Gender (<i>p/d</i>)	Dx × age/age ² × hemisphere (<i>p/d</i>)	Dx × gender × hemisphere	Dx × age/age ² × gender × hemisphere
Grey matter (GM)	N.S.	<0.001/0.47 Decrease in GM with increasing age N.S.	<0.001/0.60 More GM volume in males than in females	–	N.S. N.S.	N.S.	0.041/0.12 A small significant interaction between diagnosis and age was present in males (<i>p</i> = 0.032, <i>d</i> = 0.16), but not in females (<i>p</i> = 0.90, <i>d</i> = 0.03). A significant main effect of diagnosis was found in the oldest age group (20–27 years: <i>p</i> = 0.023, <i>d</i> = 0.43) with ASD males having on average more GM volume (<i>M</i> = 627 628.3 mm ³) than male controls (<i>M</i> = 603 607.6 mm ³).	–	–	–
White matter (WM)	N.S.	<0.001/0.25 Increase of GM with increasing age 0.001/0.22 Increase of GM with increasing age predominantly in younger subjects	<0.001/0.22 More WM volume in males than in females	–	N.S. N.S.	N.S.	0.01/0.15 A small significant interaction between gender and age was present in ASD participants (<i>p</i> = 0.013, <i>d</i> = 0.26), but not in controls (<i>p</i> = 0.20, <i>d</i> = 0.13). The effect of gender was largest in the oldest age group compared to the other two age groups (6–12 years: <i>p</i> = 0.001, <i>d</i> = 0.54; 13–19 years: <i>p</i> < 0.001, <i>d</i> = 0.54; 20–27 years: <i>p</i> < 0.001, <i>d</i> = 0.98) with males having on average more WM volume, particularly in the oldest age range (20–27 years: ASD males <i>M</i> = 570 519.8; ASD females <i>M</i> = 469 695.3 mm ³) than females	–	–	–

N.S.

Table 2 (cont.)

Brain structure	Main effect				Interaction effect					
	Dx (<i>p/d</i>)	Age (<i>p/d</i>) Age ² (<i>p/d</i>)	Gender (<i>p/d</i>)	Hemisphere (<i>p/d</i>)	Dx × age/ age ²	Dx × gender	Dx × age/age ² × Gender (<i>p/d</i>)	Dx × age/age ² × hemisphere (<i>p/d</i>)	Dx × gender × hemisphere	Dx × age/age ² × gender × hemisphere
Hippocampus	N.S.	<0.001/0.45 Increase of volume with increasing age 0.03/0.15 Increase of GM with increasing age predominantly in younger subjects	<0.001/0.28^a More volume in males than in females	0.001/0.28 Right more volume than left	N.S. N.S.	N.S.	0.006/0.19^a A small significant interaction between diagnosis and gender was present in ASD participants (<i>p</i> = 0.047, <i>d</i> = 0.20), but not in controls (<i>p</i> = 0.22, <i>d</i> = 0.12). No main effect of gender was present in the youngest group (6–12 years: <i>p</i> = 0.64, <i>d</i> = 0.08), a small gender effect was present in 13–19 year olds (<i>p</i> = 0.041, <i>d</i> = 0.31) and a large effect was present in the oldest age group (20–27 years: <i>p</i> = 0.001, <i>d</i> = 0.87) with ASD males having on average larger hippocampal volumes (<i>M</i> = 4309.2) than females (<i>M</i> = 3573.1).	N.S. N.S.	N.S. N.S.	N.S. N.S.
Caudate Nucleus	N.S.	N.S. 0.041/0.14 Increase of volume with increasing age predominantly in younger subjects	0.015/0.17 More volume in males than in females	<0.001/0.55 Right more volume than left	N.S. N.S.	N.S.	N.S.	N.S.	N.S.	N.S. N.S.
										0.022/0.16 A small main effect of age ² was present in controls (<i>p</i> = 0.019, <i>d</i> = 0.22), but not ASD patients (<i>p</i> = 0.84, <i>d</i> = 0.02).

Putamen	N.S.	<0.001/0.24 Increase of volume with increasing age 0.028/0.15 Increase of volume with increasing age predominantly in younger subjects	<0.001/0.47 More volume in males than in females	<0.001/0.23 Right more volume than left	N.S. N.S.	N.S.	0.008/0.20 A small significant interaction between diagnosis and age was present in females ($p = 0.033$, $d = 0.25$), but not in males ($p = 0.15$, $d = 0.09$). N.S.	0.026/0.15 A small significant interaction between diagnosis and age was present in ASD ($p = 0.031$, $d = 0.20$), but not in controls ($p = 0.41$, $d = 0.06$). N.S.	N.S.	0.038/0.14 In females, a moderately sized diagnosis group by age interaction for the right putamen ($p = 0.005$, $d = 0.51$) but not for the left putamen ($p = 0.22$, $d = 0.22$). No main effect of diagnosis was present in 6–16 year-old females ($p = 0.20$, $d = 0.27$), yet moderately large effect was present in the 17–27 year olds ($p = 0.05$, $d = 0.74$), with ASD females having on average smaller right putamen volumes ($M = 4515.1$) than control females ($M = 5113.4$) N.S. N.S.
Amygdala	N.S.	<0.001/0.64 Increase of volume with increasing age 0.004/0.20 Increase of volume with increasing age predominantly in younger subjects	<0.001/0.24 More volume in males than in females	0.004/0.20 Left more volume than right	N.S. N.S.	N.S.	N.S.	N.S.	N.S.	N.S. N.S.
Thalamus	N.S.	<0.001/0.49 Increase of volume with increasing age 0.012/0.17 Increase of volume with increasing age predominantly in younger subjects	<0.001/0.32 More volume in males than in females	<0.001/0.40 Left more volume than right	N.S. N.S.	N.S.	N.S.	N.S.	N.S.	N.S. N.S.
Globus pallidus	N.S.	0.003/0.21 Increase of volume with increasing age N.S.	<0.001/0.26 More volume in males than in females	N.S.	N.S. N.S.	N.S.	N.S.	N.S.	N.S.	N.S. N.S.

Table 2 (cont.)

Brain structure	Main effect		Interaction effect						
	Dx (<i>p/d</i>)	Age (<i>p/d</i>) Age ² (<i>p/d</i>)	Gender (<i>p/d</i>)	Hemisphere (<i>p/d</i>)	Dx × age/ age ²	Dx × gender	Dx × age/age ² × hemisphere (<i>p/d</i>)	Dx × gender × hemisphere	Dx × age/age ² × gender × hemisphere
Nucleus Accumbens	0.04/0.14 Less volume in ASD. ^b	<0.001/0.51 Increase of volume with increasing age	0.001/0.25 More volume in males than in females	<0.001/1.70 Left more volume than right	N.S. N.S.	N.S.	N.S. N.S.	N.S. N.S.	N.S. N.S.

Dx, diagnostic group (ASD or control group); *d*, Cohen's *d*; N.S., *p*-values above 0.05 are not significant.

The exact *p*-values (and effect size *d*) are presented in the table (significance in bold).

^a Findings became non-significant when covaried for the total brain volume.

^b The effect disappeared when covaried for FIQ.

with small to moderate effect sizes. We also found a significant main effect of age² (quadratic effect of age) on all subcortical structures except globus pallidus, as well as on total WM with small to moderate effect sizes. There was also a significant main effect of sex with males showing greater volumes than females in all structures. The main effect of hemisphere (i.e. a volume difference between the left and right structures) reached significance for all volumes (hippocampus *L*<*R*, caudate nucleus *L*<*R*, putamen *L*<*R*, amygdala *R*<*L*, *R*<*L*, nucleus accumbens *R*<*L*), except for the globus pallidus (see Table 2).

Interaction effects

For total GM, we found a three-way interaction of diagnosis by age by sex (Table 2): the decrease in volume over age was greater in male controls than in males with ASD (Fig. 2). When we stratified the male cohort into three age groups (child: 6–12 years, adolescent: 13–19 years, adult: 20–27 years), we found only a significant main effect of diagnosis in the oldest age group [20–27 years: *F* (1, 111.5)=5.29, *p*=0.023, *d*=0.43], with males with ASD showing larger GM volume (*M*=627 628.3 mm³) than male controls (*M*=603 607.6 mm³) (see Fig. 2).

For total WM, there was also a significant three-way interaction of diagnosis by age by sex (Table 2) with the difference in volume between males and females increasing with age in ASD but not in controls (Fig. 2). When the ASD cohort was stratified into three age groups (6–12 years, 13–19 years, 20–27 years), the effect of sex was largest in the oldest age group compared with the other two age groups [6–12 years: *F* (1, 161.0)=11.76, *p*=0.001, *d*=0.54; 13–19 years: *F* (1, 170.1)=12.63, *p*<0.001, *d*=0.54; 20–27 years: *F* (1, 55.7)=13.70, *p*<0.001, *d*=0.98] with ASD males having on average more WM volume than ASD females, particularly in the oldest age range (20–27 years: ASD males *M*=570 519.8; ASD females *M*=469 695.3 mm³) (Fig. 3).

For hippocampal volumes, a small three-way interaction of diagnosis by age by sex was present (Table 2). *Post hoc* analyses revealed that male adolescents and adults with ASD had larger hippocampi than females with ASD [*F* (1, 62.0)=11.60, *p*=0.001, *d*=0.87]. See Fig. 4. Findings remained identical when covaried for FIQ, but became non-significant when covaried for total brain volume.

For caudate volumes, a significant three-way interaction between diagnosis, age² and left-right hemisphere emerged (Table 2). *Post hoc* analyses revealed that the right caudate size increases with age in ASD [*F* (1, 455.1)=5.53, *p*=0.019, *d*=0.22], whereas in controls right caudate size did not increase with age

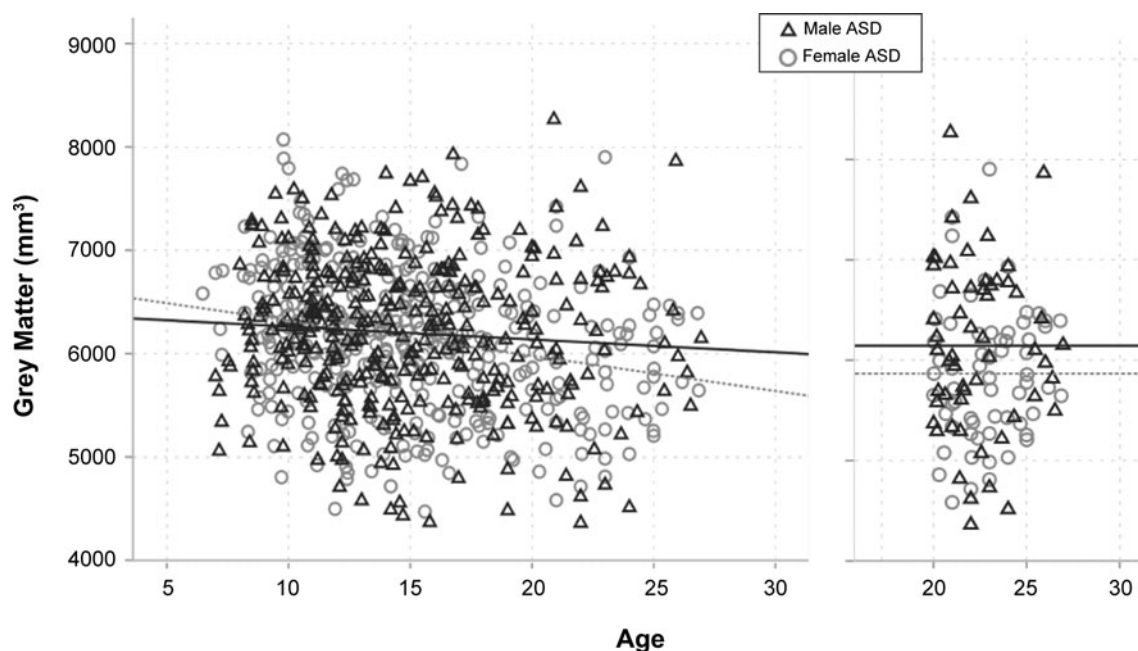


Fig. 2. Interaction between diagnosis (ASD *v.* control), age, and gender in the prediction of grey matter volume. The left panel shows all cases (male controls and male ASD) from age 6 to 27. The two regression lines illustrate that the decrease in GM volume is larger in male controls than in males with autism. The right panel shows all adult male cases (20–27 years; controls and ASD) and thus shows a subset of the cases in the left panel to illustrate that adult males with autism have more GM volume than adult controls.

[$F(1, 400.8) = 0.04, p = 0.84, d = 0.02$]. Thus, a quadratic effect of age was found on caudate volume only for participants with ASD but not for controls (Fig. 5). This effect remained significant when covaried for total brain volume and FIQ.

For putamen volumes, a small significant four-way interaction was present between diagnosis, age, sex, and left–right hemisphere (Table 2). *Post hoc* analyses revealed that only young adult females (17–27 years) with ASD had smaller right putamen volumes ($M = 4515.1$) than young adult female controls ($M = 5113.4$) [$F(1, 30.3) = 4.12, p = 0.05, d = 0.74$] (Fig. 6). Findings remained unchanged when covaried for total brain volume or FIQ.

Discussion

To facilitate future attempts to use MRI markers for stratification of ASD, this study set out to examine the main and combined effects of age and sex on cortical and subcortical GM volumes in $N = 859$ subjects (6–27 years, IQ 70–130) with and without ASD from the ABIDE dataset (Di Martino et al. 2014). Despite a relatively lenient alpha level, we found virtually no main effects of ASD on any of the volumes. This is in striking contrast with the positive findings from previous studies (McAlonan et al. 2005; Waiter et al. 2005; Hadjikhani et al. 2006; Hyde et al. 2010; Toal et al.

2010; Via et al. 2011; Ecker et al. 2012; Greimel et al. 2012) and even with several meta-analyses (Duerden et al. 2012; Nickl-Jockschat et al. 2012; Yang et al. 2016a, b). Only one small main effect of ASD diagnosis was found on the nucleus accumbens volume (being smaller in ASD compared with controls) that appeared related to full scale IQ (FSIQ) differences between the groups, since it became non-significant after correcting for FSIQ. In contrast, several diagnosis-by-age-by-sex effects were found. Male adults with ASD had more GM and WM than controls, whereas no differences for GM and WM were found for male children and adolescents with ASD, and for females with ASD. Similarly, adolescent and adult men with ASD had larger hippocampal volumes than controls, but there were no differences in male children with ASD and females with ASD. Caudate nucleus volume increased with age in controls but not in ASD, and female adolescents and adults with ASD had smaller right putamen volume than female controls, while there were no differences in men with ASD. The current findings strongly underline the sex-specific and neurodevelopmental nature of ASD (Chen et al. 2015a, b; Lai et al. 2015a, b; Wolfers et al. 2015), and suggest that MRI-based stratification of ASD will need to take account of age and sex and their interactive effects as well as should use substantially larger datasets than have been currently used for such attempts.

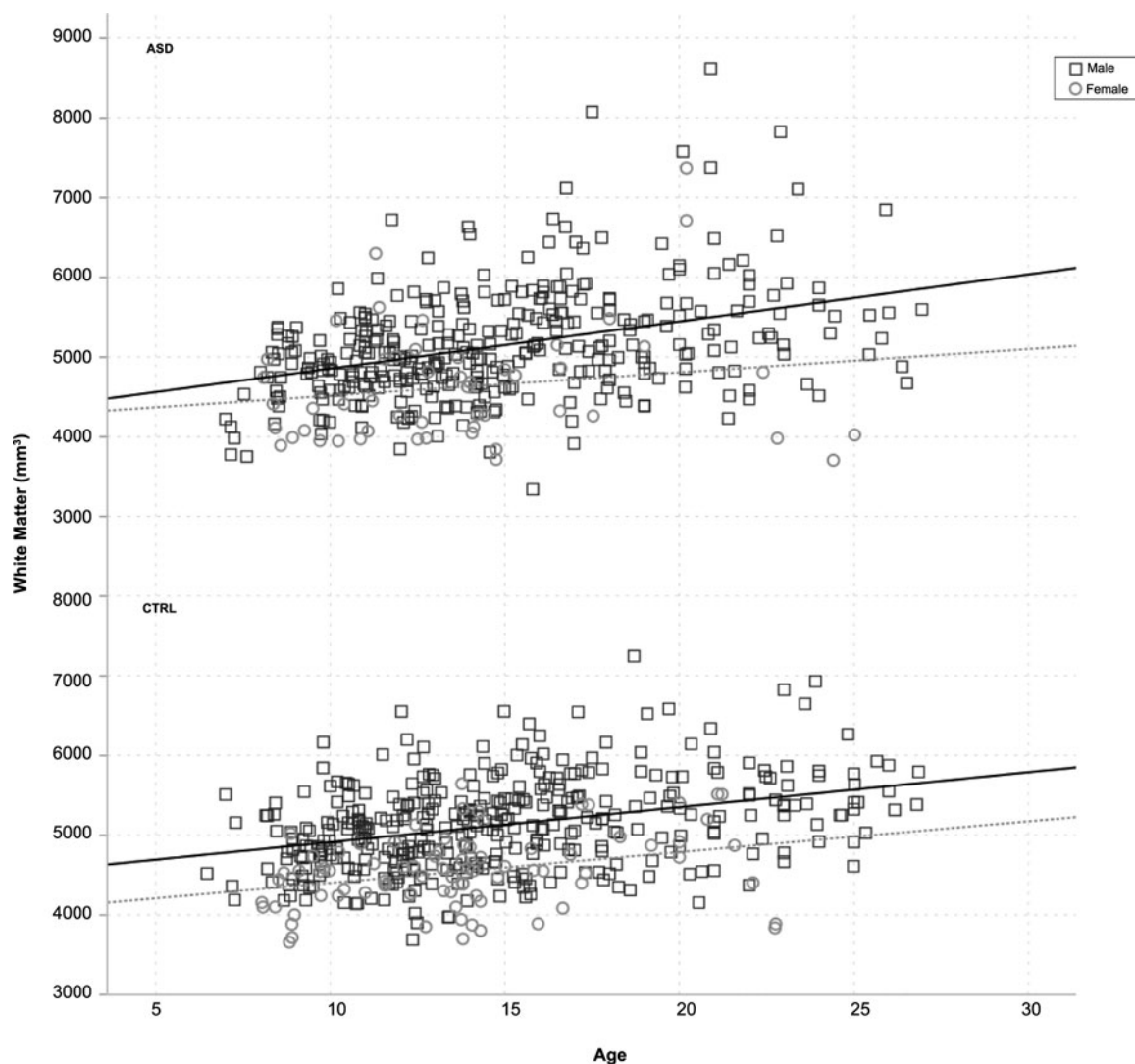


Fig. 3. Interaction between diagnosis (ASD *v.* control), age and gender in the prediction of WM volume. The WM volume is larger men with ASD than women with ASD, but not larger in control men compared with control women. The ASD males have on average more WM volume, particularly in the oldest age range (20–27 years).

Potential explanations for these age and sex moderating effects in neuroanatomy correlates of ASD may be age- and sex-specific abnormalities in gene expression. The expression of neuropathologic markers such as brain protein, or synaptic, minicolumn or neuron count in the brain may vary throughout development in ASD (Courchesne *et al.* 2011; Ramsey *et al.* 2013). Ramsey *et al.* (2013), for example, performed a multiplex immunoassay profiling analysis and found age-dependent differences in proteins involved in inflammation, growth, and hormonal signaling. The opportunity to intervene behaviorally or pharmaceutically could be age and sex specific as well: hypothetically, abnormal pruning/apoptosis or neuroinflammatory mechanisms thought to underlie neuropathological processes in ASD (Huh *et al.* 2000; Schultz & Klin,

2002; Frith, 2003; Vargas *et al.* 2005; Theoharides *et al.* 2009) may be addressed during a specific window of opportunity.

With respect to several potentially meaningful interaction effects, two prominent sex-specific findings in the current study were that adult males with ASD showed larger GM and WM volume than controls, while females with ASD did not. Since total GM and WM volume differences may well reflect the consequences of large-scale neurobiological mechanisms, such a disparity suggests that there may be different etiological mechanisms at play in males and females with ASD as proposed by the extreme male brain theory. Thus, in men, larger brains in ASD may perhaps arise from differences in sensitivity to prenatal androgens (Baron-Cohen *et al.* 2005; Lai *et al.* 2013), while

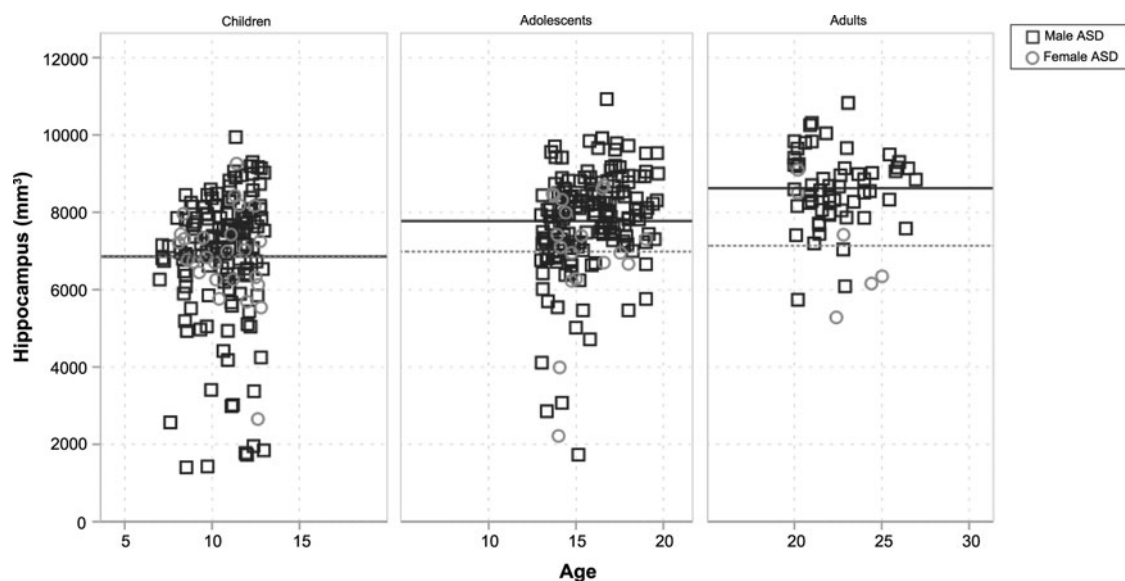


Fig. 4. Interaction between diagnosis (ASD *v.* control), age and gender in the prediction of hippocampal volume. Male adolescents and adults with ASD have larger hippocampi than females with ASD.

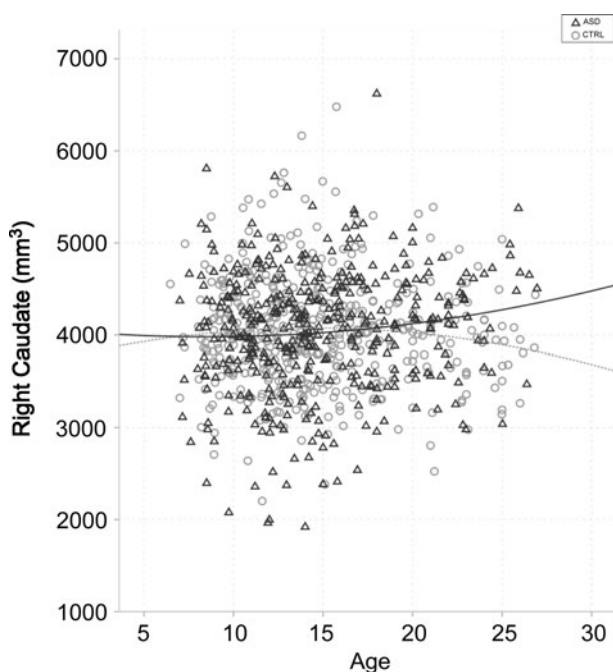


Fig. 5. Interaction between diagnosis (ASD *v.* control), age^2 and hemisphere in the prediction of caudate nucleus volume. A small main effect of age^2 was present on right caudate volume in controls, but not ASD patients.

in women with ASD, another mechanism may lead to the onset of ASD (Lai *et al.* 2015a, b).

An additional finding of interest was that the absolute hippocampus size differed between male (but not female) adolescents/adults with ASD and controls, whereas absolute amygdala, thalamus, and globus pallidus size did not. Abnormal hippocampal volume may contribute to the etiology of ASD since the

hippocampus plays an important role in emotion perception (Phillips *et al.* 2003a, b). That is, the hippocampus contributes to contextually appropriate behavior as it can regulate affective behavior elicited by emotionally salient stimuli through its inhibitory connections with the amygdala. Abnormalities of the hippocampus may thus result in impaired affective behavior and emotional regulation of ASD (Groen

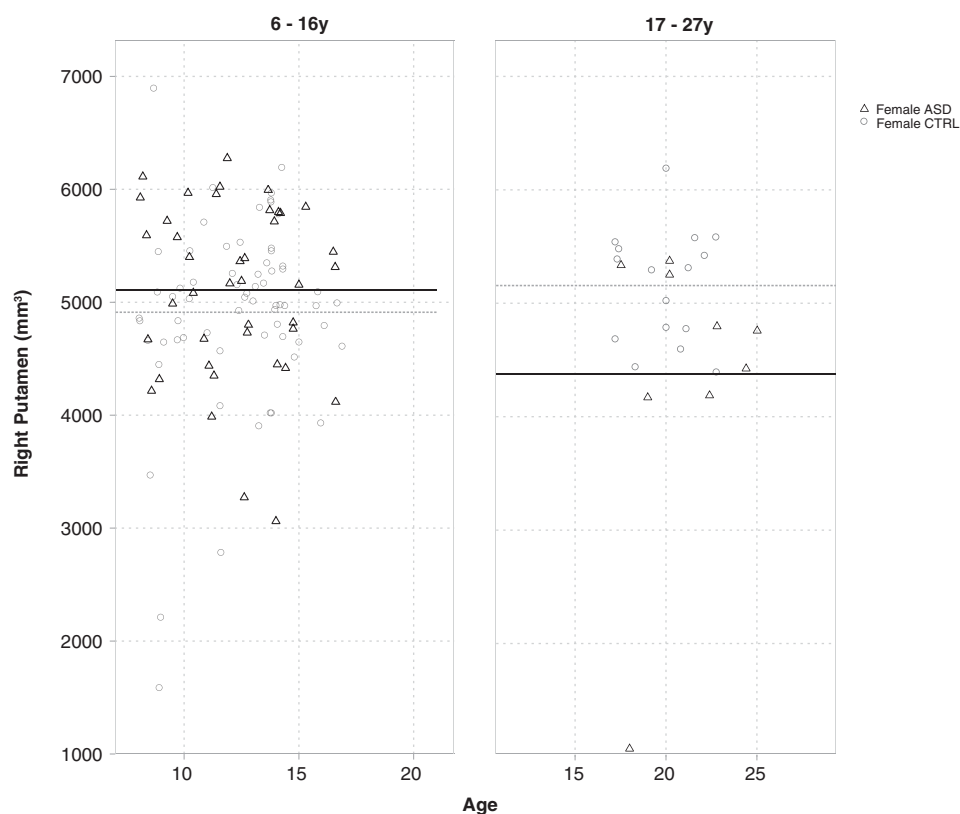


Fig. 6. Interaction between diagnosis (ASD *v.* control), age, hemisphere and gender in the prediction of putamen volume. No main effect of diagnosis was present in 6–16-year-old females, yet a moderately large effect was present in the 17–27-year olds, with ASD females having on average smaller right putamen volumes than control females.

et al. 2010). The current findings suggest that abnormal hippocampal growth may contribute to the etiology of ASD only in a specific developmental period in males but not in females, underlining age and sex specificity.

Strengths of the current study include the large sample size, uniform segmentation approach and stratified statistical analysis that corrected for center influences. Limitations are that despite the pooling of data across sites, females were underrepresented. Furthermore, no children below the age of 6 were included, limiting any examination of the developing brain before the onset of the full disorder. The current study may therefore reflect the end result or late developmental consequences of the pathology of ASD rather than the etiological changes taking place during the onset of ASD pathology. Another limitation of the current study may be the use of the automated FIRST algorithm (Patenaude *et al.* 2011) since some researchers consider the expert hand segmentation as the gold standard for subcortical segmentation (Zhang *et al.* 2001). However, FIRST results are robust, reliable and comparable with manual segmentation (Morey *et al.* 2009) and the use of this method cannot explain the absence of main effects of ASD diagnosis and presence of diagnosis by age-by-sex interaction effects.

In conclusion, despite the large sample size, we found virtually no main effects of ASD diagnosis yet several significant two- and three-way interaction effects of diagnosis by age-by-sex were found. This indicates that ‘the’ neuroanatomy of ASD does not exist. Rather, the findings underline that MRI correlates of the disorder are highly age and sex dependent. This implies that any future attempts to develop MRI diagnostic classification algorithms and MRI-based stratification of ASD into biologically more homogeneous subtypes need to take into account age, sex, and their interaction, and require substantially larger datasets than have been currently used for such attempts.

Supplementary material

The supplementary material for this article can be found at <https://doi.org/10.1017/S003329171700201X>

Acknowledgements

The study was supported by Hersenstichting Fellowship (F2010(1)-13) to WG. JB is supported by the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115300, resources of

which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007–2013) and the European Federation of Pharmaceutical Industries and Associations (EFPIA) companies.

References

- Arbabshirani MR, Plis S, Sui J, Calhoun VD (2017). Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. *Neuroimage* **145**, 137–165.
- Baron-Cohen S, Knickmeyer RC, Belmonte MK (2005). Sex differences in the brain: implications for explaining autism. *Science* **310**, 819–823.
- Blackmon K, Ben-Avi E, Wang X, Pardoe HR, Di Martino A, Halgren E, Devinsky O, Thesen T, Kuzniecky R (2015). Periventricular white matter abnormalities and restricted repetitive behaviour in autism spectrum disorder. *Neuroimage Clinical* **10**, 36–45.
- Cerliani L, Mennes M, Thomas RM, Di Martino A, Thioux M, Keysers C (2015). Increased functional connectivity between subcortical and cortical resting-state networks in autism spectrum disorder. *JAMA Psychiatry* **72**, 767–777.
- Chen CP, Keown CL, Jahedi A, Nair A, Pflieger ME, Bailey BA, Müller RA (2015a). Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism. *Neuroimage Clinical* **8**, 238–245.
- Chen H, Kelly C, Castellanos FX, He Y, Zuo XN, Reiss PT (2015b). Quantile rank maps: a new tool for understanding individual brain development. *Neuroimage* **111**, 454–463.
- Cohen J (1969). *Statistical Power Analysis for the Behavioral Sciences*. Academic Press: NY.
- Courchesne E, Campbell K, Solso S (2011). Brain growth across the life span in autism: age-specific changes in anatomical pathology. *Brain Research* **1380**, 138–145.
- Cox RW (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* **29**, 162–173.
- Di X, Biswal BB (2015). Similarly expanded bilateral temporal lobe volumes in female and male children with autism spectrum disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **1**, 178–185.
- Di Martino A, Yan C-G, Li Q, Denio E, Castellanos FX, Alaerts K, Anderson JS, Assaf M, Bookheimer SY, Dapretto M, Deen B, Delmonte S, Dinstein I, Ertl-Wagner B, Fair DA, Gallagher L, Kennedy DP, Keown CL, Keysers C, Lainhart JE, Lord C, Luna B, Menon V, Minshew NJ, Monk CS, Mueller S, Müller R-A, Nebel MB, Nigg JT, O'Hearn K, Pelphrey KA, Peltier SJ, Rudie JD, Sunaert S, Thioux M, Tyszka JM, Uddin LQ, Verhoeven JS, Wenderoth N, Wiggins JL, Mostofsky SH, Milham MP (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Journal of Molecular Psychiatry* **19**, 659–667.
- Duerden EG, Mak-Fan KM, Taylor MJ, Roberts SW (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**, 49–66.
- Ecker CC, Suckling JJ, Deoni SCS, Lombardo MVM, Bullmore ETE, Baron-Cohen SS, Catani MM, Jezzard PP, Barnes AA, Bailey AJ, Williams SC, Murphy DGM, MRC AIMS Consortium (2012). Brain anatomy and its relationship to behavior in adults with autism spectrum disorder: a multicenter magnetic resonance imaging study. *Archives of General Psychiatry* **69**, 195–209.
- Eggert LD, Sommer J, Jansen A, Kircher T, Konrad C (2012). Accuracy and reliability of automated gray matter segmentation pathways on real and simulated structural magnetic resonance images of the human brain. *PLoS ONE* **7**, e45081.
- Frith C (2003). What do imaging studies tell us about the neural basis of autism? In *Autism: Neural Basis and Treatment Possibilities: Novartis Foundation Symposium* (ed. G. Bock, J. Goode), vol. **251**, pp. 149–176. John Wiley & Sons: Chichester, UK.
- Giedd JN, Raznahan A, Mills KL, Lenroot RK (2012). Review: magnetic resonance imaging of male/female differences in human adolescent brain anatomy. *Biology of Sex Difference* **3**, 19.
- Greimel E, Nehr Korn B, Schulte-Ruther M, Fink GR, Nickl-Jockschat T, Herpertz-Dahlmann B, Konrad K, Eickhoff SB (2012). Changes in grey matter development in autism spectrum disorder. *Brain Structure and Function* **218**, 929–942.
- Groen W, Teluij M, Buitelaar J, Tendolkar I (2010). Amygdala and hippocampus enlargement during adolescence in autism. *Journal of the American Academy of Child and Adolescent Psychiatry* **49**, 552–560.
- Haar S, Berman S, Behrmann M, Dinstein I (2016). Anatomical abnormalities in autism? *Cerebral Cortex* **26**, 1440–1452.
- Hadjikhani N, Joseph RM, Synder J, Tager-Flusberg H (2006). Anatomical differences in the mirror neuron system and social cognition network in autism. *Cerebral Cortex* **16**, 1276–1282.
- Huh GS, Boulanger LM, Du H, Riquelme PA, Brotz TM, Shatz CJ (2000). Functional requirement for class I MHC in CNS development and plasticity. *Science* **290**, 2155–2159.
- Hyde KL, Samson F, Evans AC, Motttron L (2010). Neuroanatomical differences in brain areas implicated in perceptual and other core features of autism revealed by cortical thickness analysis and voxel-based morphometry. *Human Brain Mapping* **31**, 556–566.
- Jiang L, Hou XH, Yang N, Yang Z, Zuo XN (2015). Examination of local functional homogeneity in autism. *Biomed Research International* **2015**, article ID 174371.
- Kucharsky Hiess R, Alter R, Sojoudi S, Ardekani BA, Kuzniecky R, Pardoe HR (2015). Corpus callosum area and brain volume in autism spectrum disorder: quantitative analysis of structural MRI from the ABIDE database. *Journal of Autism and Developmental Disorders* **45**, 3107–3114.
- Lai MC, Baron-Cohen S, Buxbaum JD (2015a). Understanding autism in the light of sex/gender. *Molecular Autism* **6**, 24.
- Lai MC, Lombardo MV, Auyeung B, Chakrabarti B, Baron-Cohen S (2015b). Sex/gender differences and autism:

- setting the scene for future research. *Journal of the American Academy of Child and Adolescent Psychiatry* **54**, 11–24.
- Lai MC, Lombardo MV, Suckling J, Ruigrok AN, Chakrabarti B, Ecker C, Deoni SCL, Craig MC, Murphy DGM, Bullmore ET, Baron-Cohen S** (2013). Biological sex affects the neurobiology of autism. *Brain* **136**, 2799–2815.
- Marquand AF, Rezek I, Buitelaar J, Beckmann CF** (2016). Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. *Biological Psychiatry* **80**, 552–561.
- McAlonan GM, Cheung V, Cheung C, Suckling J, Lam GY, Tai KS, Yip L, Murphy DG, Chua SE** (2005). Mapping the brain in autism. A voxel-based MRI study of volumetric differences and intercorrelations in autism. *Brain* **128**, 268–276.
- Morey RA, Petty CM, Xu Y, Hayes JP, Wagner HR 2nd, Lewis DV, LaBar KS, Styner M, McCarthy G** (2009). A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage* **45**, 855–866.
- Nickl-Jockschat T, Habel U, Michel TM, Manning J, Laird AR, Fox PT, Schneider F, Eickhoff SB** (2012). Brain structure anomalies in autism spectrum disorder—a meta-analysis of VBM studies using anatomic likelihood estimation. *Human Brain Mapping* **33**, 1470–1489.
- Nielsen JA, Zielinski BA, Fletcher PT, Alexander AL, Lange N, Bigler ED, Lainhart JE, Anderson JS** (2013). Multisite functional connectivity MRI classification of autism: ABIDE results. *Frontiers in Human Neuroscience* **7**, 599.
- Patenaude B, Smith SM, Kennedy D, Jenkinson M** (2011). A Bayesian model of shape and appearance for subcortical brain. *NeuroImage* **56**, 907–922.
- Phillips ML, Drevets WC, Rauch SL, Lane R** (2003a). Neurobiology of emotion perception I: the neural basis of normal emotion perception. *Biological Psychiatry* **54**, 504–514.
- Phillips ML, Drevets WC, Rauch SL, Lane R** (2003b). Neurobiology of emotion perception II: implications for major psychiatric disorders. *Biological Psychiatry* **54**, 515–528.
- Ramsey JM, Guest PC, Broek JA, Glennon JC, Rommelse N, Franke B, Rahmoune H, Buitelaar JK, Bahn S** (2013). Identification of an agedependent biomarker signature in children and adolescents with autism spectrum disorders. *Molecular Autism* **4**, 27.
- Retico A, Giuliano A, Tancredi R, Cosenza A, Apicella F, Narzisi A, Biagi L, Tosetti M, Muratori F, Calderoni S** (2016). The effect of gender on the neuroanatomy of children with autism spectrum disorders: a support vector machine case-control study. *Molecular Autism* **7**, 5.
- Riddle K, Cascio CJ, Woodward ND** (2016). Brain structure in autism: a voxel-based morphometry analysis of the Autism Brain Imaging Database Exchange (ABIDE). *Brain Imaging and Behavior* **11**, 541–551.
- Ruigrok AN, Salimi-Khorshidi G, Lai MC, Baron-Cohen S, Lombardo MV, Tait RJ, Suckling J** (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioral Reviews* **39**, 34–50.
- Sacco R, Gabriele S, Persico AM** (2015). Head circumference and brain size in autism spectrum disorder: a systematic review and meta-analysis. *Psychiatry Research* **234**, 239–251.
- Schaer M, Kochalka J, Padmanabhan A, Supekar K, Menon V** (2015). Sex differences in cortical volume and gyrification in autism. *Molecular Autism* **6**, 42.
- Schultz R, Klin A** (2002). Genetics of childhood disorders: XLIII. Autism, part 2: neural foundations. *Journal of the American Academy of Child and Adolescent Psychiatry* **41**, 1259–1262.
- Stanfield AC, McIntosh AM, Spencer MD, Philip R, Gaur S, Lawrie SM** (2008). Towards a neuroanatomy of autism: a systematic review and meta-analysis of structural magnetic resonance imaging studies. *European Psychiatry* **23**, 289–299.
- Supekar K, Menon V** (2015). Sex differences in structural organization of motor systems and their dissociable links with repetitive/restricted behaviors in children with autism. *Molecular Autism* **6**, 50.
- Theoharides TC, Kempuraj D, Redwood L** (2009). Autism: an emerging ‘neuroimmune disorder’ in search of therapy. *Experimental Opinion on Pharmacotherapy* **10**, 2127–2143.
- Toal F, Daly EM, Page L, Deeley Q, Hallahan B, Bloemen O, Cutter WJ, Brammer MJ, Curran S, Robertson D, Murphy C, Murphy KC, Murphy DG** (2010). Clinical and anatomical heterogeneity in autistic spectrum disorder: structural MRI study. *Psychological Medicine* **40**, 1171–1181.
- Valk SL, Di Martino A, Milham MP, Bernhardt BC** (2015). Multicenter mapping of structural network alterations in autism. *Human Brain Mapping* **36**, 2364–2373.
- Van Wijngaarden-Cremers PJ, van Eeten E, Groen WB, Van Deurzen PA, Oosterling IJ, Van der Gaag RJ** (2014). Gender and age differences in the core triad of impairments in autism spectrum disorders: a systematic review and meta-analysis. *Journal of Autism and Developmental Disorders* **44**, 627–635.
- Vargas DL, Nascimbene C, Krishnan C, Zimmerman AW, Pardo CA** (2005). Neuroglial activation and neuroinflammation in the brain of patients with autism. *Annals of Neurology* **57**, 67–81.
- Venkataraman A, Duncan JS, Yang DY, Pelphrey KA** (2015). An unbiased Bayesian approach to functional connectomics implicates social-communication networks in autism. *Neuroimage Clinical* **8**, 356–366.
- 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? *Archives of General Psychiatry* **68**, 409–418.
- Waiter GD, Williams JH, Murray AD, Gilchrist A, Perrett DI, Whiten A** (2005). Structural white matter deficits in high-functioning individuals with autistic spectrum disorder: a voxel-based investigation. *Neuroimage* **24**, 455–461.
- 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. *Neuroscience and Biobehavioral Reviews* **57**, 328–349.
- Yang DY-J, Beam D, Pelphrey KA, Abdullahi S, Jou RJ** (2016a). Cortical morphological markers in children with autism: a structural magnetic resonance imaging study of

thickness, area, volume, and gyrification. *Molecular Autism* 7, 11.

Yang X, Si T, Gong Q, Qiu L, Jia Z, Zhou M, Zhao Y, Hu X, Wu M, Zhu H (2016b). Brain gray matter alterations and associated demographic profiles in adults with autism spectrum disorder: a meta-analysis of voxel-based

morphometry studies. *Australian & New Zealand Journal of Psychiatry* 50, 741–753.

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